Improving Air Transportation by Using the Fuzzy Origin–Destination Matrix

Vladimir Sudakov

Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation, 125167 Moscow, Russia; VASudakov@fa.ru

Abstract: The work is devoted to the development of new methods and algorithms to support decision making when planning air travel using uncertainties in the form of fuzzy numbers. The proposed approach makes it possible to define rational methods of choice: how to change the transport graph to better meet the needs of the population. This is particularly relevant in the context of the reduced demand for air travel caused by the pandemic and the need to switch from large to smaller aircraft types. The problem is solved by restoring the fuzzy origin–destination matrix of current statistics on air traffic between airports. The problem is that we do not know what proportion of passengers moving between the specified points are forced to use large transport hubs as intermediate destinations. To determine the validity of the origin–destination matrix, we build a number of optimization models to determine fuzzy intervals and search for correspondence with the maximum value of the membership function. Algorithmic and software search for the fuzzy origin–destination matrix and fuzzy ranking of potentially promising routes are developed. The perspective of the given approach is shown by an example of a task concerning a choice of new routes between regional airports.

Keywords: fuzzy weighted sum; fuzzy number; decision support; aircraft; origin–destination matrix

1. Introduction

Development of intelligent approaches to strategic planning of transport flows and operational decision support systems in logistics fulfils the goal of taking a leading position in the development and use of air and space, the World Ocean, the Arctic and Antarctic. It is necessary to solve multidimensional tasks of air flow planning in order to enhance the connectivity of Russian Federation regions due to long distances and a high number of small localities with poor road infrastructure. The effectiveness of present-day air transport systems is greatly enhanced by the potential of particular aircraft modifications, as well as by a reasonable approach to planning and controlling air transport complexes.

The modeling of transport flows and the evaluation of modeling results is complicated by a significant proportion of uncertain factors. These include pandemics, fundamental changes in the economy, and unpredictable political and military changes. Probabilistic estimates of these uncertainties are not always possible due to the uniqueness of situations and the lack of necessary statistical data.

The topic is important because airlines and government agencies have to make decisions on the planning of new air routes and on the assignment of certain types of aircraft to these routes. The implementation of these decisions requires a significant investment of time and money. In addition, not only do we need to offer specific recommendations to the decision maker, but also information about how much confidence we can have in estimates of the consequences of these decisions.

The research questions are:

(1) How, in the form of fuzzy numbers, we can estimate passenger flow between points in the transport network, based on information about traffic flows on separate arcs of
the graph. Since how many passengers are forced to use some arcs for their transit, we calculate fuzzy estimates of passenger flow;

(2) how, based on fuzzy estimates, decisions can be made to modify the transportation graph by adding new arcs;

(3) how to evaluate the consequences of the decisions made in the form of fuzzy multi-criteria estimates.

This paper studies:

(1) real passenger flows using the transport graph of passenger air travel in Russia for 2017–2020;

(2) various methods of reconstructing the origin–destination matrix proposed by other scientists.

The contribution is a new method for obtaining fuzzy estimates of correspondences, through formalization and solution of a number of optimization problems of linear programming.

This method offers decision support for improving the transport infrastructure, taking into account the possible degree of uncertainty of the situation, and is well scaled when implemented in the software on modern computing systems.

2. Literature Review

The primary task which it is necessary to face in the creation of transport models is the formation of an origin–destination (OD) matrix. This problem is usually solved by classical methods proposed by researchers in road transport—these are approaches based on gravity and entropy models [1]. The gravitational model attempts to find an analogy between classical physical forces of mass attraction and the attraction of people to certain areas of residence or work [2]. However, the application of this approach to air travel is questionable. To travel in a city, it is natural to assume that the attraction to easily accessible areas is greater, but at longer distances other criteria work. There is no reason to believe that the attraction of passengers to a resort city has an inverse quadratic dependence on the cost of achieving it. Another criterion for determination of correspondence is the use of the approach based on maximum entropy, which corresponds to the attempt to find an analogy with thermodynamics, when the system tries to reach the equilibrium state with maximum entropy [3]. For the “conditionally chaotic” motion of machines, such considerations may be fair, but in air transport we face greater “regularity”, due to the system of values of airlines and regulators. Nanda and Kikuchi developed a two-stage method which applied a fuzzy logic to the distribution of trips [4]. The authors believed that, if at the first stage the scheme of movement is known, at the second level it is possible to estimate the OD matrix. However, to use their model experts are needed who can specify the input and output fuzzy flow for each vertex of the transport graph. Setting fuzzy numbers by experts is a non-trivial task, and it is also problematic to ensure the expert is conscientious and competent. Other researchers tested a fuzzy logic approach to evaluate the OD matrix of air passenger flows using known production and attraction methods [5]. They found that the distribution of trips by traffic quantity scientists created a two-level output model based on fuzzy logic and entropy maximization. A general method for generating fuzzy rules from numerical data was proposed, and the same method was then applied to process the distribution of trips and modal selection [6]. However, the use of fuzzy rules also requires the involvement of experts. Rearrangement of the fuzzy rules of movements along the transport graph is certainly a simpler task for the expert than specifying fuzzy numbers in [4], but when using this approach there is still a need for a detailed survey by competent experts.

In 2008 [7], scientists proved that fuzzy logic copes well in studying discretionary travel, after various works on the distribution of trips using the fuzzy approach in 2010 [8] expanded its application to four stages of transport modeling and showed its practicality. In 2011 [9], Jassbi et al. introduced a three-phase fuzzy output model to compare social and demographic variables with OD traps. In 2012 [10], Kalic et al. reviewed fuzzy logic for
both the generation and distribution of a trip and provided evidence for the practicality of this approach. In 2016, in order to estimate the OD matrix [11], Salini et al. tried the fuzzy rule-based approach in a case study and obtained better results than those of the gravity model. Moreover, they were able to see the impact of income on travel distribution. These papers use social and economic aspects of people’s behavior, which again suggest fuzzy rules for finding fuzzy correspondences, with the authors of the papers acting as experts.

In 2013 [12], Foulds et al. considered the estimation of the OD matrix in overloaded city networks where input data are incomplete, and used an iterative linear approach. These approaches are more suitable for estimating traffic flows in urban agglomerations than for global or regional air travel.

Determining the real needs for air traffic allows us to solve the problems of scheduling and assigning crews to flights in the future. In [13], interesting formulations of aviation problems were proposed, which are solved when it is known which flights exist. In the present paper, too, an optimization problem similar to the scheduling problem is solved, whereas this is a primary problem, when there are no flights as yet and we try to determine requirements.

The importance of developing methods for multicriteria analysis of alternatives for evaluating decisions to reduce harmful emissions is well demonstrated in [14]. The definition of a fuzzy OD matrix will make it possible to obtain the required estimates, including those for evaluating the efficiency of zero-carbon transport.

The aviation industry has quite detailed statistics on air transportation between individual airports. In Russia (in the form of 14GA) and in the U.S. (in the T-100 data bank) these data are collected monthly by aviation authorities. However, it is problematic to understand how many passengers are forced to transit. For a percentage of passengers, it is possible to use boarding coupon data, but usually this is unavailable, so the sample is very limited [15]. Often an OD matrix is found through marketing surveys, but such surveys involve a fairly limited number of individuals.

Besides, classical methods for solving this problem, although they give an answer, do not allow an estimation of the degree of uncertainty of the answer.

3. Materials and Methods

Given the attached, oriented graph, for each arc there is a flow during a given unit of time: for the current applied air transport task this is the number of passengers per month between given airports, the top of the graph representing these airports. It is necessary to find an origin–destination matrix between these represented airports: how many people have moved between the given target start and target end.

Let us enter the notation: \( D = \{d\} \) set of arcs of the graph, and \( x_{ij} \) are the elements of the OD matrix to be found, i.e., a flow from point \( i \) to point \( j \) \((x_{ij} \geq 0)\). In general, this is the total flow from \( i \) to \( j \) on all routes. If \( i = j \), then \( x_{ij} = 0 \). \( V_d \) is a thread, known to us by the arc \( d \).

The solution of the problem must satisfy the system of equations:

\[
V_d = \sum_i \sum_j p^d_{ij} x_{ij} \quad \forall d \in D, \tag{1}
\]

where \( p^d_{ij} \) is the fraction of a stream from vertex \( i \) to vertex \( j \) which uses arc \( d \). Usually the fraction is determined by the economic feasibility and comfort of using arc \( d \) on the route from vertex \( i \) to vertex \( j \). In the case of air travel, expediency is the cost of air tickets and the duration of the route. For the simplest way to determine the shares, it is necessary to take the Pareto optimal routes and divide the shares between them. A more accurate solution, in case of a large number of non-dominant alternatives, is to use methods of multi-criteria analysis, including the method of fuzzy areas of preference proposed by the authors of [16].
In the current version of testing air transport models, the shortest in terms of number of route segments from the vertex \(i\) to the vertex \(j\) is taken as \(p_{ij}^d = 1\), while for the rest this is \(p_{ij}^e = 0\).

The above system from \(|D|\) equations would give an unambiguous solution if the number of equations were the same as the number of unknowns, but in practice there are usually more variables than equations. Therefore, there is no unambiguous solution.

Let \(X_{i^*j^*}\) define the flow from the vertex \(i^*\) to the vertex \(j^*\) in the form of a fuzzy number. We do not know these values, but we want to find them. Since we know the exact values of flows on each arc \(V_d\) of the current transport network, it is possible that the real value of the needs for transportation from the vertex \(i^*\) to the vertex \(j^*\) belongs to some interval \([x_{ij}^{\min}, x_{ij}^{\max}]\). \(x_{ij}^{\min}\) is the minimum possible stream from the vertex \(i^*\) to the vertex \(j^*\), \(x_{ij}^{\max}\) is the maximum possible stream from the vertex \(i^*\) to the vertex \(j^*\).

To search for a solution, it is suggested that we define the fuzzy origin–destination:

\[
X_{i^*j^*} = (x_{ij}^{\min}, x_{ij}^{\max}).
\]

The membership function will have the triangular form:

\[
\mu_{X_{i^*j^*}}(x) = \begin{cases} 
0, & x \leq x_{ij}^{\min} \\
\frac{x - x_{ij}^{\min}}{x_{ij}^{\max} - x_{ij}^{\min}}, & x_{ij}^{\min} < x \leq x_{ij}^{\max} \\
1, & x_{ij}^{\max} < x \leq x_{ij}^{\max}. 
\end{cases}
\]

Constraints (1) do not allow us to go beyond a certain interval. Some value of transportation will be the most possible. Intermediate values may not belong to a linear function, but taking into account a large degree of uncertainty, the accuracy is not so important and we can consider the membership function as linear at two intervals.

The left boundaries are found by solving linear optimization tasks (4) with limits (1) for each fixed combination of \(i^*\) and \(j^*\) values:

\[
x_{ij}^{\min} = \min_{x_{ij} : i \neq j, \{ i = 1, \ldots, n, \} \{ j = 1, \ldots, n, \}} x_{ij}. 
\]

Within the framework of one optimization problem, we need to find the minimum value for one fixed variable \(x_{ij}\) with fixed indices \(i^*j^*\); at the same time we can vary all variables \(x_{ij}\) \((i = 1, \ldots, n, j = 1, \ldots, n)\) within constraints (1). We solve such optimization problems with as many combinations of \(i^*j^*\) values as possible. For example, for three vertices \((n = 3)\), the following tasks are solved:

\[
\begin{align*}
x_{1,2}^{\min} &= \min_{\{ x_{1,2}, x_{2,3}, x_{1,3}, \}} x_{1,2} \\
x_{2,3}^{\min} &= \min_{\{ x_{2,1}, x_{3,2}, x_{2,3}, x_{3,1}, \}} x_{2,3} \\
\cdots \\
x_{3,1}^{\min} &= \min_{\{ x_{1,2}, x_{2,3}, x_{1,3}, x_{2,1}, x_{3,2}, x_{3,1}, \}} x_{3,1} 
\end{align*}
\]

\[
\begin{align*}
x_{1,2}^{\min} &= \min_{\{ x_{1,2}, x_{2,3}, x_{1,3}, \}} x_{1,2} \\
x_{2,3}^{\min} &= \min_{\{ x_{2,1}, x_{3,2}, x_{2,3}, x_{3,1}, \}} x_{2,3} \\
\cdots \\
x_{3,1}^{\min} &= \min_{\{ x_{1,2}, x_{2,3}, x_{1,3}, x_{2,1}, x_{3,2}, x_{3,1}, \}} x_{3,1} 
\end{align*}
\]
The right boundaries are found by solving linear optimization tasks (4) with the same limitations (1):
\[
\begin{align*}
x_{ij}^\text{max} &= \max_{x_{ij} : i \neq j, \ i = \frac{1}{2}} x_{ij}. \tag{6}
\end{align*}
\]

For the graph, if the number of vertices is \( n \), it is necessary to solve \( n(n - 1) \) optimization problems.

For some correspondences, it may turn out that
\[
x_{ij}^\text{min} = x_{ij}^\text{max}, \tag{7}
\]

In this case
\[
\bar{x}_{ij} = x_{ij}^\text{min} = x_{ij}^\text{max}, \tag{8}
\]

because
\[
x_{ij}^\text{min} \leq \bar{x}_{ij} \leq x_{ij}^\text{max}. \tag{9}
\]

If \( x_{ij}^\text{min} < x_{ij}^\text{max} \), it is necessary to select some value \( \bar{x}_{ij} \) from the range \([x_{ij}^\text{min}, x_{ij}^\text{max}]\) such that \( \mu(\bar{x}_{ij}) = 1 \). This choice can be conditioned by a priori information about population migration. The information can be presented in an expert form, and it is acceptable to apply approaches to coordination of the expert estimates [17]; the “experts” can also be strategic considerations which are used by the owner of the problem in the task of decision support.

Let us suppose a “generalized scenario” such as this:
\[
\bar{x}_{ij} \approx a_{ij} x_{ij}^\text{max} + \beta_{ij} x_{ij}^\text{min}. \tag{10}
\]

The sign \( \approx \) means approximate equality, because the selected \( x_{ij} \) may not be compatible in terms of the main restriction (1). \( a_{ij} \) and \( \beta_{ij} \) is the importance of shifting the highest possible estimate up (\( a_{ij} \)) or down (\( \beta_{ij} \)). It follows from condition (9) that \( a_{ij} + \beta_{ij} = 1 \). The task to improve the connectivity of the regions is to consider the scenario \( a_{ij} \to 0 \) when \( \bar{x}_{ij} \approx x_{ij}^\text{min} \) for \( j \), corresponding to a large transport hub, the load which we want to reduce, and \( a_{ij} \to 1 \) \( \bar{x}_{ij} \approx x_{ij}^\text{max} \) for the \( i \) vertex of the graph whose connectivity we want to increase. If there are no considerations, we can choose \( a_{ij} = \beta_{ij} = 1/2 \).

Let us denote the highest possible grade:
\[
\tilde{x}_{ij} = a_{ij} x_{ij}^\text{max} + \beta_{ij} x_{ij}^\text{min}. \tag{11}
\]

Final scores \( \bar{x}_{ij} \) are calculated by solving an optimization task:
\[
\min_{\bar{x}_{ij}} \sum_{i} \sum_{j} |\bar{x}_{ij} - \tilde{x}_{ij}| \tag{12}
\]
with restriction (1).

The absolute value function \(|\ldots|\) is not a linear function, but the task can be brought to a linear form, and for this purpose we will enter 2 classes of technical additional variables
\[
y_{ij}^{\hat{+}} \geq 0, y_{ij}^{\hat{-}} \geq 0. \tag{13}
\]

Let us add a restriction:
\[
\bar{x}_{ij} - \tilde{x}_{ij} = y_{ij}^{\hat{+}} - y_{ij}^{\hat{-}}, \tag{14}
\]
The target function (10) will be rewritten in the form:

\[
\min_{y_i^+, y_i^-} \sum_i \sum_j \left( y_i^+ - y_i^- \right), \tag{15}
\]

The resulting task is equivalent to the original one, and the optimization algorithm will always get either \(y_i^+\) or \(y_i^-\) equal to zero and minimize the remaining non-zero variable.

The alternative formulation is oriented to minimize, not the total, but the largest deviation from the “ideal” solution:

\[
\min_{x_i^+, x_i^-} \max_{n_j} \left\{ \max_{n_i} \left| x_i^+ - \tilde{x}_i^+ \right| \right\} \tag{16}
\]

This can also be reduced to a linear form by adding binary variables, so we can get a mixed integer linear problem. The choice of the type of target function (12) or (16) must be conditioned by the requirements of the decision maker for further application of the transport modeling results.

To estimate the passenger flow from the formal point of view, a restriction of the integer of optimization variables should be added. However, in the case of analysis of columns with hundreds and thousands of passengers on each arc, such accuracy has no practical value and the decision is simply rounded.

Software implementation of this mathematical model of fuzzy origin–destination calculation was developed in Python 3. SCIP Optimization Suite [18] was used as the optimization package. This package provides an opportunity to solve the problems of linear general algorithm for mixed integer programming. It is free for non-commercial use and comes with an open source code. The PySCIPOpt integration module with open source allows combined source data preparation using fast Python generator expressions and solution of optimization tasks in one program.

The software implementation was placed on the portal of web methods of decision support WS-DSS [19]. This portal was developed by one of the authors of the article and allows the running of software implementation of mathematical models, building sequential and parallel chains of module launch, with the transfer of parameters between them. The portal is written in Ruby to organize the call of the computational modules task manager Sidekiq, based on NoSQL. In Memory database, Redis is used, which provides high performance. In addition to modules in Python, the launch of modules in Ruby, R, C++ is implemented. With the help of Open Source Ruby on Rail framework, the ability to run models asynchronously through the RESTful API is implemented. Input and output parameters are transmitted in JSON and CSV formats. The results are saved to the PostgreSQL database.

Figure 1 shows a flowchart of the proposed method.

The general algorithm for software operation is as follows:
1. The web server WS-DSS receives an HTTP POST request with the source data for the task. In response, a unique ID of the created task is returned to the client.
2. WS-DSS forms the task for calculation in Sidekiq.
3. The Sidekiq workflow starts the calculation module in Python.
4. An array of traffic volumes \(V_d\) through all arcs \(d\) is formed.
5. With the help of the NetworkX package using Dijkstra’s algorithm, all shortest paths between all nodes are calculated.
6. To fill in \(P_d^i\) parameters, for all \(d\), whether it is included in the shortest path from node \(i\) to node \(j\) is checked. If the answer is yes, then \(P_d^i = 1/k\), where \(k\) is the total number of shortest paths from \(i\) to \(j\).
7. Formation of constraints (1).
8. Calculation \(x_{ij}^{min}\). Formation of the target function (4) solution of the obtained optimization task for all combinations of \(i^*\) and \(j^*\) values.
9. Calculation of $x_{ij}^{\text{max}}$. Formation of the target function (6) solution of the obtained optimization task for all combinations of $i^*$ and $j^*$ values.
10. Calculation of parameters $\tilde{x}_{ij'}$ by formula (11).
11. Formation of constraints (14)
12. Search $\tilde{x}_{ij'}$ by solving optimization (13) with restrictions (14) and (1).
13. Save the obtained $(x_{ij}^{\text{min}}, \tilde{x}_{ij'}, x_{ij}^{\text{max}})$ in a PostgreSQL database.
14. Returning the received solution by HTTP request from GET client with ID task. If the solution is not yet received at the time of the request, a special status of ‘waiting for the result’ will be returned.
15. If necessary, the client can make an HTTP PUT request with the modified data for recalculation. The number of tasks for an ordinary user is limited to 10.

An array of traffic volumes $V_d$ through all arcs $d$ is formed.

For all $d$ it is checked if it is included in the shortest path from node $i$ to node $j$. If the answer is yes, then $p_{ij}^d = 1/k$, where $k$ is the total number of shortest paths from $i$ to $j$.

Calculation $x_{ij'}^{\text{min}}$. Formation of the target min function solution of the obtained optimization task for all combinations of $i^*$ and $j^*$ values.

Calculation of $x_{ij'}^{\text{max}}$. Formation of the target max function solution of the obtained optimization task for all combinations of $i^*$ and $j^*$ values.

Search $\tilde{x}_{ij'}$ by solving optimization problem.

Save the obtained $(x_{ij}^{\text{min}}, \tilde{x}_{ij'}, x_{ij}^{\text{max}})$

Figure 1. Flowchart of proposed method.

4. Results
The testing of the system was carried out according to the data on air traffic in Russia between seven airports:
1. Ekaterinburg,
2. Mineralnye Vody,
3. Moscow,
4. Omsk,
5. Samara,
6. Syktyvkar,
7. Makhachkala.

City numbers $i$ correspond to the ordinal numbers in the list. July 2019 was chosen for analysis, as July is a rather busy month, and the data of 2020, due to the coronavirus pandemic, were too small for representative analysis. Figure 2 shows a graph of existing routes between cities.
2. Mineralnye Vody,  
3. Moscow,  
4. Omsk,  
5. Samara,  
6. Syktyvkar,  
7. Makhachkala.  

City numbers \(i\) correspond to the ordinal numbers in the list. July 2019 was chosen for analysis, as July is a rather busy month, and the data of 2020, due to the coronavirus pandemic, were too small for representative analysis. Figure 2 shows a graph of existing routes between cities.

As we can see from Figure 1, not all cities from this list are currently connected by direct flights. At the same time, it is known that in the 1990s there were additional direct flights, for example, between 2 and 5 and 2 and 7, which are currently missing. The information from the origin–destination matrix will allow estimation of the expediency of these flights’ revival and an understanding of what types of aircraft it is more expedient to develop.

Origin–destination between the specified cities obtained as a result of the optimization tasks is shown in Table 1.

Table 1. Fuzzy origin–destination.

| \(i^*\) | \(j^*\) | \(x_{ij}^{\min}\) | \(x_{ij}^{\max}\) | \(\bar{x}_{ij}\) |
|---|---|---|---|---|
| 1 | 2 | 1025 | 3933 | 2479 |
| 4 | 2 | 0 | 359 | 179 |
| 5 | 2 | 0 | 2105 | 1052 |
| 6 | 2 | 0 | 444 | 222 |
| 1 | 3 | 42,897 | 96,707 | 83,211 |
| 1 | 7 | 0 | 53,810 | 13,495 |
| 1 | 4 | 0 | 293 | 97 |
| 2 | 4 | 0 | 293 | 146 |
| 6 | 4 | 0 | 293 | 48 |
| 1 | 5 | 0 | 1693 | 772 |
| 2 | 5 | 0 | 1693 | 846 |
| 6 | 5 | 0 | 444 | 74 |
| 1 | 6 | 0 | 439 | 86 |
| 2 | 6 | 0 | 439 | 219 |
| 4 | 6 | 0 | 359 | 59 |
| 5 | 6 | 0 | 439 | 73 |
| 7 | 3 | 0 | 51,188 | 8531 |
| 7 | 1 | 0 | 51,188 | 12,724 |
| 7 | 2 | 0 | 51,188 | 8531 |
| 7 | 4 | 0 | 40,848 | 6808 |
| 7 | 5 | 0 | 51,188 | 8531 |
| 7 | 6 | 0 | 12,123 | 6061 |
| 2 | 1 | 0 | 4073 | 1648 |
| 2 | 3 | 0 | 71,620 | 62,651 |
| 2 | 7 | 0 | 53,810 | 8968 |
As can be seen from Table 1, missing routes 2 and 5 and 2 and 7 have some potential, but the degree of uncertainty is quite high and should be taken into account when planning management decisions based on the results obtained.

On solving $X_{i^*j^*}$ search tasks for different time periods, we can build forecasts based on fuzzy regression models and to carry out “fuzzy” smoothing of origin–destination fluctuations in search of rational models of transport and logistic systems development.

The fuzzy origin–destination matrix $X_{i^*j^*} = \langle x_{i^*j^*}^{\min}, x_{i^*j^*}, x_{i^*j^*}^{\max} \rangle$ obtained in this paper allows the solution of the following problems:

1. optimal route planning,
2. rational scheduling,
3. assessment of aircraft import substitution prospects,
4. analysis of possibilities to replace outdated aircraft equipment,
5. support for decision making on modernization of existing transport infrastructure.

Let us consider a problem in deciding which of the new routes is more appropriate to add to the existing graph.

The final score is calculated using the standard fuzzy weighted sum formula

$$Y^d = \sum_{i=1}^{n} W_i X_{i^*j^*}^d,$$

where $X_{i^*j^*}^d$ is the fuzzy evaluation of the new route $d$ by criterion $I$, $W_i$ is the fuzzy weight of criterion $i$, and $Y^d$ is the final evaluation of the new route $d$.

The rules of summation and the work of fuzzy numbers are executed on the basis of the principle of communication [20]. Here is the membership function and corresponding operation:

$$\mu(y^*) = \sup_{y_1, y_2, \ldots, y_n} \left( \Theta \mu_i(y_i) \right),$$

where $\eta$ is the operation to be applied (in this case $W_i X_{i^*j^*}$ is the multiplication and for the calculation $\sum_{i=1}^{n} W_i X_{i^*j^*}$ is the sum), $y_i$ are the values to which the required operation is applied, $\mu_i(y_i)$ is the membership function for fuzzy values, $\mu(y^*)$ is the membership function for the result of the operation $\eta$, $\Theta$—is an intersection operation for the membership functions. In this work this is min, but there are other types of this operation [21].
Let us set the triangular membership functions to the fuzzy preference of the $P_c$ development of direct air traffic with the corresponding city:

- $P_1 = <0.10, 0.50, 0.55>$,
- $P_2 = <0.30, 0.70, 0.10>$,
- $P_4 = <0.50, 0.90, 1.00>$,
- $P_5 = <0.10, 0.60, 0.70>$,
- $P_6 = <0.10, 0.40, 0.70>$,
- $P_7 = <0.00, 0.20, 0.40>$.

For the third city there is no specified preference, because it already has links with all cities. The preferences of the cities will be determined on the basis of the data on the preferred strategic development of the airport base region, using the methods of paired comparisons [17] and fuzzy preference areas [16].

For each new route, we use three fuzzy criteria: the importance of the departure city, the importance of the arrival city and the OD trip intensity on the given arc.

Let us set the weight of importance of the departure/arrival city with the triangular membership function:

- $W_1 = W_2 = <0.2, 0.25, 0.5>$.
- The OD trip intensity weight on a given arc is set in the same way: $W_3 = <0.3, 0.5, 1>$.

The final arc assessment $d = (i, j)$ is calculated by the formula:

$$Y^d = W_1 P_i + W_2 P_j + W_2 X_{ij}, \quad (19)$$

The resulting final scores are shown in Figure 3. The question arises: how to construct a route ranking? How to determine which route is better? Instead of defuzzification, which is used in many papers, this paper proposes a move towards pairwise fuzzy rank comparisons.

Let us consider two routes: $D_i$ and $D_j$. We can define a fuzzy binary relationship between them: $D_i \geq D_j$. This indicates that the object $D_i$ is no worse than the object $D_j$. 

![Figure 3. Final route scores.](image-url)
The degree of certainty that the two objects are in relation to $\geq \cap$ is given by the number $r_{ij} \in [0, 1]$. Thus, we can calculate a matrix of a fuzzy binary relation:

$$R = \begin{bmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1m} \\ r_{21} & 1 & r_{23} & \cdots & r_{2m} \\ r_{31} & r_{32} & 1 & \cdots & r_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} & 1 \end{bmatrix}. \quad (20)$$

The elements of this matrix are found by analyzing the membership functions corresponding to $D_i$ and $D_j$ objects:

$$r_{ij} = \sup_{x_i, x_j : x_i \geq x_j} \left[ \min(\mu_i(x_i), \mu_j(x_j)) \right]. \quad (21)$$

In general, this matrix is neither symmetric nor inversely symmetric. The final rankings of objects can be found using the Saaty paired comparison method [22]. The Saaty method assumes the use of a multiplicative matrix: $r_{ij}$ shows how many times one object is preferable to another, i.e., $r_{ij} r_{ji} = 1$, but in our case it is not. By performing simple arithmetical transformations, we can achieve the required condition:

$$R^* = \|r^*_ij = \frac{r_{ij}}{r_{ji}}\|.$$ \quad (22)

Further, from the equation:

$$(\lambda I - R^*)u = 0,$$ \quad (23)

we will find the maximum eigenvalue of $\lambda_{max}^*$ and its corresponding eigenvalue $u^*$. The element $u_{i}^*$ will be the final route rank $D_i$.

For this example, we have obtained the route preferences shown in Table 2.

Table 2. Results of the solution.

| Route | Route Preference |
|-------|------------------|
| 1–7   | 0.210            |
| 2–7   | 0.192            |
| 4–7   | 0.184            |
| 5–7   | 0.179            |
| 6–7   | 0.124            |
| 2–5   | 0.049            |
| 2–6   | 0.018            |
| 2–4   | 0.016            |
| 4–6   | 0.014            |
| 5–6   | 0.014            |

From the example we can see that it is reasonable to develop a transport hub in city 7. Route 1–7 is the most preferable, but from Figure 2 we can see that the degree of uncertainty in such a decision is quite high.

5. Discussion

The example has shown that the method considered in the previous section can be successfully applied to determine the correspondence between a given set of cities. The initial traffic flow along the graph edges, taken from the real statistical data, does not take into account the fact that the true initial and final vertices in the graph of some passengers may be quite different. The proposed new method does not face the problems inherent in
many other works in search of fuzzy correspondence in a graph, or the need for fuzzy expert evaluations and fuzzy rules [9–12]. However, in the future expert evaluations can be used in combination with this method for more accurate prediction of fuzzy correspondences in a graph.

A large number of works on the subject of OD matrix search show that this problem is actual both for aviation and automobile transport. The considered example of adding new transport routes as new arcs to the existing graph allows not only suggested rational variants of solutions, but also an estimation of the fuzziness of the obtained estimations.

This result is useful for developers of transport operation research models. The obtained fuzzy estimates can be considered as parameters of optimal scheduling models and can also be used in solving the problem of aircraft assignment.

6. Conclusions

This paper solves the problem of finding the fuzzy OD matrix. The method is suitable for cases when we cannot establish the real points of departure and destination of specific passengers, but we know the number of passengers transported on specific arcs of the transport graph. This is a common situation in which strategic planning tasks have to be solved not at the level of individual airlines, but at the level of the state. The defined mathematical programming task, although it formally works with discrete optimization variables, in real business cases does not require precision at the level of one passenger, so the standard solvers of linear problems, which are quite capable of dealing with high-dimensional graphs, are used.

In the conditions of availability of trustworthy aviation statistics on transportation and the absence of an OD matrix, designers of transport-logistical systems can solve the problems of planning transportation routes, choosing the airplanes, and designing new airports using this model as a source of information on potential transportation. Scientists in the field of transport operations research can verify expert judgments and rules, in order to check their proximity to possible values of OD matrix elements.

The proposed search algorithm of the fuzzy OD matrix allows the planning of strategic development of transport infrastructure. Using methods of support for decision-making in the fuzzy information environment, it is possible to define both preferences for development of new routes for air transportation, and degree of confidence in the given decision.

The main advantage of the method, compared to analogues, is that it can find the fuzzy OD matrix even when there are no expert judgments. Another advantage is the use of a model in the form of linear target functions and constraints, which are quite easy to calculate. Initially, a number of independent mathematical programming problems are solved and therefore a part of the time-consuming computational process can be easily organized in parallel threads.

Software implementations for fuzzy OD matrix search and fuzzy ranking are placed on the portal of web services of decision support WS-DSS. These are available for use by a wide range of application developers of transport and logistics systems.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data publicly available from the Federal Air Transport Agency.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Profillidis, V.A.; Botzoris, G.N. Modeling of Transport Demand—Analyzing, Calculating and Forecasting Transport; Elsevier: Amsterdam, The Netherlands, 2018. [CrossRef]

2. Ekowicaksono, I.; Bukhari, F.; Aman, A. Estimating Origin-Destination Matrix of Bogor City Using Gravity Model. IOP Conf. Ser. Earth Environ. Sci. 2016, 31, 012021. [CrossRef]
3. Wilson, A.G. A statistical theory of spatial distribution models. *Transp. Res.* **1967**, *1*, 253–270. [CrossRef]
4. Nanda, R.; Kikuchi, S. Estimation of trip O–D matrix when input and output are fuzzy. In Proceedings of the ISUMA 93, the Second International Symposium on Uncertainty Modeling and Analysis, College Park, MD, USA, 25–28 April 1993; Ayyub, B.M., Ed.; IEEE Computer Press: Piscataway, NJ, USA, 1993; pp. 104–111. [CrossRef]
5. Kalic, M.; Teodorovic, D. Trip distribution modelling using fuzzy logic and a genetic algorithm. *Transp. Plan. Technol.* **2003**, *26*, 213–238. [CrossRef]
6. Vythoulkas, P.C.; Koutsopoulos, H.N. Modeling discrete choice behavior using concepts from fuzzy set theory, approximate reasoning and neural networks. *Transp. Res. Part C Emerg. Technol.* **2003**, *11*, 51–73. [CrossRef]
7. Shafahi, Y. Fuzzy Trip Distribution Models for Discretionary Trips. In Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems, Beijing, China, 30 December 2008; pp. 557–562. [CrossRef]
8. Taylor, P.; Yaldi, G.; Yue, W.L. Examining the possibility of fuzzy set theory application in travel demand modelling. *J. East. Asia Soc. Transp. Stud.* **2010**, *8*, 579–592. [CrossRef]
9. Jassbi, J.; Makvandi, P.; Ataei, M.; Sousa, P.A. Soft system modeling in transportation planning: Modeling trip flows based on the fuzzy inference system approach. *Afr. J. Bus. Manag.* **2011**, *5*, 505–514. [CrossRef]
10. Kalić, M.; Kuljanin, J.; Dožić, S. Air Travel Demand Fuzzy Modelling: Trip Generation and Trip Distribution. In *Soft Computing in Industrial Applications*; Advances in Intelligent Systems and Computing; Snášel, V., Krömer, P., Köppen, M., Schaefer, G., Eds.; Springer: Cham, Switzerland, 2014. [CrossRef]
11. Salini, P.S.; Katti, B.K.; Paul, P.; Krishna, S.J. Developing a Fuzzy Rule Based Urban Trip Distribution Model on Income Criteria Basis at Zonal Level: A Case Study. *IOSR J. Mech. Civ. Eng. (IOSR-JMCE)* **2016**, *1*, 39–47.
12. Foulds, L.R.; Do Nascimento, H.A.; Calixto, I.C.; Hall, B.R.; Longo, H. A Fuzzy Set-Based Approach to Origin–Destination Matrix Estimation in Urban Traffic Networks with Imprecise Data. *Eur. J. Oper. Res.* **2013**, *231*, 190–201. [CrossRef]
13. Demirel, N.Ç.; Deveci, M. Novel search space updating heuristics-based genetic algorithm for optimizing medium-scale airline crew pairing problems. *Int. J. Comput. Intell. Syst.* **2017**, *10*, 1082–1101. [CrossRef]
14. Pamucar, D.; Deveci, M.; Canitez, F.; Paksoy, T.; Lukovac, V. A Novel Methodology for Prioritizing Zero-Carbon Measures for Sustainable Transport. *Sustain. Prod. Consum.* **2021**, *27*, 1093–1112. [CrossRef]
15. Park, Y.; O’Kelly, M.E. Origin–destination synthesis for aviation network data: Examining hub operations in the domestic and international US markets. *J. Adv. Transp.* **2016**, *50*, 2288–2305. [CrossRef]
16. Dutov, A.V.; Nesterov, V.A.; Sudakov, V.A.; Sypalo, K.I. Fuzzy Preference Domains and Their Use for Selecting an Electronic Flight Bag for Flight Crews. *J. Comput. Sys. Sci. Int.* **2018**, *57*, 230–238. [CrossRef]
17. Sudakov, V.; Sivakova, T. Methodology for processing expert judgments based on fuzzy weighted summation. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *927*, 012041. [CrossRef]
18. Vigerske, S.; Gleixner, A. SCIP: Global Optimization of Mixed-Integer Nonlinear Programs in a Branch-and-Cut Framework. *Optim. Methods Softw.* **2018**, *33*, 563–593. [CrossRef]
19. Sudakov, V.; Nesterov, V.; Kurennykh, A. Integration of Decision Support Systems “Kosmos” and WS-DSS with Computer Models. In Proceedings of the Tenth International Conference Management of Large-Scale System Development (MLSD), Moscow, Russia, 2–4 October 2017; pp. 1–4. [CrossRef]
20. Seresht, N.G.; Fayek, A.R. Computational method for fuzzy arithmetic operations on triangular fuzzy numbers by extension principle. *Int. J. Approx. Reason.* **2019**, *106*, 172–193. [CrossRef]
21. Abdullateef, L. On algebraic properties of fundamental group of intuitionistic fuzzy topological spaces (IFTSs). *Open J. Math. Sci.* **2020**, *4*, 34–47. [CrossRef]
22. Sudakov, V.; Kurennykh, A. Cycle-Breaking Approach to Reduce Inconsistency in Pairwise Comparisons. In *Intelligent Algorithms in Software Engineering, CSOC 2020*; Advances in Intelligent Systems and Computing; Silhavy, R., Ed.; Springer: Cham, Switzerland, 2020; pp. 395–400. [CrossRef]