A problem of airport capacity definition

Kristina V. MARINTSEVA, Firdovci ACHUN-OGLY SHARIFOV, Gennadiy N. YUN

Air Transportation Management Department, National Aviation University, Ukraine

Abstract

The problem of determining the capacity of airports is formulated as a special integer programming problem, for solution of which an effective algorithm is proposed. The initial data of the problem of passenger traffic between airports is offered to be solved by estimating the elements of the Origin-Destination Matrix.

Introduction

Air Transport System (ATS) as a communication system should function with account of all the factors that stimulate the growth and efficiency of air transportation. The growth of air traffic and, accordingly, the efficient performance of all types of services in the structure of ATS cause a number of essential organizational and technical changes in the major subsystems of the ATS, mainly, maintaining the same structure of route network and organization of passenger services at airports could initiate considerable difficulties. The main purpose of these changes is to improve the service at airports, to expand the route network geographically and to increase flight frequency. In the process of planning these changes, the problem, the solution of which would determine the optimal flight frequency, the appropriate structure of route network, and finally, planning list of different actions for passenger service and aircraft maintenance. In this case, the total cost of works, related to organizational and technical transformations result from their summing up.

Neglecting all possible interpretations of the concepts of "airport capacity" and "capacity of air route" (Harchenko and Naumenko, 2009; Hirst, 2008; Kazda and Caves, 2007; Stamatopoulos et .al.; 2004 Wells, 2004; Zaporozhets and Konovalova, 2005), in this article, these concepts are supposedly understood, respectively, as the total frequency of flights at the airport and the frequency of non-stop flights between two airports. To present airport capacity in terms of flight frequency it is also necessary to determine the size of various technical elements of the airport according to the Development Reference Manual, developed by International Air Transport Association (International Air Transport Association, 2004).

In Janić (2000), different models of determining the airport capacity are presented, which consider the size of the ground and air zones corresponding to a given period of time and the current technical parameters of the airport. The use of these models faces significant challenges, because they need specific data for their realization, as it is mentioned in the same paper.

Air routes are the basic subsystem of ATS. Thus, in determining the optimal airports capacity, all potentially possible flights should be taken into account. In some cases, we can consider the problem of determining the capacity of airports relative to the corresponding structure of the route network and the forecasted passenger traffic between the airports. Route network should be understood as a network, consisting of many airports as its nodes and of non-stop routes as its edges.

In section 1, an informal setting and a model of the general problem in terms of current and forecasted passenger traffic, a number of restrictions that is expressed as an exponent of the number of airports in route network, is formulated. Through assuming a hierarchy on a set of airports, the formulated model can be represented as a special integer programming problem. Ways of information accumulation, enabling the identification of the task are set out in section 2. In section 3, the specific task properties are studied, and a simple example illustrates the algorithm of its solution.

1. Statement of the problem and its mathematical model

Firstly, we will introduce the necessary notation. Then we will develop a mathematical model, present the methods of determining the values of the initial parameters of the problem, as well as some features of the necessary information accumulation.

V is the given set of airports, and E *is* the set of existing and potential direct routes between airports of V. Further, abstracting the "aviation" contents of the sets V and E, sometimes we will operate with sets of nodes and edges of network $G = (V, E)$ respectively. We can assume that the network G is represented by a complete graph, since non-existent direct routes between different airports pairs can be considered as potential routes. On subsets S of the set V we define a positive function $f(S)$, the values of which are the predicted or expected value of passenger traffic between airports S and $V \setminus S$. From the definition of this function it follows that $f(S) = f(V \setminus S)$, so it is symmetric. The symmetry of the function $f(S)$ means that the total number of arriving and departing passengers to and from the sets of airports S coincides with the similar amount of the passenger traffic for the set of airports $V \setminus S$. In $G = (V, E)$ the set of edges that connect the nodes from S with the nodes from $V \setminus S$, is called the cut, which is defined by non-empty subset S of V ($\emptyset \neq S \subset V$), and is denoted as $\delta(S)$. If the subset $S = \{v\}$, then for simplicity $\delta(\{v\})$ will be written down as $\delta(v)$.

The terms "passenger traffic" and "flight frequency" will be used interchangeably as synonyms in the sense that if the volume of passenger traffic and the aircraft type are known, the flight frequency could be easily determined, and vice versa. This allows consideration of the parameters of the problems in terms of passenger traffic. Therefore, while solving practical problems, it is necessary to redefine passenger traffic in terms of flight frequency based on the types of aircraft for each flight.

For some nodes (airports) $v, w \in V$ specify $l_{vw} \geq 0$, as needed (current) minimum flight frequency between airports *v* and *w*. If there is no necessity to specify the minimum frequency of service, then we can assume that $l_{vw} = 0$. Thus, without the loss of generality, we can assume that for any airport pair *v* and *w* the values $l_{vw} \geq 0$ are set.

If x_v is the unknown capacity of the airport *v*, then $c_v x_v$ is the total cost of its reconstruction, where c_v is the sum of costs per unit at all stages of reconstruction. In network *G = (V, E)* the set of possible direct flights (edges), which could be made through the airport *v*, form a cut $\delta(v) \subset E$. Therefore, if x_e is the unknown capacity of edges $(v, w) = e \in \delta(v)$ and x_v is the flight frequency, serviced by the airport *v* (airport capacity *v*), then

$$x_v = \sum_{e \in \delta(v)} x_e$$

for each airport $v \in V$. Thus, the capacity of the airport *v* is determined as the sum of capacities of edges, which belong to δ *(v).* This definition can be written briefly as $x_v = x(\delta(v))$. Then for the airport $v \in V$ we obtain

$$\sum_{v \in V} c_v x_v = \sum_{v \in V} c_v \sum_{e \in \delta(v)} x_e = \sum_{e \in E} c_e x_e,$$

where $c_e = c_v + c_w$ for all edges $(v, w) = e \in E$.

The mathematical model of general problem of the optimal airport capacity definition could be formulated as follows:

$$\min \sum_{e \in E} c_e x_e \tag{1}$$

subject to

$$\sum_{e \in \delta(S)} x_e \geq f(S) \text{for all } \emptyset \neq S \subset V, \tag{2}$$

$$x_e \geq l_e \geq 0 \text{ for all } e = (v, w) \in E, \tag{3}$$

$$x_e \text{ is integer for all } e \in E. \tag{4}$$

In section 3 we consider the simple but important properties of problem (1) - (4). In particular, we will show, that considering the dependence $x_v = x(\delta(v))$ and the assumption of hierarchy among the airports of the general problem of determining the optimal capacity of airports, the problem (1) - (4) can be reduced to the special problem of integer programming, resulting in integer solutions that determined the optimal values of the variable x_e and x_v for all edges $e = (v, w) \in E$ and all airports $v \in V$, respectively. If the edge $e = (v, w)$ refers to the existing direct flight between the airports $v, w \in V$, then it is necessary to increase the flight frequency in airports v and w to the value $\sum_{e \in \delta(v)} x_e$ and $\sum_{e \in \delta(w)} x_e$, respectively. If the edge $e = (v, w)$ refers to the potential direct flight between airports $v, w \in V$, then the new flight with the flight frequency $x_v = x_e$ and $x_w = x_e$, is created. Hence, the optimal capacity of airports (frequency) of route network, which could include some new flights between airports from the set V, is defined as a result of solution of this problem.

2. Defining parameter values of the model (1)-(4)

2.1. Evaluation of vector l components

To determine the traffic flow between any airport pair one of the known algorithms for estimating the elements of the Origin-Destination Matrix (ODM) between the points of origin and destination, which is widely used in the planning and management at transport systems, could be applied (Bussieck et. al., 1997).

In practice, estimation (finding) of ODM elements is an important solution stage of the transport network design problem (Bussieck et. al., 1997). The columns of ODM correspond to the routes, which connect points of departure and destination of passengers, rows – to the edges from the E (existing and potential non-stop flights). In a network $G = (V, E)$, a set of terminal nodes is considered as points of departure and destination of passengers, while routes between these points are represented as sequences of edges from E. To calculate the values of the vector $l = (l_e; e \in E)$ components, one of the algorithms (Bussieck et. al., 1997; Barbour and Fricker, 1994; Sherali et. al., 1994) can be applied, with the help of which the elements of ODM are evaluated. These elements represent the volume of passenger traffic between the terminal nodes over a certain period of time (day, week and etc.). Firstly, through the way of q observation the volume of t_e^q traffic is defined on each edge $e \in E$ for the evaluation of the ODM elements. Then, the vector $t^q = (t_e^q; e \in E)$ is used for static estimation of the ODM elements. For example, in Barbour and Fricker (1994) to estimate the ODM elements greedy algorithm is applied (Cook et. al., 1998). With its help the shortest path in a network *G = (V, E)* is determined. It is offered to pass the maximum flow by this path from one terminal node to another, so that the total flow to the edge *e* would not exceed t_e^q. For the estimation of the ODM elements, the corresponding linear programming problem with the condition of charging a penalty if the flow is not directed along the shortest path is formulated in Sherali et. al. (1994).

Beside these algorithms, methods of minimizing some of the nonlinear functions are used for estimating of the ODM elements. Let the routes, which are connecting certain pairs of terminal nodes, be fixed on the network *G= (V, E*). Each route *L* in the network *G= (V,E*) can be associated with $|E|$-dimensional vector $\pi(L)$, where $\pi(L_e)$ = 1, if the edge *e* is an edge of the route, and $\pi(L_e)$ = 0, if *e* is not an edge of the route *L*. It is obvious that the number of components of the vector $\pi(L)$ equal to $|E|$, and this vector is usually called the characteristic vector of the route *L*.

Considering the matrix *Q*, the columns are the characteristic vectors of the specified routes in the network *G = (V, E)*, and the rows correspond to the edges of the set *E*. Let χ - a vector, each component of which is an unknown amount of traffic on each route in the network *G = (V, E)*. Then vector χ is the solution to the linear equations $Q\chi = t^q$ system for the observed values of the vector t^q. Since the number of columns in the matrix is much greater than the number of its rows, the system $Q\chi = t^q$ has many solutions. Therefore, depending on the nature of the required flows, various objective functions, for example, $(Q\chi = t^q)^2$ are considered (Bussieck et. al., 1997). By minimizing the latter we will arrive at an appropriate solution (vector χ) to the system. Thus, according to the *q* observations the value of traffic on each edge (elements of ODM) of any route connecting the various nodes (airports), is estimated. Once the values of the vector components χ for each edge *e* ∈*E*, the value l_e is calculated as the dot product of two vectors Q_e and χ, so $l_e = Q_e\chi$, where Q_e – row of the Qmatrix, corresponding to the edge e.

2.2. Calculation of f(S) values

How is it possible to determine the value $f(S)$of an arbitrary subset *S* of the set *V*, if you know $f(\{v,w\})$for all pairs of nodes $v, w \in V$, $v \neq w$? In other words, how the projected or anticipated flows of passengers on airlines, adjacent to an arbitrary subset *S* of airports are determined? To determine the value $f(S)$ at an arbitrary subset of *S (Ø≠ S ⊂V)* it is sufficient to find the predicted values $f(\{v,w\})$for passenger traffic between all airport pairs $v, w \in V$, where $v \neq w$. Then the calculation $f(S)$ for required subsets $S \subset V$ can be carried out as follows:

Let $G_0 = (V_0, E_0)$ is the network with many nodes $V_0 = V$ and edges

$$E_0 = \{(v, w):\ f(\{v, w\}) > 0, \qquad v, w \in V\},$$

where $f(\{v, w\}) > 0$ is capacity of edge $(v, w) \in E_0$. In this network, the value of function $f(S)$ of the subset S is defined as the power or capacity of the cut, which is separating a subset S from $V \setminus S$. This definition of $f(S)$ suggests that function is symmetrical. In addition, the calculations of $f(S)$ are done using the program method for arbitrary $S \subset V$. In other words, to define the function $f(S)$ should be calculated for only required subset $S \subset V$.

3. The problem of determining airport capacity

Despite the fact that in the problem (1) - (4) the number of constraints is expressed as an exponent of a number of nodes of network *G = (V, E)* under the assumption of a hierarchy of the airports, their optimal capacities and the corresponding network structure of non-stop flights can be effectively determined by solutions of special integer programming problem. First, we recall the necessary definitions from the graph theory. A graph, which links all of the nodes of the set *V* without cycles, is called the spanning tree of network *G = (V, E).* Arbitrary spanning trees consist of $|V|$-1 edge (Cook et. al., 1998). Let $T = (V, E(T))$ is a certain spanning tree of the network $G = (V, E)$. After removal of any edge of the tree *T = (V, E (T))* it is divided into two sub trees. On the network *G = (V, E)* the cut, separating nodes of the sub tree T_1 from nodes of T_2, corresponds to the edge $e \in E(T)$. The set of cuts on the network *G = (V, E),* corresponding to the edges of the tree *T = (V, E (T)),* is denoted as $R(T)$. Since the tree *T = (V, E (T))* has $|V|$-1 attaching edges, $R(T)$ comprises $|V|$-1 various cuts. From the definition $x_v = x(\delta(v)$ it follows that a change in the capacity of edges *(v, w)* is equivalent to a change in airports *v* and *w* capacity. In order to determine the values of variables x_e for all $e \in E$ does not necessarily include all cuts of the network *G = (V, E)* in the constraint (2). It is enough to include the cuts $\delta(S)$ into the constraint (2) of problem (1) - (4), corresponding only to the edges of the tree *T = (V, E (T)).* Therefore, the problem of determining airport capacity can be formulated as a special case of problem the (1) - (4):

$$\min \sum_{e \in E} c_e z_e \tag{5}$$

subject to

$$\sum_{e \in \delta(S)} z_e = f(S) \text{ for all } \delta(S) \in R(T), \tag{6}$$

$$z_e \geq l_e \geq 0 \text{ for all } e=(v,w) \in E, \tag{7}$$

$$z_e \text{ is integer for all } e \in E. \tag{8}$$

It is clear, that (5)–(7) is a linear relaxation of this problem. After replacement $y_e = z_e - l_e$ the problem (5)-(8) can be written as:

$$\min \sum_{e \in E} c_e y_e \tag{9}$$

subject to

$$\sum_{e \in \delta(S)} y_e = f(S) - \sum_{e \in \delta(S)} l_e \quad \text{for all } \delta(S) \in R(T), \tag{10}$$

$$y_e \geq 0 \text{ for all } e=(v,w) \in E, \tag{11}$$

$$y_e \text{ is integer for all } e \in E. \tag{12}$$

Let us show that by solving the problem (9) - (11), which is a linear relaxation (9) - (12), we can find the optimal integer values of variables y_e which are the solution to the problem (9) - (12). For this purpose, we consider the matrix *A*, the rows of which are the characteristic vectors of the cuts *R(T)*, and let *b* is a vector with components $b_t = f(S) - \sum_{e \in \delta(S)} l_e$ for all cuts $\delta(S) \in R(T)$, corresponding to the edges *t* of the tree *T = (V, E (T))*. Thus, we conclude that the matrix *A* has $|E|$ columns and $|V|$-1 rows. According to this notation, we represent the problem (9) - (11) (without the constraint (12)) as follows:

$$\min \sum_{e \in E} c_e y_e \tag{13}$$

subject to

$$A\, y = b \tag{14}$$

$$y_e \geq 0 \text{ for all } e \in E, \tag{15}$$

where $y = (y_e; e \in E)$ is a vector. It is clear, that on an arbitrary edge the projected passenger traffic cannot be less than its current value. Therefore, $b_t \geq 0$ for all $t \in E(T)$. Let us consider the columns of the matrix *A*, that correspond to the edges of the skeleton tree *T = (V, E (T)).* Since each edge *t* is contained in only one of the cuts $\delta(S) \in R(T)$ rows, corresponding to the edges of cuts of the spanning tree *T = (V, E (T)),* form the unity sub matrix of matrix *A*. Therefore variables y_t , corresponding to the edges $t \in E(T)$, are the basic variables for the problem (13) - (15). Since the matrix *A* has full rank, the number of basic variables is equal to $|V|$-1. After that, having modified *(see Trubin* V.A.*, 1969*) some of the steps of the simplex method, an integer solution to the problem (13) - (15) can be found (see below for an example of solution) with its help.

In network *G = (V, E)* there is a large number of different spanning trees. There remains which an open question of choosing the spanning tree *T = (V, E (T))* on the network *G = (V, E).* To select the spanning tree, imagine, that it is possible to establish ranking among airports by any criteria, which is equivalent to some ordering of nodes in the network *G = (V, E*). Nodes of set *V* can be represented as a list $(v_1, v_2, \ldots, v_n)$, in which they are arranged in ascending ratings of airports. Here $n = |V|$. If the priority between some airports is not set, the corresponding nodes in the list can be given in random order. Considering that the full graph contains an edge, connecting arbitrary pair of nodes, we construct the spanning tree in the following way. If in the list $(v_1, v_2, \ldots, v_n)$ between v_i *and* v_{i+1} nodes the order is set, then the edge (v_i, v_{i+1}) should be included in the set *E(T).* If the order of nodes is not set, then arbitrary edges with final nodes v_i *or* v_{i+1} are included into set *E(T).* In such a way the spanning tree *T = (V, E (T))* is constructing. Sense of nodes ordering of the network *G = (V, E)* is the higher priority of the airport, the greater quantity of adjacent air routes are considered directly to determine its capacity.

To illustrate the problem (13) - (15) the simplex method considers the example in Figure 1.

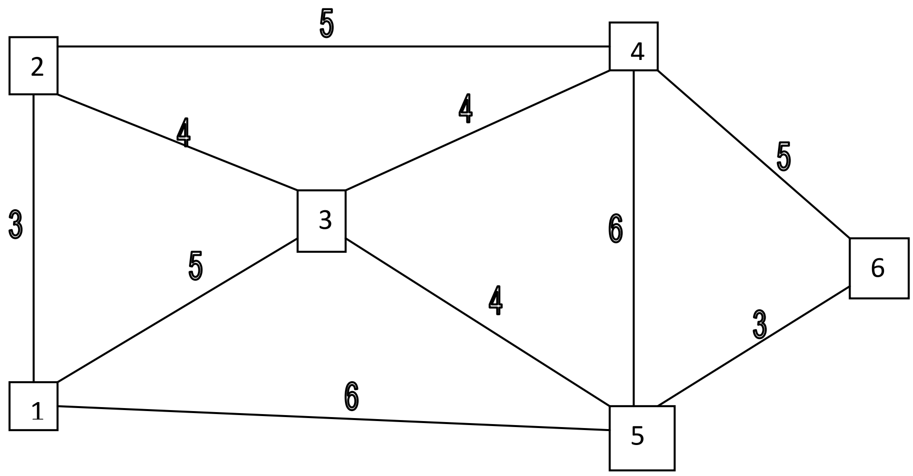

Figure 1. Network$G = (V, E)$

In this figure, nodes (airports) are presented in the form of small squares; their numbers are pointed inside of square. Edges are expressed by the line with the corresponding values of the coefficients c_e . Assume that nodes of the network are ordered according to the list (1, 2, 3, 4, 5, 6), where the order of the node 1 is greater than the unit 2, the order of the node 2 is more than the node 3. Between the nodes 4 and 5 the order is not set, but their order is greater than the order of the node 6. According to this list we build the following snapping tree of the network (Figure 2).

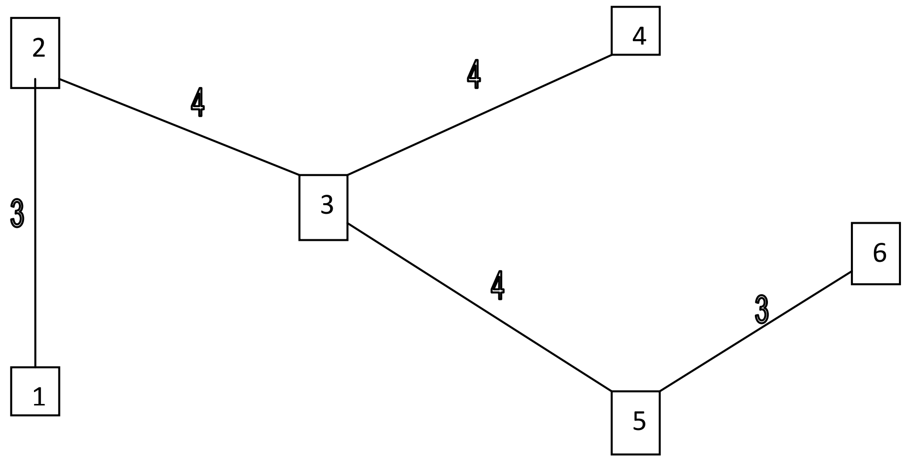

Figure 2. Snapping tree T

After choosing the snapping tree T we present the constrains of the problem (13)-(15) as:

Table 1 Constrains $A\ y=b$.

y_{12}	y_{13}	y_{15}	y_{23}	y_{24}	y_{34}	y_{35}	y_{45}	y_{46}	y_{56}		b
1	1	1	0	0	0	0	0	0	0	=	5
0	1	1	1	1	0	0	0	0	0	=	4
0	0	0	0	1	1	0	1	1	0	=	6
0	0	1	0	0	0	1	1	1	0	=	3
0	0	0	0	0	0	0	0	1	1	=	3

At the Table 1 vector components take the right-hand side of equation and the left side is represented by a matrix, where the columns correspond to the edges $e \in E$, and the rows are the characteristic vectors of the cuts corresponding to the edges (1,2), (2,3), (3,4) , (3,5), (3,6) of the snapping tree T at Figure 2. Each component is indicated on its corresponding column. Since the columns, corresponding to the variables $y_{12},\ y_{23},\ y_{34},\ y_{35},\ y_{56}$, form a unity sub-matrix , they are the basic variables in the initial basic solution to the problem. As a result of four iterations of the simplex method, the optimal solution $y^*_{12}=3$, $y^*_{13}=1$, $y^*_{15}=0$, $y^*_{23}=0$, $y^*_{24}=3$, $y^*_{34}=0$, $y^*_{35}=0$, $y^*_{45}=0$, $y^*_{46}=3$, $y^*_{56}=0$ is determined. Edges, corresponding to the optimal basic variables in the original network of airlines, are indicated by bold lines in Figure 3.

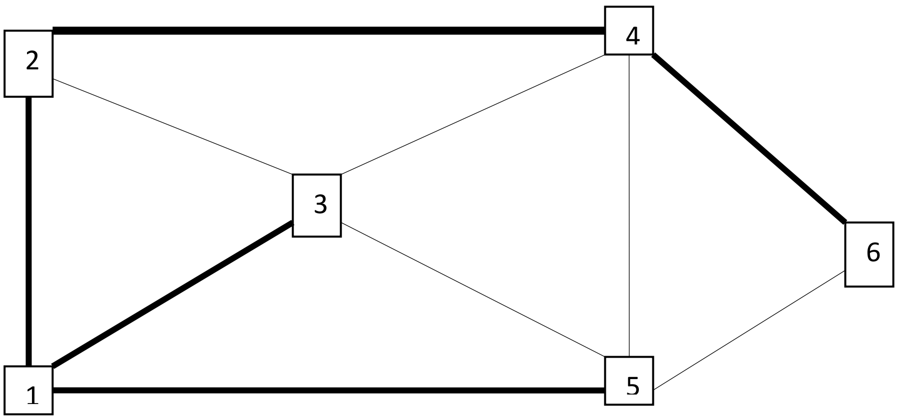

Figure 3. Optimal solution

To determine the optimal solution to the problem (5) - (8) we find the values of the variables z_e from the equation $y_e^* = z_e - l_e$. In this example, since y^*_{12} =3, y^*_{13} =1, y^*_{24}=3 and y^*_{46}=3, it is necessary to increase the capacity of airports 1, 2, 3, 4, 6. Furthermore, the flight frequency between airports 1 and 2, 1 and 3, 2 and 4, 4 and 6 increased. The frequency of other flights remained at the same level. Considering the specific constraints of the dual problem and the matrix *A* has full rank, it is easy to show that the simplex method finds the optimal solution to the problem (13) - (15) through not more than $|E|$-$|V|$+1 iterations. The proving of this statement is beyond the scope of this paper, so it is not mentioned here.

There are examples, showing that in the optimal solution to problem (13) - (15) the values of all the variables are not always integer. In case all the parameters in the constraints of (13) - (15) are integers, in the first iteration of the simplex method it finds the basic solutions, when the values of all the variables are integers too. Therefore, during the operation of the simplex method, the basic variable should be introduced that provides an integer current basic solution. This operation does not present any particular difficulties, and it can be carried in the same way as in the algorithm [5], by searching and checking all the variable-candidates to enter the basis. If the variable is not found, it is necessary to stop the operation of the simplex method. As noted above, through the solution of the problem (13) - (15), the number of simplex-algorithm iterations is not more than $|E|$-$|V|$+1 - 1. Therefore, it is necessary to sort out no more than $|E|$-$|V|$+1 variable-candidates to enter the basis.

Conclusion

1. An airport is most capital-intensive and functionally the least flexible subsystem in the air transport system from position of adaptation of airfield and terminals capacity to passengers and cargo traffic changing in the air routes network. Error in airport capacity could affect economic effectiveness deeply both the aviation and not aviation (non-core) airport's activity. A fresh example of the recent reconstruction of Ukrainian airports in connection with the EURO-2012 illustrates a shortcoming of investment distribution to some of them. Some airports have experienced overcapacity while "Odessa" airport and "Simferopol" airports operate at full capacity.

   This article showed that a certain forecast of passenger traffic on non-stop routes would offer a list of airports that need to be reconstructed. The reconstruction should be based on a capacity that corresponds to the total volume of traffic on the adjacent airlines.

2. With the help of the algorithm of estimating the ODM elements and the simplex method, it is possible to find the numerical solution to the problem of determining the capacity of airports in a given air network. The output problem next data must be assigned:
    a. The structure of air network $G = (V, E)$ of existing nonstop routs between airports. If G is incomplete graph, missing edge must be added to G.
    b. $c_v -$ reconstruction cost per unit of each airport, calculated as the sum of cost of facilitation works and service of minimum number of passengers in an airport.
    c. The minimum number of passengers in each airport can be defined as the sum of required quantity of taken seats in every nonstop active flight from this airport.
    d. $l_{vw} -$ required quantity of taken seats in every flight between different pairs of airports v and w of set V. $l_{vw} = 0$, if for the concrete flight the value of this parameter is not impotent.
    e. The forecast passengers traffic *f({v,w})* of the flight between different pairs of airports *v* and w of set V.

3. The problem of airport capacity and, hence, its reconstruction, is necessary to be solved only in combination with other airports and air routes of the given network.

References

Barbour, B., & Fricker, J. D. (1994). Estimating an origin-destination table using method based on shortest augmenting path. *Transport Research: Part B, 28*(2), 77-89.

Bussieck, M. R., Winter, T., & Zimmerman, U. T. (1997). Discrete optimization in public rail transport. *Mathematical Programming*, *79*, 415-444.

Cook, W. J., Cunningham, W. H., Pulleyblank, W., & Schrijver, S. (1998). *Combinatorial optimization*. New York: John Wiley-Sons. Inc.

Harchenko, V. P., & Naumenko, M. V. (2009). Analiz metodiv pidvyshhennja propusknoi' atnosti aeroportiv. *Naukojemni technologii', 2*, 39-42 (in Ukraine).

Hirst, M. (2008). *The air transport system*. Cambridge, UK: Woodhead Publishing Limited.

International Air Transport Association. (2004). *Airport Development Manual* (9 ed.). Montreal, Canada: IATA.

Janic, M. (2000). *Air transport system analysis and modeling: Capacity, Quality of Services and Economics*. Malaysia: Gordon and Breach Science Publishers.

Kazda, A., & Caves, R. E. (2007). *Airport Design and Operation*. United Kingdom: Elsevier.

Sherali, H. D., Sivanandan, R., & Hobeika, A. G. (1994). A linear programming approach for synthesizing origin-destination trip tables from link traffic volumes. *Transportation Research Part B: Methodological*, *28*(3), 213-233.

Stamatopoulos, M. A., Zografos, K. G., & Odoni, A. R. (2004). A Decision Support System for Airport Strategic Planning. *Transportation Research Part C: Emerging Technologies, 12*(2), 91-117.

Trubin, V. A. (1969). On the method of integer programming problem of special type. *Papers of Academy of Sciences of USSR, 180*(5), 552-554 (in Russian).

Wells, A. T. (2004). *Airport Planning and Management*. New York: McGraw Hill.

Zaporozhets, A. I., & Konovalova, E. V. (2005). Developing the algorithms of airport capacity rationalization according to restrictions on aircraft noise. *Visnyk NAU, 1*, 148-152 (in Ukraine).