INRIA

Project-Team DOLPHIN

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

Futurs

THEME NUM

Activity

Report

2007

Table of contents

1. Team

Head of the project
El-Ghazali Talbi [Professor Polytech'Lille - USTL, Research Scientist (on partial secondment) since september 2006, HdR]

Vice-head of the project
Nouredine Melab [Professor IEEA - USTL, HdR]

Administrative assistant
Corinne Davoust [Secretary INRIA]

INRIA project staff
Laetitia Vermeulen-Jourdan [Research Scientist (CR2) INRIA]

University project staff
Clarisse Dhaenens [Professor Polytech'Lille - USTL, HdR]
Francois Clautiaux [Assistant Professor IUT - USTL]
Bilel Derbel [Assistant Professor IEEA - USTL from Sept 2007]

Post-doctoral fellow
Oliver Schütze [INRIA Postdoc, up to June 2007]
Joseph El Hayek [ATER IUT - USTL, from Sept 2007]

PhD students
Jean-Charles Boisson [Research ministry support, Third year]
Jérôme Brongniart [ANR support, Second year]
Ali Khanafer [INRIA - Region support, First year]
Arnaud Liefooghe [ANR support, Second year]
Cédric Luit [Region support, Second year]
Malika Mehdi [co-tutelle with University of Luxembourg, First year]
Mohand-Said Mezmaz [Research ministry support, Fifth year, Thesis in November 2007]
Alexandru Tantar [Research ministry support, Third year]
Emilia Tantar [Region support, Third year]

Project technical staff
Thomas Legrand [Associate Engineer]
Clive Canape [Associate Engineer, from Sept 2007]

Internships
Antonio LaTorre de la Fuente [Visiting student, sept-nov 2007]
Ignacio Zunino [Visiting student, Nov 2006 - March 2007]

2. Overall Objectives

2.1. Presentation

The goal of the DOLPHIN[1] 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 approaches which generate the whole Pareto set of a given Multi-Objective Problem (MOP). For this purpose, the research actions can be resumed as follows:

- **Modeling and Analysis of MOPs:** Solving Multi-Objective Problems requires an important analysis phase to find the best suitable method to solve it. This analysis deals with the modeling of the problem and the analysis of its structure.

---

[1] Discrete multi-objective Optimization for Large scale Problems with Hybrid dIstributed techNiques.

To propose efficient models for a Multi-Objective Optimization problem, an important aspect is to integrate all the constraints of the problem. Therefore an interesting preliminary approach is to develop efficient models for the problem in its mono-objective forms in order to be able to develop methods that are taking into account characteristics of the studied problem. While studying the problem in its multi-objective form, the analysis of the structure is another interesting way.

The analysis of the structure of the Pareto front by means of different approaches (statistical indicators, meta-modeling, etc.) 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 (No Free Lunch) theorem shows that two optimization methods have the same global performance on the whole set of uniform optimization problems. Then, it is crucial to make some hypothesese 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.

Our interest in this project is to answer these questions and remarks for the multi-objective case. Another point considered is the performance evaluation of multi-objective optimization methods. We are also working on approximation algorithms with performance guarantee and the convergence properties of stochastic algorithms.

- **Cooperation of optimization methods (metaheuristics and/or exact methods):**

  The hybridization 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 the metaheuristic 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, access network design, 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/or to improve the robustness of a given method. Following this objective, we 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.

- **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 [2], 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 metaheuristics for multi-objective optimization; iii) the design of hybrid methods; iv) the development of parallel and distributed models.

  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, Globus, PThreads). The coupling with well-known frameworks for

---

[2]This framework was initially developed by Geneura TEAM (Spain), INRIA (France), LIACS (Netherlands). http://paradiseo.gforge.inria.fr.

exact methods (such as COIN) will also be considered. The exact methods for MOPs developed in this project will be integrated in those software frameworks.

The experimentation of this framework by different users and applications outside the DOLPHIN project is considered. This is done in order to validate the design and the implementation issues of ParadisEO.

- **Validation:** the designed approaches are validated on generic and real-life MOPs, such as:
    – scheduling problems: Flow-shop scheduling problem;
    – routing problems: Vehicle routing problem (VRP), 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, protein identification, docking and conformational sampling of molecules.
    – engineering design problems: Design of polymers.

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

2.2. Highlights of the year

Best Paper Award from GECCO'2007 (Genetic and Evolutionary Computation Conference) for the article "Convergence of stochastic search algorithms to gap-free Pareto Front approximations", **O. Schütze**, M. Laumanns, **E. Tantar**, C.A.C. Coello, **E-G. Talbi**.

3. Scientific Foundations

3.1. Modeling and landscape analysis

The modeling of problems, 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 problem and its landscape (roughness, convexity, etc). The notion of landscape has been first described in [92] by the way of the study of species evolution. Then, this notion has been used to analyze combinatorial optimization problems.

3.1.1. *Modeling of problems*

Generally there are several ways of modeling a given problem. First, one has to find the model that is the more suited for the type of resolution he plans to use. The choice can be made after a theoretical analysis of the model, or after computational experiments. The choice of the model depends on the type of method used. For example, a major issue in the design of exact methods is to find tight relaxations for the problem considered.

Let us note that many combinatorial optimization problems of the literature have been studied in their mono-objective form even if a lot of them are naturally of a multi-objective nature.

Therefore, in the Dolphin project, we address the modeling of MOPs in two phases. The first one consists in studying the mono-objective version of the problem, where all objectives but one are considered as constraints. In the second phase, we propose methods to adapt the mono-objective models or to create hand-tailored models for the multi-objective case. The models used may come from the first phase, or from the literature.

---

[3]Graphical User Interface for Multi-Objective Optimization (http://guimoo.gforge.inria.fr).

3.1.2. *Analysis of the structure of a problem*

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.

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 the search for the most adapted method to an instance or a set of instances of a given problem. Therefore, we are convinced 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 shown 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, that 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 treated instance.

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 [90], [91], [89], [88].

3.1.3. *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 can be adapted 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 cannot 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 [93]. The OPAC research group has proposed two indicators: the *contribution* and the *entropy* [85]. 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 solutions found. 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.

3.1.4. *Goals*

One of the main issues in the DOLPHIN project is the study of the landscape of 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 [84]. 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/her.

Population-based methods, such as evolutionary algorithms, are well fitted for multi-objective problems, as they work with a set of solutions [83], [80]. To be convinced one may refer to the list of references on Evolutionary Multi-objective Optimization maintained by Carlos A. Coello Coello [4] which contains almost 3 000 references. One of the objectives of the project is to propose advanced search mechanisms for intensification and diversification. These mechanisms have been designed in an adaptive manner, since their effectiveness is related to the landscape of the MOP and to the instance solved.

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 last 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, which are either academic (traveling salesman problem, quadratic assignment problem, scheduling problem, ...) or real-world problems. This efficiency is generally 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 [87]. It proposes to decompose those mechanisms into 4 classes:

---

[4]http://www.lania.mx/~ccoello/EMOO/EMOObib.html

- *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.

- *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.

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 a complete large scale problem. Therefore, sub search spaces have to be defined in order to be able to use exact methods. Hence, in the same manner as hybridization of metaheuristics, the cooperation of metaheuristics and exact methods is also a main issue in this project. Indeed, it allows to use the exploration capacity of metaheuristics, as well as the intensification ability of exact methods, which are able to find optimal solutions in a restricted search space. Sub search spaces 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 cooperate. 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 spaces [81].

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 which 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:

- *Design of metaheuristics for the multi-objective optimization:* To 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 hybridization of metaheuristics 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 is achieved, we will have to test and compare them in the multi-objective context.

- *Design and conception of parallel metaheuristics:* Our previous 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.

- *Validation:* In order to validate results obtained we always 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.

  Moreover, those advanced mechanisms are to be used in order to integrate the 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 improvements on 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 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.

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 returns 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 [82], [86].

3.3.2. 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 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 (Symmetric Multi-processors) machines. For GRID computing platforms, peer to peer (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.

4. Application Domains

4.1. Academic generic problems

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

- **Workshop optimization problems:**

  Workshop optimization problems deal with optimizing the production. In this project, two specific problems are under study.

  - **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 minimized objective 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. A tri-criteria flow-shop problem, minimizing in addition the maximum tardiness, is also studied. It will allow to develop and test multi-objective (and not only bi-objective) exact methods.

  - **Cutting problems:** Cutting problems occur when pieces of wire, steel, wood, or paper have to be cut from larger pieces. The objective is to minimize the quantity of lost material. Most of these problems derive from the classical one-dimensional cutting-stock problem, which have been studied by many researchers. The problem studied by the DOLPHIN project is a two-dimensional bi-objective problem, where rotating a rectangular piece has an impact on the visual quality of the cutting pattern. First we have to study the structure of the cutting-stock problem when rotation is allowed, then we will develop a method dedicated to the bi-objective version of the problem.

- **Logistics and transportation problems:**

  - **Packing problems:** In logistic and transportation fields, packing problems may be a major issue in the delivery process. They arise when one wants to minimize the size of a warehouse or a cargo, the number of boxes, or the number of vehicles used to deliver a batch of items. These problems have been the subjects of many papers, but only few of them study multi-objective cases, and to our knowledge, never from an exact point of view. Such a case occurs for example when some pairs of items cannot be packed in the same bin. The DOLPHIN project is currently studying the problem in its one-dimensional version. We plan to generalize our approach to two and three dimensional problems, and to more other conflict constraints, with the notion of *distance* between items.

  - **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. transportation, logistics, ...). 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 works that tackle a classic mono-objective routing problem by relaxing constraints and building a more general MOP.

The third studied routing problem is the Ring Star Problem (RSP). This problem consists in locating a simple cycle through a subset of nodes of a graph while optimizing two kinds of costs. The first objective is to minimize a ring cost that is related to the length of the cycle. The second one is to minimize an assignment cost from non-visited nodes to visited ones. In spite of its natural bi-criteria formulation, this problem has always been studied in a single-objective form where either both objectives are combined or one objective is treated as a constraint.

Recently, within a cooperation with SOGEP, the logistic and delivery subsidiary company of REDCATS (PINAULT PRINTEMPS REDOUTE), a new routing problem is under study. Indeed, the COLIVAD project consists in solving a logistic and transportation problem that has been reduced to a vehicle routing problem with additional constraints. First we are designing a method to solve exactly a bi-objective version of the problem in order to evaluate the interest of modifying the current process of delivery. We are also working on the resolution of a single-objective version of this problem to design an operational tool dedicated to the SOGEP problem.

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.

4.2. 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.

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.).

- 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.

4.3. Application to Bioinformatics

Bioinformatic research is a great challenge for our society and numerous research entities of different specialities (biology, medical or information technology) collaborating on specific thema.

4.3.1. *Genomic and Post-Genomic studies*

Previous studies of the DOLPHIN project mainly deal with genomic and postgenomic applications. These have been realized in collaboration with academic and industrial partners (IBL: Biology Institute of Lille; IPL: Pasteur Institute of Lille; IT-Omics firm).

First, genomic studies aim to analyze genetic factors which may explain multi-factorial diseases such as diabetes, obesity or cardiovascular diseases. The scientific goal was to formulate hypotheses describing associations that may have any influence on diseases under study.

Secondly, in the context of post-genomic, a very large amount of data are obtained thanks to advanced technologies and have to be analyzed. Hence, one of the goals of the project was to develop analysis methods in order to discover knowledge in data coming from biological experiments.

These two problems have been modeled as classical datamining tasks. First it as been modeled as an association-rule mining problem. As the combinatoric of such problems is very high and the quality criteria not unique, we proposed to model the association-rule mining problem as a multi-objective combinatorial optimization problem. An evolutionary approach has been adopted in order to cope with large scale problems. Then, in order to make more efficient such approaches, a complementary datamining task has been studied: the feature selection problem. This problem is of multi-objective nature and a multi-objective approach has been proposed.

4.3.2. *Protein identification*

Another application deals with Proteomics. Proteomics consists in the global analysis of proteins. In fact, proteomics is very important to understand the biological mechanisms in the living cells, but also how different factors can influence them. The main goal of the proteomic is to identify experimental proteins. In this domain, we collaborate with the team of C. Rollando (research Director at CNRS) head director of the proteomic platform of the genopole of Lille.

Our objective is to automatically discover proteins and new protein variants from experimental spectrum. The protein variants and new protein identification is a complex problem. In fact, it cannot be summarized as a simple scoring of an experimental protein against protein databases, it needs additional processes to explore the huge space of potential solutions. For the protein variant, there are many modifications: insertion, deletion or substitution of an amino acid and also post-traductional modifications on it. So it is not practically feasible to generate all the possibilities of combination of modifications for a given size of protein (exponential complexity). The new protein identification problem is close because we cannot generate all possible proteins (with also modifications) in order to find the experimental one. For both a method of optimization is necessary.

4.3.3. *Docking and conformational sampling*

In molecular modelling, conformational sampling and docking procedures allow to provide help for understanding the interaction mechanisms between (macro)molecules involved in physiological processes. The processes to be simulated are of a combinatorial complexity (molecule size, number of degrees of freedom) that represents an important challenge for the currently available computing power. Such challenge can be expressed by three major objectives: (1) the proposition of mathematical models of maximum simplicity that nevertheless provide a relevant description of molecular behavior, (2) the development of powerful distributed optimization algorithms (evolutionary algorithms, local search methods, hybrid algorithms) for sampling the molecular energy surface for stable, populated conformations, and (3) deploying those intrinsic distributed algorithms on computational Grids.

Within the framework of ANR DOCK and Decrypton projects, the focus is to propose with the collaboration of Institute of Biology at Lille (H. Dragos) multi-objective formulations of the conformational and docking problems. The goal is to take into account different criteria characteristics of the complex docking process. Furthermore, in order to deal with the multimodal nature of the problems it is important to define new hybrid mechanisms allowing to provide algorithms with both diversification and intensification properties. Finally, to deal with the exponential combinatory of these problems when large proteins are concerned parallel and grid computing is highly required. Using grid computing is not straightforward, so a "gridification" process is necessary. Such process allows to adapt the proposed algorithms to the characteristics of the grid. The gridification process must be exploited by the user in a transparent way. Therefore, coupling Paradiseo-Peo with a generic grid middleware such as Globus is important to provide robust and efficient algorithms to be exploited transparently.

5. Software

5.1. ParadisEO

Participants: Jean-Charles Boisson, Clive Canape, Laetitia Vermeulen-Jourdan, Thomas Legrand, Arnaud Liefooghe, Nouredine Melab, El-Ghazali Talbi [correspondent], Alexandru Tantar.

ParadisEO (PARallel and DIStributed Evolving Objects) is a C++ white-box object-oriented framework dedicated to the flexible design of metaheuristics. See web pages http://paradiseo.gforge.inria.fr/. Based on EO, a templates-based ANSI-C++ compliant evolutionary computation library, it is composed of four modules:

- Paradiseo-EO provides tools for the development of population-based metaheuristic (Genetic algorithm, Genetic programming, Particle Swarm Optimization (PSO)...)
- Paradiseo-MO provides tools for the development of single solution-based metaheuristics (Hill-Climbing, Tabu Search, Simulated annealing, Iterative Local Search (ILS), Incremental evaluation, partial neighborhood...)
- Paradiseo-MOEO provides tools for the design of Multi-objective metaheuristics (MO fitness assignment shemes, MO diversity assignment shemes, Elitism, Performance metrics, Easy-to-use standard evolutionary algorithms...)
- Paradiseo-PEO provides tools for the design of parallel and distributed metaheuristics (Parallel evaluation, Parallel evaluation function, Island model)

Furthermore, ParadisEO also introduces tools for the design of distributed, hybrid and cooperative models:

- High level hybrid metaheuristics: coevolutionary and relay model
- Low level hybrid metaheuristics: coevolutionary and relay model

The ParadisEO framework has been especially designed to best suit to the following objectives:

- Maximum design and code reuse: ParadisEO is based on a clear conceptual separation of the solution methods from the problems they are intended to solve. This separation confers to the user a maximum code and design reuse.
- Flexibility and adaptability: the fine-grained nature of the classes provided by the framework allow a higher flexibility compared to other frameworks.
- Utility: ParadisEO 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. ParadisEO is one of the rare frameworks that provide the most common parallel and distributed models. These models can be exploited in a transparent way, one has just to instantiate their associated provided classes.
- Portability: The implemented models are portable on distributed-memory machines as well as on shared-memory multiprocessors, as they use standard libraries such as MPI and PThreads.

5.1.1. ParadisEO-EO and Particle Swarm Optimization integration

ParadisEO-EO is a templates-based, ANSI-C++ compliant evolutionary computation library. It contains classes for almost any kind of evolutionary computation. EO was started by the Geneura Team at the University of Granada but the development team has been many times reinforced. Recently, we joined the developer staff to start a new contribution with ParadisEO. The goal was to create new classes and components increasing the compatibility between the framework (ParadisEO-EO) and its extensions (ParadisEO-MO, ParadisEO-MOEO and ParadisEO-PEO). Several technical features have also been improved from both sides. Furthermore, a set of classes allowing the implementation of any Particle Swarm Optimization (PSO) algorithm has been proposed to the EO community. As it was successfully integrated and tested, an extension of the sequential PSO is actually in progress and will be finalized before the end of the year. Thus, ParadisEO will propose a full support for the PSO: sequential models (including many topologies, binary flight ...), parallel and distributed models (evaluation function, island scheme...).

5.1.2. Paradiseo-MOEO

Paradiseo-MOEO (Multi-Objective Evolving Objects) is the module dedicated to multi-objective optimization. It embeds some features and techniques for scalar and Pareto-based resolution. A genuine conceptual effort has been done to provide a set of classes allowing to ease and speed up the development of efficient programs in an incremental way while having a minimum programming effort. ParadisEO-MOEO provides a broad range of fine-grained components including fitness assignment strategies (the achievement functions, the schemes used in NSGA, IBEA and more), the most common diversity preservation mechanisms (sharing, crowding), some elitist-related features as well as statistical tools. This year, the whole set of classes and templates have been completely updated in order to confer a higher genericity, flexibility, adaptability and extensibility. Some state-of-the-art evolutionary algorithms (NSGA-II, IBEA) have been added to the library aiming to be used in an easy way.

5.1.3. Paradiseo-MO

ParadisEO-MO (Moving Object) is dedicated to the design of solution-based metaheuristics. It is based on C++ template and is problem independent. The first version of ParadisEO-MO provided three algorithm schemes: the hill climbing, tabu search and simulated annealing schemes. It provided also an application example on the symmetric traveling salesman problem (TSP). During this year, this platform has been greatly improved. On the one hand a new algorithm scheme has been added: the iterated local search (ILS); and a lot of "ready-to-use" box have been provided: stopping criteria, cooling scheduler schemes... On the other hand, a complete set of documentation has been added: a source code documentation, two combinatorial concept presentations, a report that full describes the platform and four lessons, respectively for each algorithm scheme, solving the symmetric TSP problem.

5.1.4. Paradiseo-G

ParadisEO has been coupled with Globus GT4 to tackle optimization problems on Grids. The coupling of ParadisEO with Globus consisted in two major steps: design and implementation, and deployment on the Grid, in particular Grid'5000. The first step consisted in the gridification of the parallel and hybrid models provided in ParadisEO meaning their adaptation to the properties of grids (large scale, heterogeneous and dynamic nature of the resources and multi-administrative domain). The MPICH-G2 communication library has been considered. The second step consisted in building a system image for Globus 4 including MPICH-G2. This image allows to build a virtual Globus grid able to deploy and execute the parallel hybrid meta-heuristics provided by ParadisE0. This year, a new archive containing ParadisEO has been built and it can be used either under classical environments either with Globus. From the user point of view, ParadisEO consists now in a single package that can be deployed to best suit to the environment.

5.1.5. New technical features

Recently, many efforts have been made for ParadisEO to become a cross-platform easy-to-use software. All the documentation, source code, articles and resources have been gathered to build the "ParadisEO GForge

project". The INRIA GForge (http://gforge.inria.fr/) provides a set of web-interfaced utilities that allow an advanced project management. The ParadisEO project is now composed of a website, several forums and mailing-lists, a subversion repository, many announce and task publishers... Moreover, as the components (EO, MO, MOEO, PEO) were initially downloaded and installed separately, a lot of problems came-up because of dependencies and it could be difficult for a non experimented user to compile and run the whole library. Therefore, a single archive, containing the four modules (including the sources, the framework documentation, new tutorials) has been built and an installation script has been proposed to the users. The install process is now performed automatically. Another important change has also been operated to allow the compilation on any platform having a standard C++ compiler. The build process is now managed by CMake and the classical autotools (Autoconf, Automake) have been forgotten. Thanks to this migration, ParadisEO can now be used under several environments (Windows, Unix-like systems, Mac).

5.2. Ascq Mass-Engine and DeNoProtS

Participants: Jean-Charles Boisson, Laetitia Vermeulen-Jourdan [correspondent], El-Ghazali Talbi.

$ASCQ\_ME$ is a protein identification engine designed for peptide mass fingerprinting from raw MS spectra. During this year, it has been optimized and improved to be able to make protein identification from tandem mass spectrometry (mono charged MS/MS spectra).

According to the software configuration, each protein is digested into peptides. Then each peptide is fragmented into ions. Theoretical spectra are generated from these ions and compared to the experimental MS/MS spectra. The score of each peptide is based on the percentage of peptides having a "good" spectral correlation with the corresponding MS/MS spectra. The protein score is the mean of these peptide scores.

The SSO (Sequence - Shape - Order) software designed for providing a real *d*e novo protein sequencing has been also improved:
Sequence: a simplified method of de novo peptide sequencing method has been implemented. This step can be replaced by any other de novo peptide sequencing algorithm.
Shape: according to a MS spectrum, the aim of the method is to complete the missing information of each peptide sequence. Order: the two previous steps aim to find the sequence of each peptide that compounds the experimental protein. But these peptides need to be ordered to gain the experimental protein. It is the aim of this last step. According to a MS spectrum (different from the shape step), a solution-based metaheuristic is used to find the right order.

The $ASCQ\_ME$ application is available for on-line interrogation on the site https://www.genopole-lille.fr/logiciel/ascq_me/.

5.3. Docking@Grid

Participants: Jean-Charles Boisson, Laetitia Vermeulen-Jourdan, Nouredine Melab, El-Ghazali Talbi [correspondent], Alexandru Tantar.

Docking@GRID is a software dedicated to the flexible conformational sampling and docking on the computational grid. The goal of the software is to help users to perform such processes in a friendly way. In other words, the software provides a web portal for remote job submission, importation/preparation of proteins, access to protein data bank, visualization, efficient sampling and docking. The project could be later integrated into the larger platform of chemioinformatics tools under construction around the site of the "Chimiothèque Nationale"-project of the CNRS (Prof. Hibert, Strasbourg). This platform, designed as a portail for the display of the collections of molecules synthesized in French academic labs might offer predicted affinities of these compounds with respect to various biologically interesting targets, in order to facilitate compound selection.

Docking@GRID is currently available online on the Lille Genopole server and accessible at http://docking. futurs.inria.fr. The current version considers only sampling and visualization of conformations. A registration step requiring a reduced amount of information is demanded in order to access the provided resources. The software offers a hierarchical perspective, allowing to group different tasks into projects. A new project can be created by accessing the Ligands/ActiveSite section - following this initial phase, the project is displayed in a hierarchical manner. After creating the project, the user has the possibility of creating a new molecule file by employing the Msketch application (Chemaxon), which is provided in the form of a Java applet. Following this process, a conformational sampling step can be applied on the specified file. The sampling process relies on a hybrid genetic hierarchical algorithm executed in a distributed environment and making use of different parallelization strategies. The underlying framework is ParadisEO-G which is a Globus based version of the ParadisEO framework. The parallelization of the algorithm is transparently achieved by making use of the ParadisEO models - asynchronous island model, parallel evaluation of the population and parallel synchronous multi-start model. As middle-ware an MPICH-based distribution of MPI is used, the execution being performed on a dedicated set of machines.The results are displayed at the end of the sampling process, a notification mail being sent in case the processing step takes longer than 5 seconds. The obtained conformations may be visualized by using the MView tool (Chemaxon). Each conformation can be further subjected to rigid transformations (translations, rotations), animations can be constructed, etc.

6. New Results

6.1. Interactive multi-objective optimization method based on landscape analysis

Participants: Clarisse Dhaenens, El-Ghazali Talbi, Emilia Tantar.

In the context of multi-objective optimization, interactive methods are proposed in order to incorporate the user preferences during the optimization process. Several generic paths of interaction are identified for hybrid metaheuristics. Also, through means of visual guiding components, information provided by *a priori* landscape analysis is integrated in the interactive process.

Interactive methods play an important role in solving large instances for NP-hard multi-objective problems. By their possibility of including valuable topological information they provide the means of speeding-up the search process. The information concern the structure of the set of feasible solutions or of the Pareto set - the set of best compromise solutions.

The landscape analysis, seen as an *a priori* aspect, or as part of the search process, allows providing performance guarantee for the studied search spaces.

For bi-objective combinatorial problems, a technique employing ellipses for approximating the enclosing shape of the set of feasible solutions is proposed. Its generality for the overall set of bi-objective combinatorial problems is not intended, although the use of ellipses is fully justified for problem instances that respect the Lindenberg-Levi's central limit theorem. A forthcoming interactive path has been applied on the permutation flow-shop problem. The hybridation mechanism made use of reference points in guiding the search towards the ellipse (see [64]).

Other topological aspects of the landscape, both in the objective and decisional space, where studied for different types of problems.

Although thought to be an easy to handle aspect, the structure of the neighborhood of a feasible solution, raised new issues. For combinatorial problems, which mainly deal with binary variables, the occurrence of non-connectedness in the neighborhood can be the cause of failure for local search techniques. Our study is dedicated in overcoming this difficulty by keeping during the search process approximate solutions within a threshold limit. The technique not only guarantees convergence in limit but also provides the desired level of accuracy, which is of great importance for robust design.

6.2. Exact methods for multi-objective problems

Participants: Clarisse Dhaenens, Julien Lemesre, El-Ghazali Talbi.

Exact approaches in multi-objective optimization have not been widely studied. Our aim was to propose an efficient multi-objective method able to deal with difficult optimization problems. First, an effective method has been proposed for a specific flowshop bi-objective problem [20]. Then a general scheme, called PPM - Parallel Partitioning Method - has been proposed for any bi-objective combinatorial optimization problem [21]. This general scheme has been extended for multi-objective optimization problems (problems having more than two objectives) and is called k-PPM.

This method is inherently parallel. Indeed, the method is based on the splitting of the search space into several areas leading to elementary exact searches and determines all the Pareto front into three stages: 1/ Bounding the search space. 2/ Partitioning the search space into well-balanced partitions. 3/ Search the other efficient solutions in each partition.

The parallel design of the algorithm increases its performance. It has been applied on a three-objective flow-shop problem and gave good results.

6.3. Optimization Methods for Stochastic Multi-objective Problems

Participants: Laetitia Vermeulen-Jourdan, Arnaud Liefooghe, El-Ghazali Talbi.

The importance of multi-objective optimization is globally established nowadays. Furthermore, a large part of real-world optimization problems are also subject to uncertainties due to, e.g., noisy or approximated objective function(s), varying parameters or dynamic environments. Moreover, although evolutionary algorithms are commonly used to solve multi-objective optimization problems on the one hand and stochastic optimization problems on the other hand, very few approaches combine these two aspects simultaneously.

Several search methods have been designed for the particular case of stochastic multi-objective optimization problems. Contrary to other approaches, we do not consider that there is a true evaluation per solution which is blurred by noise. The resulting methods do not make any assumption on a probability distribution associated to environmental parameters or to objective functions, as they are able to handle any types of uncertainty. A preliminary study has been published in [46], but we are working on new algorithms that are currently experimented on a bi-objective flow-shop scheduling problem with stochastic processing times.

Furthermore, another fundamental aspect is the performance assessment of optimizers dealing with uncertainty. To our knowledge, no protocol fully adapted to evaluate the effectiveness of optimization methods for stochastic multi-objective problems exists by now. Different kinds of methods have been identify and are currently being tested.

6.4. New insights into the structure of cutting and packing problems

Participants: François Clautiaux, Joseph El Hayek.

We proposed new insights into the structure of several cutting and packing problems. They led to improved results for classical problems.

First we studied linear programming approaches for the classical one-dimensional cutting-stock problem. We proposed new cuts to exclude dual solutions that cannot yield improved results when a given set of solutions have been previously computed. Then we surveyed and analyzed several families of so-called *dual-feasible functions*, which are in fact solutions of the dual problem we considered. We generalized some approaches, and led to state-of-the-art lower bounds for the cutting stock problem.

Finally, we studied a two-dimensional packing problem and proposed a way of deriving tight lower bounds when the rotation of the items is allowed. Our method is based on a new relaxation of the problem, which is inspired from *langrangian relaxation* techniques. The method proposed is simple and has the complexity of the lower bound for the fixed-orientation case that is used as a subroutine. We plan to consider the bi-objective version of this problem, for which we will adapt our latest results described for the mono-objective problem.

6.5. TopTree: an interesting structure for tree optimization

Participants: Jérôme Brongniart, Clarisse Dhaenens, El-Ghazali Talbi.

Most of the NP-difficult spanning tree optimisation problems are solved using static methods, mainly based on Prims or Kruskal algorithms. This kind of methods allows to take into account some additionnal constraints and are successfuly applied with genetic algorithms, local search or constraint programming fields.

But, due to the static nature of the algorithm structure used, the information computed during the search cannot be easily updated. This drawback has an important impact on computational time required for difficult instances.

TopTree is a dynamic algorithm, firstly proposed by Alstrup et al., which allows to maintain subtrees and path information in logarithmic time. We have succesfuly adapted this method to the cycle elimination and cost based filering methods used for tree-based problems in constraint programming and this structure manage to speed up the main operations of genetic algorithms and local search due to its abillity to update complex cost functions and constraint informations [36].

6.6. Realistic formulation for UTRAN design

Participants: Jérôme Brongniart, Clarisse Dhaenens, El-Ghazali Talbi.

The Third-Generation mobile technologie provides a large number of services, from voice to multimedia transfer. This new demand requests the definition of different classes of traffic, with different delay and error-rate requirements. To fulfill these specifications, more complex scheduling schemas must be setting up.

Traditional optimisation algorithms for the access network design do not take into consideration the Quality of Service requirement imployed by these new services. Usually, in these algorithms, the influence of this important property is not taken into account in the choice of the network capacity.

We defined, in collaboration with the High Speed Networks Laboratory of Hungry, a bi-objective formulation allowing the distinction of different delay requirements. The Quality of Service is defined using network calculus and allows to consider different scheduling algorithms in a network made of an aggregation of different classes of traffic [34].

6.7. Design of multilayer shields: a multi-objective approach

Participants: Laetitia Vermeulen-Jourdan, Thomas Legrand, Oliver Schütze, El-Ghazali Talbi.

We proposed an approach to optimize multilayer shields of Polyaniline - Polyurethane conducting composites in the microwave band. Though by this method shields for different applications can be obtained which are lightweight and offer a low percolation threshold, the full potential of the design process could not be tapped since the underlying optimization problem includes only one objective.

We go one step beyond and re-formulate the design problem as a multi-objective optimization problem. To be more precise, we involve simultaneously the shielding efficiency as well as the weight and the cost of the material - i.e., all the requirements for modern shielding materials - into the optimization process. After having stated the model we present two possible ways to approximate the solution set - the so-called Pareto set - and address the related and important decision making problem. All steps are demonstrated on a particular 3-layered composite in order to show the applicability of the novel approach (see [61]).

6.8. Molecular docking problem: a new multi-objective modeling

Participants: Jean-Charles Boisson, Laetitia Vermeulen-Jourdan, El-Ghazali Talbi.

We have proposed a new tri-objective modeling for the docking molecular problem. We combined two energetic criteria and a surface criterion:

1. molecular energy: this term describe the energy of the binding site / ligand complex. It consists in six different contributions that describe the energetic interactions inside and around the complex. These contributions are the following:

   – bond,

   – angle,

   – torsion,

   – Van Der Waals,

   – Coulomb,

   – desolvatation.

2. surface: this term corresponds to the solvent accessible surface of the binding site / ligand complex. It allows to gain information about the penetration of ligand into the site.

3. free energy: the aim of this criterion is the estimation of the robustness of the current complex. Small modifications of the conformation are generated for the binding site and the ligand. Then we evaluate the corresponding complex energy variation. This criterion allows to estimate the energy well around the complex.

Thanks to the ParadisEO platform, this model has been implemented in a multi-objectif genetic algorithm making a full flexible molecular docking. Thanks to molecular rotation and translation, the algorithm try to find the best complexes binding site / ligand. Furthermore, the binding site and the ligand conformations can be modified during the evolution of the population. It is why our algorithm is a full flexible docking method. This algorithm has been included in the Docking@Grid software. It is currently in test phase.

6.9. A peer-to-peer approach for non-redundant tree exploration in scalable Branch and Bound algorithms

Participants: Nouredine Melab, Mohand Mezmaz, El-Ghazali Talbi.

We proposed (published in [51]) a grid-enabled approach for efficient tree exploration in Branch and Bound algorithms. The approach is based on a special coding of tree nodes and sub-trees. Each sub-tree is coded by a simple interval, and two operators are proposed to transform a sub-tree into an interval (folding) and an interval into a sub-tree (unfolding). The approach allows to optimize communication and storage costs involved in work distribution and checkpointing operations.

The approach has been experimented using the Farmer-Worker paradigm on the Flow-Shop scheduling problem (instance: 50 jobs on 20 machines) using the Grid5000 computational grid . The results show that the redundancy rate is 0.39% which is very low. However, for very large problems (large trees) with a CPU time intensive lower bound and objective functions like Q3AP such rate is not neglectible and can lead to decrease efficiency. To deal with such issue, we proposed (published in [49]) a peer-to-peer approach that allows to prevent redundancy. In this approach, the work distribution strategy has been re-visited. Moreover, in order to allow direct communications between peers a hybrid organization of the system is proposed combining the JXTA and Napster approaches.

The preliminary results obtained on Grid5000 using small instances of the Flow-Shop problem demonstrate the efficiency of the proposed approach. Other long-running experiments are being conducted.

6.10. Parallel hybrid metaheuristics for structure prediction of large proteins using ParadisEO on Globus

Participants: Nouredine Melab, El-Ghazali Talbi, Alexandru Tantar.

The analysis of the protein structure prediction problem (PSP) landscape and complexity has shown that the problem is highly combinatorial and multi-modal. Indeed, even if a small instance is considered the size of the corresponding search space is exponential and the number of local optima is very high. As a consequence, to deal with large proteins it is strongly recommended to use on the one hand parallel grid computing. On the hand, it is required to use mechanisms that allow to combine efficiently evolutionary algorithms and local search methods to take advantage from respectively the exploration and intensification properties of these optimization techniques.

We have proposed (published in [57]) a comparative study of parallel metaheuristics for the protein structure prediction problem (PSP) on the computational grid. The study was mainly focused on hybrid genetic algorithms (GA) and simulated annealing algorithms (SA). The different parallel metaheuristics have been implemented using ParadisEO-G and experimented on over 400 CPUs of the Grid5000 nation-wide experimental computational grid. The experimental results the GA does not outperform the SA unless it is combined with a conjugated gradient local search method. In addition, the SA gets easily trapped in local minima or becomes subject to moves only from one local optimum to another allowing less or no uphill changes.

In order to take benefits from the advantages of the two algorithms a meta-algorithm is proposed to combine the GA with the SA. The meta-algorithm is composed of a pool of copperative GAs (parallel island model) and the population of each GA is evaluated in parallel (parallel evaluation of population model). Each GA uses two hybridization mechanisms: a conjugated gradient local search is used as a mutation operator, and at the end of each generation an adaptive SA is applied to the population of the GA in a multi-start way (parallel multi-start model). The preliminary results obtained using more than 1000 CPUs of Grid5000 show that combining different parallel models and mechanisms is a powerful way to achieve high efficiency on large PSP instances.

In addition to our past collaborations in bioinformatics (IT-omics) or in telecommunications (France Telecom, Mobinets) the main collaboration for this year deals with transportation and logistics.

7. Contracts and Grants with Industry

7.1. SOGEP - REDCATS

(2006-2008): The cooperation with SOGEP, the logistic and delivery subsidiary company of REDCATS (PINAULT PRINTEMPS REDOUTE) consists in solving a logistic and transportation problem. The objective here is the design and implementation of a decision aid framework for solving complex vehicle routing problems including different constraints.

7.2. CEA

(2006-2009): The cooperation with the CEA intervenes in the ANR project "DOCK" (Docking on Grids).

8. Other Grants and Activities

8.1. Regional Actions

- COLIVAD project (Pilotage Optimal des processus de Livraison en Vente à Distance (2006-2008): This project is part of "Pole de compétitivité" PICOM (*Industries du commerce*). It deals with solving a logistic and transportation problem.

8.2. National actions

- Decrypton project, AFM-CNRS-IBM: "Conformational sampling and docking on Grids and application to neuromuscular disease" (2006-2008): collaboration with INSERM and IBL (Lille Institute of Biology).

- ANR DOCK (Docking on Grids) (2006-2009): collaboration with IBL (Institut de Biologie de Lille) and CEA (Grenoble).

- ANR CHOC (Challenges on Combinatorial Optimization on Grids) (2006-2009): collaboration with Prism (Univ. of Versailles), MOAIS (INRIA Rhones-Alpes), GILCO (Grenoble).

- PPF (Bioinformatics) (2006-2009): This national program within the university of Lille (USTL) deals with solving bioinformatics and computational biology problems using combinatorial optimization techniques.

- ACI "Masse de données" 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): The objective of the project was to build an experimental Nation-wide grid infrastructure. More exactly, Grid'5000 is a cluster of 9 clusters interconnected by Renater and one of these clusters is hosted at Lille. The DOLPHIN project team has served as a coordinator of the national project at Lille. The coordination continues through the INRIA ALADDIN grid initiative since september 2007.

- CONS-PACK project (study of constrained packing problems - 2007): collaboration with Heudiasyc lab (Compiègne) supported by GDR RO (GDR on Operational Research).

8.3. International actions

- INRIA project 3+3 Méditerrannée PERFORM (2006-2009) involving the University of Malaga (Spain), University of Constantine (Algeria), and University of Tunis (Tunisia). This project deals with multi-objective optimization.

- University of Constantine (2004-2008): CMEP program with the University of Constantine (Algeria) on "Metaheuristics for optimization of hard problems".

- COST European project GRAAL (2004-2007) on designing and experimenting multi-objective formulations for telecommunication problems.

- NEGST (NExt Grid Systems and Techniques - 2006-2009): International Collaboration and Promotion on interoperability and advanced technologies of Grid. program between CNRS (France) and Japan on optimization on Grids.

8.4. Visits and researcher invitations

The project had visitors during the year 2007:

- E. Alba (Malaga, Spain)

- A. Bendjoudi (Algiers, Algeria)

- H. Deneche (Constantine, Algeria)

- L. Fagouli (Constantine, Algeria)

- J. Figueira (Lisbonne, Portugal)

- J. Garcia-Nieto (Malaga, Spain)

- M. Khouadja (Constantine, Algeria)

- G. Luque (Malaga, Spain)

- K. Mellouli (Tunis, Tunisia)

9. Dissemination

9.1. Services to the scientific community

9.1.1. Research Management

- Co-fondator and chair of the group META (Metaheuristics: Theory and Applications, http://www.lifl. fr/~talbi/META). This group is associated with the ROADEF (French Operations Research Society), and the CNRS research groups GDR ALP and MACS.

- Chair of the group PM2O (Multi-objective Mathematical Programming, http://www.lifl.fr/PM2O). This group is associated with the ROADEF (French Operations Research Society), and the CNRS research group GDR RO (Operations Research).

- Secretary of ROADEF (the French Operational Research Society - www.roadef.org).

- Direction of the CIB (Bioinformatics Center) of the Genopole of Lille.

- Scientific Committee of the Genopole of Lille.

- Member of the Steering Committee of INRIA ALADDIN Project.

- Co-leader of an ALADDIN working group on scalability of Grid-enabled algorithms and applications.

- Member of the Scientific Committee of High-Performance Computing of Université de Lille1.

9.1.2. Participation to Working Groups

- EURO-PAREO (European working group on Parallel Processing in Operations Research).

- EURO-EU/ME (European working group on Metaheuristics).

- EURO-ESICUP (European Working Group on Cutting and Packing).

- ECCO (European Chapter on Combinatorial Optimization).

- ERCIM (European Research Consortium for Informatics and Mathematics) working group on Soft Computing.

- JET national group on evolutionary computation.

- PM2O national group on Multi-objective Mathematical Programming.

- META national group on Metaheuristics: Theory and Applications.

- KSO national group on cutting and packing.

9.1.3. Editions

- E-G. Talbi and A. Zomaya. Book on "Grid computing for bioinformatics and Computational biology" (Wiley, ISBN: 978-0-471-78409-8), 2007.

- E-G. Talbi and E. Alba and A. Zomaya. Special issue of the journal "Computer Communications"on "Nature inspired distributed computing in communication", 2007.

- E-G. Talbi and E. Alba and A. Nebro. Special issue of the journal "Journal of Heuristics" on "Latest advances in metaheuristics for multi-objective optimization", 2007.

- E-G. Talbi and E. Alba and A. Zomaya. Special issue of the journal "Journal of Mathematical Modelling and Applications (JMMA)" on "Applications of Nature Inspired Algorithms", 2007.

9.1.4. Organizations of sessions, workshops and conferences

- NIDISC Workshop organization (International Workshop on Nature Inspired Distributed Computing) organized jointly with ACM/IEEE IPDPS (International Parallel and Distributed Processing Symposium): NIDISC'07 (Long Beach, California, USA).

- Organization of a session "Parallel and Grid computing for optimization", in conference HPCS'2007, Int. Conf. on High Performance Computing and Simulation, Prague, June 2007.

- Organization of a Session "software framework for metaheuristics", in EURO European Conference on Operational Research, Prague, Jul 2007.

- Organization of sessions in ROADEF'2007, 8th conference of the French Operational Research Society, Grenoble, Feb 2007.

- **The 3rd Flow-Shop Contest:** After the 2nd Flow-Shop Contest held last year at Sophia Antipolis, we organized this year the 3rd Flow-Shop Contest in Beijing. This event is organized in collaboration with the European Research Consortium for Informatics and Mathematics (ERCIM), the INRIA Sophia Antipolis and the European Telecommunications Standards Institute (ETSI). The contest took place between 29th October 2007 and 1st November 2007. The challenge given to each team is to solve the biggest possible standard instances[5] of the Flow-Shop problem. The participating teams had more than 5.000 CPUs distributed on four grids (GRID'5000, DAS, PoweRcost and Intrigger). The winning team (TransPUTers, Poznan University of Technology, Poland) has solved all the ten instances defined with 20 jobs and 20 machines. These resolutions have required the use of about 3.200 CPUs during more than 16 minutes.

9.1.5. Reviews

- Review of journal papers:
    - IEEE Transactions on Systems Man and Cybernetics
    - IEEE transactions on Computational biology and Bioinformatics
    - IEEE Transactions on Parallel and Distributed Systems
    - IEEE Transactions on Evolutionary Computation
    - Parallel Computing
    - Calculateurs Parallèles
    - Journal of Supercomputing
    - Parallel and Distributed Computing Practices
    - Journal of Parallel and Distributed Computing
    - Genetic Programming and Evolvable Machines
    - Journal of Heuristics
    - European Journal of Operational Research
    - Annals of Operations Research
    - 4'OR
    - International Journal of Production Economics
    - Computers and Operation Research
    - Discrete Applied Mathematics
    - Journal of Computational Optimization and Applications
    - Information Processing Letters
    - Extraction de connaissances et apprentissage
    - European Physical Journal B
    - Journal of Mathematical Modelling and Algorithms
    - Bioinformatics

---

[5]http://ina2.eivd.ch/Collaborateurs/etd/problemes.dir/ordonnancement.dir/ordonnancement.html

- – Jouranl of Artificial Evolution and Applications
- – Knowledge based systems
- – ...

- Review of different projects :
  - – ECOS-Sud (Argentine, Chili, Uruguay), 2007.
  - – Dutch NOW council (Innovative Research Incentive Scheme) project, Netherlands, 2007.
  - – Expert de l'ANR "Chaire d'excellence", 2007.

9.1.6. Program Committees

- International Conferences on Evolutionary Computation:
  - – CEC (Congress on Evolutionary Computation): CEC'07 (Singapour).
  - – GECCO (Genetic and Evolutionary Computation Conference): GECCO'2007 (London).
  - – EvoCOP (European Conference on Evolutionary Computation in Combinatorial Optimization): EvoCOP'2007 (Valencia, Spain).
  - – EvoBIO (European Workshop on Evolutionary Computation and Bioinformatics): Evo-Bio'2007 (Valencia, Spain).
  - – HM (International Workshop on Hybrid Metaheuristics): HM'2007 (Dortmund, Germany).
  - – EMO (International Conference on Evolutionary Multi-Criterion Optimization): EMO2007 (Matsushima/Sendai, Japan).
  - – MIC (Metaheuristics International Conference): MIC'2007 (Montréal, Canada).
  - – Workshop PBA (Parallel Bioinspired Algorithms): PBA'2007 (London, UK).
  - – BIONETICS (Iternational Conference On Bio-Inspired Models of Network, Information and Computing Systems): BIONETICS'2007 (Budapest, Hungary).
  - – Artificial Evolution: EA'2007 (Tours, France).

- International conferences on Bio-informatics
  - – ISBRA (Int. Symposium on Bioinformatics Research and Applications): ISBRA'2007 (Atlanta, Georgia, USA).
  - – BLSC (IEEE Int. Symposium on Bioinformatics and life science computing): BLSC'2007 (Niagara Falls, Canada).
  - – CIBCB (IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology): CIBCB'2007, April 1-5 2007 (Honolulu, Hawaii, USA).

- International conferences on High-performance computing
  - – Workshop DEXA GLOBE'07 " Grid and peer-to-peer computings impacts on large scale heterogeneous distributed database systems" (Regensburg, Germany).
  - – ICPCA (International Conference on Pervasive Computing and Applications): IPCA'2007 (Birmingham, UK).
  - – Workshop HPC-GTP (Workshop on High Performance computing in Genomic Proteomic and Transcriptomic), in conjunction with ISPA (International Symposium On Parallel and distributed Processing and applications): ISPA'2007 (Sorento, Italy).
  - – Workshop PPN (Peer to peer Networks): PPN'2007 (Vilamoura, Algarve, Portugal).
  - – Intl. conf. HPC&S (High Performance Computing & Simulation Conference): HPC&S'2007 (Prague, Czech Republic).

- Intl. Conf. P2P (Intl. Conf. on Peer-to-Peer Computing): P2P'2007 (Galway, Ireland).
- Intl. Symp. ISPDC (International Symposium on Parallel and Distributed Computing): ISPDC'2007 (Hagenberg, Austria).
- IFIP NPC (International Conference on Network and Parallel Computing): NPC'2007 (Dalian, China).
- GADA (International Symposium on Grid Computing, High-Performance and Distributed Applications): GADA'2007 (Lisbonne, Portugal).

- International conferences on Operations Research and Production Management
  - LT (Logistique et Transport): LT'2007 (Sousse, Tunisie).
  - FRANCORO/ROADEF (francophone Conference on Operations Research joined with the Conference of the French Operational Research Society): FRANCORO IV/ROADEF 2007 (Grenoble, France).
  - Workshop on Multi-criteria decision making (MCDM) applications: Leveraging domain knowledge with computational intelligence, in IEEE 2007 Computational Intelligence Society Symposium, (Honolulu, Hawaii, USA).

- Other conferences
  - EGC (Journées Francophones Extraction et Gestion des Connaissances): EGC'2007 (Namur, Belgique).
  - Workshop IEEE FOCI'2007 (First IEEE Symposium on Foundations of Computational), in IEEE 2007 Computational Intelligence Society Symposium, (Honolulu, Hawaii, USA).
  - First International Conference on Multidisciplinary Design Optimization and Applications (Besançon, France).
  - IEEE ISDA (International Conference on Intelligent Systems Design and Applications): ISDA'2007 (Rio de Janeiro, Brazil).
  - SLS (Engineering Stochastic Local Search Algorithms): SLS'2007 (Brussels, Belgium).
  - IEEE APSCC (Asian-Pacific Services Computing Conference): IEEE APSCC'2007 (Tsukuba, Japan).

9.1.7. Phd and HdR committees

Pr Talbi was referee of the following thesis:

- Jan 2007, PhD of Bernabé Dorronsoro Diaz "Parallel evolutionary algorithms", University of Malaga - Espagne.
- April 2007, PhD of L. Hidri "Exact and heuristic methods for the hybrid flow shop scheduling problem", l'Institut Supérieur de Gestion, Tunis, Tunisie. Jury: M. Haouari, K. Mellouli, M-A. Laroui, M. Tagina, E-G. Talbi (président).
- Oct 2007, PhD of A. Di Constanzo intitulé "Branch-and-bound with peer-to-peer for large-scale Grids", Université de Nice Sophia-Antipolis. Jury: F. Capello, D. Caromel, M. Clergue, R. Couturier, D. Gannon, F. Laburthe, E-G. Talbi (rapporteur).
- Nov 2007, PhD of C-E. Bichot, "Partitionnement de graphe et application au découpage de l'espace aérien", Institut National Polytechnique de Toulouse. Jury: N. Durand, F. Pellegrini, P. Siarry, E-G. Talbi (rapporteur).
- 2007, PhD of M. N. Allouche, "L'ordonnancement multicritère de la production: une approche métaheuristique intégrant les préférences du gestionnaire", Institut Supérieur de Gestion, Tunis, Tunisie. Jury: B. Aouni, A. Rebai, E-G. Talbi (rapporteur).

Pr Dhaenens was referee of the following thesis:

- May 2007, PhD of K. Bouibede-Hocine, "La problématique d'énumération d'optima de Pareto en ordonnancement multicritère : application à un problème d'ordonnancement à machines parallèles.", Université de Tours. Jury: J-B. Billaut, J. Carlier, C. Dhaenens (présidente), X. Gandibleux, V. T'Kindt, B. Penz.

Pr Melab was referee of the following thesis:

- Dec 2007, PhD of J. Gossa, "Modélisation et outils génériques pour la résolution des problèmes liés à la répartition des ressources sur grilles", INSA de Lyon, Jury: M. Sibilla, F. Cappello, T. Ludwig, N. Melab, L. Brunie, J-M. Pierson.

9.2. Teaching

9.2.1. International teaching

- Postgraduate "Modern optimization techniques", University of Malaga, Spain, Jan 2007 (E-G. Talbi).
- Postgraduate "Modern optimization techniques", University of Tunis, Tunisia, May 2007 (E-G. Talbi).
- Postgraduate "Recherche opérationnelle et datamining", University of Sfax, Tunisia, May 2007 (C. Dhaenens).
- Postgraduate "Grid computing", University of Luxembourg, Luxembourg, Nov 2007 (E-G. Talbi).

9.2.2. National teaching

- Postgraduate (IEEA, USTL): "Optimization methods" (L. Vermeulen-Jourdan).
- Postgraduate (IEEA, USTL): "GRID computing", (N. Melab, B. Derbel).
- Undergraduate (IEEA, USTL): "Distributed Systems" (N. Melab, B. Derbel).
- Undergraduate (IEEA, USTL): "Operations Research" (N. Melab).
- Undergraduate (Polytech'Lille): "Operations Research" (C. Dhaenens).
- Undergraduate (Polytech'Lille): "Graphs and combinatorics" (C. Dhaenens).
- Undergraduate (Polytech'Lille): "Data mining" (L. Vermeulen-Jourdan, C. Dhaenens).
- Undergraduate (Polytech'Lille): "Advanced Optimization" (L. Vermeulen-Jourdan).
- Undergraduate (Polytech'Lille): "Production Management" (C. Dhaenens).
- Undergraduate (IUT, USTL): "Graphs and Modeling" (F. Clautiaux).

10. Bibliography

Major publications by the team in recent years

[1] F. CLAUTIAUX, A. JOUGLET, J. E. HAYEK. *A New Lower Bound for the Non-Oriented Two-Dimensional Bin-Packing Problem*, in "Operations Research Letters", vol. 35, n⁰ 3, 2007, p. 365-373.

[2] N. JOZEFOWIEZ, F. SEMET, E.-G. TALBI. *Target Aiming Pareto Search and its application to the vehicle routing problem with route balancing*, in "Journal of Heuristics", vol. 13, 2007, p. 455-469.

[3] J. LEMESRE, C. DHAENENS, E. TALBI. *An exact parallel method for a bi-objective permutation flowshop problem*, in "European Journal of Operational Research", vol. 177, n⁰ 3, 2007, p. 1641-1655.

[4] N. MELAB, S. CAHON, E.-G. TALBI. *Grid computing for parallel bioinspired algorithms*, in "Journal of Parallel and Distributed Computing (JPDC)", vol. 66, n° 8, Aug. 2006, p. 1052–1061.

[5] M. MEZMAZ, N. MELAB, E.-G. TALBI. *An efficient load balancing strategy for grid-based branch and bound algorithm*, in "Parallel Computing", vol. 33, n° 4-5, 2007, p. 302–313.

[6] J.-M. PIERSON, J. GOSSA, P. WEHRLE, Y. CARDENAS, S. CAHON, M. E. SAMAD, L. BRUNIE, C. DHAENENS, A. HAMEURLAIN, N. MELAB, M. MIQUEL, F. MORVAN, E.-G. TALBI, A. TCHOUNIKINE. *GGM: Efficient Navigation and Mining in Distributed Genomedical Data*, in "IEEE Transactions on Nanobioscience", vol. 6, n° 2, 2007, p. 110-116.

[7] E.-G. TALBI. *Parallel Combinatorial Optimization*, Wiley, 2006.

[8] A.-A. TANTAR, N. MELAB, E.-G. TALBI, B. PARENT, D. HORVATH. *A parallel hybrid genetic algorithm for protein structure prediction on the computational grid*, in "Future Gener. Comput. Syst.", vol. 23, n° 3, 2007, p. 398–409.

[9] L. VERMEULEN-JOURDAN, C. DHAENENS, E.-G. TALBI. *Linkage disequilibrium study with a parallel adaptive GA*, in "International Journal of Foundation in Computer Science", vol. 16, April 2005, p. 241–260.

[10] L. VERMEULEN-JOURDAN, C. DHAENENS, E.-G. TALBI. *Evolutionary Feature Selection for Bioinformatics*, in "Computational Intelligence in Bioinformatics", G. FOGEL, Y. PING, D. CORNE (editors), To appear Dec. 2007, IEEE CS/Wiley, 2007, p. 117–139.

Year Publications

Articles in refereed journals and book chapters

[11] F. CLAUTIAUX, J. CARLIER, A. MOUKRIM. *A new exact method for the two-dimensional bin-packing problem with fixed orientation*, in "Operations Research Letters", vol. 35, n° 3, 2007, p. 357-364.

[12] F. CLAUTIAUX, A. JOUGLET, J. E. HAYEK. *A New Lower Bound for the Non-Oriented Two-Dimensional Bin-Packing Problem*, in "Operations Research Letters", vol. 35, n° 3, 2007, p. 365-373.

[13] F. CLAUTIAUX, A. MOUKRIM, J. CARLIER. *Improving Lower Bounds for a Two-Dimensional Bin-Packing Problem by Generating New Data-Dependent Dual-Feasible Functions*, in "accepted in International Journal of Production Research", 2007.

[14] B. DERBEL, C. GAVOILLE. *Fast Deterministic Distributed Algorithms for Sparse Spanners*, in "Theoretical Computer Science (TCS)", To appear, 2007.

[15] N. JOZEFOWIEZ, F. SEMET, E.-G. TALBI. *Multi-objective vehicle routing problems*, in "European Journal of Operational Research", (available online), 2007.

[16] N. JOZEFOWIEZ, F. SEMET, E.-G. TALBI. *Target Aiming Pareto Search and its application to the vehicle routing problem with route balancing*, in "Journal of Heuristics", vol. 13, 2007, p. 455-469.

[17] N. JOZEFOWIEZ, F. SEMET, E.-G. TALBI. *The bi-objective covering tour problem*, in "Computers and Operations Research", vol. 34, 2007, p. 1929-1942.

[18] N. JOZEFOWIEZ, F. SEMET, E.-G. TALBI. *From single-objective to multi-objective vehicle routing problems: Motivations, case studies, and methods*, in "The vehicle routing problem: Latest advances and new challenges", (forthcoming), Springer, 2007.

[19] M. KHABZAOUI, C. DHAENENS, E. TALBI. *Combining Evolutionary algorithms and exact approaches for multi-objective knowledge discovery*, in "RAIRO-RO (EDP Sciences)", To appear, 2007.

[20] J. LEMESRE, C. DHAENENS, E. TALBI. *An exact parallel method for a bi-objective permutation flowshop problem*, in "European Journal of Operational Research", vol. 177, n° 3, 2007, p. 1641-1655.

[21] J. LEMESRE, C. DHAENENS, E. TALBI. *Parallel Partitionning Method (PPM): A new exact method to solve bi-objective problems*, in "Computers and Operations Research (Elsevier)", vol. 34, 2007, p. 2450-2462.

[22] M. MEZMAZ, N. MELAB, E. -G. TALBI. *An efficient load balancing strategy for grid-based branch and bound algorithm*, in "Parallel Computing", vol. 33, n° 4-5, 2007, p. 302–313.

[23] M. MEZMAZ, N. MELAB, E. -G. TALBI. *Combining Metaheuristics and Exact Methods for Solving Exactly Multi-Objective Problems on the Grid*, in "Journal of Mathematical Modelling and Algorithms", vol. 6, n° 3, 2007, p. 393–409.

[24] J.-M. PIERSON, J. GOSSA, P. WEHRLE, Y. CARDENAS, S. CAHON, M. E. SAMAD, L. BRUNIE, C. DHAENENS, A. HAMEURLAIN, N. MELAB, M. MIQUEL, F. MORVAN, E.-G. TALBI, A. TCHOUNIKINE. *GGM: Efficient Navigation and Mining in Distributed Genomedical Data*, in "IEEE Transactions on Nanobioscience", vol. 6, n° 2, 2007, p. 110-116.

[25] E.-G. TALBI, B. WEINBERG. *Breaking the search space symmetry in partitioning problems: An application to the graph coloring problem*, in "Theor. Comput. Sci.", vol. 378, n° 1, 2007, p. 78-86.

[26] A.-A. TANTAR, N. MELAB, E.-G. TALBI, B. PARENT, D. HORVATH. *A parallel hybrid genetic algorithm for protein structure prediction on the computational grid*, in "Future Gener. Comput. Syst.", vol. 23, n° 3, 2007, p. 398–409.

[27] A.-A. TANTAR, N. MELAB, E.-G. TALBI. *Molecular Docking using Grid Computing*, in "Grids for Bioinformatics and Computational Biology", John Wiley & Sons, USA, 2007, p. 179–198.

[28] L. VERMEULEN-JOURDAN, C. DHAENENS, E.-G. TALBI. *Evolutionary Feature Selection for Bioinformatics*, in "Computational Intelligence in Bioinformatics", G. FOGEL, Y. PING, D. CORNE (editors), To appear Dec. 2007, IEEE CS/Wiley, 2007, p. 117–139.

Publications in Conferences and Workshops

[29] E. ALBA, J. GARCÍA-NIETO, L. VERMEULEN-JOURDAN, E.-G. TALBI. *Gene Selection in Cancer Classification using PSO/SVM and GA/SVM Hybrid Algorithms*, in "IEEE Congress on Evolutionary Computation CEC-07, Singapore", Sep 2007.

[30] C. ALVES, F. CLAUTIAUX, J. VALÉRIO DE CARVALHO. *New ways of deriving dual cuts for the cutting stock problem*, in "OPTIMIZATION 2007, Sixth international conference on optimization", 2007.

[31] J.-C. BOISSON, O. GUERRE, L. VERMEULEN-JOURDAN, E.-G. TALBI, C. ROLANDO. *ASCQ_ME: a New Engine for Protein and Post-Translational Modification Identification from MS and MS/MS Raw Spectrum without Mass List Extraction*, in "55th ASMS Conference on Mass Spectrometry", Indianapolis, Indiana, USA, June 3-7 2007.

[32] J.-C. BOISSON, E.-G. TALBI. *Software Framework for Metaheuristic*, in "22nd European Conference on Operational Research (EURO XXII)", Prague, Czech Republic, July 8-11 2007, 45.

[33] J.-C. BOISSON, E.-G. TALBI, L. VERMEULEN-JOURDAN. *De Novo Protein Sequencing using Evolutionary Algorithm*, in "22nd European Conference on Operational Research (EURO XXII)", Prague, Czech Republic, July 8-11 2007, 98.

[34] J. BRONGNIART, T. CINKLER. *STSM: Multi-objective of tree-shaped UTRAN design Problem*, in "10th MC meeting and Discussion workshop, Cost 293", Rome, Italie, october 2007.

[35] J. BRONGNIART, C. DHAENENS, E. TALBI. *Contraintes d'optimisation pour les problèmes d'arbres couvrants sous restrictions*, in "Troisièmes Journées Francophones de Programmation par Contraintes (JFPC07)", Nantes, 2007, p. 311-320, http://hal.inria.fr/inria-00151233/fr/.

[36] J. BRONGNIART, C. DHAENENS, E. TALBI. *Defining Filtering Algorithms for Restricted Tree Problems using Balanced Trees*, in "3th International Conference on Principles and Practice of Constraint Programming (CP2007), Providence, USA", 2007.

[37] L. BRUNIE, Y. CARDENAS, C. DHAENENS, G. EVEN, J. GOSSA, H. HAMEURLAIN, N. MELAB, M. MIQUEL, F. MORVAN, J.-M. PIERSON, M. E. SAMAD, E.-G. TALBI, A. TCHOUNIKINE, P. WEHRLE. *Grid for Geno-Medecine : Architecture to manage and analyse geno-medical data distributed over computational grid*, in "Journées ouvertes Biologie Informatique Mathématiques (JOBIM'07)- Poster, Marseille", 2007.

[38] F. CLAUTIAUX, C. ALVES, J. VALÉRIO DE CARVALHO. *Une analyse comparative de fonctions dual-réalisables appliquées au problème de bin-packing*, in "8ème congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision, Grenoble", 2007.

[39] F. CLAUTIAUX, A. JOUGLET, J. CARLIER, A. MOUKRIM. *Problème de placement de rectangles*, in "8ème congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision, Grenoble", 2007.

[40] B. DERBEL, C. GAVOILLE, D. PELEG. *Deterministic Distributed Construction of Linear Stretch Spanners in Polylogarithmic Time*, in "21th Symposium on DIStributed Computing (DISC'07)", LNCS 4732, 2007, p. 179-192.

[41] G. EVEN, P. LAURENCE, C. DHAENENS, E.-G. TALBI. *RuleMining : a new analysis tool for PASE, the web-based platform for polypetide chips experiments*, in "Journées ouvertes Biologie Informatique Mathématiques (JOBIM'07)- Poster, Marseille", 2007.

[42] G. EVEN, L. VERMEULEN-JOURDAN, C. BOUTROUE, C. DHAENENS, E.-G. TALBI. *Evolutionary feature selection plugin for BASE*, in "Journées ouvertes Biologie Informatique Mathématiques (JOBIM'07)- Poster, Marseille", 2007.

[43] J. GARCÍA-NIETO, L. VERMEULEN-JOURDAN, E. ALBA, E. G. TALBI. *A Comparison of PSO and GA Approaches for Gene Selection and Classification of Microarray Data*, in "Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2007), London, UK", ACM Press, July 2007.

[44] J. E. HAYEK, F. CLAUTIAUX, A. JOUGLET. *Bornes inférieures pour le problème de bin-packing en deux dimensions avec rotation*, in "8ème congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision, Grenoble", 2007.

[45] T. LEGRAND, O. SCHÜTZE, L. VERMEULEN-JOURDAN, C. COELLO COELLO, E.-G. TALBI. *Stratégies mimétiques pour les essaims de particules et implémentation sous ParadisEO*, in "Second séminaire francophone sur l'optimisation par essaim particulaire (OEP)", April 2007.

[46] A. LIEFOOGHE, M. BASSEUR, L. VERMEULEN-JOURDAN, E.-G. TALBI. *Combinatorial Optimization of Stochastic Multi-objective Problems: an Application to the Flow-shop Scheduling Problem*, in "Evolutionary Multi-criterion Optimization (EMO 2007), Matsushima, Japan", S. OBAYASHI, ET AL. (editors), Lecture Notes in Computer Science, vol. 4403, Springer-Verlag, 2007, p. 457–471.

[47] A. LIEFOOGHE, M. BASSEUR, L. VERMEULEN-JOURDAN, E.-G. TALBI. *ParadisEO-MOEO: A Framework for Evolutionary Multi-objective Optimization*, in "Evolutionary Multi-criterion Optimization (EMO 2007), Matsushima, Japan", S. OBAYASHI, ET AL. (editors), Lecture Notes in Computer Science, vol. 4403, Springer-Verlag, 2007, p. 386–400.

[48] A. LIEFOOGHE, L. VERMEULEN-JOURDAN, E.-G. TALBI. *Software Framework for Multi-objective Optimization*, in "22nd European Conference on Operational Research (EURO XXII), Prague, Czech Republic", 2007, 45.

[49] M. MEHDI, M. MEZMAZ, N. MELAB, E.-G. TALBI. *Toward a Napster-like P2P B&B Algorithmt and Island Models for Parallel Multi-objective Optimization on the Computational Grid*, in "Proc. of the 2007 high performance computing and simulation conference (HPCS 2007) In conjunction with The 21st European conference on modelling and simulation (ECMS 2007), Prague, Czech Republic", June 2007.

[50] M. MEZMAZ, N. MELAB, E.-G. TALBI. *A Grid-based Parallel Approach of the Multi-Objective Branch and Bound*, in "Proc. of the 15th IEEE Euromicro Conf. on Parallel, Distributed and Network-based Processing, Naples, Italy", February 2007.

[51] M. MEZMAZ, N. MELAB, E.-G. TALBI. *A Grid-enabled Branch and Bound Algorithm for Solving Challenging Combinatorial Optimization Problems*, in "Proc. of 21th IEEE Intl. Parallel and Distributed Processing Symp. (IPDPS), Long Beach, California", March 2007.

[52] B. PARENT, A.-A. TANTAR, N. MELAB, E.-G. TALBI, D. HORVATH. *Grid-based Evolutionary Strategies Applied to the Conformational Sampling Problem*, in "Proc. of the IEEE Congres on Evolutionary Computation (CEC'07)", September 2007.

[53] O. SCHÜTZE, M. LAUMANNS, E. TANTAR, C. A. COELLO COELLO, E.-G. TALBI. *Convergence of stochastic search algorithms to gap-free pareto front approximations*, in "Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2007)", 2007, p. 892-901.

[54] O. SCHÜTZE, L. VERMEULEN-JOURDAN, T. LEGRAND, E.-G. TALBI, J.-L. WOJKIEWICZ. *A Multi-Objective Approach to the Design of Conducting Polymer Composites for Electromagnetic Shielding*, in "EMO 2007, Matsushima, Japan", LNCS (editor), LNCS, vol. 4403, 2007, p. 590–603.

[55] E.-G. TALBI, L. VERMEULEN-JOURDAN, J. GARCÍA-NIETO, E. ALBA. *Sélection d'attributs de puces à ADN par essaim de particules*, in "Second séminaire francophone sur l'optimisation par essaim particulaire (OEP)", April 2007.

[56] E. TANTAR, C. DHAENENS, E. TALBI. *Sur l'approximabilité du problème de flowshop de permutation bi-objectif*, in "Huitième congrès de la Société Française de Recherche Opérationnelle et Aide à la Décision (ROADEF'07)", Grenoble, 2007, p. 207-208.

[57] A.-A. TANTAR, N. MELAB, E.-G. TALBI. *A Comparative Study of Parallel Metaheuristics for Protein Structure Prediction on the Computational Grid*, in "Proc. of 21th IEEE Intl. Parallel and Distributed Processing Symp. (IPDPS)", 2007.

[58] J. VALÉRIO DE CARVALHO, F. CLAUTIAUX, C. ALVES. *A comparative analysis of superadditive functions and dual-feasible functions*, in "OPTIMIZATION 2007, Sixth international conference on optimization", 2007.

[59] J. VALÉRIO DE CARVALHO, F. CLAUTIAUX, C. ALVES. *A comparative analysis of superadditive functions and dual-feasible functions*, in "Fourth ESICUP Meeting, the EURO Working Group on Cutting and Packing, Tokyo, Japan", 2007.

[60] J. VALÉRIO DE CARVALHO, F. CLAUTIAUX, C. ALVES. *A comparative analysis of superadditive functions and dual-feasible functions*, in "22nd European Conference on Operational Research (EURO XXII)", 2007.

[61] L. VERMEULEN-JOURDAN, T. LEGRAND, O. SCHÜTZE, E.-G. TALBI, J.-L. WOJKIEWICZ. *A multiobjective genetic algorithm to optimize electromagnetic properties of conducting polymer composites in the microwave band*, in "IESM 2007, Bejing China", 2007.

[62] I. ZUNINO, N. MELAB, E.-G. TALBI. *A Grid-enabled framework for exact optimization algorithms*, in "Proc. of High Performance Computing and Simulation (HPC&S'2007), Pragua, Czeche Republic", June 2007.

Internal Reports

[63] J. BRONGNIART, T. CINKLER. *Short Term Scientific Mission, Cost293: Multi-objective resolution of tree-shaped UTRAN design problem*, Technical report, In collaboration with the High Speed Networks Laboratory, 2007.

[64] E. TANTAR, C. DHAENENS, J. FIGUEIRA, E.-G. TALBI. *Interactive Hybridization of Metaheuristics and Landscape Analysis for Multi-Objective Optimization*, Technical report, n° 10, Centro de Estudos de Gestao, Instituto Superior Tecnico, Lisbon, Portugal, 2007.

[65] A. TANTAR, N. MELAB, C. DEMAREY, E.-G. TALBI. *Building a Virtual Globus Grid in a Reconfigurable Environment - A case study : Grid5000*, Technical report, INRIA, August 2007, http://hal.inria.fr/inria-00168130.

[66] E. TANTAR, O. SCHÜTZE, J. R. FIGUEIRA, C. A. COELLO COELLO, E.-G. TALBI. *Computing and Selecting ε-Efficient Solutions of 0,1-Knapsack Problems*, Technical report, n° 14, Centro de Estudos de Gestao, Instituto Superior Tecnico, Lisbon, Portugal, 2007.

Miscellaneous

[67] C. DHAENENS. *Recherche Opérationnelle et Datamining*, 2èmes journées du doctorant en Recherche Opérationnelle et aide à la décision, Université de Sfax, Tunisie, May 2007.

[68] N. MELAB. *Grid Computing : Concepts et technologies. Etude de cas : Grid5000*, Séminaire invité du Pôle de Compétitivité "Industries du Commerce" de la Région Nord-pas-de-Calais, January 2007.

[69] N. MELAB. *Grid Computing : Concepts, technologies et applications industrielles*, Séminaire invité de l'Université de Lille I, Espace Culture, Villeneuve d'Ascq, January 2007.

[70] N. MELAB. *Grid Computing for Solving Challenging Combinatorial Optimization Problems*, Grid@Mons2007, 1st annual Intl. Conf. series on Grid Computing, May 2007.

[71] E.-G. TALBI. *A parallel evolutionary algorithm for molecule structure prediction*, Invited speaker, Symposium on Genetic Algorithms in Material Science and Engineering, Warsaw, Poland, Sept 2007.

[72] E.-G. TALBI. *Evolutionary algorithms: from design to implementation*, Invited speaker, E-MRS conference, Warsaw, Poland, Sept 2007.

[73] E.-G. TALBI. *Evolutionary molecular docking on large networks*, Invited speaker, ECCS (European Conference on Complex Systems) Satellite on Molecular Docking, Complexity and Optimization, Dresden, Germany, Oct 2007.

[74] E.-G. TALBI. *Grid computing for bioinspired algorithms*, Invited speaker, Seventh International Conference on Intelligent Systems Design and Applications, Rio de Janeiro, Brasil, Oct 2007.

[75] E.-G. TALBI. *Grid computing for optimization*, Invited presentation, NEGST France-Japan Meeting, Tokyo, Japan, May 2007.

[76] E.-G. TALBI. *Grid computing: Application to combinatorial optimization and data mining*, Invited keynote speaker, ISPS'2007 International Symposium on Programming and Systems, Algiers, Algeria, May 2007.

[77] E.-G. TALBI. *High-performance metaheuristics: Application to bioinformatics and telecommunications*, Keynote speaker, CEDI'2007 Conference, Saragossa, Spain, Sept 2007.

[78] E.-G. TALBI. *Metaheuristics on GRIDs*, Invited seminar, Fluminense Federal University, Rio de Janeiro, Brasil, Jan 2007.

[79] E.-G. TALBI. *Parallel metaheuristics: from design to implementation*, Invited tutorial, MIC'2007 Metaheuristics International Conference, Montreal, Canada, June 2007.

References in notes

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

[81] M. BASSEUR. *Design of cooperative algorithms for multi-objective optimization: Application to the Flow-shop scheduling problem*, Ph. D. Thesis, University of Sciences and Technology of Lille, France, June 2005.

[82] C. COTTA, E.-G. TALBI, E. ALBA. *Parallel hybrid approaches*, in "Parallel Metaheuristics, USA", Editors: E. Alba, J. Wiley and Sons, 2005, p. 347–370.

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

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

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

[86] A. J. NEBRO, F. LUNA, E.-G. TALBI, E. ALBA. *Parallel multi-objective optimization*, in "Parallel Meta-heuristics, USA", Editors: E. Alba, J. Wiley and Sons, 2005, p. 371–394.

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

[88] 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.

[89] 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.

[90] B. WEINBERG, E.-G. TALBI. *A cooperative parallel metaheuristic applied to the graph coloring problem*, in "Handbook of Bioinspired Algorithms and Applications, USA", ISBN: 1584884754, Editors: S. Olariu and A.Y. Zomaya, CRC Press, 2005, p. 625–638.

[91] 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.

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

[93] 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.