Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 14 (2016) 4532 – 4541

Transportation Research Procedia

www.elsevier.com/locate/procedia

6th Transport Research Arena April 18-21, 2016

The application of the artificial intelligence methods for planning of the development of the transportation network

Aleksander Król [a,*]

[a]*Faculty of Transport, University of Technology, Krasińskiego 8, 40-019 Katowice, Poland*

Abstract

A transportation network is a layout of connections, in a region, between communities of people developed in the course of interaction of economic and social, as well as natural environment factors. The current form of transportation network is the result of long term development, which started with the first settlements in the region. When in the course of history, a modernization or an expansion was done, it was to fulfil the new transport requirements, and the work was based on the current structure of the transportation network. These temporary requirements mostly arouse due to random factors which as time went by ceased to prevail. As a consequence the structure, as a rule, is not optimal. At the moment, we can observe a steep increase of communication needs associated with the intensification of economic growth, rising affluence of society and urbanization of new areas. The paper deals with the problem of the transportation network development in response to the new demands. The task is to minimize the total network impedance taking into account the constraints of the investment budget. A bi-level model of the transportation network consisting of some classes of road links was developed. At the upper level the objective function is the usage cost expressed by the total travelling time. At the lower level, at each procedure step the optimal distribution of the traffic is searched for. The main aim was to indicate the most advantageous and acceptable directions of development of the transportation network.

Peer-review under responsibility of Road and Bridge Research Institute (IBDiM)

Keywords: artificial immune system; genetic algorithm; simulated annealing; transportation network

* Corresponding author. Tel.: +48-326034120
E-mail address: aleksander.krol@polsl.pl

2352-1465
Peer-review under responsibility of Road and Bridge Research Institute (IBDiM)
doi:10.1016/j.trpro.2016.05.376

1. Introduction

The search for the optimal structure of the transportation network is called Network Design Problem (NDP). The selected performance measure of the network can be optimized by changing the capacity of existing links (Continuous NDP), adding some new links (Discrete NDP) or both possibilities are taken into account (Mixed NDP). Because the NDP is NP – hard class, the solving algorithms are computational very complex. Thus Poorzahedy and Rouhani (2007) and Chiou (2007) among many others proposed some heuristic approaches. Nie et al. (2007) and Pinninghoff et al. (2008) and some other authors have used the methods based on genetic algorithms. Kim et al. (2008), Xu et al. (2009) and Tianze (2009) applied simulated annealing and Sun et al. (2009) used some ideas taken from the immune systems. Zhang, Lu and Xiang (2008) showed a simplified approach with fixed grid of network vertices. Developing this idea Paramet et al. (2011), Changmin et al. (2012) and Shuaian et al. (2013) introduced some deterministic algorithms and applied them for the test tasks of small size.

2. Model description

As a performance measure of the network the total travel time of all the vehicles was here selected. So the NDP objective is defined as minimizing the total travel time taking into account the budget constraints.

The problem investigated in this paper can be regarded as a form of the Mixed Network Design Problem – the development of current transportation network structure is carried out by the attempts of changing the capacities of existing links or by the attempts of adding some new links. Paradoxically, in some rare cases the reduction of the capacity of a particular link can result in decreasing of the total network impedance.

2.1. Transportation network representation

A transportation network structure is modeled by a graph. The vertices correspond to the crossroads and the edges correspond to the road links. Some vertices are marked to be cities. This distinction allows to introduce the transportation needs into the model. The transportation needs are given by the origin – destination matrix. Each matrix element specifies the number of vehicles travelling in a given time period between pair of cities.

Despite of the fact, that during the optimization process the capacity of any road section can vary continuously, for best visualization each link has an additional attribute determining the its class. There is the unit construction cost and the capacity specified for each class. Five basic classes of road connection are assumed in the model, they correspond respectively to local roads, cumulative roads, main roads, express roads and motorways.

2.2. The network impedance determination

At the upper level the main aim of the optimization is searched for – the structure of the transportation network is modified to achieve minimum impedance. At the lower level, in every step the user equilibrium traffic assignment with given transportation needs is performed for actual network structure. It is assumed, that all the drivers possess complete knowledge of traffic conditions and they make rational decisions. Wardrop (1952) established, that each driver chooses by himself the fastest route according to his understanding of current network condition. The time consumed by a single driver travelling along a link can be expressed by the formula introduced by the US Bureau of Public Roads (1973):

$$T_i = L_i t_{\mathrm{cl}} \left(1 + \alpha \left(\frac{O_i}{g_{\mathrm{cl}}} \right)^{\beta} \right) \tag{1}$$

where: T_i time consumed at i-th link,
L_i length of the i-th link,
t_{cl} unit time of the free driving traffic depending on link class,
O_i current load [vehicles/h],

g_{cl} capacity depending on link class [vehicles/h],
α, β calibrating parameters (typically 0.15, 4).

The average value of the unit time of free driving traffic has been estimated for each link class (there is no distinction between the different types of vehicles).

Because of strong nonlinearity the attractiveness of a particular route depends very much on current network load. According to the required accuracy, the whole traffic volume is divided into tens (or more) parts. Starting from zero, successively for each fraction of traffic volume for each pair of cities the currently fastest routes are found using Dijkstra's algorithm. Then the links belonging to the determined routes are loaded by this fraction. As a result of this procedure the total flow of vehicles between each pair of cities can be split into several different routes. The model considered in this work includes primarily the long trips, which last for several hours and run by the interurban roads. Excluding the rush time and strict neighborhood of the large cities the roads are not overloaded, thus such approximation does not introduce significant distortions, even when the number of iterations is very low.

2.3. The network impedance determination

The unit construction cost depends on the connection class. The presented model is able to take into account the cost dependence on local natural environment or infrastructure facilities. But having in mind the extent of the analyzed area the building costs are assumed to be the same everywhere.

The cost of the whole network modification is a simple sum of the partial costs. There are following assumptions here:

- if a link capacity should be increased, the cost of the expansion is proportional to the specific construction costs of link class,
- if a link capacity should be decreased, the cost is regarded to be zero (the reduction of capacity can be done by administrative means, practically costless),
- if a link should be added, the cost is equal to whole construction cost.

3. The optimization methods

As was mentioned above, the described problem is computationally complex, so the exact optimization algorithms cannot be applied due to extreme calculation time needed. In such cases the heuristic approaches should be used to obtain a solution near the optimal one in the acceptable computation time. In the presented work some artificial intelligence methods were developed: the basic model of the transportation network is in general the same, but the optimization procedures are quite different. The approach involving the artificial intelligence has the fundamental advantage: there no need to fully understand all inner dependencies of the model or even to know some of them. It is only one demand: the model must allow to calculate the value of the objective function. The objective function is the function defined in the solution domain and determines the solution quality. The objective function does not need to be differentiable and any of its gradients does not need to be known.

Three commonly used optimization method were here used and then were compared: genetic algorithms (GA), simulated annealing (SA) and artificial immune system (AIS). The basic concepts of these methods are quite different and also the terms used to describe them differ significantly. Despite of the differences of all these methods work iteratively: the solutions from a set are evaluated and then taking into account the calculated value of the objective function the offspring set is created. Additionally, the best so far solution is hold and compared with the current at every step. In practice all this methods by browsing only a scanty part of the solution space allow to find a solution as close as wished to the optimal. A single solution represents a variant of the transportation network development: the total gain of the travel time for all drivers is the value of the objective function. Additionally, the cost of introduced changes is known and can be compared to the available budget.

During the process of optimization the items of the solution set are subjected to some operations, which can change the structure of a single item or can exchange the parts of structure between a couple of items. In such way the exploration of the solution space is performed. The implementation of these operation is here the same for all methods, but the interpretation and the effects are specific for each of them. The treatment of the costs of the transportation network expansion is also different for each method.

The structure of an item can be varied in different ways (Figure 1):

- the capacity of a randomly chosen link is changed by a random factor,
- a new vertex can be inserted by splitting the randomly chosen link,
- a new link between a pair of randomly chosen vertices is added, its class is also random.

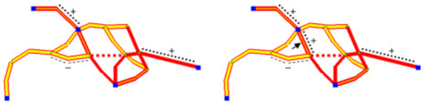

Fig. 1. An example of an item structure changes. Capacities to be increased or decreased are marked by "•" or "-" lines, links to be added are marked by dashed fill.

The exchange operation between two items applies to the randomly selected subset of links – the capacities for these links are copied to each other as shown at Figure 2.

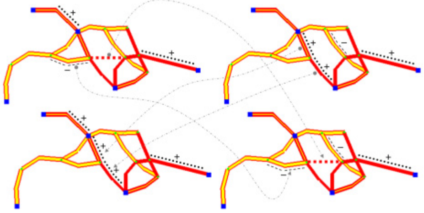

Fig. 2. The exchange operation.

3.1. Artificial immune system

According to De Castro et al. (2000, 2003) the optimization method is based on some ideas derived from the analysis of adaptive immune system. An adaptive immune system can learn and memorize the pathogens and can adjust the force of defense response. Its most important components are the lymphocytes. Each of lymphocytes is equipped with a large number of antibodies, which play the major role in its activity. There are two general kinds of lymphocytes, which significantly differ in their actions but closely cooperate (B-lymphocytes and T-lymphocytes). There is no need to distinguish both types here, and another simplification is to equate an antibody with a lymphocyte.

Individual antibodies vary in their structure, so a function called affinity takes different values for each of them (affinity describes the degree of fit to the pathogens). Depending on the value of the affinity and some other features the antibodies are subjected to the different treatment. During the process of fit increasing, the antibodies which represent the better possibilities of network development are promoted. The driving force of the evolution of the antibodies population are two kinds of transformations: mutation and crossover. Both transformations are applied with specified probability.

Under some circumstances (described below) the transformation of hyper mutation should be applied. It means that an antibody mutates very intensively – repeatedly at a time. Any antibody has an additional feature – age. It stores the number of generations while the antibody remains unchanged. Any mutation or crossover resets the age.

The affinity is defined as the gain of the total travel time of all the vehicles. The treatment of the population of all antibodies in each generation can be described by the following rules:

- antibodies with the greatest (greater than the mean) costs of modification, regardless of the value of the affinity are subject of negative selection, so they are removed from the population,
- the rest of the antibodies are subjected to the positive selection ("positive" means here "due to a desired feature"): these with the smallest (less than the mean minus the standard deviation) affinity are removed, thus only the best and the average ones remain in the population,
- antibodies with the highest (greater than the mean plus the standard deviation) affinity are subjected to the clonal selection – they are multiplied, and then the hyper mutation and crossover transformations are applied,

- the average antibodies (respectively to the total travel time gain) undergo the aging operation: those among them, which have not been improved since many generations are removed from population.

The described above selections result in some number of vacancies. The vacancies are just populated by cloning the antibodies with the highest affinity. But not all such antibodies are cloned. If simply according to the affinity value the first antibodies were chosen, the offspring individuals would be very similar to each other. It is caused by the fact, that commonly the best antibody is accompanied by a large number of close neighbors with a slightly changed structure. So the problem is to find the best individuals which are fundamentally different. To resolve the problem Król (2014) introduced a new feature – similarity. The similarity is calculated for a pair of antibodies and describes their kinship. It allows to select the set of the best but various antibodies: if an antibody with high value of affinity is too similar to all already selected will be not taken into account. Such treatment ensures the diversity of the population.

3.2. Genetic algorithm

Genetic algorithms mimic evolution processes in the living nature:

- different solution versions compete with each other (individuals),
- structure of each individual is determined by a sequence of genes – genotype,
- genotype is subjected to random changes (mutations),
- randomly chosen individuals may exchange parts of their genotypes (crossover),
- fit function being a measure of adaptation determines the probability of passing to the next generation (selection pressure),
- combining random mutations and crossover with selection pressure leads to optimal solution.

The fit function here is near identical to discussed above objective function. If the modification costs do not exceed the established budget both values are equal. If the costs are greater than the budget, the value of the objective function is multiplied by an exponential factor f_x:

$$f_x = e^{-p\frac{C-B}{B}} \tag{2}$$

where: C modification costs,
B budget,
p penalty factor, which determines the significance of the budget violation.

This factor reduces the chances of such an individual for passing to the next generation. The penalty factor p remains constant during the optimization process and should be adjusted empirically. The offspring generation is created with the roulette wheel rule – the probability of being accepted is proportional to the fit function value. A modification is here possible: the fit function value could be raised to a power to make the distribution of chances more sharp or more blurry. The exponent can be any positive real number, here it was assumed equal one. The next modification is so called elite selection – an established number of the best individuals (the elite) pass to the next generation unconditionally.

3.3. Simulated annealing

The principle simulated annealing introduced by Kirkpatrick, Gelatt and Vecchi (1983) is based on an analogy to the physical phenomena occurring during slow cooling and solidification of bodies. The process is characterized by a transition from a state of high-energy (hot fluid) to a structured state with a minimum of energy (crystal). At high temperatures the molecule, with the excess of the energy can jump to any position, as far as lowering the temperature the transition towards lower energies are preferred.

At every step the current solution is compared with randomly generated neighboring solution (only one change is introduced). If the objective function value is increased the new solution become the current one. If the objective function value is decreased the replacement is still possible, but according to the probability expressed as:

$$p_r = e^{-\frac{f_x - f_0}{T}} \tag{3}$$

where: f_0 objective function value for the current solution,
f_x objective function value for the modified solution,
T temperature.

The high value of the temperature at the beginning enables intensive exploration of the solution space, and then it diminishes with time. Thus, the probability of choosing a worse solution at the end of the process is negligible. In addition, a similar rule was applied for the budgetary restriction checking. Commonly, geometric decrease of the temperature in subsequent iterations is used:

$$T_{i+1} = qT_i \quad q < 1 \tag{4}$$

The factor q can be determined by empirical formula depending on presumed number of steps. Simulated annealing can be an efficient optimization approach for this kind of problems, where the "effective diameter" of the solutions space is relatively low. It means, that the operator generating the neighboring solution should be able to scan whole domain of each variable in relatively low number of steps. The presented issue seems to be like that, because the number of possible classes of a road connection is very small.

Standard simulated annealing algorithm was here modified by introducing the parallel mode – there are some almost independent processes running in the same time. The exchange of the information between randomly selected couple of the processes is allowed. The probability of this operation is very low.

4. Obtained results

4.1. Description of the examined example

The described methods were applied to the simplified model of the transportation network in Poland. The model contains all motorways and national roads, some important voivodeship roads are included too. While constructing the model many local roads were omitted. The model includes main towns in Poland and some smaller but important road nodes. Some big, but close to each other towns were merged. The model does not correspond strictly to the reality, because many smaller towns and all villages are not taken into account. Thus the trips undertaken by the considerable part of the population are beyond the model.

The main aim of the presented work is rather to show the possibility of application the artificial intelligence methods to such large scale problem than to use any particular result obtained.

There is no coherent and complete data concerning the transportation needs, which can be used to determine the origin – destination matrix. But, thanks to data published by Polish General Directorate of National Roads and Motorways (GDDKiA) it is possible to estimate the needed values. Every five years the General Traffic Measurement is performed on polish roads and then this data is accessible at GDDKiA website. During the specified period of time at selected road sections all the vehicles are counted and as the result the average daily traffic is determined (divided into some vehicle categories). Fortunately all of the roads included in examined model are covered by the General Traffic Measurement. This allowed to find the elements of the origin – destination matrix by matching the theoretical traffic flows with the measured ones. At the first step the geographic potential approach was used, so the population of the cities was taken into account. For the cities on the state border and border crossings the outer flows were added, this allowed to include the transit traffic. Because thanks to General Traffic Measurement data the total traffic volume for each town is known the initial values of OD matrix could be determined. Further adjustment of these values was performed with simulated annealing algorithm to match measured flows as close as possible. Figure 3 shows the actual map of the area being analyzed with the model of the transportation network and visualizes the calculated daily transportation needs. Because of significant differences between the transportation needs for particular cities the nonlinear scale was used – the diameter of the circle corresponds to square root of total number of vehicles travelling from and towards the city. The average unit cost of road construction were assumed in accordance to the data also published by GDDKiA.

Fig. 3. The model of polish road network and the transportation needs.

For the capacities and unit times of free driving for particular road classes the commonly assumed values were applied. The assumed budget is PLN 10 billion (about EUR 2.5 billion). The overall statistics of the input data for one day is presented in Table 1.

Table 1. The overall statistics of the input data.

Item	Value	Item	Value
Number of trips	677520	Average trip length [km]	234.2
Total length of trips [km]	158702523	Average trip duration [min]	185
Initial total travel time [min]	125322857	Average speed [km/h]	76

4.2. Final results

The performance of the described methods is controlled by some important parameters like the number of individuals, the number of iteration steps, the probability of mutation and crossover, etc. These values were in part already adjusted under recent works and are gathered at Table 2. The same number of individuals (30 000) is involved in the optimization process for all three methods.

Table 2. The controlling parameters.

Method	Parameter	Value	Method	Parameter	Value
GA	Number of individuals	100	AIS	Number of lymphocytes	100
	Number of steps	300		Number of steps	300
	Probability of mutation	0.1		Probability of mutation	0.7
	Probability of crossover	0.7		Probability of crossover	0.8
	Elite	5	SA	Number of processes	20
	Fit function exponent	1		Number of steps	1500
	Penalty factor	3		Probability of crossover	0.05
				Initial temperature	1000

Finally the best parameter set was applied and all the algorithms were run several times. The obtained results have shown, that despite of the similarity of applied operators all the methods behave in different way. These methods are nondeterministic, so every generated solution was slightly different. The obtained values of the objective function were also different. As was mentioned above the objective function expressed the gain of the total travel time for all drivers and is given in minutes. Table 3 contains the summary of the results.

Table 3. Summary of the results (the gain of the total travel time).

Method	Average [min]	Standard deviation [min]	Standard deviation [%]	Min [min]	Max [min]
AIS	11.54e6	6.34e5	5.5	10.38e6	12.45e6
GA	9.67e6	2.94e5	3.0	9.14e6	10.05e6
SA	4.61e6	1.71e6	37.0	2.75e6	7.75e6

As can be seen, the results provided by the simulated annealing algorithm are significantly worse than the others. The average of the objective function is more than twice smaller, where the standard deviation is much greater. This suggests that the simulated annealing algorithm is a totally inappropriate tool for such class of problems. It is a bit astonishing, because the application of this method for small tasks did not indicate so poor results. Finally only the solutions obtained with AIS and GA were considered. Figure 4 presents the progress of the optimization process for both methods.

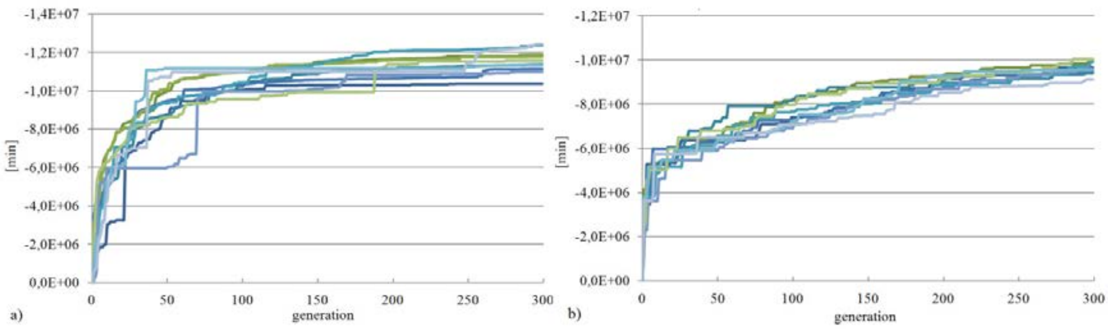

Fig. 4. The progress of the optimization process for several runs of AIS (a) and GA (b).

As can be seen at the beginning of the optimization both algorithms rapidly improve the solution quality, afterwards (about 50th generation) the progress slows down. Until the end of the optimization only minor improvements occur. AIS reaches the plateau faster, whilst GA improves the solution much more at this phase. Additionally, it can be observed that the convergence for GA is significantly better (the standard deviation of the objective function is about twice smaller).

Also the details of optimization are different for AIS and GA. When the best so far solution occurs, it rather remains the current one for the AIS. Meanwhile the optimization for GA goes in a different way: the best so far solution is rather going to be lost, then the algorithm strays and some generation must pass to find next better current solution. Figure 5 illustrates this showing the best and current solutions, curves for AIS are almost overlapped. Despite of this behavior the solutions obtained with GA are significantly more convergent. It could be explained with the presumption, that just this wandering makes the method more flexible and allows to find a better, but somewhat different solution. In the case of AIS, the sophisticated procedure of the selection assures the high quality of obtained solution, but the rules are strict and all improvements are made on the basis of just selected item. The direction once selected is rather hold. It means that the flexibility is here smaller than for GA.

The solutions obtained in successive runs differ from each other, but regardless to the minor differences, all the outputs suggested similar modifications. The best solution has provided the gain of total travel time about 10% (12.5e6 minutes ≈ 208 000 hours ≈ 8646 days). The average trip duration was shortened by almost 20 minutes.

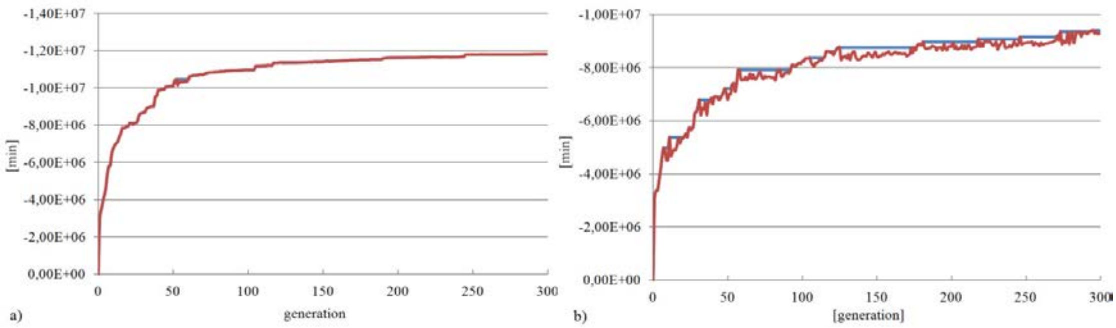

Fig. 5. An example of the best (blue) and current (red) solution for AIS (a) and GA (b). The curves for AIS are near the same.

It is difficult to clearly compare the obtained value to the assumed budget. But some very rough estimations can be done: taking into account the reported polish GDP per capita (about PLN 56 000) and the number of the work hours for a year (48 weeks * 40 hours), the value of about 30 PLN/h can be obtained. Finally, the estimated gain expressed monetary is about 6 million PLN per day, so the return of the investment will occur after about 3 years.

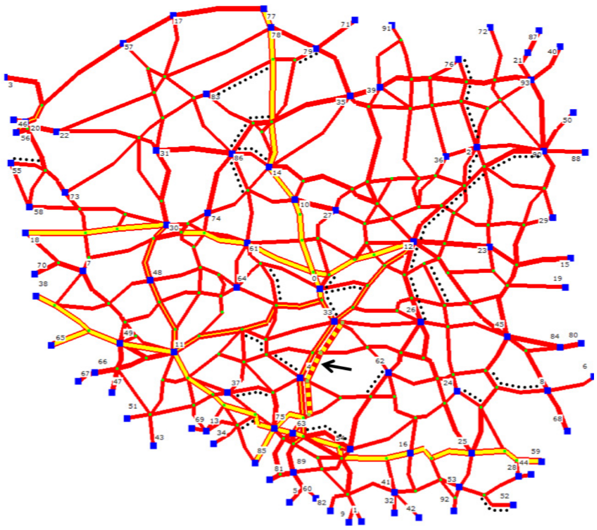

Fig. 6. The solution with most advantageous modifications of transportation network structure.

Figure 6 presents a typical solution, which was obtained with AIS algorithm. Some minor modification on local roads were omitted due to the legibility reasons. The most important suggestion is the new section of A1 motorway from Piotrków Trybunalski to Zawiercie (pointed by the arrow). Some alternative solutions included the significant increase of the capacity (over 80%) of the parallel section of E75 road. What is worth emphasizing, this is the most urgent part of the official plans of the expansion of Polish motorway network and is to be finished around year 2018. The next important tip is to increase the capacity of the route towards Warsaw from Radom and further to the north and east.

5. Conclusion

The most important method of artificial intelligence were shortly described and compared in the work. The task was to find the most promising modifications of the given transportation network structure. Also the budget constraints should be taken into account. The achieved results confirm the usefulness of AIS (Artificial Immune System) and GA (Genetic Algorithm) for solving such class of problems. Meanwhile, the SA (Simulated Annealing) algorithm has been unsuitable here. The results were clearly worse and unstable.

Some differences between AIS and GA were found. Generally, the solutions obtained by AIS were a bit better, but GA seems to be more stable. It could be a premise for further work: to join both methods in some way to take advantage of all their features.

Despite the fact that AIS and GA are non-deterministic procedures the solutions obtained in subsequent runs are nearly identical. Such behavior proves the stability of proposed methods. The computational time needed is about a few hours, what should be considered as acceptable time for such complex data set.

The obtained results match the actual intentions regarding the development of the transportation network in Poland, despite the model includes many simplifications

References

Chiou, S.W., 2007, A generalized iterative scheme for network design problem, Applied Mathematics and Computation, 188(2), pp. 1115–1123.

De Castro, L.N., Timmis, J.I., 2003, Artificial Immune Systems as a Novel Soft Computing Paradigm, Soft Computing Journal, vol. 7.

De Castro, L.N., Von Zuben, F.J., 2000, The Clonal Selection Algorithm with Engineering Applications, Workshop Proc. of GECCO, Workshop on Artificial Immune Systems and Their Applications, pp. 36–37.

http://www.gddkia.gov.pl/pl/1231/generalny-pomiar-ruchu, 2010.

Kim, B.J. , Kim, W., Song, B.H., 2008, Sequencing and scheduling highway network expansion using a discrete network design model, Ann Reg Sci 42, pp. 621–642.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983, Optimization by Simulated Annealing, “Science”, 220, pp. 671-680.

Król, A., 2014, Application of the artificial immune system to determine the priorities for the development of the transportation network, Proceedings of the First International Conference on Engineering and Applied Sciences Optimization. OPT-i, Kos Island, Greece.

Li, C., Yang, H., Zhu, D., Meng, Q., 2012, A global optimization method for continuous network design problems, Transportation Research Part B (46), pp. 1144–1158.

Luathep, P., Sumalee, A., Lam, W.H.K, Li, Z., Lo, H.K., 2011, Global optimization method for mixed transportation network design problem: A mixed-integer linear programming approach, Transportation Research Part B (45), pp. 808–827.

Nie, W., Shao, C.F., Yang, L.Y., 2007, Bi–level programming model for mixed transportation network design and genetic solution algorithm, China Civil Engineering Journal, 40(8), pp. 90–93.

Pinninghoff, M., Contreras, R., Atkinson, J., 2008, Using Genetic Algorithms to Model Road Network, IEEE Computer Society.

Poorzahedy, H., Rouhani, O.M., 2007, Hybrid meta-heuristic algorithms for solving network design problem, European Journal of Operational Research, 182(2), pp. 578–596.

Sun, Y., Song, R., He, S., Chen, Q., 2009, Mixed Transportation Network Design Based on Immune Clone Annealing Algorithm, J Transpn Sys Eng & IT, 9(3), pp. 103–108.

Tianze, X., Heng, W., Zhuan–De, W., 2009, Study on continuous network design problem using simulated annealing and genetic algorithm, Expert Systems with Applications 36, pp. 2735–2741.

US Department of Transportation, 1973, Traffic Assignment, Washington.

Wanga, S.,Meng, Q., Yang, H., 2013, Global optimization methods for the discrete network design problem, Transportation Research Part B (50), pp. 42–60.

Wardrop, J.G., 1952, Some theoretical aspects of road traffic research, Proceedings of the Institution of Civil Engineers, II 1 (2), pp. 325–378.

Xu, T., Wei, H., Huc, G.,2009, Study on continuous network design problem using simulated annealing and genetic algorithm, Expert Systems with Applications 36, pp. 1322–1328.

Zhang, G., Lu, J., Xiang, Q., 2008, Application of Genetic Algorithm to Network Design Problem, International Conference on Intelligent Computation Technology and Automation.