16th European Symposium on Computer Aided Process Engineering
and 9th International Symposium on Process Systems Engineering
W. Marquardt, C. Pantelides (Editors)

Towards a novel optimisation algorithm with simultaneous knowledge acquisition for distributed computing environments

Siyu Yang, Antonis Kokossis, Patrick Linke

Centre for Process and Information Systems Engineering, University of Surrey, Guildford, GU2 7XH, U.K.

Abstract

This paper reports on research into novel optimisation schemes for large-scale distributed computing environments that will enable data analysis and knowledge acquisition in the course of optimisation. The scheme incorporates concepts from the Simulated Annealing search strategy in order to ensure robustness. In contrast to Simulated Annealing, which is a sequential optimisation algorithm, the proposed optimisation scheme consists of a number of solution pools, each of which is associated with a system temperature which defines solution quality within the pool. The solutions in these pools are generated by performing constant temperature Markov processes on existing solutions in these pools. As the individual Markov processes are independent they can be completed in large-scale distributed computing environments, constantly producing new solutions which are stored in a central database. During the optimisation, the solutions are regularly reassigned to pools according to their performance relative to the other solutions that have been generated such that the solution quality improves towards the pool associated with the lowest temperature. This final pool accumulates the set of optimal solutions during the optimisation. The solutions of all pools are stored in a central database from which knowledge about the importance of individual solution features can be extracted in the context of the systems performance.

Keywords: Optimisation, distributed computing, knowledge acquisition.

1. Introduction

Current optimisation methods are of limited use for decision-support in complex systems due to two main short-comings. Firstly, they require long computational times to identify optimal solutions to complex problems. The algorithms are not easily parallelised for use in large-scale distributed computing environments as transitions from initial towards optimal solutions are largely sequential. Distributed environments become increasingly available with the advent of Grid Technologies and new generations of optimisation methods are required that can exploit the vast available distributed computing resources effectively. Secondly, the results obtained from optimisation runs are often difficult to interpret by the user in the context of the decisions to be taken. This is particularly true for stochastic optimisation methods, which tend to be very robust in addressing complex optimisation problems, where important solution features are often blurred by features not strongly impacting on the systems performance. An optimisation algorithm that could exploit large-scale distributed systems and provide the user with optimal solutions alongside insights into the importance of individual solution features would be very attractive for decision-support.

We have devised a novel optimisation algorithm with the aim of addressing the above shortcomings. The algorithm incorporates concepts from the stochastic optimisation strategy Simulated Annealing to enable robust optimisation, whilst doing away with the inherently sequential nature associated with this meta-heuristic-based search scheme. This sequential nature of the search has hampered previous efforts of parallelising Simulated Annealing algorithms. As a result, past developments allow only minor distribution of computations and the number of processors that can be utilised in optimisation is severely limited. The number of processors that can be deployed depends upon the length of the homogenous Markov chain to be executed at a given system temperature which is also an essential parameter to influence the performance of optimisation (Leite and Topping, 1999). The limited ability of the algorithm to exploit vast distributed computing resources presents a major deficit that prevents the exploitation of advances in computing infrastructures in the form Grids (Antonopoulos et al., 2005). The proposed new algorithm will enable the full exploitation of such resources.

Besides allowing the large-scale distribution of the optimal search, the algorithm enables the analysis of information generated during the optimal search as all intermediate and optimal solutions are stored in a central database. It is therefore possible to device information mining schemes that allow the acquisition of knowledge about the individual solution features in the context of the solution performance and to identify those solution features that strongly impact on the systems performance.

The following sections outline the new optimisation scheme. An application to a global optimisation test problem is presented to illustrate the performance of the algorithm. The development of information analysis schemes that allow efficient knowledge acquisition for a number of process systems engineering problems is the focus of current research and will be reported separately.

2. Distributed optimisation algorithm development

2.1. Architecture

The architecture of the novel optimisation algorithm is shown in Figure 1. The algorithm features a number of pools, each of which is associated with a systems temperature that controls the distribution of solution quality. The highest temperature pool (T_1) accepts almost all possible solutions to the problem, whereas the lowest temperature pool (T_N) only admits solution of the highest quality. The algorithm is initialised by assigning a number of feasible solutions to the problem to the highest temperature pool (T_1). Agents randomly select pools and solutions and perform Markov processes at the corresponding pool temperature for each selected solution. All solutions visited during the Markov process are returned into the corresponding pool. As Markov processes are random operations, some solutions generated at a high temperature will be of high quality and would warrant membership of a lower temperature pool. This calls for a dynamic update of pool memberships based on the solution quality distributions defined by the pool temperatures. This is realised through periodic distribution of solutions amongst the pools. As the search progresses, more and more solutions will penetrate the lower temperature (high performance) pools and the algorithm is terminated as a sufficient number of solutions that warrant membership of the lowest temperature pool has been generated.

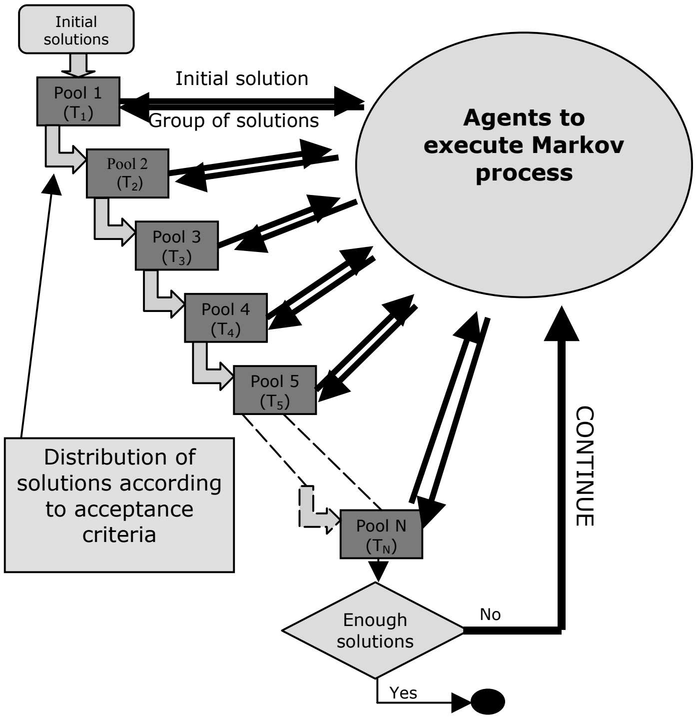

Figure 1 Novel optimisation algorithm

2.2. Pools and Agents

The concept of pools and agents allows the massive parallelisation of optimisation experiments as the agents will be able to constantly generate solutions at different pool temperatures which are stored in a central solution database that also stores information about the pool associated with a solution. In contrast to existing stochastic optimisation methods, there is no direct link to a solution from a previous iteration. This absence of successive transitions, which has hampered previous attempts to parallelise Simulated Annealing, enables massive parallelisation of the optimisation. The periodic distribution of solutions among pools can be performed in parallel to the execution of the Markov processes so that the idle times of the algorithm would be minimal.

2.3. Acceptance and termination criteria

The distribution of solutions among different pools requires acceptance criteria to decide on the membership of a solution in a given pool. For a minimisation problem, we accept a solution into a pool at temperature T if:

$$\mathrm{Exp}\left(\frac{\mathrm{Cur(So)} - \mathrm{Min(So)}}{\mathrm{T}}\right) \geq \mathrm{Rand} \qquad (1)$$

where Cur(So) is the objective function value of a candidate solution So to be distributed, Min(So) is the current best solution in the pool and Rand is a random number ($0 \leq \mathrm{Rand} \leq 1$). The acceptance criterion resembles the Metropolis criterion

employed in Simulated Annealing (Metropolis et al 1953) and has been implemented in the first instance.
The average solution quality and the quality distribution improves from the highest temperature to the lowest temperature pools. The lowest temperature pool therefore contains only the best solutions with the lowest distribution of solution quality. The more solutions are present in the lowest temperature pool, the higher will be the probability that the optimal solution has been found. The search is terminated once a specified number of solutions have entered the lowest temperature pool.

2.4. Prototype implementation

We have set up a small prototype system to test our algorithm. An SQL2000 database was set up to store the pools on our research center's central server. The agents, capable of obtaining an initital solution from a pool, executing a Markov process at the pool temperature, and returning a set of solutions into the pool, as well as the solution redistribution algorithm were coded in fortran 95 with fortransql library. The agents executed their Markov processes on a 731MHz Intel Pentium III processor. The PC and the server communicated via our local area network.

3. Illustrative example

We have tested the algorithm on five well-studied nonconvex nonlinear test problems given by Floudas et al. (1999). For lack of space, we can report on only one problem here:

$$\min\left\{\frac{(0.0039\times x7+0.0039\times x8)\times(495\times x4+385\times x5+315\times x6)}{x10}\right\}$$

subject to

$$- 0.5\times x9\times x4\times(0.8\times x7+0.33333333\ 3333333\times x8)+x1=0$$
$$- 0.5\times x9\times x5\times(0.8\times x7+0.33333333\ 3333333\times x8)+x2=0$$
$$- 0.5\times x9\times x6\times(0.8\times x7+0.33333333\ 3333333\times x8)+x3=0$$
$$\sqrt{x10 - x7} - (\sqrt{x8} - \sqrt{x9}) \geq 0$$
$$x1 - 8.46527343\ 75\times x10 \geq 0$$
$$x2 - 9.65006510\ 416667\times x10 \geq 0$$
$$x3 - 8.87167968\ 75\times x10 \geq 0$$
$$0.5\times x1\times x9 - 2.2\times(8.4652734\ 375\times x10)^{1.33333333\ 333333} \geq 0$$
$$0.5\times x2\times x9 - 2.2\times(9.6500651\ 0416667\times x10)^{1.33333333\ 333333} \geq 0$$
$$0.5\times x3\times x9 - 2.2\times(8.8716796\ 875\times x10)^{1.33333333\ 333333} \geq 0$$
$$x4 - 0.01117717\ 47883801\times x7 \geq 0.2$$
$$x5 - 0.01376553\ 60411427\times x7 \geq 0.2$$
$$x6 - 0.01556638\ 72253648\times x7 \geq 0.2$$
$$x4 - 0.01117717\ 47883801\times x8 \geq 0.2$$
$$x5 - 0.01376553\ 60411427\times x8 \geq 0.2$$
$$x6 - 0.01556638\ 72253648\times x8 \geq 0.2$$

Table 1. Effect of algorithmic parameters Poolnum and Markov on solution quality

Poolnum = 100			Markov = 100		
Markov	st	Av_Obj	Poolnum	st	Av_Obj
10	8.4E-05	-47.7063	10	2.01E-01	-46.4726
50	1.48E-04	-47.7061	50	1.52E-04	-47.7058
100	1.16E-04	-47.7061	100	1.16E-04	-47.706
500	2.81E-04	-47.7055	500	9.57E-05	-47.7061

Table 2. Comparison of the new optimisation algorithm with Simulated Annealing

Simulated Annealing			Novel optimisation algorithm				
CPU (sec)	Av_Obj	St	Markov	Poolnum	CPU (sec)	Av_Obj	St
1017.8	-47.7006	1.39E-03	500	100	102.0	-47.7045	5.86E-04

We studied the importance of the two key algorithmic parameters, the length of the Markov chains (Markov) and the number of pools employed (poolnum). The searches were terminated after at least ten solutions have penetrated the lowest temperature pool. The average objectives (Av-Obj) and the standard deviations (st) over all solutions in the lowest temperature pool are reported in Table 1. The performance of the algorithm clearly improves with the number of pools present as a result of a better equilibration of the system during cooling. However, the performance appears independent of the length of the individual Markov processes and very good performances were observed for the shortest chains studied. This behaviour has been observed for all problems studied so far and suggests that massive parallelisation of the algorithm is indeed possible.

We also solved the problem using conventional Simulated Annealing to establish a basis for comparison. The Simulated Annealing implementation employed a perturbation framework identical to the one used in our new algorithm. We developed targeting curves with increasing Markov chain lengths for sets of ten runs per case. The performance improved with the Markov chain length but the quality of the solutions did not match those obtained using of the new algorithm, even for extremely long chains. Table 2 compares the performance of new algorithm with that of Simulated Annealing for the case of the longest Markov chains studied (1000). It can be seen that the new algorithm outperforms Simulated Annealing in terms of solution quality and offers massive savings (90%) in CPU time for the case of Markov = 500 and poolnum = 100. Similar observations were made for different combinations of these two parameters.

The presented algorithm showed similar behaviour when applied to four other test problem. Most importantly, the performance was observed to be independent of the Markov parameter, which indicates the high potential for massive parallelisation. Detailed results from these tests will be published separately.

4. Conclusions

We have presented a new optimisation method that is suitable for large-scale distributed computing environments. The algorithm carries the strengths of stochastic optimisation methods such as Simulated Annealing in terms of global optimisation capabilities. A comparison with Simulated Annealing indicates that the new algorithm is also highly computationally efficient. In the absence of sequential searches, the algorithm is, in principle, not limited by the number of processors it can exploit. The algorithm will be applicable to a wide range of optimisation problems in operations as well as in design. As the solutions of all pools are stored in a database, knowledge about the importance of individual solution features can be extracted in the context of the systems performance. This is the focus of current research. We are also in the process of setting up a distributed test bed to evaluate the algorithm further. Applications to typical process and product design problems as well as problems in process operations will be the focus of future activities.

References

N. Antonopoulos, P. Linke and A. Kokossis. Chemical Engineering Communications, 28(2004) 2391

J.P.B. Leite and B.H.V. Topping, Computers and Structure Jounal 73 (1999) 545

N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, Chemical Physics, 21(1953) 1087

C.A. Floudas, P.M. Pardalos, C.S. Adjiman, W.R. Esposito, Z. Gumus, S.T. Harding, J.L. Klepeis, C.A. Meyer, and C.A. Schweiger, Handbook of Test Problems in Local and Global Optimization, Kluwer Academic Publishers, 1999.