A Meta-Heuristics Algorithm for the Global Optimization of Expensive, Simulator Evaluated, Objective Function

Przemysław Miazga

Warsaw University of Technology, Faculty of Electronics and Information Technology
Institute of Radio Electronics, Warsaw, Poland, e-mail: p.miazga@ire.pw.edu.pl

Abstract. This paper presents the idea of a meta-heuristics algorithm called Evolutionary Controlled Clustering Algorithm (ECCA) designed for implementation in simulation optimization. The method focuses on localization of function optima neighborhoods. A Evolutionary algorithm (EA) with soft selection and gene injection is used for finding basin of attraction. It operates over nodes of a grid created in a continuous parameter space. ECCA manipulates the grid density as well as the simulation accuracy. Clustered data is used for identification of the basin of attraction. Later, surrogate optimization is applied for local optima search. ECCA was optimized for operation in an uncertain and dynamically changing environment of simulation data. It was tested on the design of the shape of waveguide transition. The computer program can be executed concurrently on a multi-processor machine or on a grid of computers.

1 Introduction

Nowadays most engineers become aware of benefits resulting from evaluation of the appropriate solution in a reasonable time with reasonable accuracy. When the time taken to perform single characteristic valuation, or in other words objective function, is of the order of minutes, hours, days, … the total number of evaluations is limited by contract constraints. Also our financial and computer resources are limited. Resource restriction poses a special challenge to most global optimization methods, since those algorithms require from hundreds to some thousands of objective function evaluations [1,2]. In place of simulation data we may use the surrogate of objective function. It can be obtained by fitting the data evaluated from several results of simulation [3], available at the moment, to a given response surface. The interpolating response surface can be created with the help of the stochastic process [4,5,6]. The computer program [2,7] based on the surrogate algorithm may decrease computation time several times. But not in all cases. See the analysis of the operation of the surrogate algorithm [7] presented in section 2. Resulting from this case the conclusion is simple. Instead of simulating accurately with large cost for optimization space we may apply an algorithm which starts from coarse simulation, increasing accuracy, while getting closer to the optimal solution. This wrinkle lies as a background to the Evolutionary Controlled Clustering Algorithm (ECCA), which is presented in section 3. EECA combines particular potential of EA as a tool for global optimization of noisy function (as reported in eg. [8]) with efficiency of function estimation.

2 Application of surrogate optimization algorithm

It is very hard to find a practical example of a real-world device suitable for an accurate simulation (though we need exact data for problem analysis) in a reasonable time. The testing method was carried out on the problem of adjusting the shape of the waveguide transition [2,9]. The waveguide is a section of a rectangular (in cross-section) pipe, which is used for guiding waves in the high frequency (microwave) range. The whole device consists of three sections of waveguide: input port (23 mm * 10 mm), middle section (23 mm * w) with varying length l and width w, and output port (23 mm * 6 mm). The parameter optimal values equal to 9.5 mm and 7.8 mm for the length and the width respectively.

Objective function, being absolute value of the reflection coefficient, is evaluated by solving numerically three-dimensional Maxwell Equations set [1], using a wave simulator package [9]. The area covering the considered device is decomposed into a set of small cube cells (process which is called meshing). Then electromagnetic fields are evaluated iteratively along the mesh with a fixed time interval. Discrete and iterative character of the algorithm limits overall performance. The accuracy of simulation result is regulated by a change of mesh size. Mesh size decrement by a factor of two may decrease systematic simulation error four times. However simulation time will increase approximately 16 times. The comparison of necessary resources for waveguide transition is presented in the following table. Computation times are given for AMD Sempron 3000+.

Table 1. Simulation parameters

Mesh id.	S	A	B	Ref.	D
Mesh size	2mm	1mm	0.5mm	0.25mm	0.12 mm
Run time	0.4s	3.2s	13.2s	1m	12m
RAM size	0.5MB	1MB	4MB	23MB	460MB

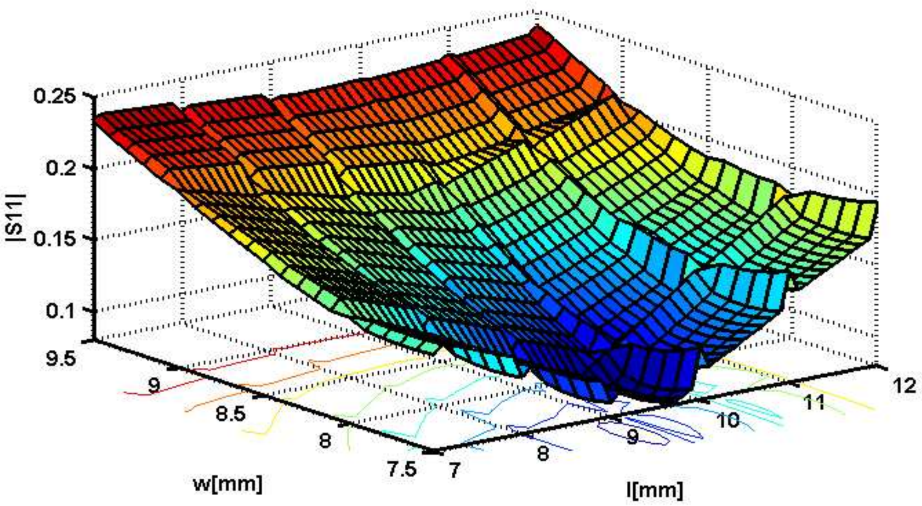

Figure 1. Objective function chart with simulation mesh size = 1 mm (A)

The theoretical shape of the objective hyper-plane is smooth and convex. The shape of the simulator evaluated objective, as shown on Figure 1, embodies plenty of "canyons". Walls of the "canyons" are caused by the systematical error. Also some parasitic optima had appeared. Anyway we may still hope that the search landscape is locally smooth.
The amplitude of the systematic noise decreases with improving simulation accuracy (compare Figure 1 and 2), but with unreasonable growth of execution time and with a need for more and more operational memory. Also, with growing amount of necessary computer resources the simulation process is more likely to hang up, crash or become unstable. So in practice we may never have certainty about simulation result.
The meshing with mesh size = 0.5 mm (B) used for creation of chart presented on Figure 2 is recognized by field simulation practitioners as the best compromise between accuracy and resource utilization for the given example. It was used for algorithm tests.

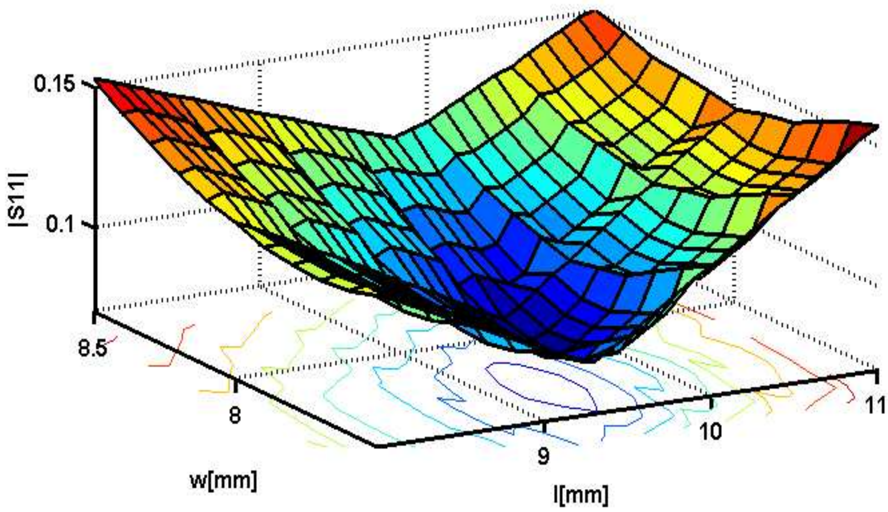

Figure 2. Enlarged optimum region with simulation mesh size = 0.5 mm (B)

The evolutionary algorithm with response surface approximation [7] was applied for optimization. The profit is a reduced execution time of about 50%, compare to the EA or non-gradient Powell method [9] . The solution with 1% accuracy was achieved quite quickly after 4^{th} generation (40 function evaluation). However with growing number of generations the approximating function apparently becomes multimodal. The quality of estimator can be seen in Figure 3.The operation of the evolutionary algorithm was disturbed by the three and four local optima (in 4^{th} and 8^{th} generation respectively). All that optima must have been verified with some computational effort. In total optimization quits after some hundreds of generations.
The level of estimation error, received after 2^{nd} generation is comparable in value to the simulation data. With growing number of simulation the overlay estimation performance gets better, but still efficiency of optimization process is not so good as it was expected. Also numerical effort, used for evaluation of estimator, increases exponentially.

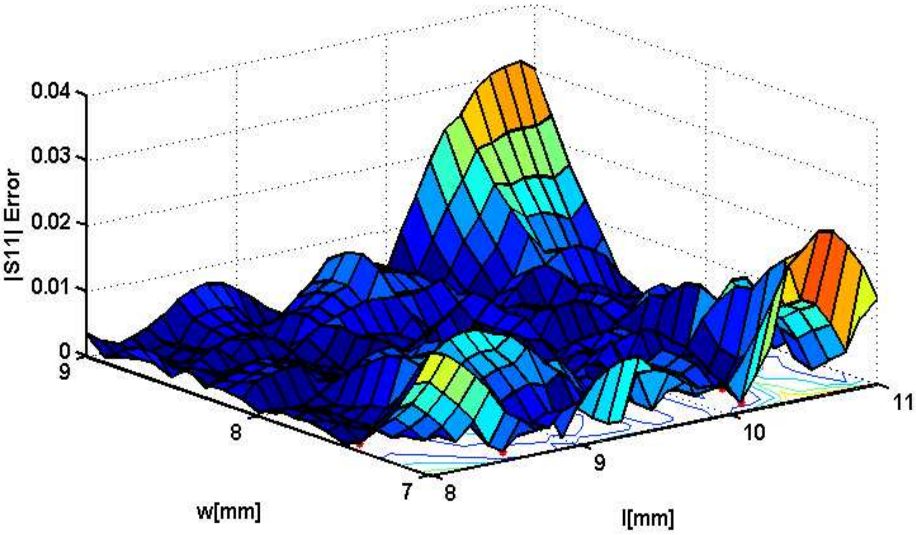

Figure 3. Estimation error after 2nd generation with simulation mesh size = 0.5 mm (B)

The ridged shape of estimator error, and poor convergence make us suspicious about problem conditions. Let's magnify objective surface and see whether all assumption concerning objective function are fulfilled. The cross section of the objective surface close to the global optima (for width w = 7.8043 mm) is shown by Figure 4.

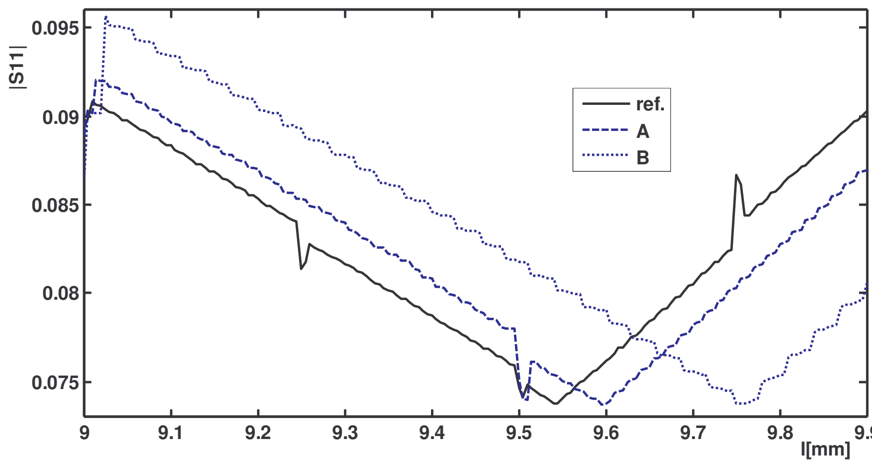

Figure 4. Comparison of the objective function for three mesh sizes.

The simulation error constitutes a quasi-periodic function with recurrent peaks. The period of the function equals the mesh size. The optima of the less accurate simulation B and A are shifted compared to the original by 2% and 0.6% respectively . Anyway we may still doubt the accuracy of our reference data. The comparison with “closer to reality” (set of points D) objective function view is presented in Figure 5. Further mesh size improvement seems to be unpractical.

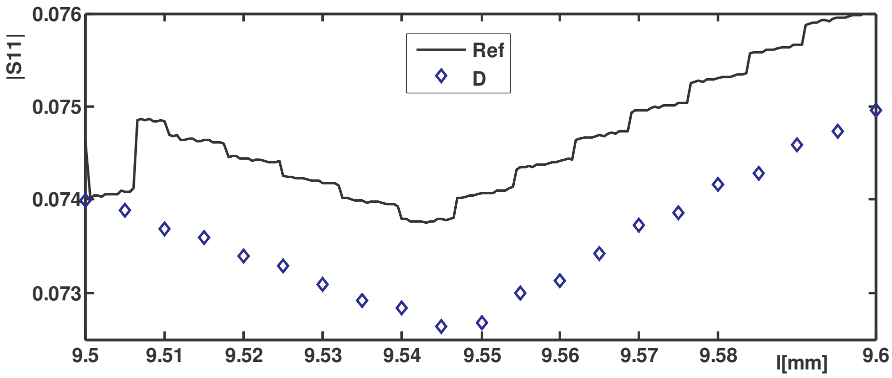

Figure 5. Magnified optima surrounding.

With magnified view we can observe that not only surface is a stepped like line distorted by large peak (see one on the left side of Figure 5), but there appear smaller ripples. So the principle constraint of the application of the response surface method, that the objective must be locally smooth and convex, is not fulfilled. The un-smoothness is presumably caused by the input/output data rounding, use of 32 bit numbers (64 and 80 bit numbers will at least double demand for computer memory) and some other unknown, at the moment, factors – lets call it numerical noise. With dense meshing (like D) we may ignore the influence of the numerical noise, but still quasi-periodic peaks may cause generation of parasitic optima on the approximation surface. Concluding, we may spend a lot of time for simulation, expend a lot of resources (that is hoped are available), with uncertain result.

3 Evolutionary Controlled Clustering Algorithm

The key issue in global optimization is the problem of finding a variety of basins of attraction of the optimum. And a side effect, which follows, is poor efficiency in the space exploitation by meta-heuristic and stochastic global optimization algorithms like simulated annealing, evolutionary algorithms, controlled random search and many others. These methods are rather capable of space exploration. Given either cluster covering optima or a set located close to one we may apply any local search algorithm for accurate location of the optima. During exploratory stage there is no need for a high simulation accuracy. Also, we may tighten function domain to a sparse grid. Later, during exploitation the grid must be dense and simulation very precise. Also we must keep in mind that most of computational effort during local optimization is consumed by the final optimum location. At that stage we can take the advantage of surrogate optimization. Flowchart of the program is shown in Figure 6.

The principle of operation of global simulation optimization algorithm, taking in consideration presented ideas, is constituted as follows:
- first locate interesting areas (clusters) using global optimization algorithm,
- then narrow observation horizon to the regular area surrounding given cluster,
- increase observation resolution by making grid denser,
- simulate function with medium accuracy over given grid,
- evaluate locally function estimator,
- find minimum of the estimation function as a starting point for further optimization,
- minimize function locally using steepest descent algorithm with estimator utilization.

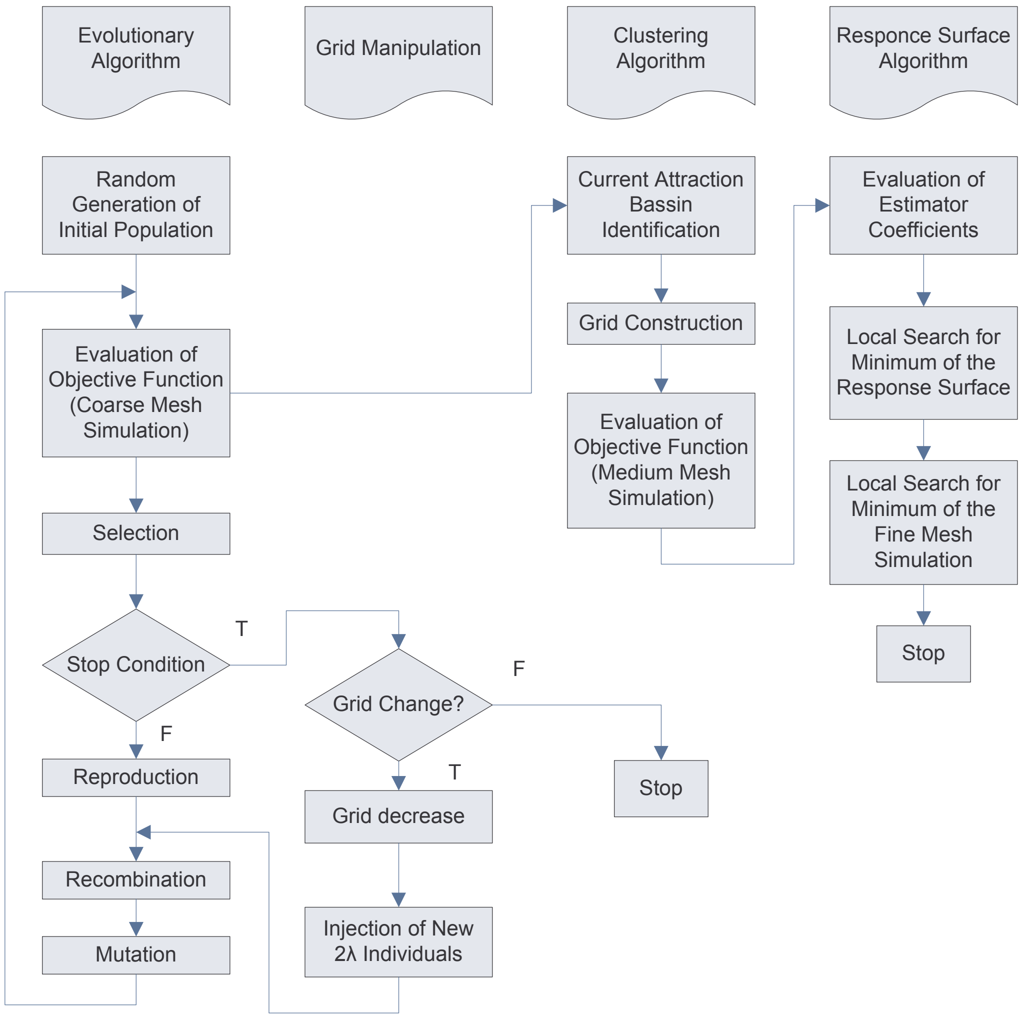

Figure 6. The flowchart of the program

For initial cluster search we can use any global optimization procedure. With problems of relatively small dimensionality, as presented here, we can evaluate all objective values over a grid. It will still cost less than a single precise simulation. In general the $\mu + \lambda$ evolutionary algorithm, with soft selection was chosen. The evolutionary operators are mutation with Gaussian distribution and linear recombination. Process starts with grid size = 0.5 mm and mesh size = 1 mm. Population size is set to $\mu = 10$ and $\lambda = 5$. Presented EA in practice saturates after a couple of generations. Therefore the stop condition is defined as exceeding a generation count (5-10) with no improvement of the best fitness. The grid is then gradually decreased until predefined limit is reached (0.1mm). Meanwhile current population is overridden with 2λ new randomly generated individuals and λ is set to 2λ. The injection with growing number of new points preserves population diversity.
The population is looked through after every generation for clusters covering optima. The clustering algorithm is designed to find convex (in a discrete manner) sets of points located within some aperture. The aperture measure depends on current grid size and mesh size. The cluster points must fall inside some regular region (hyper-cube or hyper-sphere) of function domain and its value inside an interval whose limits depend on the best fitness and estimate of maximal systematic simulation error. After recognition of the cluster the algorithm starts (concurrently) local search. Firstly, values of the cluster grid nodes are simulated with medium accuracy. Secondly, for each mate of the cluster we locate minimum of the surrogate objective function. Finally the mentioned surrogate optima is used as a starting point by steepest descent direction algorithm with precise simulation. At the final stage the space grid is set to the value of manufacturing accuracy, typically 1 μm.
The comprehensive verification and comparison of simulation optimization algorithms is very difficult due to the high computational time, huge demand for resources and uncertainty of simulation. ECCA is in fact an improvement of some of existing methods [2,6,7,10]. It inherits its advantages and drawbacks. Several tests on examples attached to the QuickWave® simulator [9] proved that ECCA is able to find optimal solution in every case. Computational efficiency is arguable. In some cases ECCA was more efficient, in some cases not. However we must emphasize that dimensionality of examples was relatively low (2-4) with no more than 4 local optima. In some cases it was observed that accuracy of the solution in the first generation was good enough to start local search. Anyway we must point out that with increasing number of variables ECCA performs more efficient and reliable than the other methods. Optimization of more sophisticated devices, with larger dimensionality, require more resources than available. Presumably a computer grid may help at the moment. However there is no available version of the simulator designed for a grid computing.

4 Summary

Evolutionary algorithm with response surface interpolation proved to be a smart and quite reliable tool for global optimization of expensive, simulator based function. However it can be efficiently applied only in cases, when systematic simulation error is neglible or can be compensated. Mentioned conditions hold only if we are able to use a very precise and reliable simulation.
The Evolutionary Controlled Clustering Algorithm implements surrogate optimization in the place where it is most efficient – for exploitation. In the initial exploration we may significantly increase number of simulations taking full advantage of evolutionary algorithm. ECCA seems to

be an efficient and reliable tool for simulation optimization. The main advantages of the method must be emphasized:

- the source of the data is mostly a "like real world" simulation process,
- simulation cost is relatively low, we can run simulation more times,
- with some data redundancy we are able to neglect uncertainty of simulation,
- simulations and estimation can run concurrently,
- synchronization of processes is not a critical issue.

Bibliography

[1] Miazga P., Gwarek W., *Improved Design of Passive Coaxial Components Using Electromagnetic 2-D Solver in an Optimization Loop*. IEEE Trans on MTT, vol 45, May, 1997, pp. 858—860

[2] Bandler J.W. et al. *Electromagnetic optimization of 3D structures.* IEEE Trans on MTT, vol 45, May 1997, pp. 770—779

[3] Carson Y., Maria A. *Simulation Optimization: Methods and Applications*. In proc. Of 1997 Winter Simulation Conference, 118-126.

[4] Currin C., Mitchell T., Morrin M., Yvisaker D. *A Bayesian Approach to the Design and Analysis of Computer Experiments.* Report : Oak Ridge National Laboratory, ORNL-6498, Sept. 1988.

[5] Jones D. R. *A Taxonomy of Global Optimization Methods Based on Response Surfaces.* Journal of Global Optimization 21, 2001, 345-383.

[6] Kohler J. B., Puchalski A. A., Simon B. *Estimating Functions Evaluated by Simulation: A Bayesian/Analytic Approach.* In The Annals of Applied Probability, 1998, Vol.8, No. 4, 1184-1215.

[7] Miazga P. *Global Optimization of Expensive Simulator-Evaluated Function Based on Response Surface Interpolation.* In Proc. of VIII KKAEiOG, Korbielów, May 2006, pp.151-156.

[8] Mathias K.,Whitley D., Kushma A., Stork C. *An Empirical Evaluation of Genetic Algorithms on Noisy Objective Functions,* Genetic Algorithms for Pattern Recognition, CRC Press, 1996, pp. 65-86.

[9] *QuickWave 3D electromagnetic simulator users manual.* „Qwed" sp. z o. o. Warszawa

[10] Knowles J. ParEGO: A Hybrid Algorithm With On-Line Landscape Approximation for Expensive Multi-objective Optimization Problems. IEEE Trans. on EC, vol.10, no.1, Feb 2005, pp. 50-66.