Modified Genetic Algorithm for Solving the Problem of Selecting Hardware and Software for Information Security

Valerii Lakhno[1], Andrii Blozva[1], Dmytro Kasatkin[1], Dmytro Tyshchenko[2], and Tamara Franchuk[2]

[1] *National University of Life and Environmental Sciences of Ukraine, Kiev, Ukraine*
[2] *Kyiv National University of Trade and Economics, Kyiv, Ukraine*

Abstract

This article proposes a modified genetic algorithm for solving the problem of hardware and software selection for information protection and information security of informatization objects. In contrast to existing solutions, it is proposed to apply a new coding method. It is also proposed to use the so-called elite strategy for the formation of new generations by a gene bank integration into the algorithm. The use of a gene bank allows one to reduce the number of generations in the process of the solution search. Therefore, it leads to a general reduction of genetic algorithm working time.

Keywords

Optimization, genetic algorithm, development of information security circuits, object of informatization.

1. Introduction

Correctly implemented selection of equipment to ensure information security (IS) of various objects of informatization (OBI) largely determines the success of modern enterprises, where a variety of information technologies (IT) have become a common component of business processes.

Nowadays, a range of equipment for solving the problems of information security (IS) provision for the objects of informatization (OBI) grants a scope of opportunities for the information security acquisition, see Fig. 1.

However, the milestones of the equipment selection for each of the IS circuits do not lose their relevance. It has become especially noticeable during recent years in the face of an increase in the number and complexity of cyber attacks against various OBIs.

The solution of problems related to the equipment selection for IS assurance of the OBI dictates the need to take into account the opposite tendencies:

- on the one hand, it is necessary to purchase information security hardware and software (H&S), which allows one to ensure a high degree of information security, both for an individual circuit and for OBI as a whole;
- on the other hand, if one does not consider the critically important objects of informatization (for which a priori reliable protection is a primary task) in the conditions of market competition, it should be taken into account that the costs of acquiring IS H&S should be minimized, and the funds invested into information security should pay off.

CPITS-II-2021: Cybersecurity Providing in Information and Telecommunication Systems, October 26, 2021, Kyiv, Ukraine

EMAIL: lva964@nubip.edu.ua (V. Lakhno); andriy.blozva@nubip.edu.ua (A. Blozva); d.kasatkin@nubip.edu.ua (D. Kasatkin); tyshchenko_d@knute.edu.ua (D. Tyshchenko); franchuk_t@knute.edu.ua (T. Franchuk)
ORCID: 0000-0001-9695-4543 (V. Lakhno); 0000-0002-4377-0916 (A. Blozva); 0000-0002-2642-8908 (D. Kasatkin); 0000-0002-2193-9012 (D. Tyshchenko); 0000-0001-7615-1276 (T. Franchuk)

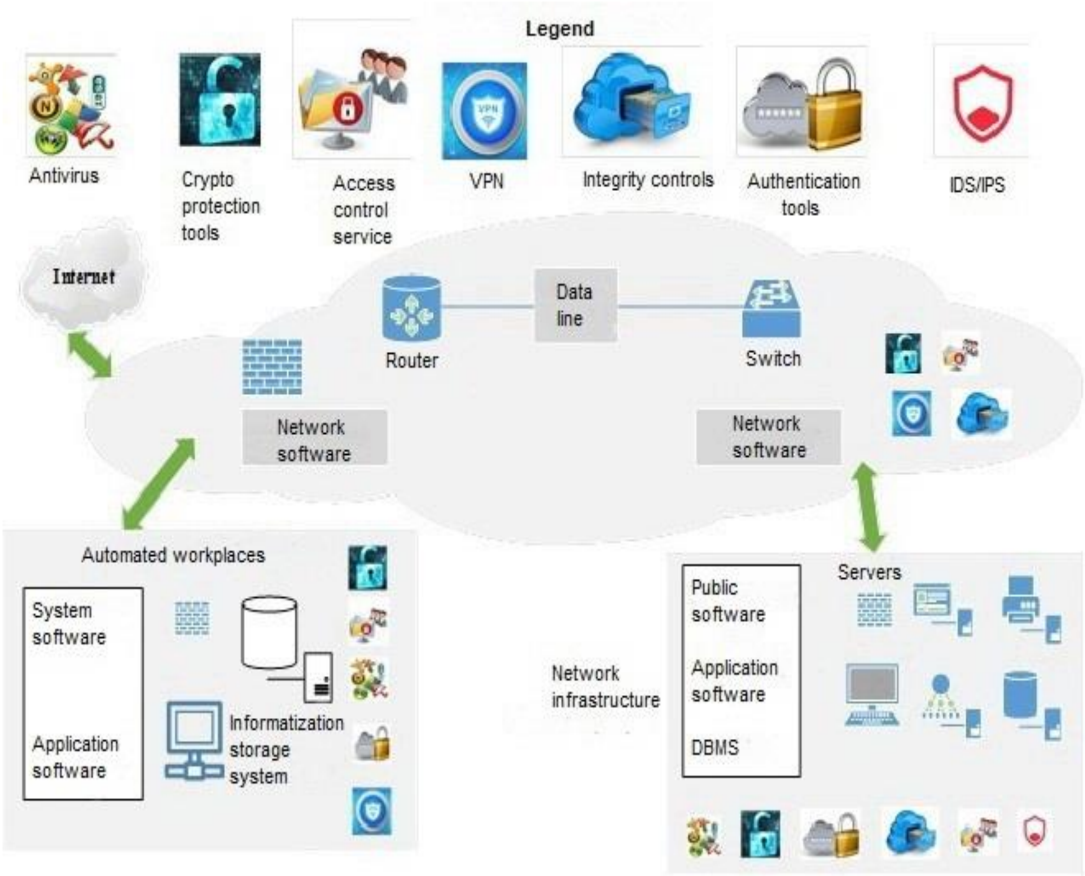

Figure 1: The structure of the object of informatization aimed to solve the optimization problem of the hardware and software information distribution in the protection circuits

Nowadays, many IS H&S have similar characteristics, moreover, some IS H&S receive broader capabilities that are overlapping characteristics of more highly specialized information security tools (IST).

Problems of IS H&S selection are especially relevant for state OBIs; since the lack or excess of equipment for various IS circuits on large commercial OBI can be compensated by other IS H&S on another protection circuit. Meanwhile, it is impossible at the state-owned enterprises due to the limited budget. Therefore, the wrong choice of IS H&S can negatively affect the OBIs information security metrics.

There are various ways to optimize the process of IS H&S selection for OBI. Quite a lot of scientific publications are devoted to this topic. A short overview of such researches is given below. However, within the scope of our study, we will focus on the possibilities of genetic algorithm (GA) usage within the problems of components of IS H&S selection along the contours of IS.

2. Review of Research Literature

The expediency of GA usage can be substantiated by the fact that the problem being solved belongs to multi-criteria and multi-extreme problems [1, 2].

In [3, 4] it is shown that GAs can be used in the course of solving multicriteria optimization problems, which are variations of evolutionary search methods.

In [5, 6] the features of GA usage for tasks related to the choice of equipment for OBI information protection systems are analyzed. However, the solution proposed by the authors is essentially a combination of the standard greedy and GA.

It is quite difficult to unambiguously algorithmize the efficiency of the H&S choice for the multi-circuit OBI due to the description of the objective function. The objective function should be multivariable, as far as it is influenced by many factors. Moreover, these factors are often probabilistic. Therefore, it is better to evaluate and model indicators in relation to the IS H&S complexes. And only after such evaluation, the impact of these complexes on the performance indicators of OBI can be evaluated, in particular using the information security metrics of the enterprise. The values of these indicators, including the information security metrics, can be used in the H&S selection for the information security.

All of the above written has determined the relevance of our research.

3. Models and Methods

The formation of the IS infrastructure contours of OBI implies the process of delimiting the tasks related to information protection and cybersecurity between these contours. As shown in Fig. 2.

The contours (perimeters) of the IB OBI in Fig. 2:

- PIS (I) – the perimeter of the information system of OBI;
- PCOI (II) - perimeter of control of information object;
- UAP (III) - User Access Perimeter;
- PNE (IV) - the perimeter of the network equipment;
- OPIO (V) - the outer perimeter of information object.

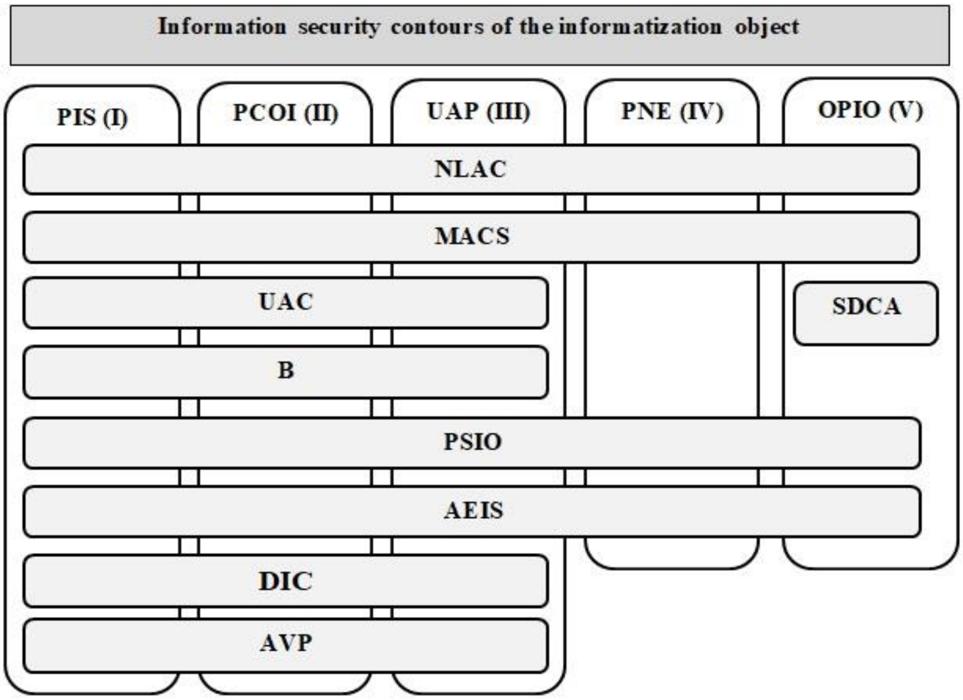

Figure 2: Typical structure of IS H&S placement along the OBI protection contours

In Fig. 2, the following designations are adopted: AVP - anti-virus software; DIC - OBI data integrity control tools; AEIS - audit of cyber incidents; PSIO - providing physical protection for OBI; B - backup of OBI information arrays; UAC - control over the actions of OBI personnel; SDCA - subsystems for detecting cyber attacks on OBI; MACS - event monitoring subsystem; NLAC - control of network activity on OBI.

The goal function for the problem of IS H&S infrastructure optimization for the OBI information security circuit will be described as the cost of an information security tools set. It can be represented as the following expression:

$$S = \sum_{l=1}^{u} \sum_{i=1}^{m} \left(c_i + \sum_{j=1}^{k} n_{ij} \cdot c_{zj} \right) \cdot cir_{il}, \tag{1}$$

where~ c_i – cost of the $i-$ the contour of IS OBI;

c_{zj} – cost of additional means of protection (IS H&S) for the $i-$ the OBI IS circuit

k – number of protection means (IS H&S) for the $i-$ the OBI IS circuit;

m – number of IB contours OBI

$u-$ the number of options for filling each of the OBI IS contours

cir_{il} – the number of IS circuits of OBI, for which specific IS H&S are needed to achieve the required indicators for IS metrics

n_{ij} – the number of IS H&S type for the go contour of the IS OBI.

Wherein

$$u = \sum_{i=1}^{m} h_i = \sum_{i=1}^{m} \frac{(k + d_i)!}{d_i!k!}, \tag{2}$$

d_i – number of IS H&S for the OBI circuit.

Optimization can be done for variables c_{zj} and cir_{il}, which are contained in the expression (1).

One can write the limitation on the number of minimum required IS H&S as follows:

$$\sum_{j=1}^{k} n_{ij} \le d_i, \quad i = 1..cir. \tag{3}$$

Restriction on the sufficiency of the integral metric [7] of information security for all OBI contours for a finite set of IS H&S, which provides a given level of information security for the analyzed object:

$$\sum_{i=1}^{k} \left(\frac{\sum_{j=1}^{k} n_{ij} \cdot met_j}{SM} \right) \ge N_M, \tag{4}$$

where met_j – IS metrics for the OBI IS contour;

SM – total IS metric for OBI;

N_M – the required number of APS IS for the safe operation of OBI.

Constraint that describes the integer nature of the current task:

$cir_{il}, n_{ij} \ge 0, \quad cir_{il}, n_{ij}$ – integral.

In the considered GA, the population is a set of decisions in the course of the IS H&S choice. Namely, these are, in fact, different combinations of IS contours sets. Therefore, the individuals in the population will contain one chromosome with the number of genes, which is equal to the number of variants of possible arrangements of the IB contour. It is determined by expression (2). In contrast to the classical GA, which uses binary coding, the study used the coding list

The list item contains such information:

- IS contour OBI in accordance with Fig. 1 and 2;
- the composition of the IS H&S for the circuit;
- general indicators for safety metrics for the circuit;
- the cost of a set of IS H&S for the circuit.

The number of genes in the chromosome (ch) is taken to be equal to the number of elements in the list of variants of the APS IB sets of the corresponding IB contour.

Expression (1) will be used as a fitness function.

Possible combinations of IS H&S sets of individual contours will constitute a population (pop). At the same time, restrictions were adopted on the number of minimum required APS IS and their total cost.

In Fig. 3 the procedure, which was applied during the creation of the initial population, is shown

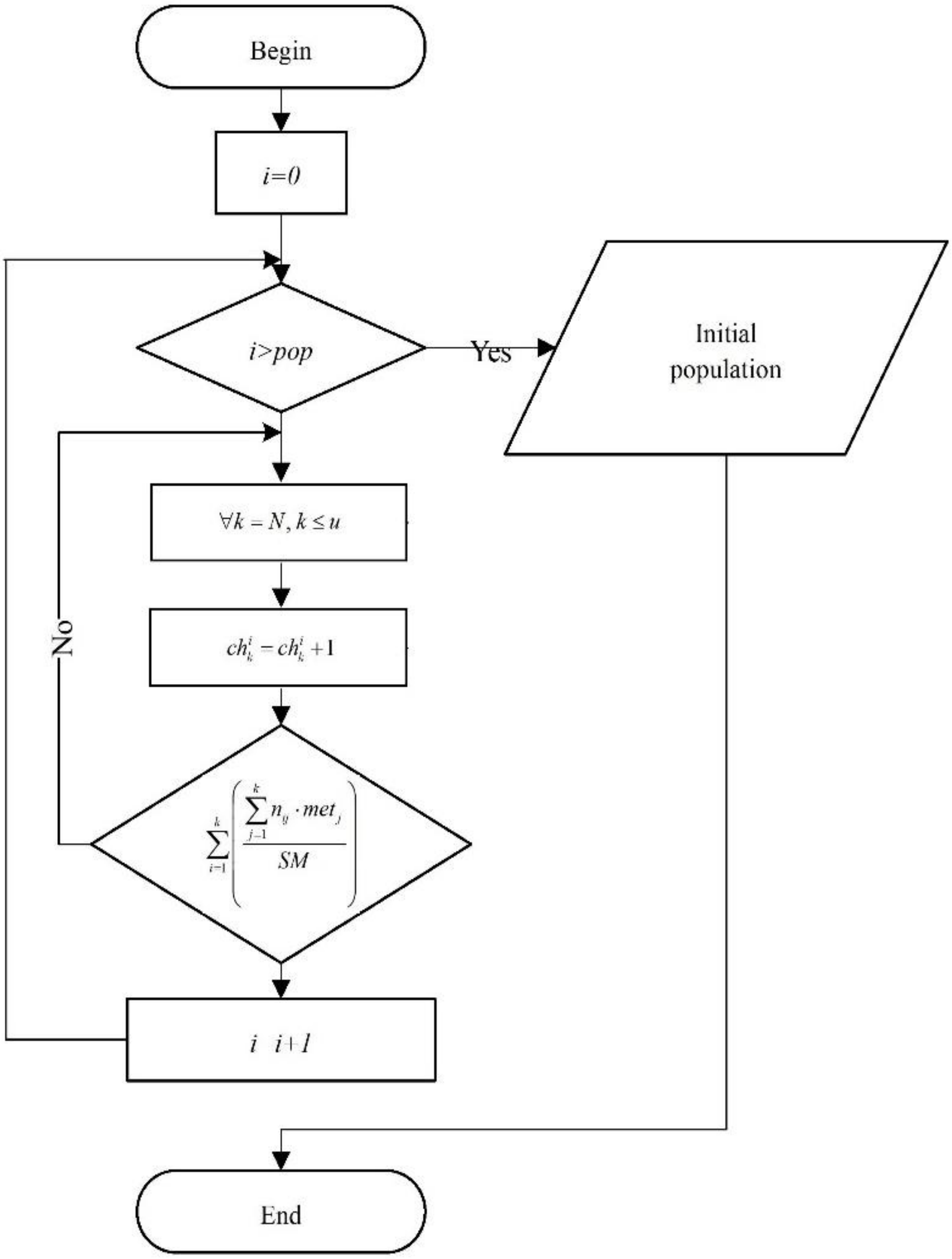

Figure 3: Block diagram of chromosome generation

We randomly select the number of the record in the list of IS H&S sets. Add 1 to the gene that corresponds to this set.

Then we check that chromosome (ch) meets constraints (3) and (4).

We repeat the procedure shown in Fig. 3 until the required indicators for the IS metrics for the analyzed OBI are achieved. In a specially created data structure we enter the numbers (NG) of generations of chromosomes (ch).

The size of the population depends on the number of chromosomes. For each chromosome (ch) in the population, fitness is assessed by calculating the fitness function. The lower the value of the fitness function will be - the quality of the chromosome will be higher. At the next step of the modified GA operation, see Fig. 4, we sort the obtained values. It is so-called rank selection [8–13].

The general scheme of the GA is shown in Fig. 4.

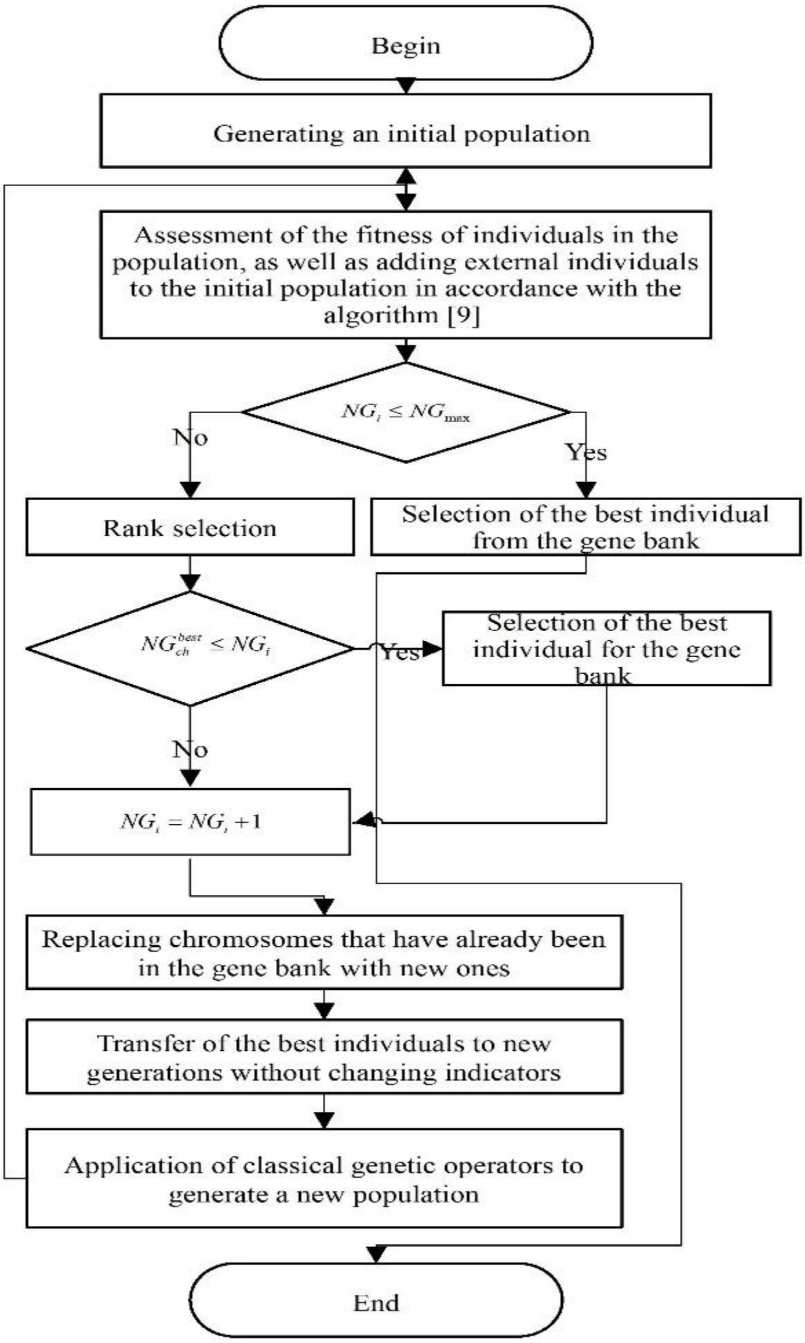
Begin
Generating an initial population
Assessment of the fitness of individuals in the population, as well as adding external individuals to the initial population in accordance with the algorithm [9]
$NG_i \leq NG_{\max}$
No
Yes
Rank selection
Selection of the best individual from the gene bank
$NG_{ch}^{best} \leq NG_i$
Yes
Selection of the best individual for the gene bank
No
$NG_i = NG_i + 1$
Replacing chromosomes that have already been in the gene bank with new ones
Transfer of the best individuals to new generations without changing indicators
Application of classical genetic operators to generate a new population
End

Figure 4: Modified GA for selection of hardware and software information security

Crossing or crossing over is the exchange of chain fragments between two parental chromosomes. In accordance with the block diagram of the algorithm shown in Fig. 4, the partition point is chosen randomly. Next, we attach the left side of the first chromosome in a pair to the right side of the second chromosome. Accordingly, we attach the left side of the second chromosome to the right side of the first chromosome.

We carry out selection for each generation.

We select "viable" individuals on the basis of constraint (4). Then the ranking is performed by the value of the fitness function (1).

The best individuals are transferred unchanged to the next generations.

The computation ends when the specified number of generations is reached. As shown by computational experiments, the convergence of the algorithm is achieved for at least fifty generations.

To check the adequacy of the model described in the work, the corresponding computational experiments were carried out, see Fig. 5.

Computational experiments were carried out for randomly generated IS H&S sets to protect OBI circuits. The efficiency of the three algorithms was compared, see Fig. 5.

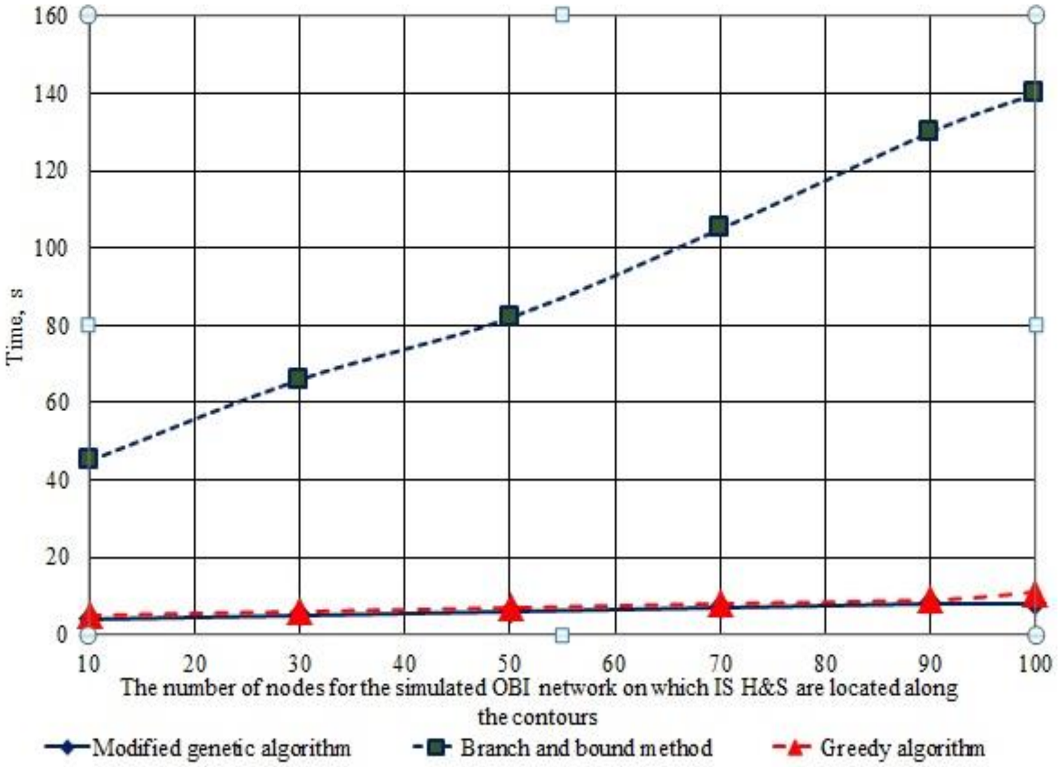

Figure 5: Results of computational experiments comparing the running time of algorithms

In the course of computational experiments, it was found that the modified GA is distinguished by a sufficiently high efficiency and speed. It was found that the time spent on solving the problem when using the modified GA described above is approximately 15–20 times less in comparison with the indicators of the branch and bound method. This circumstance allows, in the future, when finalizing the decision support system, to opt for this particular algorithm.

4. Conclusions

A modified genetic algorithm (MGA) was proposed; it can be applied in the problem of hardware and software selection for information protection and ensuring information security of informatization objects. In contrast to existing solutions, it was proposed to use a different coding method, as well as to use an elite strategy, selecting the best individuals for the gene bank. The use of a gene bank allows one to reduce the number of generations in the search for the solution. Moreover, it leads to a general reduction in the operating time of the MGA.

5. References

[1] Z. Chiba, et al., New anomaly network intrusion detection system in cloud environment based on optimized back propagation neural network using improved genetic algorithm. International Journal of Communication Networks and Information Security 11(1) (2019) 61–84.

[2] Y. Nozaki, M. Yoshikawa, Security evaluation of ring oscillator PUF against genetic algorithm based modeling attack. In: International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, 2019, pp. 338–347. doi:10.1007/978-3-030-22263-5_33.

[3] S. Dwivedi, M. Vardhan, S.Tripathi, Incorporating evolutionary computation for securing wireless network against cyberthreats. The Journal of Supercomputing (2020) 1–38. doi:10.1007/s11227-020-03161-w.

[4] F. Zhang, et al., Multilayer data-driven cyber-attack detection system for industrial control systems based on network, system, and process data, in: IEEE Transactions on Industrial Informatics 15(7) (2019) 4362–4369. doi:10.1109/tii.2019.2891261.

[5] U. Baroudi, et al., Ticket-based QoS routing optimization using genetic algorithm for WSN applications in smart grid. Journal of Ambient Intelligence and Humanized Computing (2019). doi:10.1007/s12652-018-0906-0.

[6] T. Llansó, M. McNeil, C. Noteboom, Multi-criteria selection of capability-based cybersecurity solutions. In: 52nd Hawaii International Conference on System Sciences, 2019, pp. 7322–7330. doi:10.24251/hicss.2019.879.

[7] V. Lakhno, et al., The use of a genetic algorithm in the problem of distribution of information security organizational and financial resources, in: 2020 IEEE 2nd International Conference on Advanced Trends in Information Theory, ATIT, 2020, pp. 251–254, doi:10.1109/atit50783.2020.9349310.

[8] Lakhno, V., et al., Funding Model for Port Information System Cyber Security Facilities with Incomplete Hacker Information Available, Journal of Theoretical and Applied Information Technology 96(13), 4215–4225, 2018.

[9] D. K. Proskurin, K. A. Makoviy, Modified genetic algorithm for solving the problem of selecting server, resources in building the infrastructure of virtual desktops, Voronezh State Technical University, Voronezh, Russia, 2021, pp. 6–51. doi:10.36622/vstu.2021.17.3.006

[10] V. Buriachok, et al., Invasion Detection Model using Two-Stage Criterion of Detection of Network Anomalies, Cybersecurity Providing in Information and Telecommunication Systems (CPITS), pp. 23–32, Jul. 2020.

[11] V. Lakhno, et al., Allocation of organizational and financial resources of the information protection side using a genetic algorithm, Lecture Notes in Networks and Systems (2021) 41–53. doi:10.1007/978-3-030-77448-6_5.

[12] K. Khorolska, et al. Usage of clustering in decision support system. In: Raj J.S., Palanisamy R., Perikos I., Shi Y. (eds), Intelligent Sustainable Systems. Lecture Notes in Networks and Systems, vol. 213, Springer, Singapore (2022). doi:10.1007/978-981-16-2422-3_49.

[13] B. Bebeshko, et al., Use of neural networks for predicting cyberattacks, in: Cybersecurity Providing in Information and Telecommunication Systems, vol. 2923, 2021, pp. 213–223.