INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Team DOLPHIN

Discrete multi-objective Optimization for Large scale Problems with Hybrid dIstributed techNiques

Futurs

THEME NUM

Activity Report

2004

Table of contents

1. Team

Head of the project-team
El-Ghazali Talbi [Professor, USTL]

Administrative Assistant
Axelle Magnier [INRIA Assistant project]

Research scientists
Clarisse Dhaenens [Associate Professor, USTL]
Nouredine Melab [Associate Professor, USTL]
Franck Seynhaeve [Associate Professor, USTL]

Post-doctoral fellows
Laetitia Jourdan [Lecturer at USTL and Post-Doc at University of Exeter]
Benjamin Weinberg [Lecturer, Université du Littoral]
Nicolas Jozefowiez [Lecturer, USTL]

Ph. D. students
Matthieu Basseur [Lecturer, Research ministry support, Fourth year]
Sébastien Cahon [Lecturer, France Telecom project support, Fourth year]
Mohamed Khabzaoui [Lecturer, GenHomme project support, Third year]
Julien Lemesre [France Telecom project support, Second year]
Mohand-Said Mezmaz [Research ministry support, Second year]

Partener research scientists
Frédéric Semet [Professor, ROI-LAMIH, Université de Valenciennes]

2. Overall Objectives

2.1. Motivations

Many industrial domains are concerned by large and complex optimization problems, involving important financial costs and in which decisions must be taken in an optimal way. Optimization problems encountered in practice are seldom mono-objective. In general, there are many conflicting objectives to handle: for instance, to minimize the cost and to maximize the quality (physic, mechanic, service, ...). Indeed, many diverse areas (telecommunications, genomics, engineering design, finance, chemistry and ecology among others) are concerned by multi-objective optimization problems (MOPs)[1].

Multi-objective optimization takes its root from the 19th century, with the work of Edgeworth and Pareto in economics. It has been used in economics and management science, and then gradually in engineering sciences. Nowadays, multi-objective optimization is an important area for researchers and engineers. This is due to the multi-objective nature of many real-life optimization problems, but also to the numerous interesting questions arising in this domain. The complexity of MOPs becomes more and more important in terms of the size of the problem to solve. Moreover, the search time for solving those problems has to be reasonable for most of the MOPs encountered in practice. The development of advanced hybrid optimization methods from combinatorial optimization in operations research, decision making in artificial intelligence, and parallel and distributed computing, is an important issue in solving complex and large MOPs.

The optimal solution for MOPs in not a single solution as in mono-objective optimization problems, but a set of solutions defined as Pareto optimal solutions. A solution is Pareto optimal if it is not possible to improve a given objective without deterioring at least another objective. This set of solutions represents the compromise solutions between the different conflicting objectives. The difficulty in solving MOPs lies in the following general facts: the number of Pareto optimal solutions depends on the size of the problem and mainly

[1]Note that the terms multi-objective and multi-criteria are used interchangeably throughout this report.

on the number of considered objectives. The structure of the Pareto front (continuity, convexity, modality, deceptivity, etc.) depends on the studied MOP. The Pareto optimal solutions may be localized on the frontier and inside the convex hull of feasible solutions. Moreveover, most of the MOPs are NP-hard problems.

2.2. State-of-the-art

General search and optimization techniques may be classified into two categories: exact and heuristics (or approximate) approaches. Exact methods guarantee to find optimal solutions. Heuristics provide "good" solutions within a "reasonable" search time, without a guarantee on the optimality of the obtained solutions. In the class of heuristics methods, we can find metaheuristics which are general methods that can be applied potentially to any MOP.

Exact methods (branch and bound algorithm, branch and cut algorithm, dynamic programming, constraint programming, etc.) suffer from the combinatorial explosion of NP-hard MOPs. Those techniques cannot be applied for most of real-life problems. The size of the problem is not the only criteria which limits their application. Indeed, some large instances can be optimally solved, when other instances with smaller size may not be solved in an optimal way.

For multi-objective optimization using exact methods, most of the existing work deals with continuous multi-objective programming. In the context of combinatorial optimization, a particular emphasis has been carried out on bi-objective problems using branch and bound, A* algorithm, and dynamic programming. These methods are efficient for small size problems. For large problems and/or MOPs with more than two objectives, there is no efficient exact methods.

In general, metaheuristics generate solutions of "good" quality in a reasonable search time. This is attested by the obtained results for some optimization problems[2] (best known solutions). These algorithms allow also the generation of upper bound solutions for exact methods. In the last ten years, we have seen an important interest in the literature using metaheuristics to solve different optimization problems.

Heuristic methods are inevitable to solve large size problems and MOPs with more than two objectives. They don't guarantee to find the optimal Pareto front, but an approximation of this set. Adaptations of metaheuristics have been recently proposed in the literature: simulated annealing, tabu search, and evolutionary algorithms. Most of those approaches transform the MOP into a mono-objective one (agregation methods, ϵ-constraint method, goal programming, ...). The transformation of the problem needs an A-priori knowledge of the target problem (weights, constraints, goals). These approaches modify the structure of the initial problem which loses its properties. Moreover, they are sensitive to the structure of the Pareto front (convexity, discontinuity, ...). For instance, linear aggregation methods find only supported solutions, i.e solutions which are on the convex hull of the Pareto front. The cost of the multiple executions of the algorithms to approximate the whole Pareto front is another drawback of this class of methods.

Population-based methods, such as evolutionary algorithms, are well adapted to MOPs [34] [35]. Their application to MOPs is recent and expanding in the international research community. To be convinced, we have to see the number of conferences, special issues of journals, books, workshops and sessions organized from 2001. For example, our team organized many sessions in the CEC (Congress on Evolutionary Computation) and IEEE CESA conferences, and is member of the executive committee and/or the program committee of the first conference dedicated to evolutionary multi-objective optimization (EMO conferences), and the MOPGP (Multi-objective Programmig and Goal Programming) conference.

2.3. Research Roadmap

The objective of the DOLPHIN[3] project is the modelization and resolution of large multi-criteria combinatorial problems using parallel and distributed hybrid techniques. We are interested in algorithms using Pareto

[2] We suppose here a minimization problem.
[3] Discrete multi-objective Optimization for Large scale Problems with Hybrid dIstributed techNiques.

approaches which generate the whole Pareto set of a given MOP. For this purpose, the research actions can be resumed as follows (see figure1):

- **Analysis of the structure of a MOP:** The statistical analysis of the structure of the Pareto front by means of different indicators allows the design of efficient and robust hybrid optimization techniques. In general, the current theory does not allow the complete analysis of optimization algorithms. Several questions are unanswered: i) why a given method is efficient? ii) why certain instances are difficult to solve? Some work is needed to guide the user in the design of efficient methods.
  The NFL theorem (No Free Lunch) shows that two optimization methods have the same global performances on the whole set of uniform optimization problems. Then, it is crucial to make some hypothesis on the studied problem. This may be done in two steps:
  - analyzing the target problem to identify its landscape properties,
  - including this knowledge in the proposed optimization method.

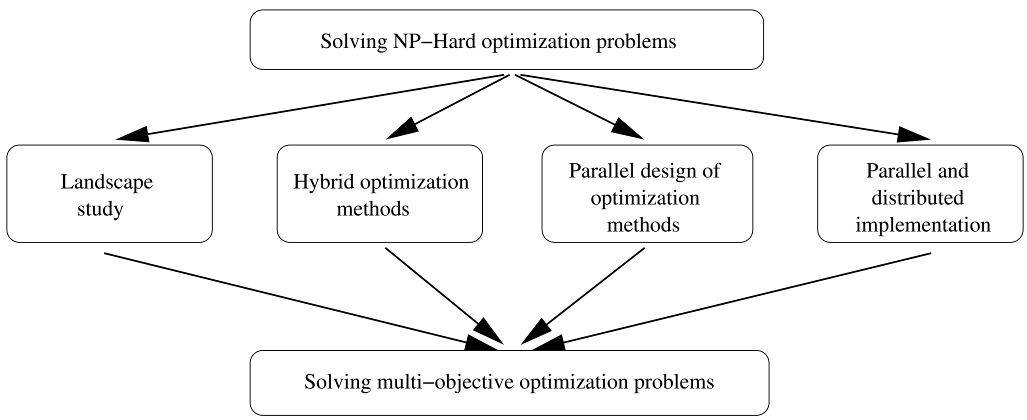

Figure 1. Research roadmap of the DOLPHIN project.

  Our interest in this project is to answer all these questions and remarks for the multi-objective case. Another point considered is the performance evaluation of multi-objective optimization methods.

- **Cooperation of optimization methods (metaheuristics and/or exact methods):**
  The hybridation of optimization methods allows the cooperation of complementary different methods. For instance, the cooperation between a metaheuristic and an exact method allows to take advantage of the intensification process of an exact method in finding the best(s) solution(s) in a sub-space, and the diversification process of metaheuristics in reducing the search space to explore.
  In this context, different types of cooperation may be proposed. Those approaches are under study in the project and we are applying them to different generic MOPs (flow-shop scheduling problem, vehicle routing problem, covering tour problem, and the association rule problem in data mining).

- **Parallel optimization methods:** Parallel and distributed computing may be considered as a tool to speedup the search to solve large MOPs and to improve the robustness of a given method. Following this objective, we will design and implement parallel metaheuristics (evolutionary algorithms, tabu search approach) and parallel exact methods (branch and bound algorithm, branch and cut algorithm) for solving different large MOPs. Moreover, the joint use of parallelism and cooperation allows the improvement of the quality of the obtained solutions.

Our experience in the domain of parallel optimization dates from the apparition of network of transputers (1988), where the mapping of the parallel program and the routing of messages were explicitely designed by the programmer.

- **Framework for parallel and distributed hybrid metaheuristics:** Our team contributes to the development of an open source framework for metaheuristics, named ParadisEO (PARAllel and DIStributed Evolving Objects). Our contribution in this project is the extension of the EO (Evolving Objects) framework [4], which consists in: i) the generalization of the framework to single solution metaheuristics such as local search, tabu search and simulated annealing; ii) the design of hybrid methods; iii) the development of parallel and distributed models, and the resolution of MOPs.
  In this project, our goal is the efficient design and implementation of this framework on different types of parallel and distributed hardware platforms: cluster of workstations (COW), networks of workstations (NOW) and grid computing platforms, using the different suited programming environments (MPI, Condor, XtremWeb). The coupling with well-known frameworks for exact methods (such as COIN and BOB++) will also be considered. The exact methods for MOPs developed in this project will be integrated in those software frameworks.
  A particular concern will be considered to the experimentation of our framework by different users and applications outside the DOLPHIN project. This is done in order to validate the design and the implementation issues of ParadisEO.
- **Validation:** The designed approaches will be validated on generic and real-life MOPs, such as:
  - Scheduling problems: Flow-shop scheduling problem;
  - Routing problems: Vehicle routing problem (VRP) and covering tour problem (CTP);
  - Mobile telecommunications: Design of mobile telecommunications networks (contract with France Telecom R&D) and design of access networks (contract with Mobinets);
  - Genomics: Association rule discovery (data mining task) for mining genomic data (contract with GenFit and IT-OMICS).

  Some benchmarks and their associated optimal Pareto fronts or the best known Pareto fronts will be available on the Web. We are also developing an open source software, named GUIMOO[5], which integrates different performance evaluation metrics and 2D/3D visualization tools of Pareto fronts.

3. Scientific Foundations

3.1. Optimization methods and landscape

The analysis of structures (landscapes) of MOPs and the performance assessment of resolution methods are significant topics in the design of optimization methods. The effectiveness of metaheuristics depends on the properties of the landscape (roughness, convexity, etc). The notion of landscape has been first described in [40] by the way of the study of species evolution. Then, this notion has been used to analyze combinatorial optimization problems.

The landscape is defined by a neighborhood operator and can be represented by a graph $G = (V, E)$. The vertices represent the solutions of the problem and an edge (e_1, e_2) exists if the solution e_2 can be obtained by an application of the neighborhood operator on the solution e_1. Then, considering this graph as the ground floor, we elevate each solution to an altitude equals to its cost. We obtain a surface, or landscape, made of peaks, valleys, plateaus, cliffs, ...The problem lies in the difficulty to have a realistic view of this landscape.

[4]This framework was initially developed by Geneura TEAM (Spain), INRIA (France), LIACS (Netherlands). http://www.lifl.fr/OPAC/paradiseo.
[5]Graphical User Interface for Multi-Objective Optimization (http://www.lifl.fr/OPAC/guimoo).

3.1.1. Analysis of the structure of a problem

As others, we believe that the main point of interest in the domain of combinatorial optimization is not the design of the best algorithm for a large number of problems but that it is to search for the most adapted method to an instance or a set of instances of a given problem. Therefore, we guess that no ideal metaheuristic, designed as a black-box, may exist.

Indeed, the first studies realized in our research group on the analysis of landscapes of different mono-objective combinatorial optimization problems (traveling salesman problem, quadratic assignment problem) have showed that not only different problems correspond to different structures but also that different instances of the same problem correspond to different structures.

For instance, we have realized a statistical study of the landscapes of the quadratic assignment problem. Some indicators, which characterize the landscape of an instance, have been proposed and a taxonomy of the instances including three classes has been deduced. Hence it is not enough to adapt the method to the problem under study but it is necessary to specialize it according to the type of instance treated.

So in its studies of mono-objective problems, the OPAC research group has introduced into the resolution methods some information about the problem to be solved. The landscapes of some combinatorial problems have been studied in order to investigate the intrinsic natures of their instances. The resulting information have been inserted into an optimization strategy and allowed the design of efficient and robust hybrid methods. The extension of these studies to multi-objective problems is a part of the DOLPHIN project [10][21][32][33].

3.1.2. Performance assessment

The DOLPHIN project is also interested in the performance assessment of multi-objective optimization methods. Nowadays, statistical techniques developed for mono-objective problems are used with adaptation to the multi-objective case. Nevertheless, specific tools are necessary in many cases: for example, the comparison of two different algorithms is relatively easy in the mono-objective case - we compare the quality of the best solution obtained in a fixed time, or the time needed to obtain a solution of a certain quality. The same idea can not be immediately transposed to the case where the output of the algorithms is a set of solutions having several quality measures, and not a single solution.

Various indicators have been proposed in the literature for evaluating the performance of multi-objective optimization methods but no indicator seems to outperform the others [41]. The OPAC research group has proposed two indicators: the *contribution* and the *entropy* [38]. The contribution evaluates the supply in term of Pareto-optimal solutions of a front compared to another one. The entropy gives an idea of the diversity of the found solutions. These two metrics are used to compare the different metaheuristics in the research group, for example in the resolution of the bi-objective flow-shop problem, and also to show the contribution of the various mechanisms introduced in these metaheuristics.

These metrics and others (generational distance, spacing, ...) are integrated in the open software GUIMOO developed within the framework of the DOLPHIN project. This software is dedicated to the visualization of landscapes (2D and 3D) for multi-objective optimization and the performance analysis by the use of special metrics.

3.1.3. Goals

One of the main issue in the DOLPHIN project is the landscape study of the multi-objective problems and the performance assessment of multi-objective optimization methods to design efficient and robust resolution methods:

- *Landscape study:* The goal here is to extend the study of landscapes of the mono-objective combinatorial optimization problems to multi-objective problems in order to determine the structure of the Pareto frontier and to integrate this knowledge about the problem structure in the design of resolution methods.
  This study has been initiated for the bi-objective flow-shop problem. We have studied the convexity of the frontiers obtained in order to show the interest of our Pareto approach compared to an

aggregation approach, which only allows to obtain the Pareto solutions situated on the convex hull of the Pareto front (supported solutions).
Our preliminary study of the landscape of the bi-objective flow-shop problem shows that the supported solutions are very closed to each other. This remark leads us to improve an exact method initially proposed for bi-objective problems. Furthermore, a new exact method able to deal with any number of objectives has been designed.

- *Performance assessment:* The goal here is to extend GUIMOO in order to provide efficient visual and metric tools for evaluating the assessment of multi-objective resolution methods.

3.2. Hybrid multi-objective optimization methods

The success of metaheuristics is based on their ability to find efficient solutions in a reasonable time [36]. But with very large problems and/or multi-objective problems, efficiency of metaheuristics may be compromised. Hence, in this context it is necessary to integrate metaheuristics in more general schemes in order to develop even more efficient methods. For instance, this can be done by different strategies such as cooperation and parallelization.

The DOLPHIN project deals with *"a posteriori"* multi-objective optimization where the set of Pareto solutions (solutions of best compromise) have to be generated in order to give to the decision maker the opportunity to choose the solution that interests him.

Population based methods, such as evolutionary algorithms, are well fitted for multi-objective problems as they work with a set of solutions [35][34]. To be convinced one may refer to the list of references on Evolutionary Multi-objective Optimization maintained by Carlos A. Coello Coello[6] that contains more than 1600 references. One of the objectives of the project is to propose advanced search mechanisms for intensification and diversification. These mechanisme have been designed in an adaptive manner, since their efficiency is related to the landscape of the MOP.

In order to assess the performances of the proposed mechanisms, we always proceed in two steps: first, experiments are carried out on academic problems, for which some best known results exist; second, we use real industrial problems to cope with large and complex MOPs. The lack of references in terms of optimal or best know Pareto set is a major problem. Therefore, the obtained results in this project and the test data sets will be available at the URL http://www.lifl.fr/OPAC at Benchmarks.

3.2.1. Cooperation of metaheuristics

In order to cope with advantages of the different metaheuristics, an interesting idea is to combine them. Indeed, the hybridization of metaheuristics allows the cooperation of methods having complementary behaviors. The efficiency and the robustness of such methods depend on the balance between the exploration of the whole search space and the exploitation of interesting areas.

Hybrid metaheuristics have received considerable interest these recent years in the field of combinatorial optimization. A wide variety of hybrid approaches have been proposed in the literature and give very good results on numerous single objective optimization problems, that are either academic (traveling salesman problem, quadratic assignment problem, scheduling problem, ...) or real-world problems. This efficiency is due to combinations of single solution based methods (iterative local search, simulated annealing, tabu search...) with population based methods (genetic algorithms, ants search, scatter search...). A taxonomy of hybridization mechanisms may be found in [39]. It proposes to decompose those mechanisms into 4 classes:

- *LRH class - Low-level Relay Hybrid*: This class groups algorithms in which a given metaheuristic is embedded into a single-solution metaheuristic. Few examples from the literature belong to this class.

[6](http://www.lania.mx/~ccoello/EMOO/EMOObib.html)

- *LTH class - Low-level Teamwork Hybrid*: In this class, a metaheuristic is embedded into a population-based metaheuristic in order to exploit strengths of single-solution and population-based metaheuristics.
- *HRH class - High-level Relay Hybrid*: Here, self contained metaheuristics are executed in a sequence. For instance, a population-based metaheuristic is executed to locate interesting regions and then a local search is performed to exploit these regions.
- *HTH class - High-level Teamwork Hybrid*: This scheme involves several self-contained algorithms performing a search in parallel and cooperating. An example will be the island model, based on GAs, where the population is partitioned into small subpopulations and a GA is executed per subpopulation. Some individuals can migrate between subpopulations.

Let us notice, that if hybrid methods have been studied in the mono-criterion case, their application in the multi-objective context is not yet widely spread. The objective of the DOLPHIN project is to integrate specificities of multi-objective optimization into the definition of hybrid models [11][23].

3.2.2. Cooperation between metaheuristics and exact methods

Until now only few exact methods have been proposed to solve multi-objective problems. They are based either on a Branch-and-bound approach, on the algorithm A^* or on dynamic programming. However, those methods are limited to two objectives and are, most of the time, not able to be used on the complete problem, which is often large scale MOP. Therefore, sub search spaces have to be defined in order to be able to use exact methods. Hence, in the same manner than hybridization of metaheuristics, the cooperation of metaheuristics and exact methods is also a main issue in this this project. Indeed, it allows to use the exploration capacity of metaheuristics as well as the intensification ability of exact methods that are able to find optimal solution(s) in a restricted search space. Sub search spaces will have to be defined along the search. Such strategies can be found in the literature, but they are only applied to mono-objective academic problems.

We have extended the previous taxonomy for hybrid metaheuristics to the cooperation between exact methods and metaheuristics. Using this taxonomy, we are investigating cooperative multi-objective methods. In this context, several types of cooperations may be considered, according to the way the metaheuristic and the exact method cooperates. For instance, a metaheuristic can use an exact method for intensification or an exact method can use a metaheuristic to reduce the search space).

Moreover, a part of the DOLPHIN project deals with studying exact methods in the multi-objective context in order: i) to be able to solve small size problems and to validate proposed heuristic approaches; ii) to have more efficient/dedicated exact methods that can be hybridized with metaheuristics. In this context, the use of parallelism will push back limits of exact methods which will be able to explore larger size search space [22][27][26].

3.2.3. Goals

Based on the previous works on multi-objective optimization, it appears that to improve metaheuristics, it becomes essential to integrate knowledge about the problem structure. This knowledge can be gained during the search. This would allow to adapt operators that may be specific for multi-objective optimization or not. The goal here is to design auto-adaptive methods that are able to react to the problem structure. Moreover, regarding the hybridization and the cooperation aspects, the objectives of the DOLPHIN project are to deepen those studies as follows:

improve metaheuristics, it becomes essential to integrate knowledge about the problem structure, that we may get during the execution. This would allow to adapt operators that may be specific for multi-objective optimization or not. The goal here is to design auto-adaptive methods that are able to react to the problem structure.

- *Design of cooperative metaheuristics:* Previous studies show the interest of hybridization for a global optimization and the importance of problem structure study for the design of efficient methods. It is now necessary to generalize metaheuristic hybridization and to propose adaptive hybrid models that may evolve during the search while selecting the appropriate metaheuristic. Multi-objective aspects have to be introduced in order to cope with specificities of multi-objective optimization.

- *Design of cooperative schemes between exact methods and metaheuristics:* Once the study on possible cooperation schemes will be ended, we will have to test and compare them in the multi-objective context.
  works on parallel metaheuristics allow us to speed up the resolution of large scale problems. It could be also interesting to study the robustness of the different parallel models (in particular in the multi-objective case) and to propose rules that determine, given a specific problem, which kind of parallelism to use.

Of course these goals are not disjoined and it will be interesting to simultaneously use hybrid metaheuristics and exact methods. Moreover, those advanced mechanisms may require the use of parallel and distributed computing in order to easily make evolve simultaneously cooperating methods and to speed up the resolution of large scale problems.

proceed in two phases: validation on academic problems, for which some best known results exist and use on real problems (industrial) to cope with problem size constraints.

distributed multi-objective aspects in the ParadisEO Platform (see the paragraph on software platform).

3.3. Parallel multi-objective optimization: models and software frameworks

Parallel and distributed computing may be considered as a tool to speedup the search to solve large MOPs and to improve the robustness of a given method. Moreover, the joint use of parallelism and cooperation allows the improvement of the quality of the obtained Pareto sets. Following this objective, we will design and implement parallel models for metaheuristics (evolutionary algorithms, tabu search approach) and for exact methods (branch and bound algorithm, branch and cut algorithm) for solving different large MOPs.

One of the goal of the DOLPHIN project is to integrate the developed parallel models into software frameworks. Several frameworks for parallel distributed metaheuristics have been proposed in the literature. Most of them focus only either on evolutionary algorithms or on local search methods. Only few frameworks are dedicated to the design of both of the two families of methods. On the other hand, existing optimization frameworks either do not provide parallelism at all or just supply at most one parallel model. In this project, a new framework for parallel hybrid metaheuristics is proposed, named *Parallel and Distributed Evolving Objects (ParadisEO)* based on EO. The framework provides in a transparent way the hybridization mechanisms presented in the previous section, and the parallel models described in the next section. Concerning the developed parallel exact methods for MOPs, we will integrate them into well-known frameworks such as COIN.

3.3.1. Parallel models

According to the family of addressed metaheuristics, we may distinguish two categories of parallel models: parallel models managing a single solution and parallel models that handle a population of solutions. The major single solution-based parallel models are the following: the *parallel neighborhood exploration model* and the *multi-start model*.

- *The parallel neighborhood exploration model* is basically a “low level" model that splits the neighborhood into partitions that are explored and evaluated in parallel. This model is particularly interesting when the evaluation of each solution is costly and/or when the size of the neighborhood is large. It has been successfully applied to the mobile network design problem (see Application section).

- *The multi-start model* consists in executing in parallel several local searches (that may be heterogeneous), without any information exchange. This model raises particularly the following question: is it equivalent to execute k local searches during a time t than executing a single local search during $k \times t$? To answer this question we tested a multi-start Tabu search on the quadratic assignment problem. The experiments have shown that the answer is often landscape-dependent. For example, the multi-start model may be well-suited for landscapes with multiple basins.

On the other hand, parallel models that handle a population of solutions are mainly: the *island model*, the *central model* and *the distributed evaluation of a single solution*. Let us notice that the last model may also be used with single-solution metaheuristics.

- In *the island model*, the population is split into several sub-populations distributed among different processors. Each processor is responsible of the evolution of one sub-population. It executes all the steps of the metaheuristic from the selection to the replacement. After a given number of generations (synchronous communication), or when a convergence threshold is reached (asynchronous communication), the migration process is activated. Then, exchanges of solutions between sub-populations are realized, and received solutions are integrated into the local sub-population.
- *The central (Master/Worker) model* allows to keep the sequentiality of the original algorithm. The master centralizes the population and manages the selection and the replacement steps. It sends sub-populations to the workers that execute the recombination and evaluation steps. The latter return back newly evaluated solutions to the master. This approach is efficient when the generation and evaluation of new solutions is costly.
- *The distributed evaluation model* consists in a parallel evaluation of each solution. This model has to be used when, for example, the evaluation of a solution requires access to very large databases (data mining applications) that may be distributed over several processors. It may also be useful in a multi-objective context, where several objectives have to be computed simultaneously for a single solution.

As these models have now been identified, our objective is to study them in the multi-objective context in order to use them advisedly. Moreover, these models may be merged to combine different levels of parallelism and to obtain more efficient methods [14][19].

3.3.2. The ParadisEO framework

In the last years, loosely and closely coupled clusters of workstations have become a real architectural alternative to traditional supercomputing environments. As a matter of fact, workstations are more and more powerful while they are less and less expensive. In addition, interconnection networks have known a great technological evolution that makes them more speedy (Myrinet, GigaEthernet, etc.). Our research group has developed a parallel adaptive programming environment named *MARS (Multi-user Adaptive Resource Scheduler)*. MARS provides for the programmers an API[7] to simplify the development of parallel applications, and a support for their scheduling on adaptive networks and clusters of heterogeneous workstations. The MARS open source has been made available on the Web at the disposal of the users for free download and use. With such an experience in the development of open source software, we are developpping the ParadisEO software. This framework is detailed further in this document.

3.3.3. Goals

Our objectives focus on these issues are the following:

- *Design of parallel models for metaheuristics and exact methods for MOPs*: We will develop parallel cooperative metaheuristics (evolutionary algorithms and local search such as Tabu search) for solving different large MOPs. Moreover, we are designing a new exact method, named PPM (Parallel

[7] Application Programming Interface.

Partition Method), based on branch and bound and branch and cut algorithms. Finally, some parallel cooperation schemes between metaheuristics and exact algorithms have to be used to solve MOPs in an efficient manner.

- *Integration of the parallel models into software frameworks*: The parallel models for metaheuristics will be integrated in the ParadisEO software framework. The proposed multi-objective exact methods must be first integrated into standard frameworks for exact methods such as COIN and BOB++. A *coupling* with ParadisEO is then needed to provide hybridization between metaheuristics and exact methods.
- *Efficient deployment of the parallel models on different parallel and distributed architecture including GRIDs*: The designed algorithms and frameworks will be efficiently deployed on non-dedicated networks of workstations, dedicated cluster of workstations and SMP (Symetric Multi-processors) machines. For GRID computing platforms, peer to peer (or P2P) middlewares (XtremWeb-Condor) will be used to implement our frameworks. For this purpose, the different optimization algorithms may be re-visited for their efficient deployment (see [28]).

4. Application Domains

4.1. Introduction

The developed algorithms and software frameworks of this project will be validated on academic generic problems and industrial real-life problems. For the industrial problems, two application domains are mainly considered: mobile telecommunications and genomics.

4.2. Academic generic problems

In this project, some well known optimization problems are re-visited in terms of multi-objective modelization and resolution:

- Flow-shop scheduling problem: The flow-shop problem is one of the most well-known scheduling problems. However, most of the works of the literature use a mono-objective model. In general, the objective minimized is the total completion time (makespan). Many other criteria may be used to schedule tasks on different machines: maximum tardiness, total tardiness, mean job flowtime, number of delayed jobs, maximum job flowtime, etc. In the DOLPHIN project, a bi-criteria model which consists in minimizing the makespan and the total tardiness is studied [11][22][23].
- Routing problems: The vehicle routing problem (VRP) is a well-known problem and it has been studied since the end of the 50's. It has a lot of practical applications in many industrial areas (ex. transport, logistic, ...). Existing studies of the VRP are almost all concerned with the minimization of the total distance only. The model studied in the DOLPHIN project introduces a second objective whose purpose is to balance the length of the tours. This new criterion is expressed as the minimization of the difference between the length of the longest tour and the length of the shortest tour. As far as we know, this model is one of the pioneer work of the literature.
  The second routing problem is a generalization of the covering tour problem (CTP). In the DOLPHIN project, this problem is solved as a bi-objective problem where a set of constraints are modeled as an objective. The two objectives are: i) minimization of the length of the tour; ii) minimization of the largest distance between a node to be covered and a visited node. As far as we know, this study is among the first work which tackle a classic mono-objective routing problem by relaxing constraints and building a more general MOP [9][16].

For all studied problems, standard benchmarks have been extended to the multi-objective case. The benchmarks and the obtained results (optimal Pareto front, best known Pareto front) are available on the Web pages associated to the project and from the MCDM (International Society on Multiple Criteria Decision Making) web site. This is an important issue to encourage comparison experiments in the research community.

4.3. Application to mobile telecommunications networks

With the extraordinary success of mobile telecommunication systems, service providers have been affording huge investments for network infrastructures. Mobile network design appears of outmost importance and then is a major issue in mobile telecommunication systems. The design of large cellular networks is a complex task with a great impact on the quality of service and the cost of the network. With the continuous and rapid growth of communication traffic, large scale planning becomes more and more difficult. Automatic or interactive optimization algorithms and tools would be very useful and helpful. Advances in this area will certainly lead to important improvements concerning the service quality and the deployment cost.

This need is now even more acute with the advent of third-generation systems, such as Universal Mobile Telecommunication System (UMTS), because of the increased complexity of the system and the number of parameters that must be considered:

- large deployment zones (many hundreds of km^2 for an urban network);
- high density (many urban networks with more than 1000 antennas);
- multi-periodic planification (important investment on long-term planification);
- introduction of many cooperating systems with different technologies (GSM, DCS, UMTS).

In this project, the solution of planification problems, in terms of modelization and resolution, are developed in a multi-criteria context associating financial criteria (cost of the network), technical criteria (coverage, availability), and marketing criteria (quality of service). Two complementary design problems are considered:

- Radio mobile network design: This work is realized in collaboration with France Telecom R&D. Engineering of radio mobile telecommunication networks involves two major problems: The design of the radio network and the frequency assignment. The design consists in positioning base stations (BS) on potential sites, in order to fulfill some objectives and constraints. The frequency planning sets up frequencies used by BS with criteria of reusing. In this project, we address the first problem. Network design is a NP-hard combinatorial optimization problem. The BS problem deals with finding a set of sites for antennas from a set of pre-defined candidate sites, determining the type and the number of antennas, and setting up the configuration of different parameters of the antennas (tilt, azimuth, power, ...). A new formulation of the problem as a multi-objective constrained combinatorial optimization problem is considered. The model deals with specific objectives and constraints due to the engineering of cellular radio network. Reducing costs without sacrificing the quality of service are issues of concern. Most of the proposed models in the literature optimize a single objective (cover, cost, linear aggregation of objectives, etc.). Moreover, other works use non realistic simplified models of the problem [29].
- Access network design: This work is realized in collaboration with Mobinets. The problem consists in minimizing the cost of the access network and maximizing its availability. Operators can only be competitive and economical if they have an optimized access network. Since the transmission costs are becoming high compared to the equipment costs, and the traffic demands are increasing with the introduction of new services, it is vital for operators to find cost-optimized transmission network solutions at higher bit rates. Many constraints dealing with technologies and service providers have to be satisfied. All deployed important technologies (ex. GSM, UMTS) will be considered.

4.4. Application to genomics

Genomic research is a great challenge for our society and numerous research entities of different specialities (biology, medical or information technology) are grouping each other to collaborate on specific thema. As far as we are concerned, we collaborate with academic and industrial partners (IBL: Biology Institute of Lille; IPL: Pasteur Institute of Lille; IT-Omics firm) to study genetic factors that may explain multi-factorial diseases such as diabete, obesity or cardiovascular diseases. The originality is to look not for a single factor, but for one or several combinations of factors (that may be of different natures: genetic factors, environmental factors...) among a very large set of potential factors (several thousands). The scientific goals are to formulate hypotheses describing associations that may have any influence on diseases under study. These hypotheses will have to be verified by biologists thanks to additional experiments.

The genomic application of the DOLPHIN project deals with post-genomic, where a very large number of data are obtained thanks to advanced technologies and have to be analyzed. Hence, one of the goals of the project is to develop methods of analysis in order to extract information from data coming from biological experiments. Our originality is to first extract knowledge discovery problems from the genomic application and then to transform these knowledge discovery problems (datamining problems) into optimization problems where an (or several) objective function(s) has(have) to be optimized. Following this, it is possible to apply to these modified problems optimization methods.

An analysis of some genomic problems lead us to model them as an association rule mining problem (classical task of datamining). The main characteristic is that we look for several associations of factors among a very large set of potential factors. Hence the combinatoric (the number of potential solutions) associated to the problem is huge. Moreover, the landscape study, which shows a flat and rough landscape, indicates that exploratory methods are required to deal with such problems. Finally a study on criteria commonly used for association rule mining shows that a lot of criteria exist and that there exists no universal criterion. Hence we propose to model the association rules mining problem as a multi-objective one.

For all these reasons, we choose to base our resolution approach on evolutionary algorithms. Now, we are working on solving these problems in a multi-objective point of view, and studying hybridization of evolutionary algorithms with exact methods able to solve sub-problems (small size problems). Another important point is that evaluation functions, for such applications, may be very time consuming, because they require complex statistical computations. Therefore a parallel implementation is required in order to allow a large exploration of the search space without degrading the computational time [15][31][30][25][24][20].

5. Software

ParadisEO [13][12] is the result of several extensions performed on the *Evolving Objects (EO)* framework. These extensions are the subject of a joint work with INRIA TAO project (Marc Schoenauer). Initially, EO has been developed within the scope of an European joint work, and was only devoted to the design of sequential evolutionary algorithms. The objective of our extensions is to allow the design of multi-objective metaheuristics including evolutionary algorithms (EA) and local search (LS) methods, to deal with parallel and distributed computing, and hybridization. Unlike most of other existing frameworks [17], it provides a wide range of reusable features and mechanisms related to parallel and distributed metaheuristics. First, it allows the design of both EA and LS and provides different reusable mechanisms for their hybridization. Second, for each family it supplies a large variety of multi-objective variation operators, multi-objective evaluation functions, etc. Third, it implements the most common parallel and distributed models presented above, and allows the user to use them in a transparent and portable way. Finally, the fine-grained evolving objects of ParadisEO allow a higher flexibility, adaptability and maximum code reuse.

A framework is normally intended to be exploited by as many users as possible. Therefore, its exploitation could be successful only if some important user criteria are satisfied. The following criteria are the major of them and constitute the main objectives of the ParadisEO framework:

- *Maximum design and code reuse*: The framework must provide for the user a whole architecture design of his/her solution method. Moreover, the programmer may redo as little code as possible. This objective requires a clear and maximal conceptual separation between the solution methods and the problems to be solved [37], and thus a deep domain analysis. The user might therefore develop only the minimal problem-specific code.
- *Flexibility and adaptability*: It must be possible for the user to easily add new features/metaheuristics or change existing ones without implicating other components. Furthermore, as in practice existing problems evolve and new others arise these have to be tackled by specializing/adapting the framework components.
- *Utility*: The framework must allow the user to cover a broad range of metaheuristics, problems, parallel distributed models, hybridization mechanisms, etc.
- *Transparent and easy access to performance and robustness*: As the optimization applications are often time-consuming the performance issue is crucial. Parallelism and distribution are two important ways to achieve high performance execution. In order to facilitate its use it is implemented so that the user can deploy his/her parallel algorithms in a transparent manner. Moreover, the execution of the algorithms must be robust to guarantee the reliability and the quality of the results. The hybridization mechanism allows to obtain robust and better solutions.
- *Portability*: In order to satisfy a large number of users the framework must support different material architectures and their associated operating systems.

In order to achieve these objectives, ParadisEO is developed as a white-box objected-oriented framework based on a heavy use of C++ templates. The conceptual separation between the metaheuristics and the problems they are intended to solve is expressed at implementation level by splitting the classes in two categories: provided classes and required classes. The provided classes constitute a hierarchy of classes implementing the invariant part of the code. Expert users can extend the framework by inheritance/specialization. The required classes coding the problem-specific part are abstract classes that have to be specialized and implemented by the user. The classes of the framework are fine-grained, and instantiated as evolving objects embodying each one only one method. This is a particular design choice adopted in ParadisEO. The heavy use of these small-size classes allows more independence and thus a higher flexibility compared to other frameworks. Changing existing components and adding new ones can be easily done without impacting the rest of the application.

ParadisEO is one of the rare frameworks that provides the most common parallel and distributed models. These models concern the island-based running of metaheuristics, the parallel evaluation of a population, and the parallel evaluation of a single solution. They are portable on distributed-memory machines and shared-memory multi-processors as they are implemented using standard libraries such as MPI, PVM and PThreads. The models can be exploited in a transparent way, one has just to instantiate their associated ParadisEO components. The user has the possibility to choose by a simple instantiation MPI or PVM for the communication layer. Furthermore, the framework provides the most common hybridization mechanisms. They can be exploited in a natural way to make cooperating metaheuristics belonging either to the same family or to different families of methods.

ParadisEO will be validated on academic and industrial problems. The experimental results demonstrate the efficiency of the parallel distributed models. Furthermore, they demonstrate the high reuse capabilities provided by the framework as the results show that the user redo little code. ParadisEO is downloadable from: http://eodev.sourceforge.net.

The developped algorithms and software frameworks will be deployed on different parallel and distributed architectures. Using standard parallel environments such as MPI, Condor and multi-threading, these frameworks will be efficiently deployed on networks of workstations, cluster of workstations and SMP (Symetric Multi-processors) machines.

For GRID platforms, peer to peer (or P2P) computing that harness massive computational resources is needed to deal with very large problems. From the algorithmical point of view, the different parallel distributed models and hybridization mechanisms have to be re-visited to manage their scaling up [18]. At the development side, ParadisEO must be coupled with a P2P middleware to scale to the Internet. In this project, we are interested in XtremWeb, a middleware developed within the framework of the INRIA Futurs Grand Large project. Building P2P-aware cooperative optimization applications is not straightforward with neither XtremWeb nor other existing P2P computing middlewares. Indeed, none allows cross-peer communications between distributed tasks in a volatile environment. As a consequence, the middleware must be extended with a coordination layer to allow cooperative optimization. Therefore, ParadisEO could be coupled with XtremWeb. As in P2P system resources are volatile, the scheduling and fault-tolerance issues must be addressed. To do that, the parallel models and hybridization mechanisms have to be adapted. From the development point of view, this means that XtremWeb has to be coupled with Condor. As a result of such coupling, ParadisEO on XtremWeb-Condor will allow users to interconnect via Internet several Condor pools of parallel hybrid exact and metaheuristics solvers.

6. Contracts and Grants with Industry

6.1. Regional Grants

- MOST Project “Methodologies for Optimization in Transport and Telecommunications Systems” (2000-2006) of the CPER (Regional Contract) TACT operation. The OPAC / DOLPHIN team leads the action “Multi-objective optimization”.

6.2. National Grants

- ACI GRID Grant (Grid Computing National Program) DOC-G “Challenges in combinatorial optimization on Grids” (2001-2004), in collaboration with Prism (Versailles) and IMAG-ID (Grenoble) laboratories.
- ACI GRID Grant (Grid Computing National Program) GRID2 (2001-2004). The goal of this project is the animation and coordination of research on Grid computing in France. The DOLPHIN team leads the section “Optimization and data mining on Grids”.
- Specific Research Action “Programming models for Grids” of the CNRS (2003-2004).
- Specific Research Action “Experimental GRID platform” of the CNRS (2003-2004).
- ACI “Masse de données” (Large dataset) Project GGM “Geno-Medical Grid” (2004-2007), in collaboration with LIRIS (Lyon) and IRIT (Toulouse) laboratories. Our concern in this project is the design and implementation of parallel multi-objective optimization techniques to extract association rules from large and distributed genomic and medical data.
- ACI GRID’5000 Grant (2004-2007). This project deals with the deployment of a national grid platform. Our laboratory will host a cluster of more than 250 PCs. The DOLPHIN project coordinated this action for Lille.

6.3. Contracts with Industry

- France Telecom R&D (2003-2005): The cooperation with France Telecom has been reinforced with this action. The objective here is the design and implementation of a framework for solving multi-objective optimization problems in the mobile telecommunications area.
- Mobinets (2004-2006): In this recent contract, the objective is to model and solve the access network design problem (GSM and UMTS technologies).
- Arcelor (2004-2007): A Phd thesis was supported by this company for the design of robust optimization algorithms for cutting problems.

7. Other Grants and Activities

- Member of the European Network of Excellence EvoNet: The goal of this network is to coordinate the research on evolutionary computation between academic research laboratories and industrials. Our team is associated mainly to the actions EvoTel (Evolutionary computation for Telecommunications) and EvoBio (Evolutionary computation for Bioinformatics).
- University of Malaga (Spain, 2003-2005): A collaboration has been initiated on optimization frameworks on Grids.
- University of Luxembourg (2004-2007): A common project has been initiated on robust optimization for cutting problems.

8. Dissemination

8.1. Services to the scientific community

8.1.1. Research Management

- Co-fondator and co-chair of the group META (Metaheuristics: Theory and Applications, http://www.lifl.fr/~talbi/META). This group is associated to the ROADEF (French Operations Research Society), and the CNRS research groups GDR ALP and MACS.
- Co-chair of the group PM2O (Multi-objective Mathematical Programming, http://www.lifl.fr/PM2O). This group is associated to the ROADEF (French Operations Research Society), and the CNRS research group GDR I3.
- Direction of the CIB (Bioinformatics Center) of the Genopole of Lille.
- Scientific Committee of the Genopole of Lille.

8.1.2. Participation to Working Groups

- EURO-PAREO (European working group on Parallel Processing in Operations Research).
- EURO-EU/ME (European working group on Metaheuristics).
- ECCO (European Chapter on Combinatorial Optimization).
- JET national group on evolutionary computation.
- ERCIM (European Research Consortium for Informatics and Mathematics) working group on Soft Computing.

8.1.3. Editions

- Special issue of the international journal “Parallel Computing” on Nature Inspired Distributed Computing (2004).
- Special issue of the international journal “International Journal on Foundations of Computer Science” on parallel computing for complex problem solving (2005).
- Book on “Parallel Combinatorial Optimization” (Wiley & Sons, USA, 2005).

8.1.4. Organization of sessions, workshops and conferences

- NIDISC Workshop (International Workshop on Nature Inspired Distributed Computing) organized jointly with ACM/IEEE IPDPS (International Parallel and Distributed Processing Symposium): NIDISC'2004 (Sante Fe, USA), NIDISC'05 (Denver, USA).
- Sessions on "Cooperative Optimization Methods" in the FRANCORO'2004 Conference, Aug 2004, Fribourg, Switzerland.
- International Conference EA'2005 (Artificial Evolution), Oct 2005, Lille, France.
- Co-organization of the International Conference MOPGP'2006 (Multi-objective Programming and Goal Programming), Jun 2006, France.

8.1.5. Reviews

- Review of journal papers: Parallel Computing, IEEE Transactions on Systems Man and Cybernetics, Annals of Operations Research, Calculateurs Parallèles, IEEE Transactions on Parallel and Distributed Systems, Journal of Supercomputing, IEEE Transactions on Evolutionary Computation, Parallel and Distributed Computing Practices, Genetic Programming and Evolvable Machines, Journal of Heuristics, European Journal of Operational Research, Journal of Computational Optimization and Applications, Information Processing Letters, Extraction de connaissances et apprentissage, European Physical Journal B, 4OR, Journal of Mathematical Modelling and Algorithms, ...
- Review of different projects : CRSNG project (Conseil de Recherches en Sciences Naturelles et en Génie Industriel, Canada), EPSRC project (Engineering and Physical Sciences Research Council, UK), ACI Masse de données project, ...

8.1.6. Program Committees

- International Conferences on Evolutionary Computation:
  - CEC (Congress on Evolutionary Computation): CEC'04 (Portland).
  - GECCO (Genetic and Evolutionary Computation Conference): GECCO'2005.
- International Conference on Parallel Problem Solving from Nature PPSN: PPSN VIII (Birmingham, England, 2004).
- Workshop "Parallel Solving of NP-complete Problems": RenPar'15 (La Colle sur Loup, 2004).
- EvoCOP European Conference on Evolutionary Computation in Combinatorial Optimization: EvoCOP'2004 (Coimbra, Portugal).
- EvoBIO European Workshop on Evolutionary Computation and Bioinformatics: EvoBio'2004 (Coimbra, Portugal).
- Scientific committee of the conference FRANCORO'4 (Operations Research), Fribourg, Switzerland, Aug 2004.
- MIC - Metaheuristics International Conference, Vienne, Austria, Aug 2005.
- SIAM International Conference on Data Mining, Newport Beach, USA, Apr 2005.

8.2. Academic Teaching

- Postgraduate: “Optimization methods” (E-G. Talbi).
- Postgraduate: “Data mining for bioinformatics” (E-G. Talbi, L. Jourdan).
- Undergraduate (Polytech’Lille): “Operations Research” (C. Dhaenens, E-G. Talbi).
- Undergraduate (Polytech’Lille): “Graphs and combinatorics” (C. Dhaenens, F. Seynhaeve, E-G. Talbi).
- Undergraduate (Polytech’Lille): “Data mining” (E-G. Talbi, C. Dhaenens).
- Undergraduate (Polytech’Lille): “Advanced Optimization” (E-G. Talbi, F. Seynhaeve).
- Undergraduate (Polytech’Lille): “Production Management” (C. Dhaenens).
- Undergraduate (Polytech’Lille): “Distributed Systems” (N. Melab, S. Cahon, E-G. Talbi).
- Undergraduate (Polytech’Lille): “Advanced Networking” (N. Melab).

9. Bibliography

Major publications by the team in recent years

[1] M. BASSEUR, F. SEYNHAEVE, E.-G. TALBI. *Design of multi-objective evolutionary algorithms: Application to the flow-shop Scheduling Problem*, in "Proceedings of the 2002 Congress on Evolutionary Computation (CEC), Honolulu, Hawaii, USA", vol. 2, IEEE Service Center, May 2002, p. 1151–1156.

[2] S. CAHON, N. MELAB, E.-G. TALBI. *Handbook of Bioinspired Algorithms and Applications*, Editors: S. Olariu and A.Y. Zomaya, chap. Frameworks for the Design of Reusable Parallel and Distributed Metaheuristics, CRC Press, USA, 2004.

[3] L. JOURDAN, C. DHAENENS, E. TALBI, S. GALLINA. *A Data Mining Approach to Discover Genetic and Environmental Factors involved in Multifactoral Diseases*, in "Knowledge Based Systems", vol. 15, n° 4, 2002, p. 235–242.

[4] L. JOURDAN, C. DHAENENS, E.-G. TALBI. *Evolutionary Computation in Bioinformatics*, Editors: G.B. Fogel and D.W. Corne, chap. Discovery of Genetic and Environmental Interactions in Disease Data using Evolutionary Computation, Morgan Kaufmann, 2002, p. 297–316.

[5] N. JOZEFOWIEZ, F. SEMET, E.-G. TALBI. *Parallel and hybrid models for multi-objective optimization: Application to the Vehicle Routing Problem*, in "Proceedings of the Seventh International Conference Parallel Problem Solving from Nature (PPSN), Granada, Spain", LNCS, vol. 2439, Springer-Verlag, September 2002, p. 271–280.

[6] N. MELAB, E.-G. TALBI. *Parallel adaptive computing on meta-systems including NOWs*, in "Parallel Computing", vol. 26, n° 2, 2000, p. 267–284.

[7] H. MEUNIER, E.-G. TALBI, P. REININGER. *A Multiobjective Genetic Algorithm for Radio Network Optimization*, in "Proceedings of the 2000 Congress on Evolutionary Computation (CEC), San Diego, USA", IEEE Service Center, July 2000, p. 317–324.

[8] E. Talbi. *A Taxonomy of Hybrid Metaheuristics*, in "Journal of Heuristics", vol. 8, n° 5, 2002, p. 541–564.

Doctoral dissertations and Habilitation theses

[9] N. Jozefowiez. *Modélisation et résolution approchées de problèmes de tournées multi-objectif*, Ph. D. Thesis, Université des Sciences et Technologies de Lille, France, December 2004.

[10] B. Weinberg. *Analyse et résolution approchées de problèmes d'optimisation combinatoire : application au problème de coloration de graphe*, Ph. D. Thesis, Université des Sciences et Technologies de Lille, France, September 2004.

Articles in referred journals and book chapters

[11] M. Basseur, F. Seynhaeve, E.-G. Talbi. *A Cooperative Metaheuristic applied to Multiobjective Flow-Shop Problem*, in "Real-world Multi-objective System Engineering, USA", Editors: N. Nedjah and L.M. Mourelle, Nova Science Publishers, 2004.

[12] S. Cahon, N. Melab, E.-G. Talbi. *Building with ParadisEO reusible parallel and distributed evolutionary algorithms*, in "Parallel Computing", vol. 30, n° 5-6, 2004, p. 677–697.

[13] S. Cahon, N. Melab, E.-G. Talbi. *ParadisEO: A Framework for the Reusable Design of Parallel and Distributed Metaheuristics*, in "Journal of Heuristics", vol. 10, n° 3, 2004, p. 357–380.

[14] C. Cotta, E.-G. Talbi, E. Alba. *Parallel hybrid approaches*, in "Parallel Metaheuristics, USA", to appear, Editors: E. Alba, J. Wiley and Sons, 2005.

[15] L. Jourdan, M. Khabzaoui, C. Dhaenens, E.-G. Talbi. *A hybrid metaheuristic for knowledge discovery in microarray experiments*, in "Handbook of Bioinspired Algorithms and Applications, USA", to appear, Editors: S. Olariu and A.Y. Zomaya, CRC Press, 2005.

[16] N. Jozefowiez, F. Semet, E.-G. Talbi. *A multi-objective evolutionary algorithm for the covering tour problem*, in "Applications of multi-objective evolutionary algorithms", Editor: C.C. Coello, World Scientific, 2004.

[17] N. Melab, E.-G. Talbi, S. Cahon. *Frameworks for the Design of Reusable Parallel and Distributed Metaheuristics*, in "Handbook of Bioinspired Algorithms and Applications, USA", to appear, Editors: S. Olariu and A.Y. Zomaya, CRC Press, 2005.

[18] N. Melab, E.-G. Talbi, M.-S. Mezmaz, B. Wei. *Parallel Hybrid Multi-objective Meta-heuristics on P2P Systems*, in "Handbook of Bioinspired Algorithms and Applications, USA", to appear, Editors: S. Olariu and A.Y. Zomaya, CRC Press, 2005.

[19] A. J. Nebro, F. Luna, E.-G. Talbi, E. Alba. *Parallel multi-objective optimization*, in "Parallel Metaheuristics, USA", to appear, Editors: E. Alba, J. Wiley and Sons, 2005.

[20] L. Vermeulen-Jourdan, E. Talbi, C. Dhaenens. *Linkage disequilibrium study with a parallel adaptive GA*, in "Intl. Journal of Foundation in Computer Science", to appear, 2005.

[21] B. WEINBERG, E.-G. TALBI. *A cooperative parallel metaheuristic applied to the graph coloring problem*, in "Handbook of Bioinspired Algorithms and Applications, USA", to appear, Editors: S. Olariu and A.Y. Zomaya, CRC Press, 2004.

Publications in Conferences and Workshops

[22] M. BASSEUR, J. LEMESRE, C. DHAENENS, E.-G. TALBI. *Cooperation between Branch and Bound and Evolutionary Approaches to Solve a Bi-objective Flow Shop Problem*, in "Proceedings of the International Workshop on Experimental and Efficient Algorithms (WEA), Angra dos Reis, Rio de Janeiro, Brazil", LNCS, vol. 3059, May 2004, p. 72–86.

[23] M. BASSEUR, F. SEYNHAEVE, E.-G. TALBI. *Coopération de méthodes d'optimisation appliqué au FlowShop MultiObjectif*, in "Quatrièmes Journées Francophones de Recherche Opérationnelle (FRANCORO), Friburg, Switzerland", August 2004.

[24] M. KHABZAOUI, C. DHAENENS, E.-G. TALBI. *A Multicriteria Genetic Algorithm to analyze DNA microarray data*, in "Proceedings of the 2004 Congress on Evolutionary Computation (CEC), Portland, Oregon, USA", vol. 2, IEEE Service Center, June 2004, p. 1874–1881.

[25] M. KHABZAOUI, C. DHAENENS, E.-G. TALBI. *Association rule discovery for DNA microarray data*, in "Proceedings of the SIAM Bioinformatics Workshop 2004 (in conjunction with the fourth SIAM International Conference on Data Mining), Orlando, USA", April 2004, p. 63–71.

[26] J. LEMESRE, C. DHAENENS, E.-G. TALBI. *A Parallel Exact Method for a Bicriteria Permutation Flow-Shop Problem*, in "Proceedings of the Ninth International Conference on Project Management and Scheduling (PMS), Nancy, France", April 2004, p. 359–362.

[27] J. LEMESRE, C. DHAENENS, E.-G. TALBI. *A Parallel Exact Scheme to solve Bicriteria Problems*, in "Sixth International Multi-Objective Programming and Goal Programming Conference (MOPGP), Hammamet, Tunisia", April 2004.

[28] M.-S. MEZMAZ, N. MELAB, E.-G. TALBI. *Towards a Coordination Model for Parallel Cooperative P2P Multi-objective Optimization*, in "Springer Verlag LNCS, Proc. of European Grid Conf. (EGC'2005), Amsterdam, The Netherlands", to appear, February 2005.

[29] E.-G. TALBI, S. CAHON, N. MELAB. *Parallel design of cellular networks*, in "Twentieth European Conference On Operational Research (EURO), Rhodes, Greece", July 2004.

[30] L. VERMEULEN-JOURDAN, C. DHAENENS, E. TALBI. *A Parallel Adaptive Genetic Algorithm for Linkage Disequilibrium in Genomics*, in "Sixth Workshop on Nature Inspired Distributed Computing (NIDISC) (in conjunction with IPDPS'2004), Santa Fe, New Mexico, USA", April 2004.

[31] L. VERMEULEN-JOURDAN, C. DHAENENS, E. TALBI. *Clustering Nominal and Numerical Data: A New Distance Concept for a Hybrid Genetic Algorithm*, in "Proceedings of the Fourth European Conference on Evolutionary Computation in Combinatorial Optimization (EvoCOP), Coimbra, Portugal", LNCS, vol. 3004, Springer Verlag, April 2004, p. 220–229.

[32] B. WEINBERG, E.-G. TALBI. *On Search Space Symmetry in Partitionning Problems*, in "Proceedings of the Fourth European Conference on Evolutionary Computation in Combinatorial Optimization (EvoCop), Coimbra, Portugal", LNCS, vol. 3004, Springer Verlag, April 2004, p. 230–240.

[33] B. WEINBERG, E.-G. TALBI. *NFL theorem is unusable on structured classes of problems*, in "Proceedings of the 2004 Congress on Evolutionary Computation (CEC), Portland, Oregon, USA", vol. 1, IEEE Service Center, June 2004, p. 220–226.

Bibliography in notes

[34] C. A. COELLO COELLO, D. A. VAN VELDHUIZEN, G. B. LAMONT (editors). *Evolutionary algorithms for solving multi-objective problems*, Kluwer Academic Press, 2002.

[35] K. DEB. *Multi-objective optimization using evolutionary algorithms*, John Wiley and sons, 2001.

[36] D. E. GOLDBERG. *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, Massachusett, USA, 1989.

[37] R. JOHNSON, B. FOOTE. *Designing Reusable Classes*, in "Journal of Object-Oriented Programming", vol. 1, n° 2, 1988, p. 22–35.

[38] H. MEUNIER. *Algorithmes évolutionnaires parallèles pour l'optimisation multi-objectif de réseaux de télécommunications mobiles*, Ph. D. Thesis, Université des Sciences et Technologies de Lille, France, June 2002.

[39] E. TALBI. *A Taxonomy of Hybrid Metaheuristics*, in "Journal of Heuristics", vol. 8, n° 5, 2002, p. 541–564.

[40] S. WRIGHT. *The roles of mutation, imbreeding, crossbreeding and selection in evolution*, in "Proc. of the sixth Congress on Genetics", 1932, p. 356–366.

[41] E. ZITZLER, L. THIELE, M. LAUMANNS, C. M. FONSECA, D. V. G. FONSECA. *Performance Assessment of Multiobjective Optimizers: an Analysis and Review*, in "IEEE Transactions on Evolutionary Computation", vol. 7, April 2003, p. 117–132.