IN PARTNERSHIP WITH:
CNRS
Université de Bordeaux

Activity Report 2013

Project-Team CEPAGE

Algorithmics for computationally intensive applications over wide scale distributed platforms

IN COLLABORATION WITH: Laboratoire Bordelais de Recherche en Informatique (LaBRI)

RESEARCH CENTER
Bordeaux - Sud-Ouest

THEME
Distributed and High Performance Computing

Table of contents

Project-Team CEPAGE

Keywords: Distributed And Parallel Algorithms, Combinatorial Optimization, Cloud Computing, Information Indexing And Retrieval, Peer-to-peer, Scheduling

Creation of the Project-Team: 2009 January 01, end of the Project-Team: 2013 December 31.

1. Members

Research Scientists
Olivier Beaumont [Team leader, Inria, Senior Researcher, HdR]
Lionel Eyraud-Dubois [Inria, Junior Researcher]
Nicolas Hanusse [CNRS, Senior Researcher, HdR]
David Ilcinkas [CNRS, Junior Researcher]
Ralf Klasing [CNRS, Senior Researcher, HdR]
Adrian Kosowski [Inria, Junior Researcher, HdR]

Faculty Members
Nicolas Bonichon [Univ. Bordeaux I, Associate Professor, HdR]
Philippe Duchon [Univ. Bordeaux I, Professor, HdR]
Cyril Gavoille [Univ. Bordeaux I, Professor, HdR]
Sofian Maabout [Univ. Bordeaux IV, Associate Professor]

Engineers
Pierre Matri [Inria, from Sep 2013]
Lucas Verdonk [Inria, until Nov 2013]

PhD Students
Christian Glacet [Inria, granted by FP7 EULER project, until Nov 2013]
Pierre Halftermeyer [Univ. Bordeaux I, until Sep 2013]
Dominik Pajak [Inria]
Przemyslaw Uznanski [Inria, until Oct 2013]
Ahmed Wade [Univ. Bordeaux I]

Post-Doctoral Fellows
Juan Angel Lorenzo Del Castillo [Inria, granted by ANR -SONGS- project, from Nov 2013]
Paul Renaud-Goud [Inria, granted by ANR -SONGS- project, until Oct 2013]

Visiting Scientist
Pawel Gawrychowski [Internship, from May 2013 until May 2013]

Administrative Assistant
Sabine Delarboulas Cusin [Inria]

Others
Rohit Kumar [Inria, Internship, from May 2013 until Aug 2013]
Thomas Lambert [ENS Lyon, Internship, until Jul 2013]
Siddharth Mandal [Inria, Internship, from May 2013 until Jul 2013]

2. Overall Objectives

2.1. General Objectives

CEPAGE project was launched (as a team) in 2006, when the development of interconnection networks had led to the emergence of new types of computing platforms, characterized by the heterogeneity of both their processing and communication resources, their geographical dispersion, the absence of centralized control, and their instability in terms of the number and performance of participating resources. At the same time, P2P file exchange systems and Volunteer Computing platforms were developing, reinforcing the idea that big platforms could be built by aggregating a large set of small distributed resources. In the context of intensive computations, the solutions were less sophisticated than in the context of file exchanges, and mostly relied on the client-server model (SETI@home, folding@home), with centralized knowledge and control of the platform, and no direct communications between peers. In this context, the general goal of CEPAGE was to come up with new strategies in order to extend the set of computation-intensive applications that could be run on large scale distributed platforms, by gathering researchers with expertize in scheduling of tasks and collective communications, graph theory, design of overlay networks, modeling of network topologies, small world networks, distributed algorithms, compact data structures, routing and randomized algorithms, later extended to include mobile agents and databases.

2.1.1. Objectives for the evaluation period

The above context lead us to identify the following research axes (extracted from the project proposal).

1. Models
   1. At a low level, to understand the underlying physical topology and to obtain realistic models whose parameters can be instantiated at runtime.
   2. At a higher level, to derive models of the dynamism of targeted platforms, both in terms of participating resources and resource performance.
2. Overlays and distributed algorithms:
   1. To understand how to augment the logical topology in order to achieve good properties of P2P systems.
   2. To build overlays dedicated to specific applications and services that achieve good performance.
   3. To understand how to dynamically adapt scheduling algorithms (in particular collective communication schemes) to changes in network performance and topology, using randomized algorithms.
3. Compact and distributed data structures:
   1. To understand how to dynamically adapt compact data structures to changes in network performance and topology.
   2. To design sophisticated labeling schemes in order to answer complex predicates using local labels only.

Regarding the 2007-2011 objectives of the projects (1)-(3), the vast majority of them has been fulfilled as shown in Sections 4.1, 4.2 and 4.3.

Nevertheless, during the last four years, several adjustments were made with respect to the initial objectives.

- First, we under-evaluated in our project proposal the importance of issues related to security and anonymity, by saying that we would rely on external mechanisms provided by distributed systems to ensure them. It turned out that the question was harder to solve and a necessary condition in order to deploy our algorithms at a large scale. Therefore, we spent a lot of effort to work on these issues using the behavior and capacities of mobile agents under adversarial models. The recruitments of David Ilcinkas (CR CNRS, 2007) and Adrian Kosowski (CR Inria, 2010) have strongly reinforced this axis.

- The picture changed due to the coming of Cloud Computing as a mean to "renting" external resources rather than aggregating and sharing privately-owned ones. This solution is particularly appealing in the context of computing-intensive applications. Our efforts to develop widely used distributed codes for molecular dynamics simulations and continuous integration were made unsuccessful by this new context. We are moving toward this direction, in particular by considering stability of resource allocation problems, reliability, and stochastic optimization issues.

2.2. Goal and Context

2.2.1. General Context

The first set of questions that we consider is related to distributed computations, where the Internet is the underlying network. Since the topology of the underlying network is unknown, the use of logical networks (overlays) is required. In turn, the choice of the overlay will have an impact on the complexity of the algorithms. In this context, only the performance of the whole chain is meaningful, which requires to collect raw data and then to propose network models and algorithms based on these models, such that the performance of the resulting algorithm is good on raw data. This also requires studying the influence of the topology of the overlay network (the underlying graph) on the complexity of fundamental questions, such as graph exploration or black-hole search.

The second set of questions is related to distributed data structures. In general, the question is related to the compromise between the size of the data structure to be stored on each node and the time to answer a request (estimate the bandwidth between two nodes, compute the closest common ancestor of two nodes in a tree) or to perform a task (route a message in a network based on the information stored at the router nodes).

In order to study these questions, our research plan is based on the following goals. Firstly, we aim both at building strong foundations for distributed algorithms (graph exploration, black-hole search,...) and distributed data structures (routing, efficient query, compact labeling...) to understand how to explore large scale networks in the context of failures and how to disseminate data so as to answer quickly to specific queries. Secondly, we aim at building simple (based on local estimations without centralized knowledge), realistic models to accurately represent resource performance and to build a realistic view of the topology of the network (based on network coordinates, geometric spanners, δ-hyperbolic spaces). Then, we aim at proving that these models are tractable by providing low complexity distributed and randomized approximation algorithms for a set of basic scheduling problems (independent tasks scheduling, broadcasting, data dissemination,...) and associated overlay networks. At last, our goal is to prove the validity of our approach through softwares dedicated to several applications (molecular dynamics simulations, continuous integration) as well as more general tools related to the model we propose (AlNEM for automatic topology discovery, SimGRID for simulations at large scale) and collections of datasets (Hubble for the continuous integration DAGs, Bedibe for latency and bandwidth measurements).

For the sake of the clarity of the presentation of our contributions during the last evaluation period, we decided to perform a synthesis, organizing our work into three main axes.

1. Resource Allocation and Scheduling (Section 4.1), with an emphasis on the interaction between network performance modeling and the design of efficient and guaranteed algorithms (covers axes 1(a), 2(a) and 2(b) of Section 2.1.1);
2. Compact Routing (Section 4.2), with an emphasis on the design of specific strategies for restricted graph classes and the design of data structures resilient to resource failures (covers axes 1(b), 2(b), 2(c), 3(a), 3(b) of Section 2.1.1);
3. Mobile Agents (Section 4.3), with en emphasis on the detection of dynamic faults and improvement of dissemination of information (covers axes 1(b), 2(b), 3(b) of Section 2.1.1);.

Nevertheless, strong relationships and collaborations exist between these axes:

- the design and analysis of overlay networks is shared by Resource Allocation and Scheduling and Mobile Agents axes (see for example [72], [96], [65], [112], [88]),
- the design of geometric spanners and geographic routing protocols is shared by Resource Allocation and Scheduling and Compact Routing axes (see for example [96], [80], [81], [116]),
- the use of specific graph classes (bounded tree-width, ...) is shared by Mobile Agents and Compact Routing axes (see for example [104], [111], [93], [98], [95], [57]).

3. Research Program

3.1. Modeling Platform Dynamics

Modeling the platform dynamics in a satisfying manner, in order to design and analyze efficient algorithms, is a major challenge. In distributed platforms, the performance of individual nodes (be they computing or communication resources) will fluctuate; in a fully dynamic platform, the set of available nodes will also change over time, and algorithms must take these changes into account if they are to be efficient.

There are basically two ways one can model such evolution: one can use a *stochastic process*, or some kind of *adversary model*.

In a stochastic model, the platform evolution is governed by some specific probability distribution. One obvious advantage of such a model is that it can be simulated and, in many well-studied cases, analyzed in detail. The two main disadvantages are that it can be hard to determine how much of the resulting algorithm performance comes from the specifics of the evolution process, and that estimating how realistic a given model is – none of the current project participants are metrology experts.

In an adversary model, it is assumed that these unpredictable changes are under the control of an adversary whose goal is to interfere with the algorithms efficiency. Major assumptions on the system's behavior can be included in the form of restrictions on what this adversary can do (like maintaining such or such level of connectivity). Such models are typically more general than stochastic models, in that many stochastic models can be seen as a probabilistic specialization of a nondeterministic model (at least for bounded time intervals, and up to negligible probabilities of adopting "forbidden" behaviors).

Since we aim at proving guaranteed performance for our algorithms, we want to concentrate on suitably restricted adversary models. The main challenge in this direction is thus to describe sets of restricted behaviors that both capture realistic situations and make it possible to prove such guarantees.

3.2. Models for Platform Topology and Parameter Estimation

On the other hand, in order to establish complexity and approximation results, we also need to rely on a precise theoretical model of the targeted platforms.

- At a lower level, several models have been proposed to describe interference between several simultaneous communications. In the 1-port model, a node cannot simultaneously send to (and/or receive from) more than one node. Most of the "steady state" scheduling results have been obtained using this model. On the other hand, some authors propose to model incoming and outgoing communication from a node using fictitious incoming and outgoing links, whose bandwidths are fixed. The main advantage of this model, although it might be slightly less accurate, is that it does not require strong synchronization and that many scheduling problems can be expressed as multi-commodity flow problems, for which efficient decentralized algorithms are known. Another important issue is to model the bandwidth actually allocated to each communication when several communications compete for the same long-distance link.

- At a higher level, proving good approximation ratios on general graphs may be too difficult, and it has been observed that actual platforms often exhibit a simple structure. For instance, many real life networks satisfy small-world properties, and it has been proved, for instance, that greedy routing protocols on small world networks achieve good performance. It is therefore of interest to prove that logical (given by the interactions between hosts) and physical platforms (given by the network links) exhibit some structure in order to derive efficient algorithms.

3.3. General Framework for Validation

3.3.1. Low level modeling of communications

In the context of large scale dynamic platforms, it is unrealistic to determine precisely the actual topology and the contention of the underlying network at application level. Indeed, existing tools such as Alnem [114] are very much based on quasi-exhaustive determination of interferences, and it takes several days to determine the actual topology of a platform made up of a few tens of nodes. Given the dynamism of the platforms we target, we need to rely on less sophisticated models, whose parameters can be evaluated at runtime.

Therefore, we propose to model each node using a small set of parameters. This is related to the theoretical notion of distance labeling [103], and corresponds to assigning labels to the nodes, so that a cheap operation on the labels of two nodes provides an estimation of the value of a given parameter (the latency or the bandwidth between two nodes, for instance). Several solutions for performance estimation on the Internet are based on this notion, under the terminology of Network Coordinate Systems. Vivaldi [94], IDES [115] and Sequoia [117] are examples of such systems for latency estimation. In the case of bandwidth estimation, fewer solutions have been proposed. We have studied the last-mile model, in which we model each node by an incoming and an outgoing bandwidth and neglect interference that appears at the core of the network (Internet), in order to concentrate on local constraints.

3.3.2. Simulation

Once low level modeling has been obtained, it is crucial to be able to test the proposed algorithms. To do this, we will first rely on simulation rather than direct experimentation. Indeed, in order to be able to compare heuristics, it is necessary to execute those heuristics on the same platform. In particular, all changes in the topology or in the resource performance should occur at the same time during the execution of the different heuristics. In order to be able to replicate the same scenario several times, we need to rely on simulations. Moreover, a metric for providing approximation results in the case of dynamic platforms necessarily requires computing the optimal solution at each time step, which can be done off-line if all traces for the different resources are stored. Using simulation rather than experiments can be justified if the simulator itself has been proven valid. Moreover, the modeling of communications, processing and their interactions may be much more complex in the simulator than in the model used to provide a theoretical approximation ratio, such as in SimGrid. In particular, sophisticated TCP models for bandwidth sharing have been implemented in SimGRID.

During the course of the USS-SimGrid ANR Arpege project, the SimGrid simulation framework has been adapted to large scale environments. Thanks to hierarchical platform description, to simpler and more scalable network models, and to the possibility to distribute the simulation of several nodes, it is now possible to perform simulations of very large platforms (of the order of 10^5 resources). This work will be continued in the ANR SONGS project, which aims at making SimGrid usable for Next Generation Systems (P2P, Grids, Clouds, HPC). In this context, simulation of exascale systems are envisioned, and we plan to develop models for platform dynamicity to allow realistic and reproducible experimentation of our algorithms.

3.3.3. Practical validation and scaling

Finally, we propose several applications that will be described in detail in Section 5. These applications cover a large set of fields (molecular dynamics, continuous integration...). All these applications will be developed and tested with an academic or industrial partner. In all these collaborations, our goal is to prove that the services that we propose can be integrated as steering tools in already developed software. Our goal is to assert the practical interest of the services we develop and then to integrate and to distribute them as a library for large scale computing.

At a lower level, in order to validate the models we propose, i.e. make sure that the predictions given by the model are close enough to the actual values, we need realistic datasets of network performance on large scale distributed platforms. Latency measurements are easiest to perform, and several datasets are available to researchers and serve as benchmarks to the community. Bandwidth datasets are more difficult to obtain, because of the measurement cost. As part of the bedibe software (see section 5.4), we have implemented a script to perform such measurements on the Planet-Lab platform [83]. We plan to make these datasets available to the community so that they can be used as benchmarks to compare the different solutions proposed.

4. Application Domains

4.1. Resource Allocation and Scheduling

4.1.1. Project-team positioning

CEPAGE has undertaken tasks related to the *high level modeling* of heterogeneous networks, both at logical level (overlay networks design) and performance level (latency, bandwidth prediction, connectivity artifacts) in order to optimize tasks such as *resource allocation* and *scheduling* of computations and communications. Objectives include replica placement, broadcasting (streaming) of large messages, independent tasks scheduling and optimization of OLAP databases. Such problems have received at lot of attention in research centers in the USA (Armherst, Colorado, ...), in Spain (Madrid), Poland (Wroclaw), Germany (Dortmund), and others. Papers on algorithmic aspects of platform modeling, scheduling and resource allocation appear at parallel processing conferences and journals in Parallel and Distributed Computing (IPDPS, EuroPar, HIPC, SPAA, IEEE TPDS, JPDC) and members of CEPAGE are strongly involved in many of these events (IPDPS, EuroPar, TPDS) as well as helping to animate well-established specialized workshops, such as HCW and HeteroPar.

Within Inria, studies on overlay networks are performed in the ASAP and GANG projects, and studies related to scheduling and resource allocation are done within the ROMA and the MOAIS projects (and to some extent within ALGORILLE).

4.1.2. Scientific achievements

The approach followed in the CEPAGE project, and our main originality, is to consider the whole chain, from gathering actual data on the networks to platform modeling and complexity analysis. Indeed, many complexity analysis studies are performed on models whose parameters cannot actually be evaluated (this applies, for instance, to all algorithms that assume that the topology of a platform running over the Internet is known in advance) and many platform models are intractable from an algorithmic perspective (this applies, for instance, to all models that represent latencies or bandwidths between all pairs of nodes as a general matrix). Our general goal is to provide models whose parameters can be evaluated at runtime using actual direct measurements, to propose algorithms whose worst-case (or average-case) behavior can be proved for this model, and finally to evaluate the whole chain (model + algorithm + implementation).

From an applicative perspective, in the framework of the PhD Thesis of Hejer Rejeb, we have considered several storage and resource allocation problems in collaboration with Cyril Banino-Rokkones at Yahoo! Trondheim (dealing with actual datasets enabled us to improve known approximation results in this specific context). We have in particular studied the modeling of TCP mechanism for handling contentions and its influence on the performance of several scheduling algorithms and advocated the use of QoS mechanisms for prescribed bandwidth sharing (IPDPS 2010 [78], ICPADS 2008 [63], AlgoTel 2009 [75], ICPADS 2009 [74], PDP 2010 [76]). In the PhD thesis of Hubert Larchevêque, we have considered the problem of aggregating resources (or placing replicas) in a distributed network (Sirocco 2008 [65], Opodis 2008 [66], ICPP 2011 [71], AlgoTel 2011 [67]) so that each group satifies some properties (in terms of aggregated memory, CPU and maximal distance in terms of latency within a group). We proved several multi-criteria approximation results for this problem, and we compared several embedding tools (Vivaldi, Sequoia) in the context of resource aggregation. For these applications, we have also provided when possible distributed algorithms based on sophisticated overlay networks, in particular in order to deal with heterogeneity (IPDPS 2008 [72]). In the PhD Thesis of Przemyslaw Uznanski, we focus on the design of efficient streaming and broadcasting strategies, in particular in presence of connectivity artifacts like firewalls (IPDPS 2010 [73], ICPADS 2011 [70]). We have also worked on establishing under the bounded multiport model several new complexity results for classical distributed computing models such as divisible load theory (HCW 2008 [68], IPDPS 2008 [118], IPDPS 2012 [69]) that have been later extended to Continuous Integration (HCW 2012 [64]).

In the context of database query optimization, materializing some queries results for optimization is a standard solution when execution time performance is crucial. In the datacube context, the problem has been studied for a long time under the storage space limit constraint. Here also, we were able to reformulate this problem by considering instead the execution time as the hard constraint while the objective is to reduce the storage space. Even if the problem turns to be NP-hard, this reformulation allowed us to provide effective approximate solutions with both space and performance bounded guarantees (EDBT 2009 [107]). Moreover, reducing the storage space tends to reduce the maintenance time since the latter is linearly proportional to the former. Finally, we characterized the minimal number of updates to be performed before performance becomes no more guaranteed and a new solution must be recomputed (ADBIS 2008 [108]). One of the key concepts we used for solving this problem was that of a *border*. It turns out that this notion is equivalent to e.g., maximal frequent itemsets or minimal functional dependencies extensively studied by data mining community. In contrast to all previous proposals, we proposed the only parallel algorithm computing these borders with a speed-up guarantee regarding the number of processing units (CIKM 2011 [106]). Besides the analytical study, its implementation in maximal frequent itemset mining outperforms state of the art implementations (see Section 5.1).

To achieve these results, our efforts have also focused on analyzing and building realistic datasets (AlgoTel 2012 [97]) and proposing data analysis results for specific distributions (ISAAC 2011 [59]). On the modeling side, in general, for bandwidth and contention modeling, we have proved that the bounded multi-port model (where each node is associated to an incoming bandwidth, an outgoing bandwidth and a maximal number of simultaneous TCP connexions) is both implementable, realistic and tractable (EuroPar 2011 [77]). In particular, we have proved in strongly different contexts (allocation of virtual machines to physical machines, overlay design for broadcasting, server allocation for volunteer computing) that the use of resource augmentation enables to obtain quasi-optimal results. All our modeling efforts and algorithms have been included into the SimGRID Software (http://simgrid.gforge.inria.fr), which enables us both to compare several algorithms under the same exact conditions and to compare the results obtained with several communication models (see Section 5.1)..

Perspectives: We believe that our approach based on sound models, approximation algorithms for these models, followed by experimental validation is a strong one and we intend to continue in this direction in the following years. Our goal of designing realistic solutions pushes towards considering average case analysis of our algorithms, as well as robust optimization techniques. Furthermore, the recent strong interest in Cloud systems from the community entices us to use our expertise in resource allocation for the optimization of Cloud systems, both from the provider and from the user points of view. We already have some interesting contacts with local companies to share start collaborating on these topics. In this context, reliability issues are very important, and we believe that robust optimization is a very relevant approach for these problems.

4.2. Compact Routing

4.2.1. Project-team positioning

In this axis, CEPAGE mainly works on the design on distributed and light data structures. One of the techniques consists in summarizing the topology and metric of the networks allowing to route or to approximate the original distances within the network. Such structures, often called *spanners*, does not require the storage of all the original network links. Then we get economic distributed data structures that can be updated without a high communication cost. Our main collaborations are done with the best specialists world-wide, in particular: Israel (Weizmann), USA (MIT, Microsoft, Chicago), Belgium (Alcatel Lucent-Bell), France (Paris, Nice).

Algorithms and Routing are also intensively studied in research labs in the USA (CAIDA). Our contributions appear regularly at all of the major conferences in Distributed Computing (PODC, DISC, SPAA), as well as at top venues with a more general algorithmic audience (STOC, SODA, ICALP, ESA). Members of CEPAGE actively participate in these events (ICALP 2010 and DISC 2009 were organized by members of CEPAGE).

Within Inria, studies of mobile agents are also performed in the GANG project and to some extent also within MASCOTTE within the european project EULER.

4.2.2. Scientific achievements

There are several techniques to manage sub-linear size routing tables (in the number of nodes of the platform) while guaranteeing almost shortest paths. Some techniques provide routes of length at most $1 + \epsilon$ times the length of the shortest one while maintaining a poly-logarithmic number of entries per routing table. However, these techniques are not universal in the sense that they apply only on some class of underlying topologies. Universal schemes exist. Typically they achieve $O(\sqrt{n})$-entry local routing tables for a stretch factor of 3 in the worst case. Some experiments have shown that such methods, although universal, work very well in practice, in average, on realistic scale-free or existing topologies.

The space lower bound of $O(\sqrt{n})$-entry for routing with *multiplicative* stretch 3 is due to the existence of dense graphs with large girth. Dense graphs can be sparsified to subgraphs (spanners), with various stretch guarantees. There are spanners with *additive* stretch guarantees (some even have constant additive stretch) but only very few additive routing schemes are known.

In (SPAA 2012 [101]), we give reasons why routing in unweighted graphs with *additive* stretch is difficult in the form of space lower bounds for general graphs and for planar graphs. On the positive side, we give an almost tight upper bound: we present the first non-trivial compact routing scheme with $o(\lg^2 n)$-bit addresses, *additive* stretch $O(n^{1/3})$, and table size $O(n^{1/3})$ bits for planar graphs.

We have recently considered the *forbidden-set* extension of distance oracles and routing schemes. Given an arbitrary set of edge/node failure F, a source s and a target t such that $s, t \notin F$, the goal is to route (or evaluate the distance) between s and t in the graph $G \setminus F$, so avoiding F. The classical problem is for $F = \varnothing$. This extension is considered as a first step toward fully dynamic data-structures, a challenging goal. For graphs of low doubling dimension we have shown in (PODC 2012 [58]) that it is possible to route from s to t in $G \setminus F$ with stretch $1 + \epsilon$, for all s, t, F, given poly-logarithmic size labels of all the nodes invoked in the query (s, t, F). This has been generalized to all planar graphs achieving similar stretch and label size performences. As a byproduct we have designed a fully dynamic algorithm for maintaining $1 + \epsilon$ approximate distances in planar graphs supporting edge/node addition/removal within update and query time $\sqrt{n}$ in the worst-case (STOC 2012 [57]).

Θ_k-graphs are geometric graphs that appear in the context of graph navigation. The shortest-path metric of these graphs is known to approximate the Euclidean complete graph up to a factor depending on the cone number k and the dimension of the space. We have introduced in (WG 2010 [79]) a specific subgraph of the Θ_6-graph defined in the 2D Euclidean space, namely the half-Θ_6-graph, composed of the even-cone edges of the Θ_6-graph. Our main contribution is to show that these graphs are exactly the TD-Delaunay graphs, and are strongly connected to the geodesic embeddings of orthogonal surfaces of coplanar points in the 3D Euclidean space. We also studied the asymptotic behavior of these spanners (*Adv. in Appl. Proba.* [116]) and in collaboration with Ljubomir Perković, we worked on the question of bounded degree planar spanner. We proposed an algorithm that computes a plane 6-spanner of degree at most 6 in (ICALP 2010 [80]). The previous best bound on the maximum degree for constant stretch plane spanners was only 14.

In order to cope with network dynamism and failures, and motivated by multipath routing, we introduce a multi-connected variant of spanners. For that purpose we introduce in (OPODIS 2011 [102]) the p-multipath cost between two nodes u and v as the minimum weight of a collection of p internally vertex-disjoint paths between u and v. Given a weighted graph G, a subgraph H is a p-multipath s-spanner if for all u, v, the p-multipath cost between u and v in H is at most s times the p-multipath cost in G. The s factor is called the stretch. Building upon recent results on fault-tolerant spanners, we show how to build p-multipath spanners of constant stretch and of $O(n^{1+1/k})$ edges, for fixed parameters p and k, n being the number of nodes of the graph. Such spanners can be constructed by a distributed algorithm running in $O(k)$ rounds. Additionally, we give an improved construction for the case $p = k = 2$. Our spanner H has $O(n^{3/2})$ edges and the p-multipath cost in H between any two node is at most twice the corresponding one in G plus $O(W)$, W being the maximum edge weight.

We also worked on compact coding in data warehouses: in order to get quick answer in large data, we have to estimate, select and materialize (store) partial data structures. We got several solutions with a prescribed guarantee in different models for the following problems: view size estimation with small samples, view selection, parallel computation of frequent itemsets. In (*Theor. Comp. Sci.* [105]) a new algorithm that allow the administrator or user of a DBMS to choose which part of the data cube to optimize (known as the *the views selection problem*), that takes as input a fact table and computes a set of views to store in order to speed up queries.

Perspectives: The compact coding activity in data-warehouse is promising since the amount of data collected keeps on increasing and being able to answer in real-time complex requests (data mining) is still challenging.

Some robust data structures already exist which, given a small number of k changes of topology or k faults, tolerate these faults, i.e., alternative routes with bounded stretch can be provided without any updates. This is a first step toward dynamic networks but the updates of these data structures are currently still quite complicated with a high communication cost.

4.3. Mobile Agents

4.3.1. Project-team positioning

CEPAGE has undertaken tasks related to the design of algorithms which control the behavior of so called *mobile agents*, moving around a network or a geometric environment, with the goal of achieving a specified objective. Objectives of central importance to the study include: exploration of unknown environments, terrain patrolling, network maintenance, and coordination of activities with other agents. Such problems have in recent years been the object of interest of numerous research teams working on Distributed Computing worldwide, in particular, at research centers in Canada (Quebec), Israel (Tel Aviv, Haifa), France (Paris, Marseille), the UK (London, Liverpool), and Switzerland (Zurich). Algorithms for mobile agents in social networking applications are also intensively studied in research labs in the USA (Stanford, Facebook). Papers on mobile agents appear regularly at all of the major conferences in Distributed Computing (PODC, DISC, SPAA), as well as at top venues with a more general algorithmic audience (SODA, ICALP, ESA). Members of CEPAGE actively participate in these events, and are also a recognizable part of the European community focused around mobile agents, helping to animate well-established specialized conferences, such as SIROCCO and OPODIS.

Within Inria, studies of mobile agents are also performed in the GANG project, and to some extent also within MASCOTTE. CEPAGE has active research links with both of these teams.

4.3.2. Scientific achievements

The work of CEPAGE has focused on contributing new decentralized algorithms for controlling mobile entities known as *agents*, deployed in unknown environments. We mainly considered the network setting, in which agents moving around the nodes of the network graph may be used to analyze the structure of the network and to perform maintenance tasks, such as detecting dynamic faults, improving/monitoring dissemination of information, etc. Our theoretical studies focused on designing new strategies for controlling the behavior of agents and answering crucial questions concerning the feasibility of solving fundamental problems, subject to different model assumptions and constraints on the knowledge and computational power of agents.

One major line of our research focused on the so called *anonymous graph model* in which an agent is unable to determine the identifier of the node of its current location, but can only see a local ordering of the links around it. Such a study is motivated e.g. by scenarios in which the identifiers of nodes may be too large for the agent to process using its bounded resources, or may change in time. In this model, we studied two of the most fundamental problems: that of traversing all of the nodes of the network (exploration) and of meeting another agent in the network (rendezvous), so as to coordinate with it. Our contributions include a precise characterization of the space requirements for agents solving both of these problems deterministically: exploration in (*Trans. Alg.* 2008 [84]) and rendezvous in (*Dist. Comp.* 2012 [92]), in a paper presented at the Best Paper Session of PODC 2010. We have also studied fast solutions for specific scenarios of the rendezvous problem (DISC 2010 [60], DISC 2011 [85], SPAA 2012 [93]) and the problem of approximate map construction within an anonymous graph (OPODIS 2010 [82]). A separate problem, intensively studied in recent years by several research teams, concerns the exploration of a network with pre-configured ports so as to assist the agent. In our work on the topic, our team has proposed several new techniques for graph decomposition, leading in particular to the shortest currently known strategies of periodic exploration for both the case of memoryless (*Algorithmica* 2012 [112]) and small-memory agents (SIROCCO 2009 [88]).

A closely related line of research was devoted to the design of network exploration strategies which guarantee a fast and fair traversal of all the nodes, making use of agents with extremely restricted capabilities. Such strategies were inspired by the random walk, but had the additional advantage of deterministic and desirable behavior in worst-case scenarios. We presented a series of results in the area at notable conferences, involving both the design of new exploration strategies (ICALP 2009 [86]) and completely new insights into previously known approaches such as the so called "rotor-router model" (DISC 2009 [61], OPODIS 2009 [62]). All of the proposed algorithms were shown to be viable alternatives to the random walk, competing in terms of such parameters as cover time, steady-state exploration frequency, and stabilization in the event of faults.

Our efforts have also focused on the theory of coordinating activities of large groups of agents. We have conducted pioneering work in the so called look-compute-move model in networks, in which extremely restricted (asynchronous and oblivious) agents, relying on snapshot views of the system, are nevertheless able to perform useful computational tasks. Our solutions to the problems of collective exploration in trees (*Theor. Comp. Sci.* 2010 [99]) and gathering agents on a ring (*Theor. Comp. Sci.* 2008 [110] and 2010 [109]) have sparked a long line of follow-up research, accumulating more than 120 citations in total (according to Google Scholar). In a slightly different scenario, we have considered computations with teams of agents whose task is to collaboratively detect and mark potentially dangerous (faulty) links of the network, called "black holes", which are capable of destroying agents which enter them. We have provided important contributions to the theory of black hole search in both undirected (SIROCCO 2008 [87], DISC 2008 [100]) and directed (*Theor. Comp. Sci.* [113]) graphs.

It is expected that the mobile agent theme of CEPAGE will give rise to 2 PhD theses. In 2013, Ahmed Wade will defend his thesis on mobile agent protocols for dynamic networks, whereas in 2014 Dominik Pajak will defend his thesis on multi-agent protocols for efficient graph exploration. Our scientific interests also include mobile agent protocols for geometric applications, more remote from the central themes of CEPAGE, but having extensive applications in robotics (providing protocols, e.g., for efficient patrolling and guarding of terrains, traversing terrains using groups of robots, etc.). We have already published several papers in this area (SIROCCO 2010 [90], SWAT 2010 [91], ESA 2011 [89]), building up the theoretical fundamentals of a new field, and already attracting the attention of a wider community of researchers working in robotics and AI.

Perspectives: Our goal is to explore applications of mobile agent techniques in domains of growing importance, namely, social networks and robotics. We are currently discussing applications of our techniques in problems of brand recognition on the web with a local industrial partner (Systonic KeepAlert), and other companies (through our research collaborators in Liverpool). We intend to undertake collaboration with European/American research labs and industrial partners.

5. Software and Platforms

5.1. SimGrid

Participants: Paul Renaud-Goud, Lionel Eyraud-Dubois [correspondant].

SimGrid (http://simgrid.gforge.inria.fr/) is a toolkit that provides core functionalities for the simulation of distributed applications in heterogeneous distributed environments. The specific goal of the project is to facilitate research in the area of parallel and distributed large scale systems, such as Grids, P2P systems and clouds. Its use cases encompass heuristic evaluation, application prototyping or even real application development and tuning. It is based on experimentally validated models, and features very high scalability, which allows to perform very large scale simulations. It is used by over a hundred academic users all over the world, and has been used in about one hundred scientific articles.

CEPAGE has contributed to this software by participating in the management of the project and in many design decisions. As a part of the SONGS project, we also participated in the development and validation of new models and interfaces for the HPC and Cloud Computing platforms.

Software assessment : A-4, SO-4, SM-4, EM-3, SDL-5
Contribution : DA-2, CD-2, MS-2, TPM-3.

5.2. Hubble

Participants: Ludovic Courtes, Nicolas Bonichon [correspondant].

Hubble is implemented in Scheme, using GNU Guile version 2. Details of the simulation, such as keeping track of processor occupation and network usage, are taken care of by SimGrid, a toolkit for the simulation of distributed applications in heterogeneous distributed environments.

The input to Hubble is an XML description of the DAG of build tasks. For each task, a build duration and the size in bytes of the build output are specified. For our evaluation purposes, we collected this data on a production system, the http://hydra.nixos.org/ build farm hosted at the Technical University of Delft. The DAG itself is the snapshot of the Nix Package Collection (Nixpkgs) corresponding to this data. Hubble has its own in-memory representation of the DAG in the form of a purely functional data structure.

The Nixpkgs DAG contains fixed-output nodes, i.e., nodes whose output is known in advance and does not require any computation. These nodes are typically downloads of source code from external web sites. The raw data collected on http://hydra.nixos.org/ specifies a non-zero duration for these nodes, which represents the time it took to perform the download. This duration info is irrelevant in our context, since they don't require any computation, and Hubble views these nodes as instantaneous.

See also the web page http://hubble.gforge.inria.fr/.

Software assessment: A-3, SO-3, SM-2, EM-1, SDL-2.
Contribution: DA-4, CD-4, MS-4, TPM-4.

5.3. Gengraph

Participant: Cyril Gavoille [correspondant].

This is a command-line tool for generating graphs. There are several output formats, includes the dot format from GraphViz. It generates also .pdf files for visualization. Several graph algorithms have been implemented (diameter, connectivity, treewidth, etc.) which can be tested on the graphs. The software has been originally designed for teaching purpose so that students can test their project algorithms on many non trivial families like random geometric graphs, graphs of given density, given treewidth. It is also used for research purpose, in particular the exhaustive search results in the Emilie Diot's thesis are based on `gengraph`. The program can filter a list of graphs based to many criteria, as for instance it can extract all graphs of a given list that are 2-connected, of diameter at least four, and that exclude some minor (or some induced subgraph).

Currently, more than 100 parametrized graph families are implemented, supporting simple operators like complementation, random edge/vertex removal, and others. The source has more than 10,000 lines including a command-line documentation of 2,000 lines. The single source file is available at http://dept-info.labri.fr/~gavoille/gengraph.c

Software assessment: A-3, SO-3, SM-2, EM-2, SDL-2.

Contribution: DA-4, CD-4, MS-4, TPM-4.

5.4. Bedibe

Participants: Lionel Eyraud-Dubois [correspondant], Przemyslaw Uznanski.

Bedibe (Benchmarking Distributed Bandwidth Estimation) is a software to compare different models for bandwidth estimation on the Internet, and their associated instantiation algorithms. The goal is to ease the development of new models and algorithms, and the comparison with existing solutions. Additionally, we have developed a measuring framework which can be deployed on PlanetLab to obtain available bandwidth datasets.

See also the web page http://bedibe.gforge.inria.fr/.

Software assessment : A-1-up2, SO-3, SM-1-up2, EM-2, SDL-1-up2.

5.5. MineWithRounds

Participants: Sofian Maabout [correspondant], Nicolas Hanusse.

The software implements a parallel algorithm aiming at computing *Borders* that's sets of maximal/minimal subsets of objects satisfying some anti-monotone condition. It is implemented in `C++` together with the `openMP` library to exploit multi-core machines. In its current status, it outperforms state of the art implementations addressing the Maximal Frequent Itemsets problem.

Software assessment: A-2, SO-4, SM-2, EM-2, SDL-2.
Contribution: DA-4, CD-4, MS-4, TPM-4.

5.6. FSM: Full Skycube Materialization

Participant: Sofian Maabout [correspondant].

The software implements a parallel algorithm aiming at fully materializing skycubes: the set of all possible skyline queries. It is implemented in C++ together with the openMP library to exploit multi-core machines. In its current status, it outperforms state of the art implementations such as QSkycube and PSkyCube (VLDB Journal '13)

Software assessment: A-2, SO-4, SM-2, EM-2, SDL-2.
Contribution: DA-4, CD-4, MS-4, TPM-4.

ADT: Pierre Matri has been hired as an engineer for two year starting from September 15th, 2013. He is in charge to implement our algorithm on top of a large scale infrastructure especially by using Map-Reduce paradigm and its Hadoop implementation. His main focus during 2013 concerned the functional dependencies extraction from large distributed data bases by considering exact and approximate solutions.

6. New Results

6.1. Resource allocation and Scheduling

6.1.1. Broadcasting on Large Scale Heterogeneous Platforms under the Bounded Multi-Port Model

Participants: Olivier Beaumont, Nicolas Bonichon, Lionel Eyraud-Dubois, Przemyslaw Uznanski.

In [17], we consider the problem of broadcasting a large message in a large scale distributed network under the multi-port communication model. We are interested in building an overlay network, with the aim of maximizing the throughput and minimizing the degree of the participating nodes. We consider a classification of participating nodes into two parts: open nodes that stay in the open-Internet and "guarded" nodes that lie behind firewalls or NATs, with the constraint that two guarded nodes cannot communicate directly. Without guarded nodes, we prove that it is possible to reach the optimal throughput with a quasi-optimal (up to a small additive increase) degree of the participating nodes. In presence of guarded nodes, we provide a closed form formula for the optimal cyclic throughput and we observe that the optimal solution may require arbitrarily large degrees. In the acyclic case, we propose an algorithm that reaches the optimal acyclic throughput with low degree. Then, we prove a worst case 5/7 ratio between the optimal acyclic and cyclic throughput and show through simulations that this ratio is on average very close to 1, what makes acyclic solutions efficient both in terms of throughput maximization and degree minimization.

6.1.2. Non Linear Divisible Load Scheduling

Participants: Olivier Beaumont, Hubert Larchevêque.

Divisible Load Theory (DLT) has received a lot of attention in the past decade. A divisible load is a perfect parallel task, that can be split arbitrarily and executed in parallel on a set of possibly heterogeneous resources. The success of DLT is strongly related to the existence of many optimal resource allocation and scheduling algorithms, what strongly differs from general scheduling theory. Moreover, recently, close relationships have been underlined between DLT, that provides a fruitful theoretical framework for scheduling jobs on heterogeneous platforms, and MapReduce, that provides a simple and efficient programming framework to deploy applications on large scale distributed platforms. The success of both have suggested to extend their framework to non-linear complexity tasks. In [32], we show that both DLT and MapReduce are better suited to workloads with linear complexity. In particular, we prove that divisible load theory cannot directly be applied to quadratic workloads, such as it has been proposed recently. We precisely state the limits for classical DLT studies and we review and propose solutions based on a careful preparation of the dataset and clever data partitioning algorithms. In particular, through simulations, we show the possible impact of this approach on the volume of communications generated by MapReduce, in the context of Matrix Multiplication and Outer Product algorithms. (Joint work with Loris Marchal from ENS Lyon)

6.1.3. Reliable Service Allocation in Clouds

Participants: Olivier Beaumont, Lionel Eyraud-Dubois, Hubert Larchevêque, Paul Renaud-Goud, Philippe Duchon.

In [30], we consider several reliability problems that arise when allocating applications to processing resources in a Cloud computing platform. More specifically, we assume on the one hand that each computing resource is associated to a capacity constraint and to a probability of failure. On the other hand, we assume that each service runs as a set of independent instances of identical Virtual Machines, and that the Service Level Agreement between the Cloud provider and the client states that a minimal number of instances of the service should run with a given probability. In this context, given the capacity and failure probabilities of the machines, and the capacity and reliability demands of the services, the question for the cloud provider is to find an allocation of the instances of the services (possibly using replication) onto machines satisfying all types of constraints during a given time period. The goal of this work is to assess the impact of the reliability constraint on the complexity of resource allocation problems. We consider several variants of this problem, depending on the number of services and whether their reliability demand is individual or global. We prove several fundamental complexity results (#P' and NP-completeness results) and we provide several optimal and approximation algorithms. In particular, we prove that a basic randomized allocation algorithm, that is easy to implement, provides optimal or quasi-optimal results in several contexts, and we show through simulations that it also achieves very good results in more general settings.

In [29], we extend this work to an energy minimisation framework, by considering two energy consumption models based on DVFS techniques, where the clock frequency of physical resources can be changed with a Dynamic Voltage and Frequency Scaling (DVFS) method. For each allocation problem and each energy model, we prove deterministic approximation ratios on the consumed energy for algorithms that provide guaranteed probability failures, as well as an efficient heuristic, whose energy ratio is not guaranteed.

In [37], we study the robustness of an allocation of Virtual Machines (VM) on a set of Physical Machines (PM) when the resource demand of the VMs can change over time. This may imply sometimes expensive "SLA violations", corresponding to some VM's consumption not satisfied because of overloaded PMs. Thus, while optimizing the global resource utilization of the PMs, it is necessary to ensure that at any moment a VM's need evolves, a few number of migrations (moving a VM from PM to PM) is sufficient to find a new configuration in which all the VMs' consumptions are satisfied. We modelize this problem using a fully dynamic bin packing approach and we present an algorithm ensuring a global utilization of the resources of 66%. Moreover, each time a PM is overloaded at most one migration is necessary to fall back in a configuration with no overloaded PM, and only 3 different PMs are concerned by required migrations that may occur to keep the global resource utilization correct. This allows the platform to be highly resilient to a great number of changes.

6.1.4. Splittable Single Source-Sink Routing on CMP Grids: A Sublinear Number of Paths Suffice

Participants: Adrian Kosowski, Przemyslaw Uznanski.

In [44], we study single chip multiprocessors (CMP) with grid topologies, where a significant part of power consumption is attributed to communications between the cores of the grid. We investigate the problem of routing communications between CMP cores using shortest paths, in a model in which the power cost associated with activating a communication link at a transmission speed of f bytes/second is proportional to f^α, for some constant exponent $\alpha > 2$. Our main result is a trade-off showing how the power required for communication in CMP grids depends on the ability to split communication requests between a given pair of node, routing each such request along multiple paths. For a pair of cores in a $m \times n$ grid, the number of available communication paths between them grows exponentially with n, m. By contrast, we show that optimal power consumption (up to constant factors) can be achieved by splitting each communication request into k paths, starting from a threshold value of $k = \Theta(n^{1/(\alpha-1)})$. This threshold is much smaller than n for typical values of $\alpha \approx 3$, and may be considered practically feasible for use in routing schemes on the grid. More generally, we provide efficient algorithms for routing multiple k-splittable communication requests between two cores in the grid, providing solutions within a constant approximation of the optimum cost. We support our results with algorithm simulations, showing that for practical instances, our approach using k-splittable requests leads to a power cost close to that of the optimal solution with arbitrarily splittable requests, starting from the stated threshold value of k.

6.1.5. Maximum matching in multi-interface networks

Participants: Adrian Kosowski, Dominik Pajak.

In [26], we consider the standard matching problem in the context of multi-interface wireless networks. In heterogeneous networks, devices can communicate by means of multiple wireless interfaces. By choosing which interfaces to switch on at each device, several connections might be established. That is, the devices at the endpoints of each connection share at least one active interface. In the studied problem, the aim is to maximize the number of parallel connections without incurring in interferences. Given a network $G = (V, E)$, nodes V represent the devices, edges E represent the connections that can be established. If node x participates in the communication with one of its neighbors by means of interface i, then another neighboring node of x can establish a connection (but not with x) only if it makes use of interface $j \neq i$. The size of a solution for an instance of the outcoming matching problem, that we call *Maximum Matching in Multi-Interface* networks (MMMI for short), is always in between the sizes of the solutions for the same instance with respect to the standard matching and its induced version problems. However, we prove that MMMI is NP-hard even for proper interval graphs and for bipartite graphs of maximum degree $\Delta \geq 3$. We also show polynomially solvable cases of MMMI with respect to different assumptions.

6.1.6. Parallel scheduling of task trees with limited memory

Participant: Lionel Eyraud-Dubois.

In a paper submitted to ACM TOPC, we have investigated the execution of tree-shaped task graphs using multiple processors. Each edge of such a tree represents some large data. A task can only be executed if all input and output data fit into memory, and a data can only be removed from memory after the completion of the task that uses it as an input data. Such trees arise, for instance, in the multifrontal method of sparse matrix factorization. The peak memory needed for the processing of the entire tree depends on the execution order of the tasks. With one processor the objective of the tree traversal is to minimize the required memory. This problem was well studied and optimal polynomial algorithms were proposed. We have extended the problem by considering multiple processors, which is of obvious interest in the application area of matrix factorization. With multiple processors comes the additional objective to minimize the time needed to traverse the tree, i.e., to minimize the makespan. Not surprisingly, this problem proves to be much harder than the sequential one. We study the computational complexity of this problem and provide inapproximability results even for unit weight trees. We design a series of practical heuristics achieving different trade-offs between the minimization of peak memory usage and makespan. Some of these heuristics are able to process a tree while keeping the memory usage under a given memory limit. The different heuristics are evaluated in an extensive experimental evaluation using realistic trees.

6.1.7. Point-to-point and congestion bandwidth estimation: experimental evaluation on PlanetLab

Participants: Lionel Eyraud-Dubois, Przemyslaw Uznanski.

In large scale Internet platforms, measuring the available bandwidth between nodes of the platform is difficult and costly. However, having access to this information allows to design clever algorithms to optimize resource usage for some collective communications, like broadcasting a message or organizing master/slave computations. In [54], we analyze the feasibility to provide estimations, based on a limited number of measurements, for the point-to-point available bandwidth values, and for the congestion which happens when several communications take place at the same time. We present a dataset obtained with both types of measurements performed on a set of nodes from the PlanetLab platform. We show that matrix factorization techniques are quite efficient at predicting point-to-point available bandwidth, but are not adapted for congestion analysis. However, a LastMile modeling of the platform allows to perform congestion predictions with a reasonable level of accuracy, even with a small amount of information, despite the variability of the measured platform.

6.1.8. Parallel Mining of Functional Dependencies

Participants: Sofian Maabout, Nicolas Hanusse.

The problem of extracting functional dependencies (FDs) from databases has a long history dating back to the 90's. Still, efficient solutions taking into account both material evolution, namely the advent of multicore machines, and the amount of data that are to be mined, are still needed. In [46] we propose a parallel algorithm which, upon small modifications, extracts (i) the minimal keys, (ii) the minimal exact FDs, (iii) the minimal approximate FDs and (iv) the Conditional functional dependencies (CFDs) holding in a table. Under some natural conditions, we prove a theoretical speed up of our solution with respect to a baseline algorithm which follows a depth first search strategy. Since mining most of these dependencies require a procedure for computing the number of distinct values (NDV) which is a space consuming operation, we show how sketching techniques for estimating the exact value of NDV can be used for reducing both memory consumption as well as communications overhead when considering distributed data while guaranteeing a certain quality of the result. Our solution is implemented in both shared, using C++ and OpenMP, and distributed memory, using Hadoop implementation of Map-Reduce. The experimental results showthe efficiency and scalability of our proposal. Most notably, the theoretical speed ups are confirmed by the experiments.

6.1.9. Fast Skyline Query Evaluation with Skycuboids Materialization based on Functional Dependencies

Participants: Sofian Maabout, Nicolas Hanusse.

Ranking multidimensional data via different Skyline queries gives rise to the so called skycube structure. Most of previous work on optimizing the subspaces skyline queries have concentrated on full materialization of the skycube. Due to the exponential number of skylines one must pre-compute, the full materialization is unfeasible in practice. However, due to the non monotonic nature of skylines, there is no immediate inclusion relationship between the skycuboids when we have an inclusion of the dimensions. This makes the partial materialization harder. In this paper, we identify sufficient conditions for establishing inclusions between skycuboids thanks to the functional dependencies that hold in the underlying data. This leads to the characterization of a *minimal* set of skycuboids to be materialized in order to answer all the possible skyline queries without resorting to the underlying data. We conduct an extensive set of experiments showing that with the help of a small fraction of the skycube, we can efficiently answer all the possible skyline queries. In addition, our proposal turns to be helpful even in the full materialization setting. Indeed, thanks to the inclusions we identify, we devise a full materialization algorithm which outperforms state of the art skycube computation algorithms especially when data and dimensions get large. The results are reported in the technical report submitted to SIGMOD'14.

6.2. Compact Routing

6.2.1. On the Communication Complexity of Distributed Name-Independent Routing Schemes

Participants: Cyril Gavoille, Nicolas Hanusse, David Ilcinkas.

In [38], we present a distributed asynchronous algorithm that, for every undirected weighted n-node graph G, constructs name-independent routing tables for G. The size of each table is $\widetilde{O}(\sqrt{n}\,)$, whereas the length of any route is stretched by a factor of at most 7 w.r.t. the shortest path. At any step, the memory space of each node is $\widetilde{O}(\sqrt{n}\,)$. The algorithm terminates in time $O(D)$, where D is the hop-diameter of G. In synchronous scenarios and with uniform weights, it consumes $\widetilde{O}(m\sqrt{n} + n^{3/2} \min D, \sqrt{n}\,)$ messages, where m is the number of edges of G.

In the realistic case of sparse networks of poly-logarithmic diameter, the communication complexity of our scheme, that is $\widetilde{O}(n^{3/2})$, improves by a factor of $\sqrt{n}$ the communication complexity of *any* shortest-path routing scheme on the same family of networks. This factor is provable thanks to a new lower bound of independent interest.

6.2.2. There are Plane Spanners of Maximum Degree 4

Participant: Nicolas Bonichon.

Let $\mathcal{E}$ be the complete Euclidean graph on a set of points embedded in the plane. Given a fixed constant $t \geq 1$, a spanning subgraph G of $\mathcal{E}$ is said to be a *t-spanner* of $\mathcal{E}$ if for any pair of vertices u, v in $\mathcal{E}$ the distance between u and v in G is at most t times their distance in $\mathcal{E}$. A spanner is *plane* if its edges do not cross.

We consider the question: “What is the smallest *maximum degree* that can be achieved for a *plane* spanner of $\mathcal{E}$?” Without the planarity constraint, it is known that the answer is 3 which is thus the best known lower bound on the degree of any plane spanner. With the planarity requirement, the best known upper bound on the maximum degree is 6, the last in a long sequence of results improving the upper bound. In this work we show that there is a constant $t \geq 1$ such that the complete Euclidean graph always contains a plane t-spanner of maximum degree 4 and make a big step toward closing the question. Our construction leads to an efficient algorithm for obtaining the spanner from Chew's L_1-Delaunay triangulation.

6.3. Mobile Agents

6.3.1. Collision-Free Network Exploration

Participants: Ralf Klasing, Adrian Kosowski, Dominik Pajak.

A set of mobile agents is placed at different nodes of a n-node network. The agents synchronously move along the network edges in a *collision-free* way, i.e., in no round may two agents occupy the same node. In each round, an agent may choose to stay at its currently occupied node or to move to one of its neighbors. An agent has no knowledge of the number and initial positions of other agents. We are looking for the shortest possible time required to complete the collision-free *network exploration*, i.e., to reach a configuration in which each agent is guaranteed to have visited all network nodes and has returned to its starting location. In [34], we first consider the scenario when each mobile agent knows the map of the network, as well as its own initial position. We establish a connection between the number of rounds required for collision-free exploration and the degree of the minimum-degree spanning tree of the graph. We provide tight (up to a constant factor) lower and upper bounds on the collision-free exploration time in general graphs, and the exact value of this parameter for trees. For our second scenario, in which the network is unknown to the agents, we propose collision-free exploration strategies running in $O(n^2)$ rounds for tree networks and in $O(n^5 \log n)$ rounds for general networks.

6.3.2. Deterministic Rendezvous of Asynchronous Bounded-Memory Agents in Polygonal Terrains

Participant: Adrian Kosowski.

In [22], we deal with a more geometric variant of the rendezvous problem. Two mobile agents, modeled as points starting at different locations of an unknown terrain, have to meet. The terrain is a polygon with polygonal holes. We consider two versions of this rendezvous problem: exact RV, when the points representing the agents have to coincide at some time, and ϵ-RV, when these points have to get at distance less than ϵ in the terrain. In any terrain, each agent chooses its trajectory, but the movements of the agent on this trajectory are controlled by an adversary that may, e.g., speed up or slow down the agent. Agents have bounded memory: their computational power is that of finite state machines. Our aim is to compare the feasibility of exact and of ϵ-RV when agents are anonymous vs. when they are labeled. We show classes of polygonal terrains which distinguish all the studied scenarios from the point of view of feasibility of rendezvous. The features which influence the feasibility of rendezvous include symmetries present in the terrains, boundedness of their diameter, and the number of vertices of polygons in the terrains.

6.3.3. Optimal Patrolling of Fragmented Boundaries

Participant: Adrian Kosowski.

Mobile agents in geometric scenarios are also studied in [33], where a set of mobile robots is deployed on a simple curve of finite length, composed of a finite set of vital segments separated by neutral segments. The robots have to patrol the vital segments by perpetually moving on the curve, without exceeding their maximum speed. The quality of patrolling is measured by the idleness, i.e., the longest time period during which any vital point on the curve is not visited by any robot. Given a configuration of vital segments, our goal is to provide algorithms describing the movement of the robots along the curve so as to minimize the idleness. Our main contribution is a proof that the optimal solution to the patrolling problem is attained either by the cyclic strategy, in which all the robots move in one direction around the curve, or by the partition strategy, in which the curve is partitioned into sections which are patrolled separately by individual robots. These two fundamental types of strategies were studied in the past in the robotics community in different theoretical and experimental settings. However, to our knowledge, this is the first theoretical analysis proving optimality in such a general scenario.

6.3.4. Fast Collaborative Graph Exploration

Participants: Adrian Kosowski, Dominik Pajak, Przemyslaw Uznanski.

In [35], we study the following scenario of online graph exploration. A team of k agents is initially located at a distinguished vertex r of an undirected graph. At every time step, each agent can traverse an edge of the graph. All vertices have unique identifiers, and upon entering a vertex, an agent obtains the list of identifiers of all its neighbors. We ask how many time steps are required to complete exploration, i.e., to make sure that every vertex has been visited by some agent. We consider two communication models: one in which all agents have global knowledge of the state of the exploration, and one in which agents may only exchange information when simultaneously located at the same vertex. As our main result, we provide the first strategy which performs exploration of a graph with n vertices at a distance of at most D from r in time $O(D)$, using a team of agents of polynomial size $k = Dn^{1+\epsilon} < n^{2+\epsilon}$, for any $\epsilon > 0$. Our strategy works in the local communication model, without knowledge of global parameters such as n or D. We also obtain almost-tight bounds on the asymptotic relation between exploration time and team size, for large k. For any constant $c > 1$, we show that in the global communication model, a team of $k = Dn^c$ agents can always complete exploration in $D(1 + \frac{1}{c-1} + o(1))$ time steps, whereas at least $D(1 + \frac{1}{c} - o(1))$ steps are sometimes required. In the local communication model, $D(1 + \frac{2}{c-1} + o(1))$ steps always suffice to complete exploration, and at least $D(1 + \frac{2}{c} - o(1))$ steps are sometimes required. This shows a clear separation between the global and local communication models.

6.3.5. A $\widetilde{O}(n2)$ *Time-Space Trade-off for Undirected s-t Connectivity*

Participant: Adrian Kosowski.

The work [43] makes use of the Metropolis-type walks due to Nonaka et al. (2010) to provide a faster solution to the S-T-connectivity problem in undirected graphs (USTCON). As the main result of this research, we propose a family of randomized algorithms for USTCON which achieves a time-space product of $S \cdot T = \widetilde{O}(n^2)$ in graphs with n nodes and m edges (where the $\widetilde{O}$-notation disregards poly-logarithmic terms). This improves the previously best trade-off of $\widetilde{O}(nm)$, due to Feige (1995). Our algorithm consists in deploying several short Metropolis-type walks, starting from landmark nodes distributed using the scheme of Broder et al. (1994) on a modified input graph. In particular, we obtain an algorithm running in time $\widetilde{O}(n + m)$ which is, in general, more space-efficient than both BFS and DFS. Finally, we show how to fine-tune the Metropolis-type walk so as to match the performance parameters (e.g., average hitting time) of the unbiased random walk for any graph, while preserving a worst-case bound of $\widetilde{O}(n^2)$ on cover time.

6.3.6. *The multi-agent rotor-router on the ring: a deterministic alternative to parallel random walks*

Participants: Ralf Klasing, Adrian Kosowski, Dominik Pajak.

The *rotor-router mechanism* was introduced as a deterministic alternative to the random walk in undirected graphs. In this model, an agent is initially placed at one of the nodes of the graph. Each node maintains a cyclic ordering of its outgoing arcs, and during successive visits of the agent, propagates it along arcs chosen according to this ordering in round-robin fashion. In [42], we consider the setting in which multiple, indistinguishable agents are deployed in parallel in the nodes of the graph, and move around the graph in synchronous rounds, interacting with a single rotor-router system. We propose new techniques which allow us to perform a theoretical analysis of the multi-agent rotor-router model, and to compare it to the scenario of parallel independent random walks in a graph. Our main results concern the n-node ring, and suggest a strong similarity between the performance characteristics of this deterministic model and random walks.

We show that on the ring the rotor-router with k agents admits a cover time of between $\Theta(n^2/k^2)$ in the best case and $\Theta(n^2/\log k)$ in the worst case, depending on the initial locations of the agents, and that both these bounds are tight. The corresponding expected value of cover time for k random walks, depending on the initial locations of the walkers, is proven to belong to a similar range, namely between $\Theta(n^2/(k^2/\log^2 k))$ and $\Theta(n^2/\log k)$.

Finally, we study the limit behavior of the rotor-router system. We show that, once the rotor-router system has stabilized, all the nodes of the ring are always visited by some agent every $\Theta(n/k)$ steps, regardless of how the system was initialized. This asymptotic bound corresponds to the expected time between successive visits to a node in the case of k random walks. All our results hold up to a polynomially large number of agents ($1 \leq k < n^{1/11}$).

6.3.7. *Efficient Exploration of Anonymous Undirected Graphs*

Participant: Ralf Klasing.

In [41], we consider the problem of exploring an anonymous undirected graph using an oblivious robot. The studied exploration strategies are designed so that the next edge in the robot's walk is chosen using only local information. We present some current developments in the area. In particular, we focus on recent work on *equitable strategies* and on the *multi-agent rotor-router*.

6.3.8. Gathering radio messages in the path

Participant: Ralf Klasing.

In [19], we address the problem of gathering information in one node (sink) of a radio network where interference constraints are present: when a node transmits, it produces interference in an area bigger than the area in which its message can actually be received. The network is modeled by a graph; a node is able to transmit one unit of information to the set of vertices at distance at most dt in the graph, but when doing so it generates interferences that do not allow nodes at distance up to di ($di \geq dt$) to listen to other transmissions. We are interested in finding a gathering protocol, that is an ordered sequence of rounds (each round consists of non-interfering simultaneous transmissions) such that $w(u)$ messages are transmitted from any node u to a fixed node called the sink. Our aim is to find a gathering protocol with the minimum number of rounds (called *gathering time*). In [19], we focus on the specific case where the network is a path with the sink at an end vertex of the path and where the traffic is unitary ($w(u) = 1$ for all u); indeed this simple case appears to be already very difficult. We first give a new lower bound and a protocol with a gathering time that differ only by a constant independent of the length of the path. Then we present a method to construct incremental protocols. An incremental protocol for the path on $n + 1$ vertices is obtained from a protocol for n vertices by adding new rounds and new calls to some rounds but without changing the calls of the original rounds. We show that some of these incremental protocols are optimal for many values of dt and di (in particular when dt is prime). We conjecture that this incremental construction always gives optimal protocols. Finally, we derive an approximation algorithm when the sink is placed in an arbitrary vertex in the path.

6.3.9. Computing Without Communicating: Ring Exploration by Asynchronous Oblivious Robots

Participant: David Ilcinkas.

In [24], we consider the problem of exploring an anonymous unoriented ring by a team of k identical, oblivious, asynchronous mobile robots that can view the environment but cannot communicate. This weak scenario is standard when the spatial universe in which the robots operate is the two-dimensional plane, but (with one exception) has not been investigated before for networks. Our results imply that, although these weak capabilities of robots render the problem considerably more difficult, ring exploration by a small team of robots is still possible. We first show that, when k and n are not co-prime, the problem is not solvable in general, e.g., if k divides n there are initial placements of the robots for which gathering is impossible. We then prove that the problem is always solvable provided that n and k are co-prime, for $k \geq 17$, by giving an exploration algorithm that always terminates, starting from arbitrary initial configurations. Finally, we consider the minimum number $\rho(n)$ of robots that can explore a ring of size n. As a consequence of our positive result we show that $\rho(n)$ is $O(\log n)$. We additionally prove that $\Omega(\log n)$ robots are necessary for infinitely many n.

6.3.10. Worst-case optimal exploration of terrains with obstacles

Participant: David Ilcinkas.

A mobile robot represented by a point moving in the plane has to explore an unknown flat terrain with impassable obstacles. Both the terrain and the obstacles are modeled as arbitrary polygons. We consider two scenarios: the *unlimited vision*, when the robot situated at a point p of the terrain explores (sees) all points q of the terrain for which the segment pq belongs to the terrain, and the *limited vision*, when we require additionally that the distance between p and q is at most 1. All points of the terrain (except obstacles) have to be explored and the performance of an exploration algorithm, called its complexity, is measured by the length of the trajectory of the robot.

For unlimited vision we show in [21] an exploration algorithm with complexity $O(P + D\sqrt{k})$, where P is the total perimeter of the terrain (including perimeters of obstacles), D is the diameter of the convex hull of the terrain, and k is the number of obstacles. We do not assume knowledge of these parameters. We also prove a matching lower bound showing that the above complexity is optimal, even if the terrain is known to the robot. For limited vision we show exploration algorithms with complexity $O(P + A + \sqrt{Ak})$, where A is the area of the terrain (excluding obstacles). Our algorithms work either for arbitrary terrains (if one of the parameters A or k is known) or for c-fat terrains, where c is any constant (unknown to the robot) and no additional knowledge is assumed. (A terrain $\mathcal{T}$ with obstacles is c-fat if $R/r \leq c$, where R is the radius of the smallest disc containing $\mathcal{T}$ and r is the radius of the largest disc contained in $\mathcal{T}$.) We also prove a matching lower bound $\Omega(P + A + \sqrt{Ak})$ on the complexity of exploration for limited vision, even if the terrain is known to the robot.

6.3.11. Exploration of the T-Interval-Connected Dynamic Graphs: the Case of the Ring

Participants: David Ilcinkas, Ahmed Wade.

In [40], we study the T-interval-connected dynamic graphs from the point of view of the time necessary and sufficient for their exploration by a mobile entity (agent). A dynamic graph (more precisely, an evolving graph) is T-interval-connected ($T \geq 1$) if, for every window of T consecutive time steps, there exists a connected spanning subgraph that is stable (always present) during this period. This property of connection stability over time was introduced by Kuhn, Lynch and Oshman (STOC 2010). We focus on the case when the underlying graph is a ring of size n, and we show that the worst-case time complexity for the exploration problem is $2n - T - \Theta(1)$ time units if the agent knows the dynamics of the graph, and $n + \frac{n}{\max\{1,T-1\}}(\delta - 1) \pm \Theta(\delta)$ time units otherwise, where δ is the maximum time between two successive appearances of an edge.

6.3.12. Time vs. space trade-offs for rendezvous in trees

Participant: Adrian Kosowski.

In [23], we consider the rendezvous problem, in which two identical (anonymous) mobile agents start from arbitrary nodes of an unknown tree and have to meet at some node. Agents move in synchronous rounds: in each round an agent can either stay at the current node or move to one of its neighbors. We consider deterministic algorithms for this rendezvous task. We obtain a tight trade-off between the optimal time of completing rendezvous and the size of memory of the agents. For agents with k memory bits, we show that optimal rendezvous time is $\Theta(n + n^2/k)$ in n-node trees. More precisely, if $k \geq c \log n$, for some constant c, we design agents accomplishing rendezvous in arbitrary trees of size n (unknown to the agents) in time $O(n + n^2/k)$, starting with arbitrary delay. We also show that no pair of agents can accomplish rendezvous in time $o(n + n^2/k)$, even in the class of lines of known length and even with simultaneous start. Finally, we prove that at least logarithmic memory is necessary for rendezvous, even for agents starting simultaneously in a n-node line.

7. Partnerships and Cooperations

7.1. National Initiatives

- **ANR ALADDIN** (Algorithm Design and Analysis for Implicitly and Incompletely Defined Interaction Networks; GANG and CEPAGE project-teams): the members of Cepage have been participating to the ANR project "blanc" (i.e. fundamental research) about the fundamental aspects of large interaction networks enabling massive distributed storage, efficient decentralized information retrieval, quick inter-user exchanges, and/or rapid information dissemination. The project is mostly oriented towards the design and analysis of algorithms for these (logical) networks, by taking into account proper ties inherent to the underlying infrastructures upon which they are built. The infrastructures and/or overlays considered in this project are selected from different contexts, including communication networks (from Internet to sensor networks), and societal networks (from the Web to P2P networks).

- **ANR SONGS** (Simulation of Next Generation Systems; participants: AlGorille (LORIA, Nancy), MESCAL (Grenoble), GRAAL (ENS Lyon), IN2P3 (Lyon), CEPAGE, HiePACS, RUNTIME (Bordeaux), LSIIT (Strasbourg), ASCOLA (Nantes), MASCOTTE, MODALIS (Sophia Antipolis)). This project started in 2012 as a follow-up of the USS-SIMGRID project. The aim is to further extend the domain of SimGrid, by designing a unified simulation framework for the four application domains: Grids, Peer-to-Peer systems, High Performance Computing, and Cloud systems. Achieving this goal mandates careful representation and modeling of the underlying concepts presented by each domain (memory, disks, energy, network and volatility) and of the interfaces specific to each domain. It also requires a transversal work on the simulation framework itself. CEPAGE is actively involved in this project, both for the peer-to-peer use cases and for the coordination of the modeling effort of the project.
- **ANR Displexity** (Calcul DIStribué: calculabilité et comPLEXITé; participants: CEPAGE, GANG and ASAP projects). The main goal of DISPLEXITY is to establish the scientific foundations of a theory of calculability and complexity for distributed computing. Displexity started in 2012.
- **ANR IDEA** ANR program “defis”: project IDEA (2009-2012). The goal of this ANR is the study of identifying codes in evolving graphs. Ralf Klasing is the overall leader of the project.
- **ANR “Jeunes chercheurs” EGOS - Embedded Graphs and their Oriented Structures** (2012-2014) (see http://www.lirmm.fr/egos/)

  Participants: CEPAGE/LaBRI(Bordeaux) LIRMM(Montpellier), LIX(Palaiseau) The goal of this project is the study oriented structures on graphs of arbitrary genus.
- **AMADEUS** (CNRS funding on “BIG DATA”: 2012-): Analysis of MAssive Data in Earth and Universe Sciences. This a multidisciplinary research project between computer science teams (LIRMM: University of Montpellier, LIF: University of Marseille) and CEPAGE), earth and climate science (CEREGE: Montpellier and IRD: Aix) and astronomy (LAM: University of Marseille). The aim of the project is to propose effective techniques for mining large data by essentially using distributed computing, visualization, summarization and approximation.

7.2. European Initiatives

7.2.1. FP7 Projects

7.2.1.1. EULER

EULER

- Title: EULER (Experimental UpdateLess Evolutive Routing)
- Type: COOPERATION (ICT)
- Defi: Future Internet Experimental Facility and Experimentally-driven Research
- Instrument: Specific Targeted Research Project (STREP)
- Duration: October 2010 - September 2013
- Coordinator: ALCATEL-LUCENT (Belgium)
- Others partners:
  Alcatel-Lucent Bell, Antwerpen, Belgium
  3 projects from Inria: CEPAGE, GANG and MASCOTTE, France
  Interdisciplinary Institute for Broadband Technology (IBBT),Belgium
  Laboratoire d’Informatique de Paris 6 (LIP6), UniversitÃÂ© Pierre Marie Curie (UPMC), France
  Department of Mathematical Engineering (INMA) UniversitÃÂ© Catholique de Louvain, Belgium
  RACTI, Research Academic Computer Technology Institute University of Patras, Greece

CAT, Catalan Consortium: Universitat Politecnica de Catalunya, Barcelona and University of Girona, Spain

- See also: http://www-sop.inria.fr/mascotte/EULER/wiki/
- Abstract: The title of this study is "Dynamic Compact Routing Scheme". The aim of this projet is to develop new routing schemes achieving better performances than current BGP protocols. The problems faced by the inter-domain routing protocol of the Internet are numerous:

  The underlying network is dynamic: many observations of bad configurations show the instability of BGP;

  BGP does not scale well: the convergence time toward a legal configuration is too long, the size of routing tables is proportional to the number of nodes of network (the network size is multiplied by 1.25 each year);

  The impact of the policies is so important that the many packets can oscillated between two Autonomous Systems.

  In this collaboration, we mainly focus on the scalability properties that a new routing protocol should guarantee. The main measures are the size of the local routing tables, and the time (or message complexity) to update or to generate such tables. The design of schemes achieving sub-linear space per routers, say in n where n is the number of AS routers, is the main challenge. The target networks are AS-network like with more than 100,000 nodes. This projet, in collaboration with the MASCOTE Inria-project in Nice Sophia-Antipolis, makes the use of simulation, developed at both sites.

7.2.2. Collaborations in European Programs, except FP7

- Program: European COST
- Project acronym: Complex HPC IC0805.
- Project title: Open Network for High-Performance Computing on Complex Environments
- Duration: 2010-2013
- Coordinator: Inria
- Other partners: 26 countries, see list at http://www.cost.eu/domains_actions/ict/Actions/IC0805?parties
- Abstract: The main objective of this COST action is to coordinate European groups working on the use of heterogeneous and hierarchical systems for HPC as well as the development of collaborative activities among the involved research groups (http://complexhpc.org/index.php).

7.3. International Initiatives

7.3.1. Inria International Partners

- **Royal Society Grant with the University of Liverpool.** International Joint Project, 2011-2013, entitled “SEarch, RENdezvous and Explore (SERENE)”, on foundations of mobile agent computing, in collaboration with the Department of Computer Science, University of Liverpool. Funded by the Royal Society, U.K. Principal investigator on the UK side: Leszek Gasieniec. Ralf Klasing is the principal investigator on the French side.

  Participants: Nicolas Hanusse, David Ilcinkas, Ralf Klasing, Adrian Kosowski.
- **Spanish program CLOUDS**: Cloud Computing for Scalable, Reliable and Ubiquitous Services (http://lsd.ls.fi.upm.es/clouds). This is a large scale program which aims at advancing research in the area of Cloud Computing. CEPAGE is more particularly in contact with the LaDyr team of Univ. Rey Juan Carlos in Madrid, on the topic of resource allocation problems for Cloud providers.

  Participants: Olivier Beaumont, Lionel Eyraud-Dubois.

- **Collaboration with Canada.**

  Members of CEPAGE have a long-standing collaboration with researchers from the Chair of Distributed Computing at the University of Quebec in Outaouais and the Department of Computer Science at Carleton University. Sources of financing include: personal NSERC grants of Canadian professors (Prof. Andrzej Pelc, Prof. Jurek Czyzowicz, Prof. Evangelos Kranakis), funding from other Canadian grant agencies (a travel grant from Mitacs Inc.), and University of Bordeaux funding (a 3-month invited professorship for Prof. Jurek Czyzowicz).

  Participants: David Ilcinkas, Ralf Klasing, Adrian Kosowski.

- **Collaboration with Chile.**

  Adrian Kosowski is a foreign partner of the Chilean ministry grant (ANILLO CONICYT programme) entitled "Mathematical modeling for industrial and management science applications: a multidisciplinary approach". The Project Director is Eric Goles from Universidad Adolfo Ibañez, and collaborating researchers on the Chilean side include Karol Suchan and Ivan Rappaport. The collaboration has led to 2 joint papers.

  Participants: Adrian Kosowski.

7.4. International Research Visitors

7.4.1. Visits of International Scientists

Tomasz Radzik, King's College London, UK, 02/12-06/12/2013

Miroslaw Korzeniowski, TU Wroclaw, Poland, 09/13-10/2013

Petra Berenbrink, Simon Fraser University, Burnaby, Canada, 22/10-26/10/2013

Joseph G. Peters, Simon Fraser University, Burnaby, Canada (Invited professor Bdx1) 24/01-24/02/2013

Carlos Ordonez, the University of Houston, USA (06-07/2013) supported by CNRS.

Dariusz Dereniowski, Gdansk University of Technology, Poland, 26/04-31/05/2013

Lukasz Kuszner, Gdansk University of Technology, Poland, 24/04-02/06/2013

Marcin Markiewicz, University of Gdansk, Poland, 02/09-15/09/2013

Leszek Gasieniec, University of Liverpool, UK, 24/09-27/09/2013

Jakub Lacki, University of Warsaw, Poland, 25/11-30/11/2013

Przemyslaw Uznanski, Universite de Marseille, France, 25/11-30/11/2013

7.4.1.1. Internships

Siddharth Mandal

Subject: Reliability Issues in Cloud Computing

Date: from May 2013 until Jul 2013

Institution: IIT Delhi (India)

Rohit Kumar

Subject: Robust Dynamic Allocation in Cloud Computing

Date: from May 2013 until Aug 2013

Institution: IIT Delhi (India)

8. Dissemination

8.1. Scientific Animation

8.1.1. Editorial Work

- Ralf Klasing is Associate Editor for
  - Algorithmic Operations Research (since May 2007),
  - Parallel Processing Letters (since August 2007),
  - Networks (since September 2007),
  - Computing and Informatics (since January 2008),
  - Theoretical Computer Science (since December 2009),
  - Fundamenta Informaticae (since January 2010),
  - Discrete Applied Mathematics (since February 2010),
  - Wireless Networks (since May 2010),
  - Journal of Interconnection Networks (since November 2010).
- Olivier Beaumont is Associate Editor for
  - IEEE Transactions on Parallel and Distributed Systems (since June 2010).

8.1.2. Steering Committees

- Ralf Klasing is a member of the Steering Commitee of the *International Colloquium on Structural Information and Communication Complexity (SIROCCO).*

8.1.3. Program Committees

- Olivier Beaumont: ICPP 2013 (Chair, Algorithm Track), IPDPS 2013, HCW 2013, ISCIS 2013, IPDPS 2014, EuroPar 2014.
- Ralf Klasing: SSS 2013 (Co Chair, Ad-hoc, sensor, robot and opportunistic networks Track) WALCOM 2015, IWOCA 2014,ADHOC-NOW 2014,ALGOSENSORS 2013, ADHOC-NOW 2013, FCT 2013, IWOCA 2013,vSEA 2013.
- David Ilcinkas: SIROCCO 2013, ADHOC-NOW 2013, AlgoTel 2013.
- Lionel Eyraud-Dubois: IPDPS 2013, HiPC 2013, IPDPS 2014, HiPC 2014.
- Adrian Kosowski: COCOA 2013, SSS 2013, DISC 2013, ADHOC-NOW 2013, FCT 2015 (PC co-chair).
- Sofian Maabout: DaWAK 2013, DOLAP 2013, BDA 2013, EDA 2013, DaWAK 2014. PC co-chair of MEDI 2013.
- Cyril Gavoille: WG 2014, PODC 2014, IPDPS 2014, PODC 2013, ICALP 2013, IPDPS 2013, SODA 2013.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Master : Communication and Routing (last year of engineering school ENSEIRB, 2012) O. Beaumont, N. Bonichon, L. Eyraud, N. Hanusse, R. Klasing, A. Kosowski (16h)

Master : Communication Algorithms in Networks (2nd year MASTER "Algorithms and Formal Methods", University of Bordeaux, 2012) R. Klasing (24h)

Master : Search Engines (2nd year of engineering school ENSEIRB, 2012) O. Beaumont (20h)

Master : Distributed Computing (2nd year MASTER "Réseaux, Systèmes et Mobilité"), (24h) C. Gavoille

8.2.2. Supervision

HdR : Nicolas Bonichon, Université de Bordeaux 1, 2013

HdR : Adrian Kosowski, Université de Bordeaux 1, 2013

PhD : Ahmed Wade, Université de Bordeaux, 2014

PhD : Przemyslaw Uznanski, Université de Bordeaux 1, 2013

PhD : Pierre Halftermeyer, Université de Bordeaux 1, 2013

PhD : Quentin Godfroy , Université de Bordeaux 1, 2013

PhD : Christian Glacet , Université de Bordeaux 1, 2013

9. Bibliography

Major publications by the team in recent years

[1] O. BEAUMONT, H. REJEB. *On the Importance of Bandwidth Control Mechanisms for Scheduling on Large Scale Heterogeneous Platforms*, in "24th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2010)", Atlanta, United States, April 2010, http://hal.inria.fr/inria-00444585/en

[2] N. BONICHON, C. GAVOILLE, N. HANUSSE, L. PERKOVIC. *Plane Spanners of Maximum Degree Six*, in "Automata, Languages and Programming 37th International Colloquium, ICALP 2010, Bordeaux, France, July 6-10, 2010, Proceedings, Part I", France, 2010, vol. 6198, pp. 19-30 [*DOI :* 10.1007/978-3-642-14165-2_3], http://hal.archives-ouvertes.fr/hal-00534212/en/

[3] C. COOPER, D. ILCINKAS, R. KLASING, A. KOSOWSKI. *Derandomizing random walks in undirected graphs using locally fair exploration strategies*, in "Distributed Computing", 2011, vol. 24, n° 2, pp. 91-99 [*DOI :* 10.1007/s00446-011-0138-4], http://hal.archives-ouvertes.fr/hal-00638229/en/

[4] J. CZYZOWICZ, A. KOSOWSKI, A. PELC. *How to meet when you forget: log-space rendezvous in arbitrary graphs*, in "29th Annual ACM Symposium on Principles of Distributed Computing (PODC)", ACM Press, July 2010, pp. 450-459

[5] P. DUCHON, N. HANUSSE, E. LEBHAR. *Towards small world emergence*, in "Proceedings of the eighteenth annual ACM symposium on Parallelism in algorithms and architectures", ACM Press New York, NY, USA, 2006, pp. 225–232

[6] P. DUCHON, N. HANUSSE, N. SAHEB, A. ZEMMARI. *Broadcast in the rendezvous model*, in "Information and Computation", 2006, vol. 204, n° 5, pp. 697–712

[7] L. EYRAUD-DUBOIS, G. MOUNIÉ, D. TRYSTRAM. *Analysis of Scheduling Algorithms with Reservations*, in "IEEE International Parallel and Distributed Processing Symposium, IPDPS 2007", Long Beach, California, USA, IEEE, 2007, pp. 1-8, http://dx.doi.org/10.1109/IPDPS.2007.370304

[8] P. FLOCCHINI, D. ILCINKAS, N. SANTORO. *Ping Pong in Dangerous Graphs: Optimal Black Hole Search with Pure Tokens*, in "Proceedings of the 22nd International Symposium on Distributed Computing", France, G. TAUBENFELD (editor), Lecture Notes in Computer Science, Springer Berlin / Heidelberg, September 2008, vol. 5218, pp. 227-241 [*DOI :* 10.1007/978-3-540-87779-0_16], http://hal.archives-ouvertes.fr/hal-00341523/en/

[9] P. FRAIGNIAUD, C. GAVOILLE, D. ILCINKAS, A. PELC. *Distributed Computing with Advice: Information Sensitivity of Graph Coloring*, in "ICALP", 2007, pp. 231-242

[10] P. FRAIGNIAUD, C. GAVOILLE, A. KOSOWSKI, E. LEBHAR, Z. LOTKER. *Universal augmentation schemes for network navigability: overcoming the $\sqrt{n}$-barrier*, in "Proceedings of the nineteenth annual ACM symposium on parallelism and architectures", San Diego, California, États-Unis, ACM, 2007, pp. 1-7, COST Action 295 "Dynamo", ACI project "Fragile", Inria project "Grand Large", http://hal.archives-ouvertes.fr/hal-00155186/en/

[11] E. GARNAUD, S. MAABOUT, M. MOSBAH. *Using Functional Dependencies for Reducing the Size of a Data Cube*, in "Proceedings of the 7th Int. Symp. on Foundations of Information and Knowledge Systems (FoIKS)", LNCS, Springer, 2012, 20 p.

[12] L. GASIENIEC, R. KLASING, M. RUSSELL, A. NAVARRA, X. ZHANG. *Fast Periodic Graph Exploration with Constant Memory*, in "Journal of Computer and System Sciences", 2008, vol. 74, n^{o} 5, pp. 808–822 [*DOI :* 10.1016/J.JCSS.2007.09.004], http://hal.archives-ouvertes.fr/hal-00306995/en/

[13] J. HROMKOVIČ, P. KANAREK, R. KLASING, K. LORYS, W. UNGER, H. WAGENER. *On the Size of Permutation Networks and Consequences for Efficient Simulation of Hypercube Algorithms on Bounded-Degree Networks*, in "SIAM Journal on Discrete Mathematics", 2009, vol. 23, n^{o} 3, pp. 1612–1645, http://hal.archives-ouvertes.fr/hal-00402764/en/

Publications of the year

Doctoral Dissertations and Habilitation Theses

[14] N. BONICHON. , *Quelques algorithmes entre le monde des graphes et les nuages de points*, Université Sciences et Technologies - Bordeaux I, April 2013, Habilitation à Diriger des Recherches, http://hal.inria.fr/tel-00922501

[15] A. KOSOWSKI. , *Time and Space-Efficient Algorithms for Mobile Agents in an Anonymous Network*, Université Sciences et Technologies - Bordeaux I, September 2013, Habilitation à Diriger des Recherches, http://hal.inria.fr/tel-00867765

[16] P. UZNANSKI. , *Large scale platform : Instantiable models and algorithmic design of communication schemes*, Université Sciences et Technologies - Bordeaux I, October 2013, http://hal.inria.fr/tel-00878837

Articles in International Peer-Reviewed Journals

[17] O. BEAUMONT, N. BONICHON, L. EYRAUD-DUBOIS, P. UZNANSKI. *Broadcasting on Large Scale Heterogeneous Platforms under the Bounded Multi-Port Model*, in "IEEE Transactions on Parallel and Distributed Systems", 2013, http://hal.inria.fr/hal-00861830

[18] A. BENOIT, R. MELHEM, P. RENAUD-GOUD, Y. ROBERT. *Assessing the performance of energy-aware mappings*, in "Parallel Processing Letters", 2013, vol. 23, n^{o} 2 [*DOI :* 10.1142/S0129626413400033], http://hal.inria.fr/hal-00926105

[19] J.-C. BERMOND, R. KLASING, N. MORALES, S. PÉRENNES, P. REYES. *Gathering radio messages in the path*, in "Discrete Mathematics, Algorithms and Applications", March 2013, vol. 5, n^{o} 1, pp. 1-28 [*DOI :* 10.1142/S1793830913500043], http://hal.inria.fr/hal-00907494

[20] O. BODINI, A. JACQUOT, P. DUCHON, L. R. MUTAFCHIEV. *Asymptotic Analysis and Random Sampling of Digitally Convex Polyominoes*, in "Discrete Applied Mathematics", June 2013, pp. 1-23, http://hal.inria.fr/hal-00832016

[21] J. CZYZOWICZ, D. ILCINKAS, A. LABOUREL, A. PELC. *Worst-case optimal exploration of terrains with obstacles*, in "Information and Computation", April 2013, vol. 225, pp. 16-28 [*DOI :* 10.1016/J.IC.2013.02.001], http://hal.inria.fr/hal-00863496

[22] J. CZYZOWICZ, A. KOSOWSKI, A. PELC. *Deterministic Rendezvous of Asynchronous Bounded-Memory Agents in Polygonal Terrains*, in "Theory of Computing Systems", 2013, vol. 52, n^{o} 2, pp. 179-199 [*DOI :* 10.1007/S00224-011-9379-7], http://hal.inria.fr/hal-00940803

[23] J. CZYZOWICZ, A. KOSOWSKI, A. PELC. *Time vs. space trade-offs for rendezvous in trees*, in "Distributed Computing", December 2013 [*DOI :* 10.1007/S00446-013-0201-4], http://hal.inria.fr/hal-00646912

[24] P. FLOCCHINI, D. ILCINKAS, A. PELC, N. SANTORO. *Computing Without Communicating: Ring Exploration by Asynchronous Oblivious Robots*, in "Algorithmica", March 2013, vol. 65, n^{o} 3, pp. 562-583 [*DOI :* 10.1007/S00453-011-9611-5], http://hal.inria.fr/hal-00912499

[25] E. GARNAUD, S. MAABOUT, M. MOSBAH. *Functional dependencies are helpful for partial materialization of data cubes*, in "Annals of Mathematics and Artificial Intelligence", August 2013, pp. 1-20 [*DOI :* 10.1007/S10472-013-9375-5], http://hal.inria.fr/hal-00940776

[26] A. KOSOWSKI, A. NAVARRA, D. PAJAK, C. PINOTTI. *Maximum matching in multi-interface networks*, in "Theoretical Computer Science", 2013, vol. 507, pp. 52-60 [*DOI :* 10.1016/J.TCS.2013.01.018], http://hal.inria.fr/hal-00940802

[27] C. ORDONEZ, S. MAABOUT, D. S. MATUSEVICH, W. CABRERA. *Extending ER models to capture database transformations to build data sets for data mining*, in "Data and Knowledge Engineering", December 2013, vol. 89, pp. 38-54 [*DOI :* 10.1016/J.DATAK.2013.11.002], http://hal.inria.fr/hal-00940778

International Conferences with Proceedings

[28] G. AUPY, A. BENOIT, R. MELHEM, P. RENAUD-GOUD, Y. ROBERT. *Energy-aware checkpointing of divisible tasks with soft or hard deadlines*, in "IGCC - 4th International Green Computing Conference - 2013", Arlington, United States, February 2013, http://hal.inria.fr/hal-00857244

[29] O. BEAUMONT, P. DUCHON, P. RENAUD-GOUD. *Approximation Algorithms for Energy Minimization in Cloud Service Allocation under Reliability Constraints*, in "HIgh Performance Computing", Bengalore, India, February 2013, 20 p. , http://hal.inria.fr/hal-00788964

[30] O. BEAUMONT, L. EYRAUD-DUBOIS, H. LARCHEVÊQUE. *Reliable Service Allocation in Clouds*, in "IPDPS - 27th IEEE International Parallel & Distributed Processing Symposium", Boston, United States, 2013 [*DOI :* 10.1109/IPDPS.2013.64], http://hal.inria.fr/hal-00743524

[31] O. BEAUMONT, L. EYRAUD-DUBOIS, P. PESNEAU, P. RENAUD-GOUD. *Reliable Service Allocation in Clouds with Memory and Capacity Constraints*, in "Resilience 2013, in conjunction with EuroPar 2013", Aachen, Germany, September 2013, http://hal.inria.fr/hal-00850125

[32] O. BEAUMONT, H. LARCHEVÊQUE, L. MARCHAL. *Non Linear Divisible Loads: There is No Free Lunch*, in "IPDPS 2013, 27th IEEE International Parallel & Distributed Processing Symposium", Boston, United States, IEEE, 2013, http://hal.inria.fr/hal-00771640

[33] A. COLLINS, J. CZYZOWICZ, L. GASIENIEC, A. KOSOWSKI, E. KRANAKIS, D. KRIZANC, R. MARTIN, O. MORALES PONCE. *Optimal Patrolling of Fragmented Boundaries*, in "SPAA - 25th Symposium on Parallelism in Algorithms and Architectures", Montreal, Canada, ACM, 2013, pp. 241-250 [*DOI :* 10.1145/2486159.2486176], http://hal.inria.fr/hal-00686270

[34] J. CZYZOWICZ, D. DERENIOWSKI, L. GASIENIEC, R. KLASING, A. KOSOWSKI, D. PAJAK. *Collision-Free Network Exploration*, in "Proceedings of the 11th Latin American Theoretical Informatics Symposium (LATIN 2014)", Uruguay, Lecture Notes in Computer Science, March 2014, vol. 8392, pp. 342-354, http://hal.inria.fr/hal-00927151

[35] D. DERENIOWSKI, Y. DISSER, A. KOSOWSKI, D. PAJAK, P. UZNANSKI. *Fast Collaborative Graph Exploration*, in "ICALP - 40th International Colloquium on Automata, Languages and Programming", Riga, Latvia, Lecture Notes in Computer Science, Springer, 2013, vol. 7966, pp. 520-532 [*DOI :* 10.1007/978-3-642-39212-2_46], http://hal.inria.fr/hal-00802308

[36] D. DERENIOWSKI, Y. DISSER, A. KOSOWSKI, D. PAJAK, P. UZNANSKI. *Fast Collaborative Graph Exploration*, in "15èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel)", France, N. NISSE, F. ROUSSEAU, Y. BUSNEL (editors), May 2013, pp. 1-4, http://hal.inria.fr/hal-00818451

[37] L. EYRAUD-DUBOIS, H. LARCHEVÊQUE. *Optimizing Resource allocation while handling SLA violations in Cloud Computing platforms*, in "IPDPS - 27th IEEE International Parallel & Distributed Processing Symposium", Boston, United States, 2013 [*DOI :* 10.1109/IPDPS.2013.67], http://hal.inria.fr/hal-00772846

[38] C. GAVOILLE, C. GLACET, N. HANUSSE, D. ILCINKAS. *On the Communication Complexity of Distributed Name-Independent Routing Schemes*, in "DISC 2013", Jérusalem, Israel, LNCS, 2013, vol. 8205, pp. 418-432, http://hal.inria.fr/hal-00851217

[39] D. ILCINKAS, A. WADE. *Exploration des graphes dynamiques T-intervalle-connexes : le cas de l'anneau*, in "15èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel)", Pornic, France, N. NISSE, F. ROUSSEAU, Y. BUSNEL (editors), 2013, pp. 73-76, http://hal.inria.fr/hal-00818450

[40] D. ILCINKAS, A. WADE. *Exploration of the T-Interval-Connected Dynamic Graphs: the Case of the Ring*, in "SIROCCO 2013", Ischia, Italy, LNCS, 2013, to appear, http://hal.inria.fr/hal-00847771

[41] R. KLASING. *Efficient Exploration of Anonymous Undirected Graphs*, in "Proceedings of the 24th International Workshop on Combinatorial Algorithms (IWOCA 2013)", Rouen, France, Lecture Notes in Computer Science, Springer, July 2013, vol. 8288, pp. 7-13 [*DOI :* 10.1007/978-3-642-45278-9_2], http://hal.inria.fr/hal-00927154

[42] R. KLASING, A. KOSOWSKI, D. PAJAK, T. SAUERWALD. *The multi-agent rotor-router on the ring: a deterministic alternative to parallel random walks*, in "Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC 2013)", Montréal, Canada, ACM, July 2013, pp. 365-374 [*DOI :* 10.1145/2484239.2484260], http://hal.inria.fr/hal-00927152

[43] A. KOSOWSKI. *A $tildeO(n2)$ Time-Space Trade-off for Undirected s-t Connectivity*, in "SODA - 24th ACM-SIAM Symposium on Discrete Algorithms", New Orleans, United States, SIAM, January 2013, pp. 1873-1883, Version 3 makes use of the Metropolis-Hastings walk, http://hal.inria.fr/hal-00685373

[44] A. KOSOWSKI, P. UZNANSKI. *Splittable Single Source-Sink Routing on CMP Grids: A Sublinear Number of Paths Suffice*, in "Euro-Par - 19th International Conference on Parallel Processing", Aachen, Germany, Lecture Notes in Computer Science, Springer, August 2013, vol. 8097, pp. 671-683 [*DOI :* 10.1007/978-3-642-40047-6_67], http://hal.inria.fr/hal-00737611

National Conferences with Proceedings

[45] L. EYRAUD-DUBOIS, A. LÈBRE, P. MARTINEAU, A. SOUKHAL, V. T'KINDT, D. TRYSTRAM. *A Server Consolidation Problem: Definition and Model*, in "Proceedings of the 14th conference ROADEF", Troyes, France, 2013, http://hal.inria.fr/hal-00921363

[46] E. GARNAUD, N. HANUSSE, S. MAABOUT, N. NOVELLI. *Calcul parallèle de dépendances*, in "Bases de Données Avancées", Nantes, France, October 2013, pp. 1-20, http://hal.inria.fr/hal-00940779

[47] C. GAVOILLE, C. GLACET, N. HANUSSE, D. ILCINKAS. *Algorithme distribué de routage compact en temps optimal*, in "15èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel)", Pornic, France, N. NISSE, F. ROUSSEAU, Y. BUSNEL (editors), May 2013, pp. 117-120, http://hal.inria.fr/hal-00818445

Scientific Books (or Scientific Book chapters)

[48] , *Proceedings of the third MEDI international conference*, Springer, October 2013, 291 p. , http://hal.inria.fr/hal-00940783

[49] A. BEN ZAKOUR, S. MAABOUT, M. MOSBAH, M. SISTIAGA. *Relaxing Time Granularity for Mining Frequent Sequences*, in "Advances in Knowledge and Data Mining", Springer, December 2013, pp. 53-76, http://hal.inria.fr/hal-00940782

Other Publications

[50] O. BEAUMONT, L. EYRAUD-DUBOIS, P. RENAUD-GOUD. , *Efficient and Robust Allocation Algorithms in Clouds under Memory Constraints*, 2013, http://hal.inria.fr/hal-00874936

[51] O. BEAUMONT, L. MARCHAL. , *What Makes Affinity-Based Schedulers So Efficient ?*, 2013, http://hal.inria.fr/hal-00875487

[52] D. DERENIOWSKI, A. KOSOWSKI, D. PAJAK. , *Distinguishing Views in Symmetric Networks: A Tight Lower Bound*, October 2013, http://hal.inria.fr/hal-00875370

[53] D. DERENIOWSKI, A. KOSOWSKI, D. PAJAK, P. UZNANSKI. , *Bounds on the Cover Time of Parallel Rotor Walks*, 2013, http://hal.inria.fr/hal-00865065

[54] L. EYRAUD-DUBOIS, P. UZNANSKI. , *Point-to-point and congestion bandwidth estimation: experimental evaluation on PlanetLab*, March 2013, http://hal.inria.fr/hal-00802307

[55] A. Kosowski, M. Markiewicz. , *Lower Bounds on the Communication Complexity of Binary Local Quantum Measurement Simulation*, October 2013, http://hal.inria.fr/hal-00871120

[56] A. Kosowski, D. Pajak. , *Does Adding More Agents Make a Difference? A Case Study of Cover Time for the Rotor-Router*, February 2014, http://hal.inria.fr/hal-00950743

References in notes

[57] I. Abraham, S. Chechik, C. Gavoille. *Fully dynamic approximate distance oracles for planar graphs via forbidden-set distance labels*, in "44th Annual ACM Symposium on Theory of Computing (STOC)", New-York, États-Unis, ACM (editor), 2012, pp. 1199-1217, http://hal.archives-ouvertes.fr/hal-00725839

[58] I. Abraham, S. Chechik, C. Gavoille, D. Peleg. *Forbidden-Set Distance Labels for Graphs of Bounded Doubling Dimension*, in "29th Annual ACM Symposium on Principles of Distributed Computing (PODC)", ACM Press, July 2010, pp. 192-200 [*DOI :* 10.1145/1835698.1835743]

[59] I. Atsonios, O. Beaumont, N. Hanusse, Y. Kim. *On Power-Law Distributed Balls in Bins and its Applications to View Size Estimation*, in "ISAAC", Yokohama, Japon, 2011, http://hal.inria.fr/inria-00618785

[60] E. Bampas, J. Czyzowicz, L. Gasieniec, D. Ilcinkas, A. Labourel. *Almost Optimal Asynchronous Rendezvous in Infinite Multidimensional Grids*, in "DISC 2010", États-Unis, Springer Berlin / Heidelberg, 2010, vol. 6343, pp. 297–311, http://hal.archives-ouvertes.fr/hal-00534263

[61] E. Bampas, L. Gasieniec, N. Hanusse, D. Ilcinkas, R. Klasing, A. Kosowski. *Euler Tour Lock-in Problem in the Rotor-Router Model*, in "DISC 2009", Espagne, Springer Berlin / Heidelberg, 2009, vol. 5805, pp. 421–433, http://hal.archives-ouvertes.fr/hal-00402749

[62] E. Bampas, L. Gasieniec, R. Klasing, A. Kosowski, T. Radzik. *Robustness of the rotor-router mechanism*, in "OPODIS 2009", France, Springer Berlin / Heidelberg, 2009, vol. 5923, pp. 345–358, http://hal.archives-ouvertes.fr/hal-00417639

[63] C. Banino-Rokkones, O. Beaumont, H. Rejeb. *Scheduling Techniques for Effective System Reconfiguration in Distributed Storage Systems*, in "IEEE Intl Conference on Parallel and Distributed Systems Conferences", Melbourne, Australie, 2008, http://hal.inria.fr/inria-00336242

[64] O. Beaumont, N. Bonichon, L. Courtès, X. Hanin, E. Dolstra. *Mixed Data-Parallel Scheduling for Distributed Continuous Integration*, in "Heterogeneity in Computing Workshop, in IPDPS 2012", Shangaï, Chine, IEEE (editor), 2012, http://hal.inria.fr/hal-00684220

[65] O. Beaumont, N. Bonichon, P. Duchon, H. Larchevêque. *Distributed Approximation Algorithm for Resource Clustering*, in "SIROCCO 2008", Suisse, 2008, pp. 61–73, http://hal.archives-ouvertes.fr/hal-00323068

[66] O. Beaumont, N. Bonichon, P. Duchon, H. Larchevêque. *Use of Internet Embedding Tools for Heterogeneous Resources Aggregation*, in "Heterogeneity in Computing Workshop", Anchorage, États-Unis, 2011, 10 p. , http://hal.inria.fr/inria-00562655

[67] O. Beaumont, N. Bonichon, P. Duchon, H. Larchevêque. *Utilisation d'outils de plongement d'Internet pour l'agrégation de ressources hétérogènes*, in "Conférence : 13es Rencontres Francophones sur

les Aspects Algorithmiques de Télécommunications (AlgoTel)", Cap Estérel, France, B. DUCOURTHIAL, P. FELBER (editors), 2011, http://hal.inria.fr/inria-00585254

[68] O. BEAUMONT, N. BONICHON, L. EYRAUD-DUBOIS. *Scheduling divisibleworkloads on heterogeneous platforms under bounded multi-port model*, in "Heterogeneity in Computing Workshop, in IEEE International Symposium on Parallel and Distributed Processing, IPDPS 2008", Miami, FL, États-Unis, 2008, http://hal.inria.fr/inria-00336189

[69] O. BEAUMONT, N. BONICHON, L. EYRAUD-DUBOIS, L. MARCHAL. *Minimizing Weighted Mean Completion Time for Malleable Tasks Scheduling*, in "IPDPS 2012, 26th IEEE International Parallel & Distributed Processing Symposium", Shangai, Chine, IEEE (editor), 2012, http://hal.inria.fr/inria-00564056

[70] O. BEAUMONT, N. BONICHON, L. EYRAUD-DUBOIS, P. UZNANSKI. *Broadcasting on Large Scale Heterogeneous Platforms with connectivity artifacts under the Bounded Multi-Port Model*, in "ICPADS 2011 - 17th Internation Conference on Parallel and Distributed Systems", Taïnan, Taïwan, Province De Chine, IEEE (editor), 2011, vol. Proceeding of ICPADS 2011, pp. 173-180, http://hal.inria.fr/inria-00618784

[71] O. BEAUMONT, N. BONICHON, H. LARCHEVÊQUE. *Modeling and Practical Evaluation of a Service Location Problem in Large Scale Networks*, in "Internation Conference on Parallel Computing 2011", Taipei, Taïwan, Province De Chine, 2011, 10 p. , http://hal.inria.fr/inria-00580969

[72] O. BEAUMONT, P. DUCHON, M. KORZENIOWSKI. *Heterogenous dating service with application to rumor spreading*, in "IEEE International Symposium on Parallel and Distributed Processing, IPDPS 2008", Miami, FL, États-Unis, IEEE, 2008, pp. 1–10, http://hal.inria.fr/inria-00142778

[73] O. BEAUMONT, L. EYRAUD-DUBOIS, A. KUMAR. *Broadcasting on Large Scale Heterogeneous Platforms under the Bounded Multi-Port Model*, in "24th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2010)", Atlanta, États-Unis, 2010, http://hal.inria.fr/inria-00444586

[74] O. BEAUMONT, L. EYRAUD-DUBOIS, H. REJEB, C. THRAVES. *Allocation of Clients to Multiple Servers on Large Scale Heterogeneous Platforms*, in "The Fifteenth International Conference on Parallel and Distributed Systems (ICPADS'09)", Shenzhen, Chine, 2009, http://hal.inria.fr/inria-00413437

[75] O. BEAUMONT, L. EYRAUD-DUBOIS, H. REJEB, C. THRAVES. *Online Allocation of Splitable Clients to Multiple Servers on Large Scale Heterogeneous Platforms*, in "AlgoTel", Carry-Le-Rouet, France, A. CHAINTREAU, C. MAGNIEN (editors), 2009, http://hal.inria.fr/inria-00384569

[76] O. BEAUMONT, L. EYRAUD-DUBOIS, H. REJEB, C. THRAVES. *On-line Allocation of Clients to Multiple Servers on Large Scale Heterogeneous Systems*, in "PDP 2010 - The 18th Euromicro International Conference on Parallel, Distributed and Network-Based Computing", Pisa, Italie, 2010, http://hal.inria.fr/inria-00444584

[77] O. BEAUMONT, L. EYRAUD-DUBOIS, J. WON. *Using the Last-mile Model as a Distributed Scheme for Available Bandwidth Prediction*, in "EuroPar 2011", Bordeaux, France, SPRINGER (editor), 2011, http://hal.inria.fr/inria-00588651

[78] O. BEAUMONT, H. REJEB. *On the Importance of Bandwidth Control Mechanisms for Scheduling on Large Scale Heterogeneous Platforms*, in "24th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2010)", Atlanta, États-Unis, 2010, http://hal.inria.fr/inria-00444585

[79] N. BONICHON, C. GAVOILLE, N. HANUSSE, D. ILCINKAS. *Connections between Theta-Graphs, Delaunay Triangulations, and Orthogonal Surfaces*, in "WG 2010", Grèce, Springer Berlin / Heidelberg, 2010, vol. 6410, pp. 266–278, http://hal.archives-ouvertes.fr/hal-00536710

[80] N. BONICHON, C. GAVOILLE, N. HANUSSE, L. PERKOVIC. *Plane Spanners of Maximum Degree Six*, in "ICALP", France, 2010, vol. 6198, pp. 19-30, http://hal.archives-ouvertes.fr/hal-00534212

[81] N. BONICHON, C. GAVOILLE, N. HANUSSE, L. PERKOVIĆ. *The Stretch Factor of L_1- and L_∞-Delaunay Triangulations*, in "20th Annual European Symposium on Algorithms (ESA)", L. EPSTEIN, P. FERRAGINA (editors), Springer-Verlag, September 2012, vol. 7501 of Lecture Notes in Computer Science, pp. 205-216

[82] J. CHALOPIN, S. DAS, A. KOSOWSKI. *Constructing a Map of an Anonymous Graph: Applications of Universal Sequences*, in "OPODIS - 14th International Conference On Principles Of DIstributed Systems - 2010", Tozeur, Tunisie, Springer, 2010, vol. 6490, pp. 119-134, http://hal.inria.fr/hal-00726071

[83] B. CHUN, D. CULLER, T. ROSCOE, A. BAVIER, L. PETERSON, M. WAWRZONIAK, M. BOWMAN. *PlanetLab: an overlay testbed for broad-coverage services*, in "ACM SIGCOMM Comput. Commun. Rev.", July 2003, vol. 33, pp. 3–12, http://doi.acm.org/10.1145/956993.956995

[84] R. COHEN, P. FRAIGNIAUD, D. ILCINKAS, A. KORMAN, D. PELEG. *Label-Guided Graph Exploration by a Finite Automaton*, in "ACM Transactions on Algorithms", 2008, vol. 4, n^{o} 4, Article 42, http://hal.archives-ouvertes.fr/hal-00341609

[85] A. COLLINS, J. CZYZOWICZ, L. GASIENIEC, A. KOSOWSKI, A. RUSSELL. *Synchronous Rendezvous for Location-Aware Agents*, in "DISC 2011 - 25th International Symposium on Distributed Computing", Rome, Italie, D. PELEG (editor), Springer, 2011, vol. 6950, pp. 447-459, http://hal.inria.fr/hal-00646907

[86] C. COOPER, D. ILCINKAS, R. KLASING, A. KOSOWSKI. *Derandomizing Random Walks in Undirected Graphs Using Locally Fair Exploration Strategies*, in "ICALP 2009", Grèce, Springer Berlin / Heidelberg, 2009, vol. 5556, pp. 411-422, http://hal.archives-ouvertes.fr/hal-00374071

[87] C. COOPER, R. KLASING, T. RADZIK. *Locating and repairing faults in a network with mobile agents*, in "Proceedings of the 15th Colloquium on Structural Information and Communication Complexity (SIROCCO 2008)", Suisse, 2008, vol. 5058, pp. 20–32, http://hal.archives-ouvertes.fr/hal-00307025

[88] J. CZYZOWICZ, S. DOBREV, L. GASIENIEC, D. ILCINKAS, J. JANSSON, R. KLASING, I. LIGNOS, R. MARTIN, K. SADAKANE, W.-K. SUNG. *More Efficient Periodic Traversal in Anonymous Undirected Graphs*, in "SIROCCO 2009", Slovénie, Springer Berlin / Heidelberg, 2009, vol. 5869, pp. 174–188, http://hal.archives-ouvertes.fr/hal-00371489

[89] J. CZYZOWICZ, L. GASIENIEC, A. KOSOWSKI, E. KRANAKIS. *Boundary Patrolling by Mobile Agents with Distinct Maximal Speeds*, in "ESA 2011 - 19th Annual European Symposium on Algorithms", Saarbruecken, Allemagne, C. DEMETRESCU, M. M. HALLDÓRSSON (editors), Springer, 2011, vol. 6942, pp. 701-712, http://hal.inria.fr/hal-00646910

[90] J. CZYZOWICZ, D. ILCINKAS, A. LABOUREL, A. PELC. *Asynchronous Deterministic Rendezvous in Bounded Terrains*, in "SIROCCO 2010", Şirince, Turquie, Springer Berlin / Heidelberg, 2010, vol. 6058, pp. 72-85, http://hal.archives-ouvertes.fr/hal-00516048

[91] J. CZYZOWICZ, D. ILCINKAS, A. LABOUREL, A. PELC. *Optimal Exploration of Terrains with Obstacles*, in "SWAT 2010", Bergen, Norvège, Springer Berlin / Heidelberg, 2010, vol. 6139, pp. 1-12, http://hal.archives-ouvertes.fr/hal-00516061

[92] J. CZYZOWICZ, A. KOSOWSKI, A. PELC. *How to meet when you forget: log-space rendezvous in arbitrary graphs*, in "Distributed Computing", 2012, vol. 25, n^{o} 2, pp. 165-178, http://hal.inria.fr/hal-00725875

[93] J. CZYZOWICZ, A. KOSOWSKI, A. PELC. *Time vs. space trade-offs for rendezvous in trees*, in "SPAA - 24th Symposium on Parallelism in Algorithms and Architectures - 2012", Pittsburgh, États-Unis, 2012, pp. 1-10, http://hal.inria.fr/hal-00725924

[94] F. DABEK, R. COX, F. KAASHOEK, R. MORRIS. *Vivaldi: a decentralized network coordinate system*, in "SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications", New York, NY, USA, ACM, 2004, pp. 15–26, http://doi.acm.org/10.1145/1015467.1015471

[95] E. DIOT, C. GAVOILLE. *On the Path Separability of Planar Graphs*, in "EuroComb 2009, Bordeaux, France", Bordeaux, France, 2009, vol. 34C, pp. 549-552, http://hal.archives-ouvertes.fr/hal-00408228

[96] P. DUCHON, H. LARCHEVÊQUE. *On the search path length of random binary skip graphs*, in "ANALCO 2010", Austin, États-Unis, SIAM (editor), SIAM, 2010, pp. 1-8, http://hal.inria.fr/inria-00547935

[97] L. EYRAUD-DUBOIS, P. UZNANSKI. *Bedibe: Datasets and Software Tools for Distributed Bandwidth Prediction*, in "AlgoTel'2012 - Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications", La Grande Motte, France, N. HANUSSE, F. MATHIEU (editors), 2012, http://hal.inria.fr/hal-00690861

[98] P. FLOCCHINI, D. ILCINKAS, A. PELC, N. SANTORO. *Remembering without Memory: Tree Exploration by Asynchronous Oblivious Robots*, in "SIROCCO 2008", Villars-sur-Ollon, Suisse, A. A. SHVARTSMAN, P. FELBER (editors), Springer Berlin / Heidelberg, 2008, vol. 5058, pp. 33-47, http://hal.archives-ouvertes.fr/hal-00341465

[99] P. FLOCCHINI, D. ILCINKAS, A. PELC, N. SANTORO. *Remembering Without Memory: Tree Exploration by Asynchronous Oblivious Robots*, in "Theoretical Computer Science", 2010, vol. 411, n^{o} 14-15, pp. 1583-1598, http://hal.archives-ouvertes.fr/hal-00464588

[100] P. FLOCCHINI, D. ILCINKAS, N. SANTORO. *Ping Pong in Dangerous Graphs: Optimal Black Hole Search with Pure Tokens*, in "DISC 2008", France, G. TAUBENFELD (editor), Springer Berlin / Heidelberg, 2008, vol. 5218, pp. 227-241, http://hal.archives-ouvertes.fr/hal-00341523

[101] C. GAVOILLE, S. CHRISITAN. *Sparse Spanners vs. Compact Routing*, in "23rd Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA)", San Jose, États-Unis, ACM, 2011, pp. 225-234, http://hal.archives-ouvertes.fr/hal-00651836

[102] C. GAVOILLE, Q. GODFROY, L. VIENNOT. *Node-Disjoint Multipath Spanners and their Relationship with Fault-Tolerant Spanners*, in "OPODIS'11 - 15th International Conference on Principles of Distributed Systems", Toulouse, France, A. F. ANTA, G. LIPARI, M. ROY (editors), Springer, 2011, vol. 7109, pp. 143-158, http://hal.archives-ouvertes.fr/hal-00651825

[103] C. GAVOILLE, D. PELEG, S. PÉRENNES, R. RAZ. *Distance labeling in graphs*, in "Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms", Society for Industrial and Applied Mathematics, 2001, 219 p.

[104] N. HANUSSE, D. ILCINKAS, A. KOSOWSKI, N. NISSE. *Locating a Target with an Agent Guided by Unreliable Local Advice*, in "PODC 2010", Zurich, Suisse, ACM New York, NY, USA, 2010, pp. 355-364, http://hal.archives-ouvertes.fr/hal-00516695

[105] N. HANUSSE, D. KAVVADIAS, E. KRANAKIS, D. KRIZANC. *Memoryless search algorithms in a network with faulty advice*, in "Theoretical Computer Science", 2008, vol. 402, n^{o} 2-3, pp. 190-198, http://hal.archives-ouvertes.fr/hal-00341468

[106] N. HANUSSE, S. MAABOUT. *A parallel Algorithm for Computing Borders*, in "International Conference on Information and Knowledge Managment (CIKM)", Glasgow, Royaume-Uni, ACM, 2011, pp. 1639-1648, http://hal.archives-ouvertes.fr/hal-00644517

[107] N. HANUSSE, S. MAABOUT, R. TOFAN. *A view selection algorithm with performance guarantee*, in "International Conference on Extending Database Technology", St Petersburg, Russie, Fédération De, 2009, http://hal.archives-ouvertes.fr/hal-00341475

[108] N. HANUSSE, S. MAABOUT, R. TOFAN. *Revisiting the Partial Data Cube Materialization*, in "Conference on Advances in Data Bases and Information Systems (ADBIS)", Vienne, Autriche, Springer, 2011, n^{o} 6909, pp. 70-83, http://hal.archives-ouvertes.fr/hal-00644526

[109] R. KLASING, A. KOSOWSKI, A. NAVARRA. *Taking Advantage of Symmetries: Gathering of many Asynchronous Oblivious Robots on a Ring*, in "Theoretical Computer Science", 2010, vol. 411, n^{o} 34–36, pp. 3235–3246, http://hal.archives-ouvertes.fr/hal-00519069

[110] R. KLASING, E. MARKOU, A. PELC. *Gathering asynchronous oblivious mobile robots in a ring*, in "Theoretical Computer Science", 2008, vol. 390, n^{o} 1, pp. 27–39, http://hal.archives-ouvertes.fr/hal-00307237

[111] A. KOSOWSKI, B. LI, N. NISSE, K. SUCHAN. *k-Chordal Graphs: from Cops and Robber to Compact Routing via Treewidth*, in "14èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel)", La Grande Motte, France, F. MATHIEU, N. HANUSSE (editors), 2012, http://hal.archives-ouvertes.fr/hal-00687120

[112] A. KOSOWSKI, A. NAVARRA. *Graph Decomposition for Memoryless Periodic Exploration*, in "Algorithmica", 2012, vol. 63, n^{o} 1-2, pp. 26-38, http://hal.inria.fr/hal-00725922

[113] A. KOSOWSKI, A. NAVARRA, C. PINOTTI. *Synchronous Black Hole Search in Directed Graphs*, in "Theoretical Computer Science", 2011, vol. 412, n^{o} 41, pp. 5752-5759, http://hal.inria.fr/inria-00614476

[114] A. LEGRAND, F. MAZOIT, M. QUINSON. , *An Application-level network mapper*, LIPENS Lyon, France, feb 2003, n^{o} RR-2003-09

[115] Y. MAO, L. SAUL, J. SMITH. *IDES: An Internet Distance Estimation Service for Large Networks*, in "IEEE JSAC", Dec. 2006, vol. 24, n^{o} 12, pp. 2273 -2284

[116] J.-F. MARCKERT, N. BONICHON. *Asymptotics of geometrical navigation on a random set of points of the plane*, in "Advances in Applied Probability", 2011, vol. 43, n^{o} 4, pp. 899-942, http://hal.inria.fr/hal-00649901

[117] V. RAMASUBRAMANIAN, D. MALKHI, F. KUHN, M. BALAKRISHNAN, A. GUPTA, A. AKELLA. *On the treeness of internet latency and bandwidth*, in "SIGMETRICS '09: Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems", New York, NY, USA, ACM, 2009, pp. 61–72

[118] H. WATANABE, H. NAKAZATO, A. GHATPANDE, O. BEAUMONT. *Divisible Load Scheduling with Result Collection on Heterogeneous Systems*, in "IEEE International Symposium on Parallel and Distributed Processing, IPDPS 2008", Miami, FL, États-Unis, 2008, http://hal.inria.fr/inria-00336186