Jordan Journal of Electrical Engineering ISSN (Print): 2409-9600, ISSN (Online): 2409-9619

Using the Clonal Selection Algorithm for the Synthesis of the Topological Structure for Data Network

Mohammad Alhihi

Department of Communications and Electronic Engineering, Philadelphia University, Amman, Jordan
e-mail: Alhihimohammad25@hotmail.com

Received: November 1, 2017 *Accepted: December 2, 2017*

Abstract— A Genetic Algorithm (GA) is a high-level procedure applied to find, generate or select a good solution for an optimization problem, usually in imperfect or limited computation capacity, based on the Evolutionary Algorithms (EA). In this paper, we propose a clonal selection algorithm for the synthesis of a data transmission network topology, which unlike former works, depends on the model cost to determine the optimality value which solves the problem of distributing information flows in a network, and abilities of communication channels (available channel capacity). The proposed genetic synthesis algorithm solves the optimizing criteria, taking into account the optimal trade-off between the network dependability and cost. Finally, the algorithm was applied over the Jordanian cities where each city represents an algorithm node. The outcome of this application shows the gained improvement of the proposed algorithm.

Keywords— Clonal selection algorithm, Evolutionary algorithms, Genetic algorithm, Optimizing criteria.

I. INTRODUCTION

In modern and developed societies, communication systems become a life style and a true necessity that the whole life issues depend on. In fact, communication systems are regarded as important as the electrical systems. Accordingly, any failure in communication systems is considered as serious as the failure in the electrical systems; and it hence leads to equal harmful consequences, mainly, over the national security, defense and law enforcement agencies or transportation systems which could be catastrophic. That is why the problem of establishing dependable and cost-efficient telecommunication systems is considered essential.
The cost efficiency and telecommunication systems' dependency are directly related and well-linked to the data-communication network typological structure, which justifies the true need to develop effective synthesis typological structure methods to ensure the optimum trade-off in cost and dependability criteria.
A Genetic Algorithm (GA) is defined in [1] and [2] as a high-level of searching, generating or selecting a good solution for a certain problem by using the Evaluation Algorithms (EA) [1], usually in imperfect or limited computational capacity, based on the Evolutionary Algorithms (EA) [2].
The synthesis of the topological structure in distributed networks includes searching for a topology which minimizes the expenditures of subsystem communication channels, based on the restriction of dependability and delay at the available cost limit, which is the optimization problem that [3] solves by the heuristic method.
In [4], the search for the optimal solution is characterized by complexity and uncertainty. Hence, it acquires new information to solve the uncertainty problem and increase the average performance.
In [5], Hybrid GAs are described for the graph partitioning problem, which results in a fast heuristic local improvement, such as the schema preprocessing phase that improves GAs' space searching capability.

Corresponding author's e-mail: Alhihimohammad25@hotmail.com

However, both [4] and [5] did not take capacity into account when searching for the optimal solution for the proposed problems.

The research for GA in [6] was published in the 1970s, and showed that GA was successfully used to solve a large number of difficult formalized problems and obtain the Nondeterministic Polynomial (NP) optimization solutions. However, more recent work was introduced in [7]-[12] which used different methods to achieve the same targets for different network parameters.

The work presented in [13]-[15] contains different GA topologies for the search of optimal solutions to the data-communication network, where each reference tackles optimization according to certain criteria.

Shortening the path as a method to optimize network performance was recently shown in [16]-[18], where in [16], a weighted oriented graph as a structural model was used to achieve the required quality of service (QoS). In [17] and [18], the authors presented the optimization problem over network traffic across multiple independent paths [17], and a practical routing protocol model to improve the network performance and adequacy [18].

However, the presented work in [16]-[18] does not solve the problem of the algorithms immunity when searching for the optimal solution, which is what this paper does.

In [19] it is experimentally and theoretically indicated that the immune (clonal) algorithm is effective as a problem-solving procedure in multimodal functions; and it deduced that the proposed GA can be considered as a high-level immune algorithm. However, [19] did not propose any variety in enabling mechanism in GA under investigation.

In addition, in [19], immune algorithms of optimization have a set of important capabilities, including the repeatability of decision candidates assigned at all parameter spaces; a massive production through trial and error of different search directions around the most appropriate initial points, and effective parallel search which shows the number of points registered by the training function.

The rest of the paper is organized as follows: Section II introduces the objective of the proposed AG; Section III shows the Algorithm clonal. Section IV introduces the algorithm coding and decoding. Section V shows the criterion end of algorithm limit; and it applies the algorithm over Jordanian cities which represent the algorithm nodes. Finally, section VI concludes the paper.

II. THE PROPOSED WORK OBJECTIVES

The main objective of this paper is to design an immune clonal algorithm for an optimal topology synthesis based on the minimal cost of the communication channel occupancy.

The topology synthesis, in general terms, consists of a great number of switching centers $X = (x_j)_{j=1 \ldots N}$, the distance matrix between all the switching centers $L = \|l_{ij}\|_{i,j=1 \ldots N}$ and the requirement matrix at the information exchange between all the switching centers $H = \|h_{ij}\|_{i,j= \ldots N}$, where h_{ij} is the flow value which must be transferred from the unit *i* to the unit *j* in unit time; $D = (d_i)_{i=1 \ldots n}$ is the communication channel available capacity, $C(d,l)$ is the channel cost function and *l* is the path length.

The assigned topology for the data-communication network is $M^0 = \{(r,s)\}$. Accordingly, it is required to find the communication capacities of all communication channels $\{d_{rs}^0\}$ and the flow distribution at the communication channels $\{f_{rs}^{0}\}$. The appropriate required matrix in such a manner is provided in (1), [13]:

$$\min \sum_{(r,s)\in M^0} C(d_i, l_i) \tag{1}$$

Under the restriction of three conditions which are:

- Channel flow: $f_{rs} < d_{rs}$, $(r,s)\in M$
- Average time of delivery: $T_{aver} \le t_{\text{init}}$
- Connection coefficient: $k_{rel} \ge \grave{k}_{rel}$

III. ALGORITHM CLONAL

The algorithm clonal is represented in [13] as:

$$CLONALG = (P^l, G^k, l, k, m_{Ab}, \delta, f, I, \tau, AG, AB, S, C, M, n, d) \tag{2}$$

where P^l is the space of search (shape-space); G^k is the space representation; l is the length of the attributes vector (dimension of the research space); k is the length of antibody receptor; m_{Ab} is the dimension of the antibodies population; δ is the expression function; f is the affinity function; I is the initial population function of the antibodies population; τ is the condition of completion of algorithm work; AG is the subset of antigens; AB is population of antibodies; S is the operator of selection; C is the operator of cloning; M is the mutation operator; n is the number of the best antibodies selected for cloning; d is the number of the worst antibodies subjected to substitution for new ones.

The application of the clonal algorithm as a method for solving optimization problems shows that it possesses the properties of both gradient and stochastic methods. The combination of gradient and stochasticity makes these algorithms effective methods for solving optimization problems in the multi-external landscape of the objective function. It can be assumed that they are promising for solving the applied problem of synthesizing the topological structure of switching centers and communication channels

Algorithm clonal selection can be summarized in the following steps as follows:

1. *Initialization:* Creation (usually by random generation) of the initial population of antibodies AB.
2. *Determination of affinity:* Every antibody AB_i, $AB_j \in AB$ determines its affinity relative to every antigen Ag_i , $Ag_i \in AG$. Write the result into the matrix of affinities $D : D = \left[\left|AG\right| \times m_{Ab}\right]$, and $d_{ij} = f\left(Ab_j, Ag_i\right)$, $d_{ij} \in D$.
3. *Clonal selection and propagation:* Select from population n of each the best antibodies for every row of the matrix D, and place them into separate population of clones $AB_C, \left|AB_C\right| = n \cdot \left|AG\right|$. It is necessary to generate the clones of the population elements AB_C proportionally to their affinity, i.e., the greater the affinity, the greater is the number of the generated clones and vice versa.
4. *Affinity maturation:* Subject to mutation all the clones of population AB_C with probability inversely proportional to affinities, i.e., the greater the probability of mutation, the lower is its affinity. Determine a new affinity of each antibody AB_j ,

$AB_j \in AB_C$ like in step 2 and obtain the matrix of affinities D_C. Select n antibodies from the population AB_C for which the corresponding vector-column of the matrix D_C gives the best generalized result of the affinity, and transfers it into the population of memory cells M_R.

5. *Metadynamic:* Substitute the worst d antibodies of the population AB by new random individuals.
6. *Substitute n antibodies* of the population AB by memory cells from M_R and then go back to step 2. Repeat the work from step 2 to step 5 till reaching the desired criterion that should be reached in this proposed algorithm.

Fig. 1 shows the flow chart for the proposed clonal selection algorithm:

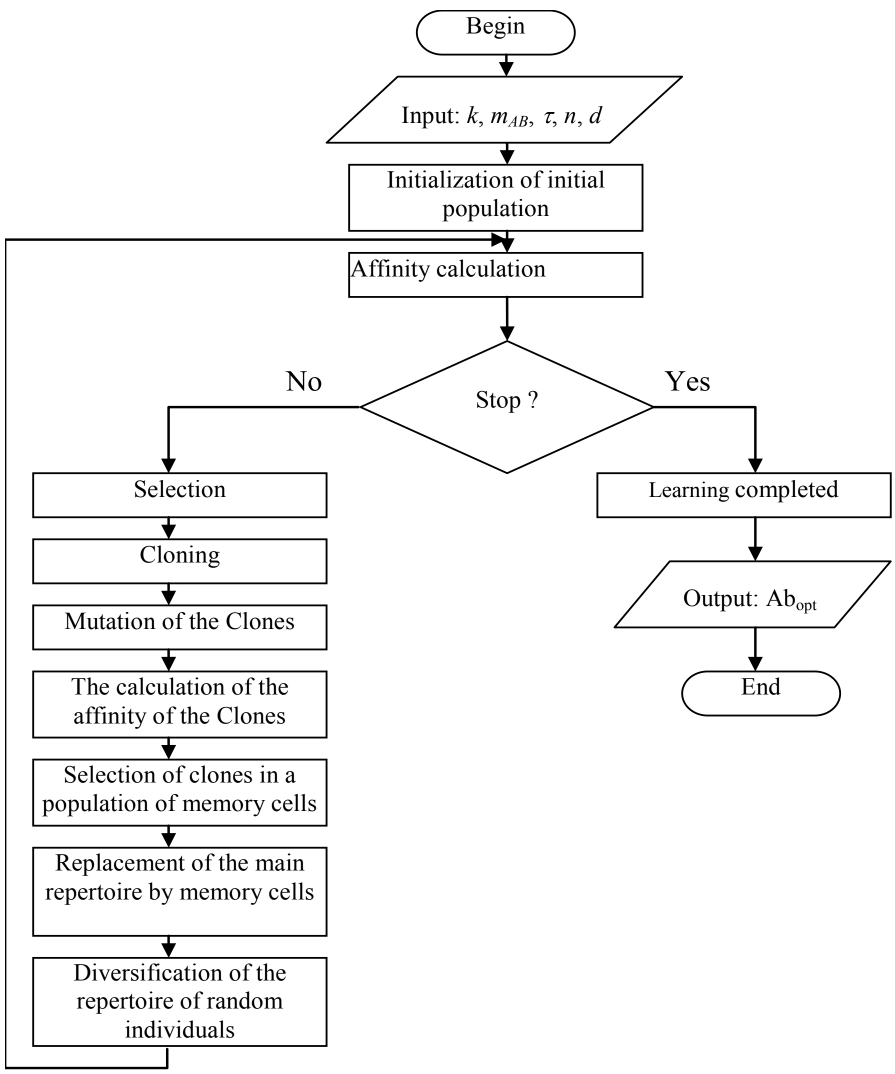


Fig. 1. Flow chart of clonal selection algorithm

The flow chart in Fig. 1 shows that the complexity of the proposed algorithm maintains the same level of complexity for the genetic algorithm.

Peculiarities of the clonal selection algorithm are based on a fact that contrasts network immunity, because it assumes constant size of antibodies population [19], [20].

Moreover, in [18], the mechanism of antibodies variety is described; and the hypothesis is suggested based on searching for antibody Ab_1 which is selected during the primary response and the further point hypermutations before moving to the following stage. The system investigates local zones only around Ab_1 by small stages towards the antibody with a high affinity to find the local optimum of Ab_1^*. Searching the local optimum leads to losing the mutations with no great level. Antibodies cannot search the local minimum/maximum. To avoid this shortage in [18], the receptor is edited by including the parameter d to realize antibodies in a great region of affinity searching steps, to be stopped at the place where the affinity becomes lower than the chosen optimization level at any random step (Ab_2). However, a random step could result in antibody on the side of the vertex, where region of ascent is more promising (Ab_3) i.e., in different step, which makes it possible to reach practically arbitrary global maximum, resulting to the ability to bring antibody hypermutations place to the most top vertex (Ab_3^*). It is necessary to know that hypermutations are useful for investigating and searching the local domains, while editing (i.e., removing antibodies with low affinity) can save antibodies, which hit the local optimum as shown in Fig. 2.

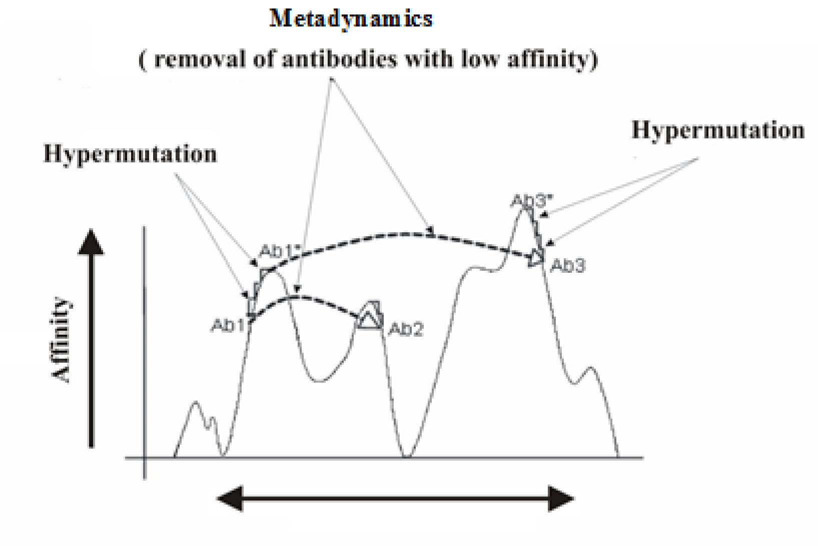


Fig. 2. Hypothetical representation of the process metadynamics

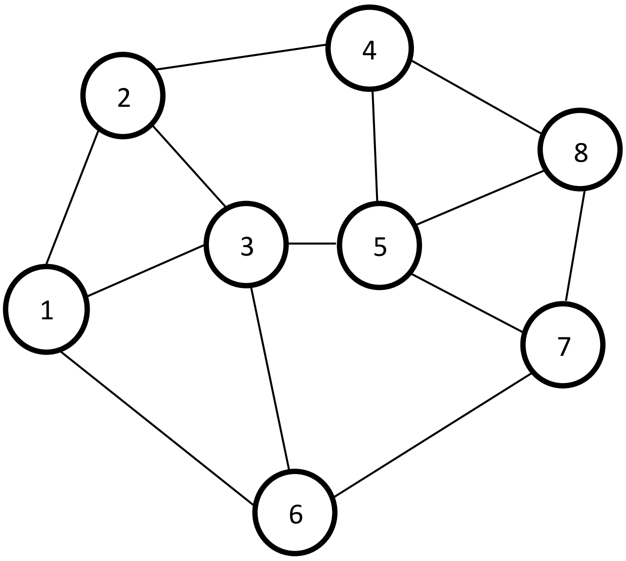
IV. ALGORITHM CODING AND DECODING:

The graphical representation of the node matrix is regarded as a complicated problem. In fact, it is considered as one of the most demonstrable and compact problems in topologies. The difficulties in the topology result from the complexity matrixes, where the rows and columns are marked by the graph nodes. There is 1 or 0 at the intersection of column and row depending on whether there is edge or not, that comes from the row label to the column label. By virtue of the fact that co-set relation is symmetrical in the undirected graphs, there is symmetry of a connectivity matrix in respect of the principal diagonal. That is why it is sufficient to define the upper triangular matrix for having all the information about the graph topology. The number of elements in such a matrix is defined by the following formula:

$$l = \frac{N \cdot (N-1)}{2} \tag{3}$$

where N is the number of the graph nodes.

The graph topology antibody encoded consists of the genes, the number of which is constant and is defined by (1). Each gene encodes the edge availability between the relevant knot pair as shown in Fig. 3.



$$S = \left| \begin{array}{ccccccc} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ & 1 & 1 & 0 & 0 & 0 & 0 \\ & & 0 & 1 & 1 & 0 & 0 \\ & & & 1 & 0 & 0 & 1 \\ & & & & 0 & 1 & 1 \\ & & & & & 1 & 0 \\ & & & & & & 1 \\ & & & \delta & & & \end{array} \right|$$

1-2	1-3	1-4	1-5	1-6	1-7	1-8	2-3	2-5	2-6	2-7	2-8
1	1	0	0	1	0	0	1	1	0	0	0

3-4	3-5	3-6	3-7	3-8	4-5	4-6	4-7	4-8	5-6	5-7	5-8	6-7	6-8	7-8
0	1	1	0	0	1	0	0	1	0	1	1	1	0	1

Fig. 3. Gene encode algorithm

V. Stop Criterion and Objective Function and Real: Example from Jordanian Cities

In this clonal sorting algorithm, the finish of the finding solutions is the absence of the amending of the average suitability of population over a period of three generations in a row. However, the generation period could be modified when required to have better but more complex performance in the algorithm. However, the average suitability of population in this paper is applied over just a period of two generations in a row to maintain the complexity.

The capacity of the objective function used the same cost model of data-communication network in [13] to solve the allocation problem of information flows in the network through the shortest routes. It as well solves the problem of choosing the optimal flow and communication channels capacity.

A. Topology Synthesis

In the collected experimental results, the following parameters, shown in Table 1, were selected for the clonal sorting algorithm:

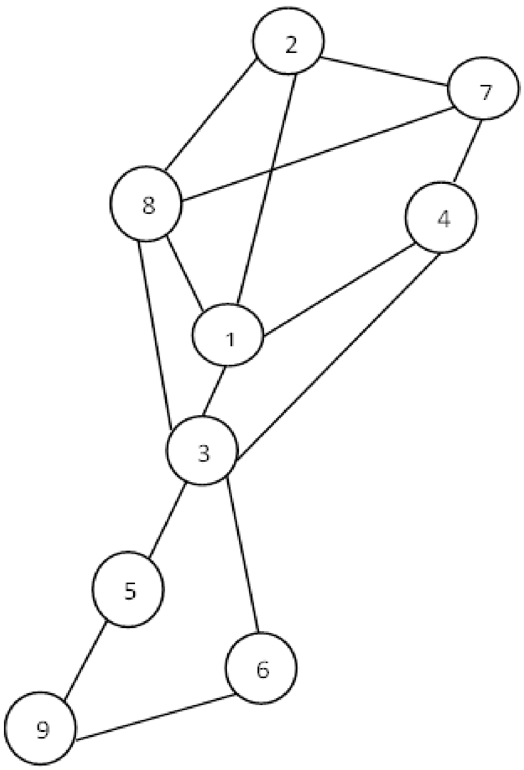
TABLE 1
PARAMETERS OF THE CLONAL SORTING ALGORITHM

Parameter Name	Parameter Value
Size of the main population	100
Size of the population of clones	600
Number of generations	900
Selection coefficient	0.7 (*n*=70)
Size of tournament	5
Level of mutation	0.01
Precision	8 bit
Time step	0.25
Data number of a row	20

Where the size of the population is 300; the probability of mutation is 0.06; the probability of crossing procedure application is 0.80; the crossover is simple.

B. Real Example in Jordanian Cities

This algorithm was applied for the topology synthesis of data-communication network with the nine units which are located at the administrative center of the Hashemite Kingdom of Jordan. The rental cost of digital communication channels was evaluated according to the maximum tariffs which are estimated according to the committee of Jordanian communication and information network. The flow value h_{ij} between units for k=3 of all pairs comprises 50 bps. The maximum allowed delay time limitation is 0.005sec. As a result of the GA functioning, we achieved a topology of data-communication network with the rental cost of digital communication channels of 8.5$/p.h as shown in Fig. 4.



1. Amman	2. Irbid	3. Karak
4. Zarqa	5. Tafila	6. Ma'an
7. Mafraq	8. Jerash	9. Aqaba

Fig. 4. Topology of data-communication network

The convergence dynamics of GA is shown in Fig. 5.

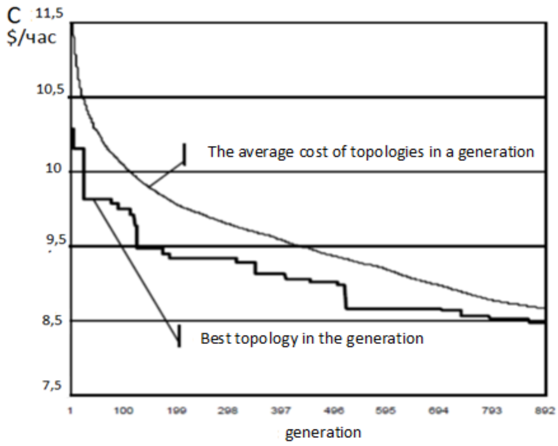


Fig. 5. Dynamics of the clonal selection algorithm functioning

VI. Conclusion

The performed researches show the perspective of applying Genetic Algorithms (GA) as methods of communication data network synthesis of optimal topologies, which are derived from the well-known heuristic methods. The proposed algorithm has a higher probability to identify a global optimum due to the fact that there is a simultaneous presence of stochasticity and gradient in the GA. The stochasticity is present at the mutation step; and the gradient is reflected in the selection rules and generation of new species. High efficiency of finding solutions can be measured when varying these methods of contribution in proportion to the GA due to the change of mutation probability, determination of the average number and the power of mutating solutions. The main disadvantage of GA is its requirement of big computing resources, both from the memory side and from the side of the processor performance, the result of which is a delay in finding solutions. In this case, it takes 76 hours for the system on Pentium 4 processor 2400 GHz and 256 MB. In the case of GA usage for design problem solutions, this disadvantage is not critical.

Future work is intended to be directed to apply the Network Coding (NC) technique proposed in [21]-[23] to find the shortest path for the goal of reaching optimal bandwidth and channel capacity by choosing the shortest path.

References

[1] M. Melanie, *An Introduction to Genetic Algorithms*, MIT Press, 1996.

[2] R. Poli, W. Langdon, and N. McPhee, *A Field Guide to Genetic Programming*, England & Wales License, 2008.

[3] J. Leeuwen, *Algorithms and Complexity*, Elsevier B.V., 1990.

[4] J. Holland, "Genetic algorithms and the optimal allocation of trials," *SIAM Journal on Computing*, vol. 2, no. 2, pp. 88-105, 1973.

[5] T. Bui and B. Moon, "Genetic algorithm and graph partitioning," *IEEE Transactions on Computers*, vol. 45, no. 7, pp. 841-855, 1996.

[6] J. Holland and A. Arbor, *Adaptation in Natural and Artificial Systems*, Michigan Press, 1975.

[7] S. Samoylenko, "Application of heuristic problem-solving methods in computer communication networks," *Machine Intelligence*, vol. 3, no. 4, pp. 197-210, 1985.

[8] A. Kershenbaum, P. Kermani, and G. Mentro, "An algorithm for mesh network topological optimization and routing," *IEEE Transactions on Communications*, vol. 39, no. 4, pp. 503-513, 1991.

[9] C. Rose, "Low mean inter nodal distance network topologies and simulated annealing," *IEEE Transactions on Communications*, vol. 40, no. 8, pp. 1319-1326, 1992.

[10] A. Dutta and S. Mitra, "Integrating heuristic knowledge and optimization models for communication network design," *IEEE Transactions on Knowledge and Data Engineering*, vol. 5, no. 6, pp. 999-1017, 1993.

[11] S. Pierre, "A learning-by-example method for improving performance of network topologies," *Engineering Applications of Artificial Intelligence*, vol. 7, no. 3, pp. 279-288, 1994.

[12] B. Dengiz and C. Alabas, "A simulated annealing algorithm for design of computer communication networks," *Proceedings of World Multiconference on Systemics, Cybernetics and Informatics*. vol. 5, pp. 34-39, 2001.

[13] S. Pierre and G. Legault, "A genetic algorithm for designing distributed computer network topologies," *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 28, no. 2, pp. 249-258, 1998.

[14] T. Suda, "Topology reconfiguration of an IP network embedded over an ATM network," *IEICE Transactions Communications*, vol. E84-B, no. 11, pp. 1-4, 2001.

[15] D. Reichelt, F. Rothlauf, and P. Gmilkowsky, "Designing reliable communication networks with a genetic algorithm using a repair heuristic," *Proceedings of the European Conference*, pp. 177-187, 2004.

[16] M. Alhihi, H. Attar, M. Samour, and A. Akulynichev, "Researching the impact of parameters of the developed routing models on network performance," *Studies in Engineering and Technology*, vol. 4, no. 1, pp. 61-69, 2017.

[17] M. Alhihi, "Setting an optimization problem of distributing traffic across multiple independent paths," *Transactions on Networks and Communications*, vol. 5, no. 4, pp. 40-46, 2017.

[18] M. Alhihi, "Practical routing protocol models to improve network performance and adequacy," *Computer and Communications*, vol. 5, no. 6, pp. 114-124, 2017.

[19] L. Castro and J. Timmis, *Artificial Immune Systems: A New Computational Intelligence Approach*, Springer, 2002.

[20] T. Fukuda, K. Mori, and M. Tsukiama, "Parallel search for multimodal function optimization with diversity and learning of immune algorithm," *Artificial Immune Systems and Their Applications*, pp.210-220, 1998.

[21] H. Attar, D. Vukobratovic, L. Stankovic and V. Stankovic, "Performance analysis of node cooperation with network coding in wireless sensor networks," *Proceedings of IFIP International Conference on New Technologies, Mobility and Security*, pp. 1-4, 2011.

[22] H. Attar, L. Stankovic, and V. Stankovic, "Cooperative network-coding system for wireless sensor networks," *IET Communications*, vol. 6, no. 3, pp. 344-352, 2011.

[23] M. El-Hihi, H. Attar, A. Solyman, and L. Stankovic, "Network coding cooperation performance analysis in wireless network over a lossy channel, M users and a destination scenario," *Communications and Network*, vol. 8, no. 4, pp. 257-280, 2016.