Jarosław JAJCZYK

Poznan University of Technology, Institute of Electrical Engineering and Electronics

Optimization calculations with the genetic algorithm method on a computer cluster

Abstract. *The work presents a way to execute optimization calculations with the genetic algorithm method on a parallel computer of the cluster type. The scope of electromagnetic calculations which should be performed while determining the objective function was outlined. Sample indicators describing the quality of paralleling the calculation process were provided.*

Streszczenie. *W pracy przedstawiono sposób realizacji obliczeń optymalizacyjnych metodą algorytmu genetycznego na komputerze równoległym typu klaster. Zakreślono zakres obliczeń elektromagnetycznych, które należy wykonać podczas wyznaczania funkcji celu. Podano przykładowe wskaźniki opisujące jakość zrównoleglenia procesu obliczeniowego (**Obliczenia optymalizacyjne metodą algorytmu genetycznego na klastrze komputerowym**).*

Keywords: optimization, genetic algorithm, computer cluster, parallel computing.
Słowa kluczowe: optymalizacja, algorytm genetyczny, klaster komputerowy, obliczenia równoległe.

doi:10.12915/pe.2014.04.56

Introduction

Using computers in technical applications have become common. Contemporary computing machines are used in solving various problems regarding the design of equipment, selection of components, or determining optimal operation parameters. Thanks to computer technology, calculation results are obtained faster and with a higher level of precision [1,2,3,4,5]. What is more, considerable computation power of mobile devices makes it possible to apply them more and more commonly. Thanks to that, more mobility of the calculation stands may be obtained and wireless communication allows for cooperation with other devices. The possibility to use computers for numeric calculations becomes particularly helpful in the context of complex problems. In such cases, an analytical solution often cannot be achieved and numeric calculations, although they do lead to the result, mean a long waiting time before the solution is obtained due to the complexity of the calculations themselves. Using parallel computers becomes an advantageous solution in such situations. This approach is particularly helpful while executing optimization during which calculations are repeated multiple times in order to identify the optimal solution. This, of course, has a negative influence on the calculation time.

Paralleling the calculations, including optimization calculations, is often difficult to implement due to the complexities resulting from the need to exchange data between the calculation units. Often the time costs connected with the communication between computers (processors) lead to a reverse effect instead of shortening the waiting time before the solution is obtained. An efficient solution of this problem in optimization calculations is a combination of parallel calculations and optimization with the use of a genetic algorithm (GA). In many cases, proper configuration of such an algorithm limits the data exchange between particular calculation units to a level that has practically no influence on the waiting time before the solution is obtained [2,5,6,7].

The work presents the results which reflect the quality of paralleling optimization calculations of a three-phase busduct. The subject of optimization was its geometrical dimensions. The genetic algorithm method was used as the optimization method. A computer cluster formed by PC computers connected with a broadband computer network was used to run the calculations. The efficiency of the implementation of the optimization problem analyzed with the use of a genetic algorithm on a computer cluster was examined.

Organization of optimization calculations

Optimization with the use of the GA method involves defining the quality factors for the solutions (individuals) that form a specific group (population) of solutions on the basis of which the next population is built [6,7]. What is characteristic of the GA method is the fact that the calculations of the fitness of particular individuals are performed independently of one another. This fact can successfully be used to parallel the algorithm.

The following could be distinguished among different methods of paralleling a genetic algorithm: a centralized synchronous organization, a centralized semi-synchronous organization, a distributed asynchronous organization, a network organization, a community organization, and a pollen organization [7]. The use of some of those methods results in the need to assure uninterrupted communication and sending large amounts of data between the calculation nodes. The first two of the methods listed above are characterized by the lowest time costs connected with the communication among particular computers.

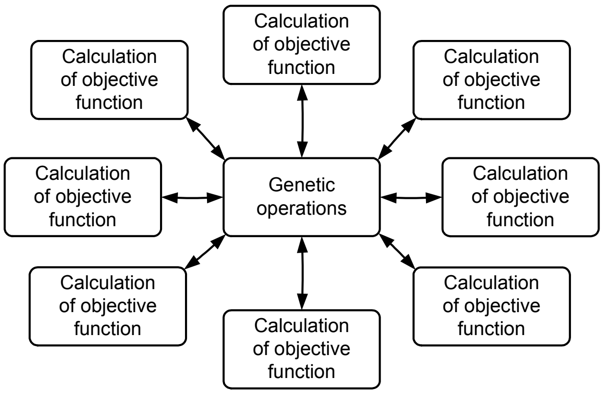

Fig.1. The organization of the parallel genetic algorithm

A synchronous centralized organization of a parallel GA algorithm was used in the present work (Fig. 1). In this case, one computer on which the main process is run is responsible for all genetic operations and for distributing the tasks to other calculation units which are responsible for the identification of the fitness characteristics of particular individuals [2,5,6,7].

Thanks to the fact that computation power consumption of the main process is low in comparison to the subordinate processes, it is possible to run fitness function calculations on the main computer as well.

A block diagram of the parallel genetic algorithm is presented on figure 2.

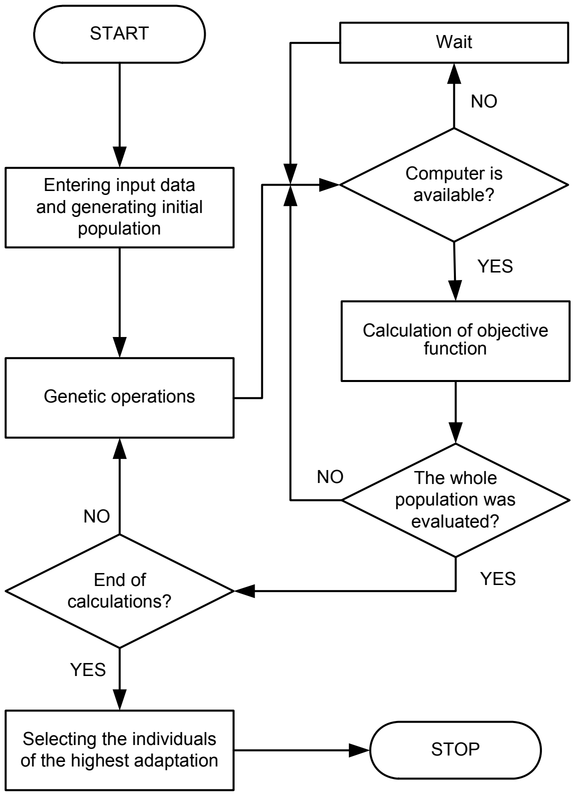

Fig.2. Block diagram of the parallelized genetic algorithm

Most of the algorithm is performed on the main computer managing the calculations and responsible for all the genetic operations. The tasks connected with calculating the objective function for particular individuals are assigned to subordinate computers. In such an approach, communication between the main node and the calculation units is limited to the transfer of new task parameters and the collection of the calculation result which is a single number (the fitness factor).

Parallel quality measures

The quality of paralleling the optimization process performed with the use of the genetic algorithm method can be determined with one of the basic parameters which characterize parallel algorithms – the task acceleration factor (1) [1]. Its value is equal to the relation of the time needed to complete the task of the size n on one processor to the time needed to complete the same task on p processors and it is calculated on the basis of the following dependency(1):

(1) $$S(n,p)=\frac{T(n,1)}{T(n,p)}$$

where: $T(n,p)$ – the time needed to complete the task of the size n on p processors, n – task size, p – number of processors.

What is more, while evaluating the quality of parallel calculation, its effectiveness (efficiency) expressed as a percentage value and equal to the relation of the acceleration factor of the task with the size of n on p processors to the number of processors p is determined (2) [1].

(2) $$S_{\%}(n,p)=\frac{S(n,p)}{p}\cdot 100\%$$

The efficiency of parallel calculation never reaches 100% due to the cost of communication between the calculation units.

The optimized object and the objective function

The purpose of using a parallel genetic algorithm was tested on the example of the optimization of the geometrical dimensions of a three-phase unshielded busduct with solid insulation. Its cross section is presented on figure 3.

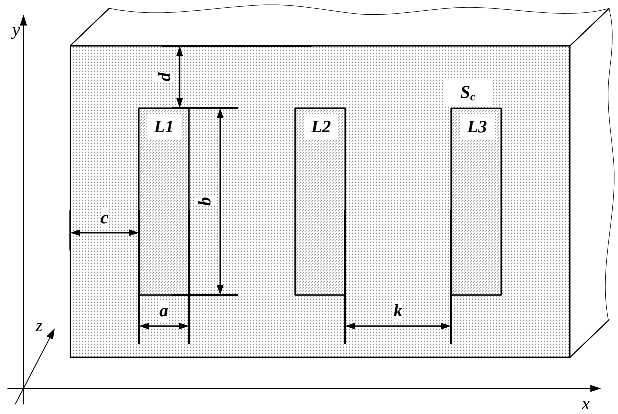

Fig.3. Cross-section of the high current busway with permanent insulation

Phase conductors, each of a S_c cross-section area, are embedded in solid insulation, made out of a component of epoxide resins. The geometry of the system is conditioned by five variables: a, b – dimensions of the cross-section of a phase conductor; and c, d, k – dimensions determining the distribution of conductors in the insulation.

All electrodynamics calculations start with defining the distribution of current density $J(x,y)$ in live working conductors with specified phase currents [2]. It can be obtained by solving the system of integral equations (3).

(3) $$\underline{J}(x,y)=-\frac{1}{2\pi}j\omega\mu\gamma\int_{S_c}\underline{J}(x',y')\ln\frac{1}{\sqrt{(x-x')^2+(y-y')^2}}dx'dy'$$
$$\int_{S_c}\underline{J}(x',y')dx'dy'=\underline{I}_c$$

where: μ – magnetic permeability of the conductor material; ω – pulsation; γ - electrical conductivity of the conductor material; (x, y) – the observation point; (x', y') – the source point; S_c - cross-section area of the conductor.

Knowing the distribution of current density makes it possible to determine power losses in phase wires on the basis of Joule's law. Temperature distribution in the system is determined on that basis. It is of crucial influence on the geometrical dimensions of the busduct which determine its ability to emit heat. Also electrical stress and the forces operating in the system are included in the calculations. Details regarding the solution of the simultaneous equations (3) and the identification of other electrodynamic parameters are provided in the work [2].

Minimizing the costs of manufacturing and use of the object over a set period of time while satisfying a set of limitations was assumed as the optimization criterion. The objective function (4) is of financial character and it is a function of geometrical variables which influence the size of the busduct cross section (investment costs) as well as the value of active power losses (exploitation costs).

(4) $$S(\mathbf{u}) = k_{invest.} + k_{exploit.}$$

where: $\mathbf{u}$ – decision variable vector; $k_{inwest.}$ – investment costs, $k_{exploit.}$ – exploitation costs

The objective function $S(\mathbf{u})$ that is minimized in the optimization process must meet a series of limitations. The most important of them include: the maximum temperature of the service wire and of the insulator, maximum electrical stress, maximum forces exerted under stable and short-circuit conditions and standard requirements regarding, for example, the consequences of short-circuit current [2].

Calculation results

As part of the work, optimization of the geometrical dimensions of an unshielded three-phase busduct in solid insolation was performed with the purpose of identifying a busduct whose dimensions would constitute a compromise between exploitation costs (transmission losses) and material costs (investment costs). Details regarding the optimization of a busduct of this type were provided in [1,2]. The following values were assumed in the calculations: phase current – 6 kA; wire voltage – 15 kV; exploitation time – 10 years; the wires are made of copper and the insulator is made of epoxy resin.

Calculations were performed on a computer cluster formed of 12 computers connected with a broadband network. The hardware characteristics of the computers used are provided in table 1.

Table1. Characteristics of the computers used to build the cluster

Processor	Operating system	RAM memory	Number of units in the cluster
Intel i7-2600, 3,4 GHz	MS Windows 7 (64bit)	16 GB	3
Intel i5-3450, 3,1 GHz	MS Windows 7 (64bit)	8GB	9

Calculations with the use of a genetic algorithm were performed for 40 generations. Every population consisted of 50 individuals. The calculation time (Fig. 4) and the fluctuations of the acceleration factor and of the paralleling efficiency depending on the number of nodes (Fig. 5) were analyzed.

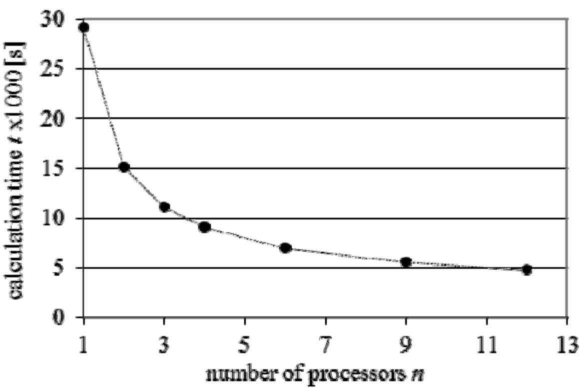

Fig.4. Calculation time depending on the number of computers

Firstly, computers equipped with the Intel i7 processor (also the main process was run on such a processor) and then the remaining computers were connected.

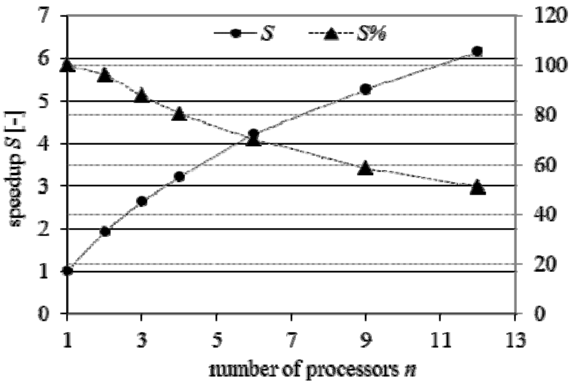

Fig.5. Calculation acceleration and paralleling efficiency depending on the number of computers

Conclusions

A genetic algorithm is an efficient method of searching for the optimum in the global sense. Thanks to the application of parallel calculations, calculation time reduction from over 8 hours in the case of 1 computer to 1,3 hours for 12 computers was obtained (Fig.4.).

Increasing the number of computers comprising the cluster accelerates the optimization process; however, the acceleration level does not increase proportionally to the number of computers (Fig. 5). It was observed that when 12 computers were used, the acceleration level slightly exceeded the rate of 6. This is reflected in the investigation of the paralleling efficiency factor whose value reached about 50% for twelve computers taking part in the process of solving the optimization task. Such a low value of the factor is a consequence of the use of a synchronous, centralized organization of the parallel genetic algorithm. High time loss resulting from the need to synchronize the calculations between particular generations leads to the decrease of the quality of paralleling. Further work should be focused on improving this factor.

REFERENCES

[1] Bednarek K., Kasprzyk L., Speeding up of electromagnetic and optimization calculations by the use of the parallel algorithms, *Przegląd Elektrotechniczny*, 85 (2009), nr 12, 65-68
[2] Jajczyk J., Optimization of the geometry of high currency busducts with solid insulation with the use of the modified genetic algorithm method, *PhD thesis, Poznan University of Technology*, (2008)
[3] Ramirez J. A., Saldanha R. R., Takahashi R. H. C., Vasconcelos J. A., Improvements in Genetic Algorithms, *IEEE Transactions on Magnetics*, 37 (2001), No. 5, 3414-3417
[4] Skowronek K., Trzmiel G., The method for identification of fotocell in real time, *Przegląd Elektrotechniczny*, 83 (2007), nr 11, 108-110
[5] Wyrzykowski R., PC computer clusters and multi-core architectures. Structure and use, *Exit*, (2009)
[6] Fogel D. B., Evolutionary Computation: Toward a New Philosophy of Machine Intelligence, *IEEE Press/Wiley-Interscience*, (2006)
[7] Goldberg D.E., Genetic Algorithms in Search, Optimization and Machine Learning, *Addison-Wesley Publishing Company*, (1989)

***Author**: dr inż. Jarosław Jajczyk, Politechnika Poznańska, Instytut Elektrotechniki i Elektroniki Przemysłowej, ul. Piotrowo 3a, 60-965 Poznań, e-mail: jaroslaw.jajczyk@put.poznan.pl*