



IN PARTNERSHIP WITH:
**Institut national des sciences
appliquées de Rennes**
Université Rennes 1

Activity Report 2016

Project-Team ASAP

**As Scalable As Possible: Foundations of large
scale dynamic distributed systems**

IN COLLABORATION WITH: Institut de recherche en informatique et systèmes aléatoires (IRISA)

RESEARCH CENTER
Rennes - Bretagne-Atlantique

THEME
Distributed Systems and middleware

Table of contents

1. Members	1
2. Overall Objectives	2
2.1. Decentralized personalization	2
2.2. Scalability: Cloud computing meets p2p	2
2.3. Privacy-aware decentralized computations	3
2.4. Information dissemination over social networks	3
2.5. Computability and efficiency of distributed systems	4
3. Research Program	5
3.1. Theory of distributed systems	5
3.2. Peer-to-peer overlay networks	5
3.3. Epidemic protocols	5
3.4. Malicious process behaviors	6
3.5. Online social networks and recommender systems	6
4. Highlights of the Year	6
5. New Software and Platforms	7
5.1. Asapknn (MediEgo)	7
5.2. Brow2Brow	7
5.3. Dashboard	7
5.4. GossipLib	8
5.5. HEAP	8
5.6. HyRec	8
5.7. WebGC	8
5.8. WhatsUp	9
5.9. YALPS	9
5.10. P2P-AllYours	10
6. New Results	10
6.1. Theory of Distributed Systems	10
6.1.1. t -Resilient Immediate Snapshot is Impossible	10
6.1.2. Two-Bit Messages are Sufficient to Implement Atomic Read/Write Registers in Crash-Prone Systems	10
6.2. Network and Graph Algorithms	11
6.2.1. Vertex Coloring with Communication and Local Memory Constraints in Synchronous Broadcast Networks	11
6.2.2. Optimal Collision/Conflict-Free Distance-2 Coloring in Wireless Synchronous Broadcast/Receive Tree Networks	11
6.2.3. Efficient Plurality Consensus, or: The Benefits of Cleaning Up from Time to Time	12
6.2.4. Bounds on the Voter Model in Dynamic Networks	12
6.2.5. How Asynchrony Affects Rumor Spreading Time	12
6.2.6. Amplifiers and Suppressors of Selection for the Moran Process on Undirected Graphs	13
6.3. Scalable Systems	13
6.3.1. Cache locality is not enough: High-Performance Nearest Neighbor Search with Product Quantization Fast Scan	13
6.3.2. Toward an Holistic Approach of Systems-of-Systems	14
6.3.3. Speed for the Elite, Consistency for the Masses: Differentiating Eventual Consistency in Large-Scale Distributed Systems	14
6.3.4. Bringing Secure Bitcoin Transactions to your Smartphone	14
6.3.5. Multithreading Approach to Process Real-Time Updates in KNN Algorithms	14
6.3.6. The Out-of-Core KNN Awakens: The Light Side of Computation Force on Large Datasets	15

6.3.7.	Partial Replication Policies for Dynamic Distributed Transactional Memory in Edge Clouds	15
6.3.8.	Being Prepared in a Sparse World: The Case of KNN Graph Construction	15
6.3.9.	Exploring the Use of Tags for Georeplicated Content Placement	15
6.3.10.	Mignon: A Fast Decentralized Content Consumption Estimation in Large-Scale Distributed Systems	16
6.4.	Privacy in User Centric Applications	16
6.4.1.	Hybrid Recommendations with Dynamic Similarity Measure	16
6.4.2.	Lightweight Privacy-Preserving Averaging for the Internet of Things	16
6.4.3.	Collaborative Filtering Under a Sybil Attack: Similarity Metrics do Matter!	17
6.4.4.	Privacy-Preserving Distributed Collaborative Filtering	17
7.	Bilateral Contracts and Grants with Industry	17
8.	Partnerships and Cooperations	17
8.1.	National Initiatives	17
8.1.1.	ANR project SocioPlug	17
8.1.2.	DeScenT CominLabs	18
8.1.3.	ANR Blanc project Displexity	18
8.1.4.	ANR project PAMELA	18
8.1.5.	ANR project OBrowser	18
8.1.6.	ANR project DESCARTES	19
8.1.7.	ANR-ERC Tremplin project NDFUSION	19
8.2.	International Initiatives	19
8.3.	International Research Visitors	19
8.3.1.	Visits of International Scientists	19
8.3.2.	Visits to International Teams	20
8.3.2.1.	Research Stays Abroad	20
8.3.2.2.	Internships	20
9.	Dissemination	20
9.1.	Promoting Scientific Activities	20
9.1.1.	Scientific Events Organisation	20
9.1.2.	Scientific Events Selection	20
9.1.3.	Journal	21
9.1.4.	Invited Talks	21
9.1.5.	Leadership within the Scientific Community	22
9.1.6.	Research Administration	22
9.2.	Teaching - Supervision - Juries	22
9.2.1.	Teaching	22
9.2.2.	Supervision	23
9.2.3.	Juries	23
10.	Bibliography	24

Project-Team ASAP

Creation of the Project-Team: 2007 July 01

Keywords:

Computer Science and Digital Science:

- 1.1.1. - Multicore
- 1.1.6. - Cloud
- 1.1.7. - Peer to peer
- 1.1.9. - Fault tolerant systems
- 1.2.9. - Social Networks
- 1.3. - Distributed Systems
- 1.5.2. - Communicating systems
- 2.1.6. - Concurrent programming
- 2.1.7. - Distributed programming
- 2.6.2. - Middleware
- 3.1.3. - Distributed data
- 3.1.8. - Big data (production, storage, transfer)
- 3.5.1. - Analysis of large graphs
- 3.5.2. - Recommendation systems
- 4.8. - Privacy-enhancing technologies
- 7.1. - Parallel and distributed algorithms
- 7.9. - Graph theory

Other Research Topics and Application Domains:

- 6.3.1. - Web
- 6.3.3. - Network Management
- 6.3.4. - Social Networks
- 6.4. - Internet of things
- 6.5. - Information systems
- 9.4.1. - Computer science
- 9.8. - Privacy

1. Members

Research Scientists

- Anne-Marie Kermarrec [Inria, Senior Researcher, **on secondment at Mediego since Apr 2015**, HDR]
- Daide Frey [Inria, Researcher]
- George Giakkoupis [Inria, Researcher]

Faculty Members

- Francois Taiani [Team leader, Univ. Rennes I, Professor, HDR]
- Marin Bertier [INSA Rennes, Associate Professor]
- David Bromberg [Univ. Rennes I, Professor, HDR]
- Michel Raynal [Univ. Rennes I, Professor, HDR]

Engineer

Alexandre Van Kempen [Inria, from Mar 2016]

PhD Students

Fabien Andre [Technicolor, until Nov 2016, granted by CIFRE]

Simon Bouget [Univ. Rennes I]

Adrien Luxey [Univ. Rennes I, from Oct 2016]

Nupur Mittal [Inria, until Nov 2016, granted by Conseil Régional de Bretagne]

Javier Olivares [Inria, until Dec 2016, granted by Google Inc]

Antoine Rault [Inria, until Jun 2016]

Pierre-Louis Roman [Univ. Rennes I]

Olivier Ruas [Inria, granted by Google Inc]

Resmi Ariyattu Chandrasekharannair [Univ. Rennes I, granted by UEB]

Stephane Delbruel [Univ. Rennes I]

Post-Doctoral Fellows

Zohir Bouzid [Univ. Rennes I]

Hicham Lakhlef [Univ. Rennes I, until Aug 2016]

Visiting Scientist

Arvid Bosk [KTH Royal Institute of Technology, Guest PhD Student, from Dec 2016]

Administrative Assistant

Virginie Desroches [Inria]

Others

Julien Lepiller [Inria, Student, from Feb to Jun 2016]

Florestan de Moor [Univ. Rennes I, Student, from May 2016 until July 2016]

2. Overall Objectives

2.1. Decentralized personalization

Our first objective is to offer full-fledged personalization in notification systems. Today, almost everyone is suffering from an overload of information that hurts both users and content providers. This suggests that not only will notification systems take a prominent role but also that, in order to be useful, they should be personalized to each and every user depending on her activity, operations, posts, interests, etc. In the GOSSPLE implicit instant item recommender, through a simple interface, users get automatically notified of items of interest for them, without explicitly subscribing to feeds or interests. They simply have to let the system know whether they like the items they receive (typically through a like/dislike button). Throughout the system's operation the personal data of users is stored on their own machines, which makes it possible to provide a wide spectrum of privacy guarantees while enabling cross-application benefits.

Our goal here is to provide a fully decentralized solution without ever requiring users to reveal their private preferences.

2.2. Scalability: Cloud computing meets p2p

Our second objective is to move forward in the area of **scalable infrastructures** for data intensive applications. In this context, we focus significant efforts on personalization systems, which represent one of the biggest challenges addressed by most large stake holders.

Hybrid infrastructures for personalisation. So far, social filtering techniques have mainly been implemented on centralized architectures relying on smart heuristics to cope with an increasing load of information. We argue however that, no matter how smart these heuristics and how powerful the underlying machines running them, a fully centralized approach might not be able to cope with the exponential growth of the Internet and, even if it does, the price to be paid might simply not be acceptable for its users (privacy, ecological footprint, etc.).

At the other end of the spectrum, lie fully decentralized systems where the collaborative filtering system is implemented by the machines of the users themselves. Such approaches are appealing for both scalability and privacy reasons. With respect to scalability, storage and computational units naturally grow with the number of users. Furthermore, a p2p system provides an energy-friendly environment where every user can feel responsible for the ecological foot-print of her exploration of the Internet. With respect to privacy, users are responsible for the management of their own profiles. Potential privacy threats therefore do not come from a big-brother but may still arise due to the presence of other users.

We have a strong experience in devising and experimenting with such kinds of p2p systems for various forms of personalization. More specifically, we have shown that personalization can be effective while maintaining a reasonable level of privacy. Nevertheless, frequent connections/disconnections of users make such systems difficult to maintain while addressing privacy attacks. For this reason, we also plan to explore hybrid approaches where the social filtering is performed by the users themselves, as in a p2p manner, whereas the management of connections-disconnections, including authentication, is managed through a server-based architecture. In particular, we plan to explore the trade-off between the quality of the personalization process, its efficiency and the privacy guarantees.

2.3. Privacy-aware decentralized computations

Gossip algorithms have also been studied for more complex global tasks, such as computation of network statistics or, more generally, aggregation functions of input values of the nodes (e.g., sum, average, or max). We plan to pursue this research direction both from a theoretical and from a practical perspective. We provide two examples of these directions below.

Computational capabilities of gossip. On the theoretical side, we have recently started to study gossip protocols for the assignment of unique IDs from a small range to all nodes (known as the *renaming* problem) and computing the rank of the input value of each node. We plan to further investigate the class of global tasks that can be solved efficiently by gossip protocols.

Private computations on decentralized data. On a more practical track, we aim to explore the use of gossip protocols for decentralized computations on privacy sensitive data. Recent research on private data bases, and on homomorphic encryption, has demonstrated the possibility to perform complex operations on encrypted data. Yet, existing systems have concentrated on relatively small-scale applications. In the coming years, we instead plan to investigate the possibility to build a framework for querying and performing operations for large-scale decentralized data stores. To achieve this, we plan to disseminate queries in an epidemic fashion through a network of data sources distributed on a large scale while combining privacy preserving techniques with decentralized computations. This would, for example, enable the computation of statistical measures on large quantities of data without needing to access and disclose each single data item.

2.4. Information dissemination over social networks

While we have been studying information dissemination in practical settings (such as WhatsUp in GOSSPLE), modeling such dynamic systems is still in its infancy. We plan to complement our practical work on gossip algorithms and information dissemination along the following axes:

Rumour spreading is a family of simple randomized algorithms for information dissemination, in which nodes contact (uniformly) random neighbours to exchange information with them. Despite their simplicity these protocols have proved very efficient for various network topologies. We are interested in studying their properties in specific topologies such as social networks be they implicit (interest-based as in GOSSPLE) or explicit (where users choose their friends as in Facebook). Recently, there has been some work on bounding the speed of rumour spreading in terms of abstract properties of the network graph, especially the graph's expansion properties of conductance and vertex expansion. It has been shown that high values for either of these guarantees fast rumour spreading—this should be related to empirical observations that social networks have high expansion. Some works established increasingly tighter upper bounds for rumour spreading in term of conductance or vertex expansion, but these bounds are not tight.

Our objective is to prove the missing tight upper bound for rumour spreading with vertex expansion. It is known that neither conductance nor vertex expansion are enough by themselves to completely characterize the speed of rumour spreading: are there graphs with bad expansion in which rumours spread fast?

Overcoming the dependence on expansion: Rumour spreading algorithms have very nice properties such as their simplicity, good performances for many networks but they may have very poor performance for some networks, even though these networks have small diameter, and thus it is possible to achieve fast information dissemination with more sophisticated protocols. Typically nodes may choose the neighbours to contact with some non-uniform probabilities that are determined based on information accumulated by each node during the run of the algorithm. These algorithms achieve information dissemination in time that is close to the diameter of the network. These algorithms, however, do not meet some of the other nice properties of rumour spreading, most importantly, robustness against failures. We are investigating algorithms that combine the good runtime of these latest protocols with the robustness of rumour spreading.

Competing rumours: Suppose now that two, or more, conflicting rumours (or opinions) spread in the network, and whenever a node receives different rumours it keeps only one of them. Which rumour prevails, and how long does it take until this happens? Similar questions have been studied in other contexts but not in the context of rumour spreading. The *voter* model is a well studied graph process that can be viewed as a competing rumour process that follows the classic PULL rumour spreading algorithm. However, research has only recently started to address the question of how long it takes until a rumour prevails. An interesting variant of the problem that has not been considered before is when different rumours are associated with different weights (some rumours are more convincing than others). We plan to study the above models and variations of them, and investigate their connection to the standard rumour spreading algorithms. This is clearly related to the dissemination of news and personalization in social networks.

2.5. Computability and efficiency of distributed systems

A very relevant challenge (maybe a Holy Grail) lies in the definition of a computation model appropriate to dynamic systems. This is a fundamental question. As an example there are a lot of peer-to-peer protocols but none of them is formally defined with respect to an underlying computing model. Similarly to the work of Lamport on "static" systems, a model has to be defined for dynamic systems. This theoretical research is a necessary condition if one wants to understand the behavior of these systems. As the aim of a theory is to codify knowledge in order it can be transmitted, the definition of a realistic model for dynamic systems is inescapable whatever the aim we have in mind, be it teaching, research or engineering.

Distributed computability: Among the fundamental theoretical results of distributed computing, there is a list of problems (e.g., consensus or non-blocking atomic commit) that have been proved to have no deterministic solution in asynchronous distributed computing systems prone to failures. In order such a problem to become solvable in an asynchronous distributed system, that system has to be enriched with an appropriate oracle (also called failure detector). We have been deeply involved in this research and designed optimal consensus algorithms suited to different kind of oracles. This line of research paves the way to rank the distributed computing problems according to the "power" of the additional oracle they required (think of "additional oracle" as "additional assumptions"). The ultimate goal would be the statement of a distributed computing hierarchy, according to the minimal assumptions needed to solve distributed computing problems (similarly to the Chomsky's hierarchy that ranks problems/languages according to the type of automaton they need to be solved).

Distributed computing abstractions: Major advances in sequential computing came from machine-independent data abstractions such as sets, records, etc., control abstractions such as while, if, etc., and modular constructs such as functions and procedures. Today, we can no longer envisage not to use these abstractions. In the "static" distributed computing field, some abstractions have been promoted and proved to be useful. Reliable broadcast, consensus, interactive consistency are some examples of such abstractions. These abstractions have well-defined specifications. There are both a lot of theoretical results on them (mainly decidability and lower bounds), and numerous implementations. There is no such equivalent for dynamic distributed systems, i.e. for systems characterized by nodes that may join and leave, or that may change

their characteristics at runtime. Our goal is to define such novel abstractions, thereby extending the theory of distributed systems to the dynamic case.

3. Research Program

3.1. Theory of distributed systems

Finding models for distributed computations prone to asynchrony and failures has received a lot of attention. A lot of research in this domain focuses on what can be computed in such models, and, when a problem can be solved, what are its best solutions in terms of relevant cost criteria. An important part of that research is focused on distributed computability: what can be computed when failure detectors are combined with conditions on process input values for example. Another part is devoted to model equivalence. What can be computed with a given class of failure detectors? Which synchronization primitives is a given failure class equivalent to? These are among the main topics addressed in the leading distributed computing community. A second fundamental issue related to distributed models is the definition of appropriate models suited to dynamic systems. Up to now, the researchers in that area consider that nodes can enter and leave the system, but do not provide a simple characterization, based on properties of computation instead of description of possible behaviors [58], [51], [53]. This shows that finding dynamic distributed computing models is today a "Holy Grail", whose discovery would allow a better understanding of the essential nature of dynamic systems.

3.2. Peer-to-peer overlay networks

A standard distributed system today is related to thousands or even millions of computing entities scattered all over the world and dealing with a huge amount of data. This major shift in scalability requirements has led to the emergence of novel computing paradigms. In particular, the peer-to-peer communication paradigm imposed itself as the prevalent model to cope with the requirements of large scale distributed systems. Peer-to-peer systems rely on a symmetric communication model where peers are potentially both clients and servers. They are fully decentralized, thus avoiding the bottleneck imposed by the presence of servers in traditional systems. They are highly resilient to peers arrivals and departures. Finally, individual peer behavior is based on a local knowledge of the system and yet the system converges toward global properties.

A peer-to-peer overlay network logically connects peers on top of IP. Two main classes of such overlays dominate, structured and unstructured. The differences relate to the choice of the neighbors in the overlay, and the presence of an underlying naming structure. Overlay networks represent the main approach to build large-scale distributed systems that we retained. An overlay network forms a logical structure connecting participating entities on top of the physical network, be it IP or a wireless network. Such an overlay might form a structured overlay network [59], [60], [61] following a specific topology or an unstructured network [56], [62] where participating entities are connected in a random or pseudo-random fashion. In between, lie weakly structured peer-to-peer overlays where nodes are linked depending on a proximity measure providing more flexibility than structured overlays and better performance than fully unstructured ones. Proximity-aware overlays connect participating entities so that they are connected to close neighbors according to a given proximity metric reflecting some degree of affinity (computation, interest, etc.) between peers. We extensively use this approach to provide algorithmic foundations of large-scale dynamic systems.

3.3. Epidemic protocols

Epidemic algorithms, also called gossip-based algorithms [55], [54], constitute a fundamental topic in our research. In the context of distributed systems, epidemic protocols are mainly used to create overlay networks and to ensure a reliable information dissemination in a large-scale distributed system. The principle underlying technique, in analogy with the spread of a rumor among humans via gossiping, is that participating entities continuously exchange information about the system in order to spread it gradually and reliably. Epidemic algorithms have proved efficient to build and maintain large-scale distributed systems in the context of many applications such as broadcasting [54], monitoring, resource management, search, and more generally in building unstructured peer-to-peer networks.

3.4. Malicious process behaviors

When assuming that processes fail by simply crashing, bounds on resiliency (maximum number of processes that may crash, number of exchanged messages, number of communication steps, etc.) are known both for synchronous and augmented asynchronous systems (recall that in purely asynchronous systems some problems are impossible to solve). If processes can exhibit malicious behaviors, these bounds are seldom the same. Sometimes, it is even necessary to change the specification of the problem. For example, the consensus problem for correct processes does not make sense if some processes can exhibit a Byzantine behavior and thus propose an arbitrary value. In this case, the validity property of consensus, which is normally "a decided value is a proposed value", must be changed to "if all correct processes propose the same value then only this value can be decided." Moreover, the resilience bound of less than half of faulty processes is at least lowered to "less than a third of Byzantine processes." These are some of the aspects that underlie our studies in the context of the classical model of distributed systems, in peer-to-peer systems and in sensor networks.

3.5. Online social networks and recommender systems

Social Networks have rapidly become a fundamental component of today's distributed applications. Web 2.0 applications have dramatically changed the way users interact with the Internet and with each other. The number of users of websites like Flickr, Delicious, Facebook, or MySpace is constantly growing, leading to significant technical challenges. On the one hand, these websites are called to handle enormous amounts of data. On the other hand, news continue to report the emergence of privacy threats to the personal data of social-network users. Our research aims to exploit our expertise in distributed systems to lead to a new generation of scalable, privacy-preserving, social applications.

We also investigate approaches to build implicit social networks, connecting users sharing similar interests. At the heart of the building of such similarity graphs lie k-nearest neighbor (KNN) algorithms. Our research in this area is to design and implement efficient KNN algorithms able to cope with a huge volume of data as well as a high level of dynamism. We investigate the use of such similarity graphs to build highly scalable infrastructures for recommendation systems.

4. Highlights of the Year

4.1. Highlights of the Year

Anne Marie Kermarrec has been named an ACM Fellow "for her contributions to large-scale distributed computing."

George Giakkoupis was the General Chair of the 35th ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC 2016).

Michel Raynal renewed his appointment as an Adjunct Professor at the University of Hong Kong.

4.1.1. Awards

BEST PAPERS AWARDS:

[30]

S. DELBRUEL, D. FREY, F. TAÏANI. *Exploring The Use of Tags for Georeplicated Content Placement*, in "IEEE IC2E'16", Berlin, Germany, April 2016, <https://hal.inria.fr/hal-01257939>

[27]

N. CHILUKA, A.-M. KERMARREC, J. OLIVARES. *The Out-of-core KNN Awakens: The light side of computation force on large datasets*, in "The International Conference on Networked Systems NETYS", Marrakech, Morocco, May 2016, <https://hal.inria.fr/hal-01336673>

5. New Software and Platforms

5.1. Asapknn (MediEgo)

KEYWORDS: Widget web - Social network - Recommendation

FUNCTIONAL DESCRIPTION

Asapknn (MediEgo) is a solution for content recommendation based on the users navigation history. The solution 1) collects the usages of the Web users and store them in a profile, 2) uses this profile to associate to each user her most similar users, 3) leverages this implicit network of close users in order to infer their preferences and recommend advertisements and recommendations. MediEgo achieves scalability using a sampling method, which provides very good results at a drastically reduced cost.

- Participants: Antoine Boutet, Jacques Falcou, Jean Francois Verdonck, Anne Marie Kermarrec, Sebastien Campion, Rachid Guerraoui, Davide Frey and Arnaud Jegou
- Partner: EPFL - Ecole Polytechnique Fédérale de Lausanne
- Contact: Sebastien Campion

5.2. Brow2Brow

Browser-to-browser serverless toolboxes

FUNCTIONAL DESCRIPTION

Brow2Brow is an “Action de Development Technologique”, i.e. a collaborative development project that aims at providing a middleware and software library for browser-to-browser applications. Brow2Brow involves the ASAP team as well as the DICE Team from Inria Grenoble (Antenne de Lyon). The project seeks to provide an alternative to the current model followed by Web2.0 applications by exploiting the recently introduced WebRTC standard. Existing Web 2.0 applications collect data on browsers and send it to servers that store and process it. The goal of Brow2Brow is to provide an alternative approach where browsers can themselves proceed to collaborative data processing. This will make it possible avoid data concentration at a single server. The project has resulted so far in the development of WebGC, a library for gossip-based applications on browsers.

- Participants: Anne-Marie Kermarrec, Davide Frey and Raziel Carvajal Gomez
- Contact: Raziel Carvajal Gomez

5.3. Dashboard

MediEgo Dashboard: A personalized news dashboard

KEYWORDS: Recommender system - Personalized stream of news - Dashboard

FUNCTIONAL DESCRIPTION

This work has led to the development of MEDIEGO Dashboard, a personalized news recommendation system. In MEDIEGO Dashboard, users benefit from a personalized stream of news matching their interests. Additionally, users can use explicit subscriptions as well as post content and navigate through tags. MEDIEGO Dashboard is available through a web interface and a mobile-based Android application. To provide personalization, MEDIEGO Dashboard exploits the users’ opinions regarding their received news to identify users with similar interests. MEDIEGO Dashboard is centralized and it allows us to test and evaluate different recommendation schemes. In collaboration with EIT/ICT Lab, an experiment has been conducted with a set of users at Trento (Italie). This experiment allowed us to collect traces and to perform a user survey to assess and improve our solution. This solution will soon be interconnected to P2P-AllYours.

- Participants: Anne-Marie Kermarrec, Antoine Boutet, Yuri Barssi and Jean Francois Verdonck
- Contact: Anne-Marie Kermarrec
- URL: <http://www.mediego.com>

5.4. GossipLib

KEYWORDS: Nat traversal - Epidemic protocols - Gossip protocols - Overlay maintenance - Peer-to-peer - Dissemination

FUNCTIONAL DESCRIPTION

GossipLib is a library consisting of a set of Java classes aimed to facilitate the development of gossip-based application in a large-scale setting. It provides developers with a set of support classes that constitute a solid starting point for building any gossip-based application. GossipLib is designed to facilitate code reuse and testing of distributed application and as thus also provides the implementation of a number of standard gossip protocols that may be used out of the box or extended to build more complex protocols and applications. These include for example the peer-sampling protocols for overlay management.

GossipLib also provides facility for the configuration and deployment of applications as final-product but also as research prototype in environments like PlanetLab, clusters, network emulators, and even as event-based simulation. The code developed with GossipLib can be run both as a real application and in simulation simply by changing one line in a configuration file.

- Participants: Davide Frey, Ribeiro Heverson, Anne Marie Kermarrec, Imane Al Ifdal and Ilham Ikbal
- Contact: Davide Frey
- URL: <http://gossiplib.gforge.inria.fr/>

5.5. HEAP

Heterogeneity-aware gossip protocol

FUNCTIONAL DESCRIPTION

A video streaming platform based on HEAP. The platform is particularly suited for environment characterized by heterogeneous bandwidth capabilities such as those comprising ADSL edge nodes. HEAP is, in fact, able to dynamically leverage the most capable nodes and increase their contribution to the protocol, while decreasing by the same proportion that of less capable nodes. During the last few months, we have integrated HEAP with the ability to dynamically measure the available bandwidth of nodes, thereby making it independent of the input of the user.

- Participants: Davide Frey, Arnaud Jegou, Anne-Marie Kermarrec, Vivien Quema, Maxime Monod and Rachid Guerraoui
- Contact: Davide Frey

5.6. HyRec

Hybrid Recommender System

FUNCTIONAL DESCRIPTION

The motivation of this work is to explore solutions that could in some sense democratize personalization by making it accessible to any content provider company without generating huge investments. HyRec implements a user-based collaborative filtering scheme and offloads CPU-intensive recommendation tasks to front-end client browsers, while retaining storage and orchestration tasks within back-end servers. HyRec seeks to provide the scalability of p2p approaches without forcing content providers to give up the control of the system.

- Participants: Antoine Boutet, Davide Frey, Anne Marie Kermarrec, Arnaud Jegou and Rachid Guerraoui
- Contact: Davide Frey

5.7. WebGC

Web-based Gossip Communication

FUNCTIONAL DESCRIPTION

WebGC is a library for gossip-based communication between web-browsers. It has been developed in collaboration with Mathieu Simonin in the context of the Brow2Brow ADT project. WebGC builds on the recent WebRTC standard as well as on PeerJS, an open-source project that provides primitives for data transfer on top of WebRTC.

- Participants: Raziel Carvajal Gomez, Davide Frey and Anne-Marie Kermarrec
- Contact: Davide Frey

5.8. WhatsUp

KEYWORD: Recommender system

FUNCTIONAL DESCRIPTION

WhatsUp is a distributed recommendation system aimed to distribute instant news in a large scale dynamic system. WhatsUp has two parts, an embedded application server in order to exchange with others peers in the system and a fully dynamic web interface for displaying news and collecting opinions about what the user reads. Underlying this web-based application lies Beep, a biased epidemic dissemination protocol that delivers news to interested users in a fast manner while limiting spam. Beep is parametrized on the fly to manage the orientation and the amplification of news dissemination. Every user forwards the news of interest to a randomly selected set of users with a preference towards those that have similar interests (orientation). The notion of interest does not rely on any explicit social network or subscription scheme, but rather on an implicit and dynamic overlay capturing the commonalities between users with respect to they are interested in. The size of the set of users to which a news is forwarded depends on the interest of the news (amplification). A centralized version of WhatsUp is already up and running and the decentralized one is still in beta version.

- Participants: Davide Frey, Ribeiro Heverson, Antoine Boutet, Anne Marie Kermarrec, Arnaud Jegou, Rachid Guerraoui and Jean Francois Verdonck
- Contact: Davide Frey

5.9. YALPS

KEYWORDS: Traffic-shaping - Nat traversal - Experimentation - Peer-to-peer - Simulator - Deployment

FUNCTIONAL DESCRIPTION

YALPS is an open-source Java library designed to facilitate the development, deployment, and testing of distributed applications. Applications written using YALPS can be run both in simulation and in real-world mode without changing a line of code or even recompiling the sources. A simple change in a configuration file will load the application in the proper environment. A number of features make YALPS useful both for the design and evaluation of research prototypes and for the development of applications to be released to the public. Specifically, YALPS makes it possible to run the same application as a simulation or in a real deployment. Applications communicate by means of application-defined messages which are then routed either through UDP/TCP or through YALPS's simulation infrastructure. In both cases, YALPS's communication layer offers features for testing and evaluating distributed protocols and applications. Communication channels can be tuned to incorporate message losses or to constrain their outgoing bandwidth. Finally, YALPS includes facilities to support operation in the presence of NATs and firewalls using relaying and NAT-traversal techniques. The implementation of YALPS includes approximately 16K lines of code, and is used in several projects by ASAP, including HEAP, P2P-AllYours, and Behave.

- Participants: Davide Frey, Maxime Monod, Heverson Borba Ribeiro, Anne Marie Kermarrec and Arnaud Jegou
- Contact: Davide Frey
- URL: <http://yalps.gforge.inria.fr/>

5.10. P2P-AllYours

Peer-to-Peer AllYours

FUNCTIONAL DESCRIPTION

P2P-AllYours is customization of WhatsUp developed in the context of the EIT/ICT-Labs AllYours project. In addition to WhatsUp (the distributed recommender engine), P2P-AllYours comprises the following features:

- A new web interface, which users can access through a local web-server integrated in P2P-AllYours.
- A set of automatic nodes (BOTs) that can extract news items from RSS feeds and insert them into the recommender system.
- A content-bootstrap that solves the issues related to bootstrapping the recommender system when a user connects for the first time.
- An experiment management server that allows users to register for the application in the context of the testing program.
- Participants: Davide Frey, Heverson Borba Ribeiro, Raziel Carvajal Gomez, Arnaud Jegou and Anne-Marie Kermarrec
- Contact: Davide Frey

6. New Results

6.1. Theory of Distributed Systems

6.1.1. *t-Resilient Immediate Snapshot is Impossible*

Participant: Michel Raynal.

Immediate snapshot is the basic communication object on which relies the read/write distributed computing model made up of n crash-prone asynchronous processes, called iterated distributed model. Each iteration step (usually called a round) uses a new immediate snapshot object, which allows the processes to communicate and cooperate. More precisely, the x -th immediate snapshot object can be used by a process only when it executes the x -th round. An immediate snapshot object can be implemented by an $(n-1)$ -resilient algorithm, i.e. an algorithm that tolerates up to $(n-1)$ process crashes (also called wait-free algorithm). Considering a t -crash system model (i.e. a model in which up to t processes are allowed to crash), this work [46] is on the construction of an extension of immediate snapshot objects to t -resiliency. In the t -crash system model, at each round each process may be ensured to get values from at least $n-t$ processes, and t -immediate snapshot has the properties of classical immediate snapshot (1-immediate snapshot) but ensures that each process will get values from at least $n-t$ processes. Its main result is the following. While there is a (deterministic) t -resilient read/write-based algorithm implementing t -immediate snapshot in a t -crash system when $t = n-1$, there is no t -resilient algorithm in a t -crash model when $t \in [1..(n-2)]$. This means that the notion of t -resilience is inoperative when one has to implement immediate snapshot for these values of t : the model assumption “at most $t < n-1$ processes may crash” does not provide us with additional computational power allowing for the design of genuine t -resilient algorithms (genuine meaning that such a t -resilient algorithm would work in the t -crash model, but not in the $(t+1)$ -crash model). To show these results, the paper relies on well-known distributed computing agreement problems such as consensus and k -set agreement.

This work was done in collaboration with Carole Delporte, Hugues Fauconnier, and Sergio Rajsbaum, and appeared at SIROCCO 2016.

6.1.2. *Two-Bit Messages are Sufficient to Implement Atomic Read/Write Registers in Crash-Prone Systems*

Participant: Michel Raynal.

Atomic registers are certainly the most basic objects of computing science. Their implementation on top of an n -process asynchronous message-passing system has received a lot of attention. It has been shown that $t < n/2$ (where t is the maximal number of processes that may crash) is a necessary and sufficient requirement to build an atomic register on top of a crash-prone asynchronous message-passing system. Considering such a context, this work [49] presents an algorithm which implements a single-writer multi-reader atomic register with four message types only, and where no message needs to carry control information in addition to its type. Hence, two bits are sufficient to capture all the control information carried by all the implementation messages. Moreover, the messages of two types need to carry a data value while the messages of the two other types carry no value at all. As far as we know, this algorithm is the first with such an optimality property on the size of control information carried by messages. It is also particularly efficient from a time complexity point of view.

This work was done in collaboration with Achour Mostefaoui, and appeared at PODC 2016.

6.2. Network and Graph Algorithms

6.2.1. *Vertex Coloring with Communication and Local Memory Constraints in Synchronous Broadcast Networks*

Participants: Hicham Lakhlef, Michel Raynal, Francois Taiani.

This work [41] considers the broadcast/receive communication model in which message collisions and message conflicts can occur because processes share frequency bands. (A collision occurs when, during the same round, messages are sent to the same process by too many neighbors. A conflict occurs when a process and one of its neighbors broadcast during the same round.) More precisely, this work considers the case where, during a round, a process may either broadcast a message to its neighbors or receive a message from at most m of them. This captures communication-related constraints or a local memory constraint stating that, whatever the number of neighbors of a process, its local memory allows it to receive and store at most m messages during each round. This work defines first the corresponding generic vertex multi-coloring problem (a vertex can have several colors). It focuses then on tree networks, for which it presents a lower bound on the number of colors K that are necessary (namely, $K = \lceil \frac{\Delta}{m} \rceil + 1$, where Δ is the maximal degree of the communication graph), and an associated coloring algorithm, which is optimal with respect to K .

6.2.2. *Optimal Collision/Conflict-Free Distance-2 Coloring in Wireless Synchronous Broadcast/Receive Tree Networks*

Participants: Davide Frey, Hicham Lakhlef, Michel Raynal.

We studied the problem of decentralized distance-2 coloring in message-passing systems where communication is (a) synchronous and (b) based on the “broadcast/receive” pair of communication operations. “Synchronous” means that time is discrete and appears as a sequence of time slots (or rounds) such that each message is received in the very same round in which it is sent. “Broadcast/receive” means that during a round a process can either broadcast a message to its neighbors or receive a message from one of them. In such a communication model, no two neighbors of the same process, nor a process and any of its neighbors, must be allowed to broadcast during the same time slot (thereby preventing message collisions in the first case, and message conflicts in the second case). From a graph theory point of view, the allocation of slots to processes is known as the distance-2 coloring problem: a color must be associated with each process (defining the time slots in which it will be allowed to broadcast) in such a way that any two processes at distance at most 2 obtain different colors, while the total number of colors is “as small as possible”. In this context, we proposed a parallel message-passing distance-2 coloring algorithm suited to trees, whose roots are dynamically defined. This algorithm, which is itself collision-free and conflict-free, uses $\Delta + 1$ colors where Δ is the maximal degree of the graph (hence the algorithm is color-optimal). It does not require all processes to have different initial identities, and its time complexity is $O(d\Delta)$, where d is the depth of the tree. As far as we know, this is the first distributed distance-2 coloring algorithm designed for the broadcast/receive round-based communication model, which owns all the previous properties. We published these results in [39].

6.2.3. Efficient Plurality Consensus, or: The Benefits of Cleaning Up from Time to Time

Participant: George Giakkoupis.

Plurality consensus considers a network of n nodes, each having one of k opinions. Nodes execute a (randomized) distributed protocol with the goal that all nodes adopt the *plurality* (the opinion initially supported by the most nodes). Communication is realized via the random phone call model. A major open question has been whether there is a protocol for the complete graph that converges (w.h.p.) in polylogarithmic time and uses only polylogarithmic memory per node (local memory). We answered this question affirmatively.

In [22], we propose two protocols that need only mild assumptions on the bias in favor of the plurality. As an example of our results, consider the complete graph and an arbitrarily small constant multiplicative bias in favor of the plurality. Our first protocol achieves plurality consensus in $O(\log k \cdot \log \log n)$ rounds using $\log k + O(\log \log k)$ bits of local memory. Our second protocol achieves plurality consensus in $O(\log n \cdot \log \log n)$ rounds using only $\log k + 4$ bits of local memory. This disproves a conjecture by Becchetti et al. (SODA'15) implying that any protocol with local memory $\log k + O(1)$ has worst-case runtime $\Omega(k)$. We provide similar bounds for much weaker bias assumptions. At the heart of our protocols lies an *undecided state*, an idea introduced by Angluin et al. (Distributed Computing'08).

This work was done in collaboration with Petra Berenbrink (SFU), Tom Friedetzky (Durham University), and Peter Kling (SFU).

6.2.4. Bounds on the Voter Model in Dynamic Networks

Participants: George Giakkoupis, Anne-Marie Kermarrec.

In the *voter model*, each node of a graph has an opinion, and in every round each node chooses independently a random neighbour and adopts its opinion. We are interested in the *consensus time*, which is the first point in time where all nodes have the same opinion. In [23], we consider dynamic graphs in which the edges are rewired in every round (by an adversary) giving rise to the graph sequence G_1, G_2, \dots , where we assume that G_i has conductance at least ϕ_i . We assume that the degrees of nodes don't change over time as one can show that the consensus time can become super-exponential otherwise. In the case of a sequence of d -regular graphs, we obtain asymptotically tight results. Even for some static graphs, such as the cycle, our results improve the state of the art. Here we show that the expected number of rounds until all nodes have the same opinion is bounded by $O(m/(\delta \cdot \phi))$, for any graph with m edges, conductance ϕ , and degrees at least δ . In addition, we consider a *biased* dynamic voter model, where each opinion i is associated with a probability P_i , and when a node chooses a neighbour with that opinion, it adopts opinion i with probability P_i (otherwise the node keeps its current opinion). We show for any regular dynamic graph, that if there is an $\epsilon > 0$ difference between the highest and second highest opinion probabilities, and at least $\Omega(\log n)$ nodes have initially the opinion with the highest probability, then all nodes adopt w.h.p. that opinion. We obtain a bound on the convergence time, which becomes $O(\log n/\phi)$ for static graphs.

This work was done in collaboration with Petra Berenbrink (SFU), and Frederik Mallmann-Trenn (SFU).

6.2.5. How Asynchrony Affects Rumor Spreading Time

Participant: George Giakkoupis.

In standard randomized (push-pull) rumor spreading, nodes communicate in synchronized rounds. In each round every node contacts a random neighbor in order to exchange the rumor (i.e., either push the rumor to its neighbor or pull it from the neighbor). A natural asynchronous variant of this algorithm is one where each node has an independent Poisson clock with rate 1, and every node contacts a random neighbor whenever its clock ticks. This asynchronous variant is arguably a more realistic model in various settings, including message broadcasting in communication networks, and information dissemination in social networks.

In [35] we study how asynchrony affects the rumor spreading time, that is, the time before a rumor originated at a single node spreads to all nodes in the graph. Our first result states that the asynchronous push-pull rumor spreading time is asymptotically bounded by the standard synchronous time. Precisely, we show that for any graph G on n -nodes, where the synchronous push-pull protocol informs all nodes within $T(G)$ rounds with high probability, the asynchronous protocol needs at most time $O(T(G) + \log n)$ to inform all nodes with high probability. On the other hand, we show that the expected synchronous push-pull rumor spreading time is bounded by $O(\sqrt{n})$ times the expected asynchronous time.

These results improve upon the bounds for both directions shown recently by Acan et al. (PODC 2015). An interesting implication of our first result is that in regular graphs, the weaker push-only variant of synchronous rumor spreading has the same asymptotic performance as the synchronous push-pull algorithm.

This work was done in collaboration with Yasamin Nazari and Philipp Woelfel from the University of Calgary.

6.2.6. Amplifiers and Suppressors of Selection for the Moran Process on Undirected Graphs

Participant: George Giakkoupis.

In [47] we consider the classic Moran process modeling the spread of genetic mutations, as extended to structured populations by Lieberman et al. (Nature, 2005). In this process, individuals are the vertices of a connected graph G . Initially, there is a single mutant vertex, chosen uniformly at random. In each step, a random vertex is selected for reproduction with a probability proportional to its fitness: mutants have fitness $r > 1$, while non-mutants have fitness 1. The vertex chosen to reproduce places a copy of itself to a uniformly random neighbor in G , replacing the individual that was there. The process ends when the mutation either reaches fixation (i.e., all vertices are mutants), or gets extinct. The principal quantity of interest is the probability with which each of the two outcomes occurs.

A problem that has received significant attention recently concerns the existence of families of graphs, called strong amplifiers of selection, for which the fixation probability tends to 1 as the order n of the graph increases, and the existence of strong suppressors of selection, for which this probability tends to 0. For the case of directed graphs, it is known that both strong amplifiers and suppressors exist. For the case of undirected graphs, however, the problem has remained open, and the general belief has been that neither strong amplifiers nor suppressors exist. In this work we disprove this belief, by providing the first examples of such graphs. The strong amplifier we present has fixation probability $1 - \tilde{O}(n^{-1/3})$, and the strong suppressor has fixation probability $\tilde{O}(n^{-1/4})$. Both graph constructions are surprisingly simple. We also prove a general upper bound of $1 - \tilde{\Omega}(n^{-1/3})$ on the fixation probability of any undirected graph. Hence, our strong amplifier is existentially optimal.

6.3. Scalable Systems

6.3.1. Cache locality is not enough: High-Performance Nearest Neighbor Search with Product Quantization Fast Scan

Participants: Fabien Andre, Anne-Marie Kermarrec.

Nearest Neighbor (NN) search in high dimension is an important feature in many applications (e.g., image retrieval, multimedia databases). Product Quantization (PQ) is a widely used solution which offers high performance, i.e., low response time while preserving a high accuracy. PQ represents high-dimensional vectors (e.g., image descriptors) by compact codes. Hence, very large databases can be stored in memory, allowing NN queries without resorting to slow I/O operations. PQ computes distances to neighbors using cache-resident lookup tables, thus its performance remains limited by (i) the many cache accesses that the algorithm requires, and (ii) its inability to leverage SIMD instructions available on modern CPUs. In this paper, we advocate that cache locality is not sufficient for efficiency. To address these limitations, in [19] we design a novel algorithm, PQ Fast Scan, that transforms the cache-resident lookup tables into small tables, sized to fit SIMD registers. This transformation allows (i) in-register lookups in place of cache accesses and (ii) an efficient SIMD implementation. PQ Fast Scan has the exact same accuracy as PQ, while having 4 to 6 times lower response time (e.g., for 25 million vectors, scan time is reduced from 74ms to 13ms).

6.3.2. *Toward an Holistic Approach of Systems-of-Systems*

Participants: Simon Bouget, David Bromberg, Francois Taiani.

Large scale distributed systems have become ubiquitous, from on-line social networks to the Internet-of-Things. To meet rising expectations (scalability, robustness, flexibility,...) these systems increasingly espouse complex distributed architectures, that are hard to design, deploy and maintain. To grasp this complexity, developers should be allowed to assemble large distributed systems from smaller parts using a seamless, high-level programming paradigm. We present in [24] such an assembly-based programming framework, enabling developers to easily define and realize complex distributed topologies as a construction of simpler blocks (e.g. rings, grids). It does so by harnessing the power of self-organizing overlays, that is made accessible to developers through a high-level Domain Specific Language and self-stabilizing run-time. Our evaluation further shows that our approach is generic, expressive, low-overhead and robust.

6.3.3. *Speed for the Elite, Consistency for the Masses: Differentiating Eventual Consistency in Large-Scale Distributed Systems*

Participants: Davide Frey, Pierre-Louis Roman, Francois Taiani.

Eventual consistency is a consistency model that emphasizes liveness over safety; it is often used for its ability to scale as distributed systems grow larger. Eventual consistency tends to be uniformly applied to an entire system, but we argue that there is a growing demand for differentiated eventual consistency requirements.

We address this demand with UPS [34], a novel consistency mechanism that offers differentiated eventual consistency and delivery speed by working in pair with a two-phase epidemic broadcast protocol. We propose a closed-form analysis of our approach's delivery speed, and we evaluate our complete mechanism experimentally on a simulated network of one million nodes. To measure the consistency trade-off, we formally define a novel and scalable consistency metric that operates at runtime. In our simulations, UPS divides by more than 4 the inconsistencies experienced by a majority of the nodes, while reducing the average latency incurred by a small fraction of the nodes from 6 rounds down to 3 rounds.

This work was done in collaboration with Achour Mostefaoui and Matthieu Perrin from the LINA laboratory in Nantes.

6.3.4. *Bringing Secure Bitcoin Transactions to your Smartphone*

Participants: Davide Frey, Pierre-Louis Roman, Francois Taiani.

To preserve the Bitcoin ledger's integrity, a node that joins the system must download a full copy of the entire Bitcoin blockchain if it wants to verify newly created blocks. At the time of writing, the blockchain weights 79 GiB and takes hours of processing on high-end machines. Owners of low-resource devices (known as thin nodes), such as smartphones, avoid that cost by either opting for minimum verification or by depending on full nodes, which weakens their security model.

In this work [33], we propose to harden the security model of thin nodes by enabling them to verify blocks in an adaptive manner, with regards to the level of targeted confidence, with low storage requirements and a short bootstrap time. Our approach exploits sharing within a distributed hash table (DHT) to distribute the storage load, and a few additional hashes to prevent attacks on this new system.

This work was done in collaboration with Marc X. Makkes and Spyros Voulgaris from Vrije Universiteit Amsterdam (The Netherlands).

6.3.5. *Multithreading Approach to Process Real-Time Updates in KNN Algorithms*

Participants: Anne-Marie Kermarrec, Nupur Mittal, Javier Olivares.

K-Nearest Neighbors algorithm is the core of a considerable amount of online services and applications, like recommendation engines, content-classifiers, information retrieval systems, etc. The users of these services change their preferences and evolve with time, aggravating the computational challenges of KNN more with the ever evolving data to process. In this work [48], we present *UpKNN*: an efficient thread-based approach to take the updates of users preferences into account while it computes the KNN efficiently, keeping a check on the wall-time.

By using an efficient thread-based approach, *UpKNN* processes millions of updates online, on a single commodity PC. Our experiments confirm the scalability of *UpKNN*, both in terms of the number of updates processed and the threads used. *UpKNN* achieves speedups ranging from 13.64X to 49.5X in the processing of millions of updates, with respect to the performance of a non-partitioned baseline. These results are a direct consequence of reducing the number of disk operations, roughly speaking, only 1% disk operations are performed as compared to the baseline.

6.3.6. *The Out-of-Core KNN Awakens: The Light Side of Computation Force on Large Datasets*

Participants: Anne-Marie Kermarrec, Javier Olivares.

K-Nearest Neighbors (KNN) is a crucial tool for many applications, e.g. recommender systems, image classification and web-related applications. However, KNN is a resource greedy operation particularly for large datasets. We focus on the challenge of KNN computation over large datasets on a single commodity PC with limited memory. We propose a novel approach [27] to compute KNN on large datasets by leveraging both disk and main memory efficiently. The main rationale of our approach is to minimize random accesses to disk, maximize sequential accesses to data and efficient usage of only the available memory.

We evaluate our approach on large datasets, in terms of performance and memory consumption. The evaluation shows that our approach requires only 7% of the time needed by an in-memory baseline to compute a KNN graph.

6.3.7. *Partial Replication Policies for Dynamic Distributed Transactional Memory in Edge Clouds*

Participant: Francois Taiani.

Distributed Transactional Memory (DTM) can play a fundamental role in the coordination of participants in edge clouds as a support for mobile distributed applications. DTM emerges as a concurrency mechanism aimed at simplifying distributed programming by allowing groups of operations to execute atomically, mirroring the well-known transaction model of relational databases. In spite of recent studies showing that partial replication approaches can present gains in the scalability of DTMs by reducing the amount of data stored at each node, most DTM solutions follow a full replication scheme. The few partial replicated DTM frameworks either follow a random or round-robin algorithm for distributing data onto partial replication groups. In order to overcome the poor performance of these schemes, this work [36] investigates policies to extend the DTM to efficiently and dynamically map resources on partial replication groups. The goal is to understand if a dynamic service that constantly evaluates the data mapped into partial replicated groups can contribute to improve DTM based systems performance.

This work was performed in collaboration with Diogo Lima and Hugo Miranda from the University of Lisbon (Portugal).

6.3.8. *Being Prepared in a Sparse World: The Case of KNN Graph Construction*

Participants: Anne-Marie Kermarrec, Nupur Mittal, Francois Taiani.

Work [25] presents KIFF, a generic, fast and scalable KNN graph construction algorithm. KIFF directly exploits the bipartite nature of most datasets to which KNN algorithms are applied. This simple but powerful strategy drastically limits the computational cost required to rapidly converge to an accurate KNN solution, especially for sparse datasets. Our evaluation on a representative range of datasets show that KIFF provides, on average, a speed-up factor of 14 against recent state-of-the art solutions while improving the quality of the KNN approximation by 18

This work was done in collaboration with Antoine Boutet from CNRS, Laboratoire Hubert Curien, Saint-Etienne, France.

6.3.9. *Exploring the Use of Tags for Georeplicated Content Placement*

Participants: Stephane Delbruel, Davide Frey, Francois Taiani.

A large portion of today's Internet traffic originates from streaming and video services. Such services rely on a combination of distributed datacenters, powerful content delivery networks (CDN), and multi-level caching . In spite of this infrastructure, storing, indexing, and serving these videos remains a daily engineering challenge that requires increasing efforts on the part of providers and ISPs. In this work [30], we explore how the tags attached to videos by users could help improve this infrastructure, and lead to better performance on a global scale. Our analysis shows that tags can be interpreted as markers of a video's geographic diffusion, with some tags strongly linked to well identified geographic areas. Based on our findings, we demonstrate the potential of tags to help predict distribution of a video's views, and present results suggesting that tags can help place videos in globally distributed datacenters. We show in particular that even a simplistic approach based on tags can help predict a minimum of 65.9% of a video's views for a majority of videos, and that a simple tag-based placement strategy is able to improve the hit rate of a distributed on-line video service by up to 6.8% globally over a naive random allocation.

6.3.10. Mignon: A Fast Decentralized Content Consumption Estimation in Large-Scale Distributed Systems

Participants: Stephane Delbruel, Davide Frey, Francois Taiani.

Although many fully decentralized content distribution systems have been proposed, they often lack key capabilities that make them difficult to deploy and use in practice. In this work [31], we look at the particular problem of content consumption prediction, a crucial mechanism in many such systems. We propose a novel, fully decentralized protocol that uses the tags attached by users to on-line content, and exploits the properties of self-organizing kNN overlays to rapidly estimate the potential of a particular content without explicit aggregation.

6.4. Privacy in User Centric Applications

6.4.1. Hybrid Recommendations with Dynamic Similarity Measure

Participants: Anne-Marie Kermarrec, Nupur Mittal.

This project aims to combine the classical methods of content based and collaborative filtering recommendations, in addition to dynamic similarity computations. The objective is to exploit the varied item-data available from the world wide web, to overcome trivial problems like that of cold-start. In this work, we have designed a new similarity metric inspired from the existing DICE similarity that takes into account changing item/user behavior to compute updated similarity values for the purpose of recommendations. The work leverages the idea of content based recommendations as a first step to create vivid user and item profiles that are iteratively updated.

This work was done in collaboration with Rachid Guerraoui (EPFL, Switzerland), Rhicheck Patra (EPFL, Switzerland).

6.4.2. Lightweight Privacy-Preserving Averaging for the Internet of Things

Participants: Davide Frey, George Giakkoupis, Julien Lepiller.

The number of connected devices is growing continuously, and so is their presence into our everyday lives. From GPS-enabled fitness trackers, to smart fridges that tell us what we need to buy at the grocery store, connected devices—things—have the potential to collect and make available significant amounts of information. On the one hand, this information may provide useful services to users, and constitute a statistical gold mine. On the other, its availability poses serious privacy threats for users. In this work, we designed two new protocols that make it possible to aggregate personal information collected by smart devices in the form of an average, while preventing attackers from learning the details of the non-aggregated data. The first protocol exploits randomness and decomposition into shares as techniques to obfuscate the value associated with each node and lightweight encryption techniques to withstand eavesdropping attacks. The second exploits only randomness and encryption. We carried out a preliminary evaluation and published the results related to the first protocol in [18].

This work was done in collaboration with Tristan Allard from the DRUID Team at IRISA, Rennes.

6.4.3. Collaborative Filtering Under a Sybil Attack: Similarity Metrics do Matter!

Participants: Davide Frey, Anne-Marie Kermarrec, Antoine Rault, Florestan de Moor.

Whether we are shopping for an interesting book or selecting a movie to watch, the chances are that a recommendation system will help us decide what we want. Recommendation systems collect information about our own preferences, compare them to those of other users, and provide us with suggestions on a variety of topics. But is the information gathered by a recommendation system safe from potential attackers, be them other users, or companies that access the recommendation system? And above all, can service providers protect this information while still providing effective recommendations? In this work, we analyze the effect of Sybil attacks on collaborative-filtering recommendation systems, and discuss the impact of different similarity metrics in the trade-off between recommendation quality and privacy. Our results, on a state-of-the-art recommendation framework and on real datasets show that existing similarity metrics exhibit a wide range of behaviors in the presence of Sybil attacks. Yet, they are all subject to the same trade off: Sybil resilience for recommendation quality. We therefore propose and evaluate a novel similarity metric that combines the best of both worlds: a low RMSE score with a prediction accuracy for Sybil users of only a few percent. A preliminary version of this work was published at EuroSec 2015 [57]. This year, we significantly extended the work during the summer internship of Florestan De Moor. Specifically, we considered new attacks that specifically target our novel similarity metric and showed that regardless of the attack configuration, our metric can preserve the privacy of users without hampering recommendation quality. A new paper with these new results was submitted to PETS 2017.

6.4.4. Privacy-Preserving Distributed Collaborative Filtering

Participants: Davide Frey, Anne-Marie Kermarrec.

In this work, we propose a new mechanism to preserve privacy while leveraging user profiles in distributed recommender systems. Our mechanism relies on (i) an original obfuscation scheme to hide the exact profiles of users without significantly decreasing their utility, as well as on (ii) a randomized dissemination protocol ensuring differential privacy during the dissemination process.

We compare our mechanism with a non-private as well as with a fully private alternative. We consider a real dataset from a user survey and report on simulations as well as planetlab experiments. We dissect our results in terms of accuracy and privacy trade-offs, bandwidth consumption, as well as resilience to a censorship attack. In short, our extensive evaluation shows that our twofold mechanism provides a good trade-off between privacy and accuracy, with little overhead and high resilience.

This work was done with Antoine Boutet and Arnaud Jegou when they were part of the team, and in collaboration with Rachid Guerraoui from EPFL. But the complete results were published this year in [15].

7. Bilateral Contracts and Grants with Industry

7.1. Contract with Technicolor

Participants: Fabien Andre, Anne-Marie Kermarrec.

We had a contract with Technicolor for collaboration on large-scale infrastructure for recommendation systems. In this context, Anne-Marie Kermarrec was the PhD advisor of Fabien Andre until Nov 2016. In his Phd, Fabien Andre worked on efficient algorithms for heterogeneous data on large-scale platforms.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR project SocioPlug

Participants: Davide Frey, Anne-Marie Kermarrec, Pierre-Louis Roman, Francois Taiani.

SocioPlug is a collaborative ANR project involving Inria (ASAP team), the Univ. Nantes, and LIRIS (INSA Lyon and Univ. Claude Bernard Lyon). The project emerges from the observation that the features offered by the Web 2.0 or by social media do not come for free. Rather they bring the implicit cost of privacy. Users are more or less consciously selling personal data for services. SocioPlug aims to provide an alternative for this model by proposing a novel architecture for large-scale, user centric applications. Instead of concentrating information of cloud platforms owned by a few economic players, we envision services made possible by cheap low-end plug computers available in every home or workplace. This will make it possible to provide a high amount of transparency to users, who will be able to decide their own optimal balance between data sharing and privacy.

8.1.2. *DeScenT CominLabs*

Participants: Resmi Ariyattu Chandrasekharannair, Davide Frey, Michel Raynal, Francois Taiani.

The DeScenT project aims to ease the writing of distributed programs on a federation of plug computers. Plug computers are a new generation of low-cost computers, such as Raspberry pi (25\$), VIA- APC (49\$), and ZERO Devices Z802 (75\$), which offer a cheap and readily available infrastructure to deploy domestic on-line software. Plug computers open the opportunity for everyone to create cheap nano-clusters of domestic servers, host data and services and federate these resources with their friends, colleagues, and families based on social links. More particularly we will seek in this project to develop novel decentralized protocols than can encapsulate the notion of privacy-preserving federation in plug-based infrastructures. The vision is to use these protocols to provide a programming toolkit that can support the convergent data types being developed by our partner GDD (Gestion de Données Distribuées) at Univ. Nantes.

8.1.3. *ANR Blanc project Displexity*

Participants: George Giakkoupis, Anne-Marie Kermarrec, Michel Raynal.

The Displexity project started in 2011. The aim of this ANR project that also involves researchers from Paris and Bordeaux is to establish the scientific foundations for building up a consistent theory of computability and complexity for distributed computing. One difficulty to be faced by DISPLEXITY is to reconcile two non necessarily disjoint sub-communities, one focusing on the impact of temporal issues, while the other focusing on the impact of spatial issues on distributed algorithms.

8.1.4. *ANR project PAMELA*

Participants: Davide Frey, George Giakkoupis, Francois Taiani.

PAMELA is a collaborative ANR project involving ASAP, Inria Lille, UMPC, Mediego and Snips. The project aims at developing machine learning theories and algorithms in order to learn local and personalized models from data distributed over networked infrastructures. This project seeks to provide first answers to modern information systems built by interconnecting many personal devices holding private user data in the search of personalized suggestions and recommendations. More precisely, we will focus on learning in a collaborative way with the help of neighbors in a network. We aim to lay the first blocks of a scientific foundation for these new types of systems, in effect moving from graphs of data to graphs of data and learned models. We argue that this shift is necessary in order to address the new constraints arising from the decentralization of information that is inherent to the emergence of big data. We will in particular focus on the question of learning under communication and privacy constraints. A significant asset of the project is the quality of its industrial partners, SNIPS and MEDIEGO, who bring in their expertise in privacy protection and distributed computing as well as use cases and datasets. They will contribute to translate this fundamental research effort into concrete outcomes by developing personalized and privacy-aware assistants able to provide contextualized recommendations on small devices and smartphones.

8.1.5. *ANR project OBrowser*

Participants: David Bromberg, Davide Frey, Francois Taiani.

OBrowser is a collaborative ANR project involving Inria (ASAP team), the Univ. Nantes, the Bretagne Sud. University, and Orange. The project emerges from the vision of designing and deploying distributed application on millions of machines using web-enabled technologies without relying on a cloud or a central authority. OBrowser proposes to build collaborative applications through a decentralized execution environment composed of users' browsers that autonomously manages issues such as communication, naming, heterogeneity, and scalability. The introduction of browser-to-browser communication with WebRTC's Datachannel has made these scenarios closer, but today only experts can afford to tackle the technical challenges associated with large-scale browser-based deployments such as decentralized instant-messaging (Firechat) and Infrastructure-less Mission Critical Push To Talk. OBrowser aims to solve these challenges by means of a novel programming framework.

8.1.6. ANR project *DESCARTES*

Participants: George Giakkoupis, Michel Raynal, Francois Taiani.

DESCARTES is a collaborative ANR project involving ASAP, Labri (U. Bordeaux), Lafia (U. Paris Diderot), Vérimag (Grenoble), LIF (Marseilles), and LINA (Nantes). Despite the practical interests of reusable frameworks for implementing specific distributed services, many of these frameworks still lack solid theoretical bases, and only provide partial solutions for a narrow range of services. In this project, we argue that this is mainly due to the lack of a generic framework that is able to unify the large body of fundamental knowledge on distributed computation that has been acquired over the last 40 years. The DESCARTES project aims at bridging this gap, by developing a systematic model of distributed computation that organizes the functionalities of a distributed computing system into reusable modular constructs assembled via well-defined mechanisms that maintain sound theoretical guarantees on the resulting system. DESCARTES arises from the strong belief that distributed computing is now mature enough to resolve the tension between the social needs for distributed computing systems, and the lack of a fundamentally sound and systematic way to realize these systems.

8.1.7. ANR-ERC Tremplin project *NDFUSION*

Participant: George Giakkoupis.

NDFUSION is an 18-month ANR project awarded to the PI to support his preparation for his upcoming ERC grant application. The idea of intervening in a network diffusion process to enhance or retard its spread has been studied in various contexts, e.g., to increase the spread or speed of diffusion by choosing an appropriate set of seed nodes (a standard goal in viral marketing by word-of-mouth), or achieve the opposite effect either by choosing a small set of nodes to remove (a goal in immunization against diseases), or by seeding a competing diffusion (e.g., to limit the spread of misinformation in a social network). The aim of this project is to consolidate existing work under a single, comprehensive framework, and using this framework to develop new, efficient algorithms for optimizing (maximizing or minimizing) the spread of diffusion processes. Novel aspects of the project involve issues of scalability, multiple concurrent diffusions, and the use of multistage online strategies to optimize diffusions. Results from this project are likely to be relevant to many different disciplines, from network optimization in computing to disease containment in medicine.

8.2. International Initiatives

8.2.1. Inria International Labs

Anne-Marie Kermarrec has been scientific collaborator at EPFL, Lausanne, since Feb 2014.

Anne-Marie Kermarrec has been the scientific coordinator of the EPFL/Inria International Lab since Feb 2015.

8.3. International Research Visitors

8.3.1. Visits of International Scientists

Antonio Carzaniga, Università della Svizzera italiana (USI), Apr 29.

Evangelos Bampas, Aix-Marseille Université, Mar 23.

Fábio Moreira Costa, Institute of Informatics, Federal University of Goiás, Goiânia-GO, Brazil, Sep 12–16.

Ricardo Couto Antunes da Rocha, Department of Computer Science, Federal University of Goiás, Catalão-GO, Brazil, Sep 12–16.

Rachid Guerraoui, EPFL, Switzerland, invited Professor at Univ. Rennes I / ISTIC, since September 2016.

Arvid Bosk, KTH Royal Institute of Technology, guest PhD Student, from Dec 2016.

8.3.1.1. Internships

Florestan De Moor, Collaborative Filtering Under a Sybil Attack, Univ. Rennes I, May to Jul 2016, supervised by Davide Frey.

Julien Lepiller, Private Decentralized Aggregation, Inria, Feb to Jun 2016, supervised by Davide Frey and Francois Taiani.

8.3.2. Visits to International Teams

8.3.2.1. Research Stays Abroad

David Bromberg visited UFG, Goiania, Brazil, from Jun to Jul 2016 (CONFAP CNRS project)

8.3.2.2. Internships

Pierre-Louis Roman did an internship at Vrije Universiteit Amsterdam, The Netherlands, from Jun to Aug 2016 under the supervision of Spyros Voulgaris, with a grant from Université Bretagne Loire.

Simon Bouget did an internship at Centre for Complex Systems and Big Data at University of Neuchatel, Switzerland, from May to Aug 2016, under the supervision of Etienne Rivière, with a grant from the “Outgoing Mobility for Doctoral Students” program of Rennes Métropole.

Stéphane Delbruel did an internship at Università della Svizzera Italiana, Switzerland, from Jun to Jul 2016, under the supervision of Antonio Carzaniga.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

George Giakkoupis was the General Chair of the 35th ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC 2016), Chicago, IL, USA, Jul 2016.

9.1.2. Scientific Events Selection

9.1.2.1. Member of Conference Program Committees

Pierre-Louis Roman served in the program committee of the 6th IEEE International Symposium on Cloud and Service Computing (SC2), Nadi, Fiji, Dec 2016.

Francois Taiani served in the program committee of the:

17th ACM/IFIP/USENIX International Conference on Middleware (Middleware), Trento, Italy, Dec 2016.

15th Workshop on Adaptive and Reflective Middleware (ARM'16), Trento, Italy, Dec 2016.

4th workshop on Distributed Cloud Computing (DCC), Chicago, USA, Jul 2016

International Workshop on Assurance in Distributed Systems and Networks (ADSN 2016), Nara, Japan, Jun 2016.

IEEE International Conference on Cloud Engineering 2016 (IC2E 2016), Berlin, Germany, Apr 2016.

David Bromberg served in the program committee of the:

17th ACM/IFIP/USENIX International Conference on Middleware (Middleware), Trento, Italy, Dec 2016.

IEEE International Conference on Cloud Engineering 2016 (IC2E 2016), Berlin, Germany, Apr 2016.

Davide Frey served in the program committee of the

11th ACM International Conference on Distributed and Event-Based Systems (DEBS 2017), Barcelona, Spain, Jun 2017.

17th ACM/IFIP/USENIX International Conference on Middleware (Middleware), Trento, Italy, Dec 2016.

George Giakkoupis served in the program committee of the:

17th International Conference on Distributed Computing and Networking (ICDCN 2017), Singapore, Jan 2016.

3rd ACM-W Europe Celebration of Women in Computing (womENCourage 2016), Linz, Austria, Sep 2016.

31st IEEE International Parallel & Distributed Processing Symposium (IPDPS 2017), Orlando, FL, USA, May 2017.

9.1.3. Journal

9.1.3.1. Member of Editorial Boards

Anne-Marie Kermarrec is an associate editor of IEEE Internet Computing.

Anne-Marie Kermarrec is an associate editor of the Springer Computing Journal.

9.1.4. Invited Talks

Michel Raynal, A Look at Basics of Distributed Computing, Keynote Talk in the 36th International Conference on Distributed Computing Systems (ICDCS 2016), Nara, Japan, June 2016.

Pierre-Louis Roman, Bringing secure Bitcoin transactions to your smartphone, Technicolor, Rennes, Nov 24 2016.

Francois Taiani, Being prepared in a sparse world: The case of KNN graph construction, Invited Seminar of the Département Informatique et télécommunications (DIT) of ENS Rennes, Sep 2016.

Francois Taiani, Being prepared in a sparse world: The case of KNN graph construction, Invited Seminar of the Inria Project Team TAO, Inria Saclay, May 2016.

David Bromberg, Towards microservices architecture to transcode videos in the large at low costs, Invited seminar of the computer science department of UFG - Instituto de Informatica, Goiânia, Brasil, Jun 2016.

George Giakkoupis, Randomized adversary models. CMO-BIRS Workshop on Complexity and Analysis of Distributed Algorithms, Oaxaca, Mexico, Dec 1 2016.

George Giakkoupis, Amplifiers and suppressors for the Moran process on undirected graphs. ART-Oberseminar, University of Hamburg, Germany, Nov 8 2016.

George Giakkoupis, Amplifiers and suppressors of evolutionary dynamics on undirected graphs. Workshop on Random Processes in Discrete Structures, University of Warwick, UK, Aug 31 2016.

George Giakkoupis, Recent advances in randomized rumor spreading. Nexus of Information and Computation Theories, Distributed Computation and Communication Theme, The Henri Poincare Institute (IHP), Paris, France, Feb 1 2016.

Davide Frey gave an Invited (Keynote) talk at the W-PSDS workshop co-located with SRDS 2016 in Budapest, Hungary, on Sep 26, 2016.

Davide Frey gave an Invited talk at the SG-2 meeting of the BDVA Summit in Valencia, Spain, on Dec 1, 2016.

9.1.5. Leadership within the Scientific Community

Anne-Marie Kermarrec has been a member of the Academia Europea since 2013.

Anne-Marie Kermarrec is a member of the scientific committee of the Société Informatique de France.

George Giakkoupis was a member of the Steering Committee of the 35th ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC 2016), Chicago, IL, USA, Jul 2016.

Francois Taiani has been a member of the Steering Committee of IFIP WG 6.1 International Conference on Distributed Applications and Interoperable Systems (DAIS) since 2013.

Francois Taiani has been a member of the Steering Committee of the ACM/IFIP/USENIX International Conference on Middleware (Middleware) since 2014.

Francois Taiani served on the hiring committee for the position of Maître de Conférence a the University of Perpignan – Via Domitia in 2016.

9.1.6. Research Administration

David Bromberg is member of the scientific committee of the Media & Networks competitiveness cluster.

David Bromberg is scientific correspondent for ICT Digital in Rennes.

Davide Frey is “correspondant scientifique Europe” at the DPEI for Inria Rennes.

Davide Frey is an associate member of the COST-GTRI of Inria.

Anne-Marie Kermarrec has been a member of the ERC panel for Consolidator Grants since 2013.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence : Olivier Ruas, Initiation au Génie Logiciel (GEN), 40, L2, Univ. Rennes I (ISTIC), France

Licence : Olivier Ruas, Algorithmique Fonctionnelle (APF), 24, L1, Univ. Rennes I (ISTIC), France

Engineering School : Simon Bouget, SYS1 (Introduction aux Systèmes d’exploitation), 6 HEQTD (3x 2 heures de TP), ESIR1 (L3), Univ. Rennes I (ESIR), France

Master : Simon Bouget, TLC (Techniques Logicielles pour le Cloud computing), 18 HEQTD (7x 2 heures de TP + 2x 2 heures de TD), M2 Génie Logiciel Alternance, Univ. Rennes I (ISTIC), France

Engineering School: Francois Taiani, Synchronization and Parallel Programming, 34h, 2nd year of Engineering School (M1), ESIR / Univ. Rennes I, France.

Engineering School: Francois Taiani, Distributed Systems, 24h, 3rd year of Engineering School (M2), ESIR / Univ. Rennes I, France.

Engineering School: Francois Taiani, Parallel Algorithms for Big Data, 24h, 3rd year of Engineering School (M2), ESIR / Univ. Rennes I, France.

Engineering School: Francois Taiani, Introduction to Operating Systems, 24h, 1st year of Engineering School (L3), ESIR / Univ. Rennes I, France.

Master: Francois Taiani, Programming Technologies for the Cloud, 14h, M2, Univ. Rennes I, France.

Engineering School: David Bromberg, tools and programming languages for the Web, 48h, 2nd year of Engineering School (M1), ESIR / Univ. Rennes I, France.

Engineering School: David Bromberg, Distributed software architecture, 24h, 2nd year of Engineering School (M1), ESIR / Univ. Rennes I, France.

Engineering School: David Bromberg, Network security, 48h, 2nd year of Engineering School (M1), ESIR / Univ. Rennes I, France.

Engineering School: David Bromberg, Cloud for the Internet of Things, 24h, 3rd year of Engineering School (M2), ESIR / Univ. Rennes I, France.

Master: Davide Frey, Scalable Distributed Systems, 10 hours, M1, EIT/ICT Labs Master School, Univ. Rennes I, France.

9.2.2. Supervision

PhD: Antoine Rault, User privacy in collaborative filtering systems [13], Univ. Rennes I, Jun 23, 2016, Anne-Marie Kermarrec and Davide Frey.

PhD: Fabien Andre, Exploiting Modern Hardware for Large Scale Nearest Neighbor Search [52],¹ Nov 25 2016, Anne-Marie Kermarrec (and Nicolas Le Scouarnec–Technicolor).

PhD: Nupur Mittal, Data, Learning & Privacy in Recommendation Systems [11], Nov 25, 2016, George Giakkoupis and Anne-Marie Kermarrec.

PhD: Javier Olivares, Scaling Out-of-Core K-Nearest Neighbors Computation on Single Machines [12], Dec 19 2016, Anne-Marie Kermarrec.

PhD in progress : Pierre-Louis Roman, Epidemic Distributed Convergence for Decentralized Social Networking, Oct 2014, Francois Taiani and Davide Frey.

PhD in progress : Stéphane Delbruel, Towards a Decentralized Embryomorphic Storage System, Oct 2013, Francois Taiani and Davide Frey (since Oct 2014).

PhD in progress : Resmi Ariyattu Chandrasekharannair, Towards Decentralized Federations for Plug-based Decentralized Social Networks, Dec 2013, Francois Taiani.

PhD in progress : Olivier Ruas, Dynamic Learning and Recommendations in Very Large Distributed Computing Infrastructures, Oct 2015, Anne-Marie Kermarrec and Francois Taiani.

PhD in progress: Simon Bouget, EMILIO: Emergent Middleware for Extra-Large-Scale Self Adaptation, Sep 2014, Francois Taiani and David Bromberg (since Sep 2015).

PhD in progress: Adrien Luxey, Towards New Solutions to Build Large Scale Distributed Applications in the Cloud, Oct 2016, David Bromberg.

9.2.3. Juries

Francois Taiani was an internal examiner of the PhD of Jackie Bourgeois. Interactive Demand Shifting in the Context of Domestic Micro-Generation. Open University (UK) / Univ. Rennes I, Jun 30 2016.

Francois Taiani was the chair of the PhD panel of Anca Iordache. Performance-Cost Trade-Offs in Heterogeneous Clouds. Univ. Rennes I, Sep 9 2016.

Francois Taiani was a reviewer of the PhD of Milan Kabac. A Design-Driven Methodology for the Development of Large-Scale Orchestrating Applications. Univ. Bordeaux, rapporteur, Sep 26 2016.

Francois Taiani was a reviewer of the PhD of Thomas Hartmann. Enabling Model-Driven Live Analytics For Cyber-Physical Systems: The Case of Smart Grids. Univ. of Luxembourg (Luxembourg), Nov 8 2016.

Francois Taiani was a reviewer of the PhD of Lucas Perrone. Vers des protocoles de tolérance aux fautes Byzantines efficaces et robustes. Univ. Grenoble Alpes, Dec 8 2016.

¹The thesis was not yet available at HAL at the time when this report was submitted.

Francois Taiani was the chair a of the PhD panel of Javier Olivares. Scaling out-of-core K-Nearest Neighbors Computation on Single Machines. Univ. Rennes I, Dec 19 2016.

Francois Taiani was a reviewer of the HDR of Fabien Huet. From HPC to Big Data : Models and Tools for Large Scale Middleware. Université de Nice, Sophia Antipolis, Feb 12 2016.

Francois Taiani was the chair of the HDR panel of Erwan Le Merrer. Structuring Overlays for Distributed Computing: From Edge-Devices to the Clouds. Univ. Rennes I, Nov 23 2016.

George Giakkoupis was a member of the PhD committee of Maryam Helmi. The Space Complexity of Distributed Tasks in the Shared Memory Model. University of Calgary, Jun 7 2016.

George Giakkoupis was a member of the PhD committee member of Scott Roche. Robust Local Algorithms for Communication and Stability in Distributed Networks. Northeastern University, Nov 14 2016.

George Giakkoupis was an external reviewer of the PhD dissertation of Emanuele Natale. On the Computational Power of Simple Dynamics. Sapienza University of Rome, Dec 15 2016.

10. Bibliography

Major publications by the team in recent years

- [1] F. ANDRÉ, A.-M. KERMARREC, E. LE MERRER, N. LE SCOUARNEC, G. STRAUB, A. VAN KEMPEN. *Archiving Cold Data in Warehouses with Clustered Network Coding*, in "EuroSys 2014", Amsterdam, Netherlands, ACM New York, NY, USA, April 2014 [DOI : 10.1145/2592798.2592816], <https://hal.inria.fr/hal-00994660>
- [2] M. BERTIER, D. FREY, R. GUERRAOU, A.-M. KERMARREC, V. LEROY. *The Gossple Anonymous Social Network*, in "ACM/IFIP/USENIX 11th International Middleware Conference", India Bangalore, November 2010
- [3] S. BOUGET, A.-M. KERMARREC, H. KERVADEC, F. TAÏANI. *Polystyrene: the Decentralized Data Shape That Never Dies*, in "The 2014 IEEE 34th International Conference on Distributed Computing Systems (ICDCS 2014)", Madrid, Spain, June 2014, pp. 288 - 297 [DOI : 10.1109/ICDCS.2014.37], <https://hal.inria.fr/hal-01080608>
- [4] J. CAO, M. RAYNAL, X. YANG, W. WU. *Design and Performance Evaluation of Efficient Consensus Protocols for Mobile Ad Hoc Networks*, in "IEEE Transactions on Computers", 2007, vol. 56, n^o 8, pp. 1055–1070
- [5] K. CENSOR-HILLEL, M. GHAFARI, G. GIAKKOUPIS, B. HAEUPLER, F. KUHN. *Tight Bounds on Vertex Connectivity Under Vertex Sampling*, in "26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2015)", San Diego, CA, United States, January 2015, pp. 2006-2018 [DOI : 10.1137/1.9781611973730.133], <https://hal.inria.fr/hal-01250519>
- [6] G. GIAKKOUPIS, P. WOELFEL. *Randomized Mutual Exclusion with Constant Amortized RMR Complexity on the DSM*, in "55th IEEE Symposium on Foundations of Computer Science (FOCS)", Philadelphia, PA, United States, October 2014, <https://hal.inria.fr/hal-01097144>
- [7] M. JELASITY, S. VOULGARIS, R. GUERRAOU, A.-M. KERMARREC, M. VAN STEEN. *Gossip-Based Peer Sampling*, in "ACM Transactions on Computer Systems", August 2007, vol. 41, n^o 5

- [8] A.-M. KERMARREC, F. TAÏANI, J. M. TIRADO MARTIN. *Scaling Out Link Prediction with SNAPLE*, in "16th Annual ACM/IFIP/USENIX Middleware Conference", Vancouver, Canada, December 2015, 12 p. [DOI : 10.1145/2814576.2814810], <https://hal.inria.fr/hal-01244663>
- [9] A. MOSTÉFAOUI, H. MOUMEN, M. RAYNAL. *Signature-Free Asynchronous Binary Byzantine Consensus with $t < n/3$, $O(n^2)$ Messages, and $O(1)$ Expected Time*, in "Journal of the ACM (JACM)", December 2015, vol. 62, pp. 1000-1020 [DOI : 10.1145/2785953], <https://hal.archives-ouvertes.fr/hal-01176110>
- [10] F. TAÏANI, S. LIN, G. S. BLAIR. *GOSSIPKIT: A Unified Component Framework for Gossip*, in "IEEE Transactions on Software Engineering", March 2014, vol. 40, n^o 2, pp. 123-136 [DOI : 10.1109/TSE.2013.50], <https://hal.inria.fr/hal-01080198>

Publications of the year

Doctoral Dissertations and Habilitation Theses

- [11] N. MITTAL. *Data, learning & privacy in recommandation systems*, Inria Rennes - Bretagne Atlantique ; ASAP, November 2016, <https://hal.inria.fr/tel-01419886>
- [12] J. OLIVARES. *Scaling out-of-core K-Nearest Neighbors Computation on Single Machines*, Inria Rennes, December 2016, <https://hal.inria.fr/tel-01421362>
- [13] A. RAULT. *User privacy in collaborative filtering systems*, Université Rennes 1, June 2016, <https://tel.archives-ouvertes.fr/tel-01385503>

Articles in International Peer-Reviewed Journals

- [14] M. BERTIER, M. PERRIN, C. TEDESCHI. *On the Complexity of Concurrent Multiset Rewriting*, in "International Journal of Foundations of Computer Science", 2016, vol. 27, n^o 1 [DOI : 10.1142/S0129054116500052], <https://hal.inria.fr/hal-01326849>
- [15] A. BOUTET, D. FREY, R. GUERRAOUI, A. JÉGOU, A.-M. KERMARREC. *Privacy-Preserving Distributed Collaborative Filtering*, in "Computing", August 2016, vol. 98, n^o 8, <https://hal.inria.fr/hal-01251314>
- [16] B. LE GAL, Y.-D. BROMBERG, L. RÉVEILLÈRE, J. SOLANKI. *A Flexible SoC and Its Methodology for Parser-Based Applications*, in "ACM Transactions on Reconfigurable Technology and Systems (TRETTS)", 2016, vol. 10, n^o 1, 4 p. [DOI : 10.1145/2939379], <https://hal.archives-ouvertes.fr/hal-01415653>

Invited Conferences

- [17] M. RAYNAL. *A Look at Basics of Distributed Computing **, in "IEEE ICDCS 2016 - 36th International Conference on Distributed Computing Systems", Nara, Japan, June 2016, <https://hal.inria.fr/hal-01337523>

International Conferences with Proceedings

- [18] T. ALLARD, D. FREY, G. GIAKKOUPIS, J. LEPILLER. *Lightweight Privacy-Preserving Averaging for the Internet of Things*, in "M4IOT 2016 - 3rd Workshop on Middleware for Context-Aware Applications in the IoT", Trento, Italy, ACM, December 2016, pp. 19 - 22 [DOI : 10.1145/3008631.3008635], <https://hal.inria.fr/hal-01421986>

- [19] F. ANDRÉ, A.-M. KERMARREC, N. LE SCOUARNEC. *Cache locality is not enough: High-Performance Nearest Neighbor Search with Product Quantization Fast Scan*, in "42nd International Conference on Very Large Data Bases", New Delhi, India, September 2016, vol. 9, n^o 4, 12 p. , <https://hal.inria.fr/hal-01239055>
- [20] O. BARAIS, J. BOURCIER, Y.-D. BROMBERG, C. DION. *Towards microservices architecture to transcode videos in the large at low costs*, in "TEMU 2016 - International Conference on Telecommunications and Multimedia", Heraklion, Greece, July 2016, pp. 1 - 6 [DOI : 10.1109/TEMU.2016.7551918], <https://hal.inria.fr/hal-01427277>
- [21] Y. BENKAOUZ, M. ERRADI, A.-M. KERMARREC. *Nearest Neighbors Graph Construction: Peer Sampling to the Rescue*, in "4th International Conference, NETYS 2016, Marrakech, Morocco, May 18-20, 2016, Revised Selected Papers", Marrakech, Morocco, Networked Systems, Springer, May 2016, vol. 9944, pp. 48 - 62 [DOI : 10.1007/978-3-319-46140-3_4], <https://hal.inria.fr/hal-01407514>
- [22] P. BERENBRINK, T. FRIEDETZKY, G. GIAKKOUPIS, P. KLING. *Efficient plurality consensus, or: The benefits of cleaning up from time to time*, in "43rd International Colloquium on Automata, Languages and Programming (ICALP 2016)", Rome, Italy, July 2016 [DOI : 10.4230/LIPIcs.ICALP.2016.136], <https://hal.archives-ouvertes.fr/hal-01353690>
- [23] P. BERENBRINK, G. GIAKKOUPIS, A.-M. KERMARREC, F. MALLMANN-TRENN. *Bounds on the Voter Model in Dynamic Networks*, in "ICALP 2016 - 43rd International Colloquium on Automata, Languages and Programming", Rome, Italy, July 2016 [DOI : 10.4230/LIPIcs.ICALP.2016.146], <https://hal.archives-ouvertes.fr/hal-01353695>
- [24] S. BOUGET. *Position paper: Toward an holistic approach of Systems of Systems*, in "Middleware 2016 – Doctoral Symposium", Trento, Italy, December 2016 [DOI : 10.1145/3009925.3009935], <https://hal.inria.fr/hal-01419712>
- [25] A. BOUTET, A.-M. KERMARREC, N. MITTAL, F. TAÏANI. *Being prepared in a sparse world: the case of KNN graph construction*, in "International Conference on Data Engineering (ICDE)", Helsinki, Finland, May 2016, <https://hal.inria.fr/hal-01251010>
- [26] A. CASTAÑEDA, C. DELPORTE-GALLET, H. FAUCONNIER, S. RAJSBAUM, M. RAYNAL. *Making Local Algorithms Wait-Free: The Case of Ring Coloring*, in "SSS", Lyon, France, Stabilization, Safety, and Security of Distributed Systems, November 2016, vol. 10083, 16 p. , <https://hal.inria.fr/hal-01416522>
- [27] *Best Paper*
N. CHILUKA, A.-M. KERMARREC, J. OLIVARES. *The Out-of-core KNN Awakens: The light side of computation force on large datasets*, in "The International Conference on Networked Systems NETYS", Marrakech, Morocco, May 2016, <https://hal.inria.fr/hal-01336673>.
- [28] C. DELPORTE-GALLET, H. FAUCONNIER, M. RAYNAL, S. RAJSBAUM. *Implementing Snapshot Objects on Top of Crash-Prone Asynchronous Message-Passing Systems*, in "Algorithms and Architectures for Parallel Processing (ICA3PP)", Granada, France, Algorithms and Architectures for Parallel Processing Volume 10048 of the series Lecture Notes in Computer Science pp 341-355, Springer, December 2016, vol. 10048, 15 p. , <https://hal.inria.fr/hal-01416498>

- [29] C. DELPORTE-GALLET, H. FAUCONNIER, M. RAYNAL, S. RAJSBAUM. *t-Resilient Immediate Snapshot Is Impossible*, in "SIROCCO", Helsinki, France, Structural Information and Communication Complexity, July 2016, vol. 9988, 15 p. , <https://hal.inria.fr/hal-01416515>
- [30] *Best Paper*
S. DELBRUEL, D. FREY, F. TAÏANI. *Exploring The Use of Tags for Georeplicated Content Placement*, in "IEEE IC2E'16", Berlin, Germany, April 2016, <https://hal.inria.fr/hal-01257939>.
- [31] S. DELBRUEL, D. FREY, F. TAÏANI. *Mignon: a Fast Decentralized Content Consumption Estimation in Large-Scale Distributed Systems*, in "16th IFIP International Conference on Distributed Applications and Interoperable Systems DAIS'16", Heraklion, Greece, June 2016, <https://hal.inria.fr/hal-01301230>
- [32] D. DIDONA, N. DIEGUES, R. GUERRAOU, A.-M. KERMARREC, R. NEVES, P. ROMANO. *ProteusTM: Abstraction Meets Performance in Transactional Memory*, in "Twenty First International Conference on Architectural Support for Programming Languages and Operating Systems", Atlanta, United States, April 2016 [DOI : 10.1145/2872362.2872385], <https://hal.inria.fr/hal-01250459>
- [33] D. FREY, M. X. MAKKES, P.-L. ROMAN, F. TAÏANI, S. VOULGARIS. *Bringing secure Bitcoin transactions to your smartphone*, in "Proceedings of the 15th International Workshop on Adaptive and Reflective Middleware (ARM 2016)", Trento, Italy, ACM, December 2016, pp. 3:1-3:6 [DOI : 10.1145/3008167.3008170], <https://hal.inria.fr/hal-01384461>
- [34] D. FREY, A. MOSTEFAOUI, M. PERRIN, P.-L. ROMAN, F. TAÏANI. *Speed for the elite, consistency for the masses: differentiating eventual consistency in large-scale distributed systems*, in "Proceedings of the 2016 IEEE 35th Symposium on Reliable Distributed Systems (SRDS 2016)", Budapest, Hungary, IEEE Computer Society, September 2016, pp. 197-206 [DOI : 10.1109/SRDS.2016.30], <https://hal.inria.fr/hal-01344138>
- [35] G. GIAKKOUPIS, Y. NAZARI, P. WOELFEL. *How asynchrony affects rumor spreading time*, in "35th ACM Symposium on Principles of Distributed Computing (PODC 2016)", Chicago, United States, July 2016 [DOI : 10.1145/2933057.2933117], <https://hal.archives-ouvertes.fr/hal-01353698>
- [36] D. LIMA, H. MIRANDA, F. TAÏANI. *Partial Replication Policies for Dynamic Distributed Transactional Memory in Edge Clouds*, in "1st Workshop on Middleware for Edge Clouds & Cloudlets (MECC'16)", Trento, Italy, December 2016, pp. 1 - 6 [DOI : 10.1145/3017116.3022872], <https://hal.inria.fr/hal-01420708>

National Conferences with Proceedings

- [37] J. L. RODAS, J. OLIVARES, J. A. GALINDO DUARTE, D. BENAVIDES. *Hacia el uso de sistemas de recomendación en sistemas de alta variabilidad*, in "CEDI 2016", Salamanca, Spain, September 2016, <https://hal.inria.fr/hal-01342353>

Conferences without Proceedings

- [38] R. ARIYATTU, F. TAÏANI. *Filament : a cohort construction service for decentralized collaborative editing platforms*, in "Compas 2016", Lorient, France, July 2016, <https://hal.archives-ouvertes.fr/hal-01333513>
- [39] H. LAKHLEF, D. FREY, M. RAYNAL. *Optimal Collision/Conflict-Free Distance-2 Coloring in Wireless Synchronous Broadcast/Receive Tree Networks*, in "45th International Conference on Parallel Processing",

Philadelphia, PA, United States, August 2016 [DOI : 10.1109/ICPP.2016.47], <https://hal.archives-ouvertes.fr/hal-01396940>

Research Reports

- [40] D. FREY, A. MOSTEFAOUI, M. PERRIN, A.-M. KERMARREC, C. MADDOCK, A. MAUTHE, P.-L. ROMAN, F. TAÏANI. *D.1.3 – Protocols for emergent localities*, IRISA ; LINA-University of Nantes ; Inria Rennes Bretagne Atlantique ; Lancaster University, July 2016, n^o D1.3, This report has been written by members of the consortium from IRISA and LINA partner, <https://hal.archives-ouvertes.fr/hal-01343348>
- [41] H. LAKHLEF, M. RAYNAL, F. TAÏANI. *Vertex Coloring with Communication and Local Memory Constraints in Synchronous Broadcast Networks*, IRISA ; Université de Rennes 1, 2016, n^o 2035, 23 p. , <https://hal.inria.fr/hal-01300095>
- [42] A. MOSTEFAOUI, M. RAYNAL. *Time-Efficient Read/Write Register in Crash-prone Asynchronous Message-Passing Systems*, IRISA, January 2016, 14 p. , <https://hal.inria.fr/hal-01256067>
- [43] M. RAYNAL. *Distributed Universal Constructions: a Guided Tour*, IRISA, November 2016, n^o 2040, 23 p. , <https://hal.inria.fr/hal-01397265>

Other Publications

- [44] A. CASTAÑEDA, C. DELPORTE, H. FAUCONNIER, S. RAJSBAUM, M. RAYNAL. *Wait-freedom and Locality are not Incompatible (with Distributed Ring Coloring as an Example)*, January 2016, working paper or preprint, <https://hal.inria.fr/hal-01265958>
- [45] C. DELPORTE, H. FAUCONNIER, S. RAJSBAUM, M. RAYNAL. *Implementing Snapshot Objects on Top of Crash-Prone Asynchronous Message-Passing Systems*, May 2016, working paper or preprint, <https://hal.inria.fr/hal-01313584>
- [46] C. DELPORTE, H. FAUCONNIER, S. RAJSBAUM, M. RAYNAL. *t-Resilient Immediate Snapshot is Impossible*, May 2016, working paper or preprint, <https://hal.inria.fr/hal-01315342>
- [47] G. GIAKKOUPIS. *Amplifiers and Suppressors of Selection for the Moran Process on Undirected Graphs*, November 2016, working paper or preprint, <https://hal.inria.fr/hal-01393044>
- [48] A.-M. KERMARREC, N. MITTAL, J. OLIVARES. *Multithreading approach to process real-time updates in KNN algorithms*, December 2016, working paper or preprint, <https://hal.inria.fr/hal-01415495>
- [49] A. MOSTEFAOUI, M. RAYNAL. *Two-Bit Messages are Sufficient to Implement Atomic Read/Write Registers in Crash-prone Systems*, February 2016, working paper or preprint, <https://hal.inria.fr/hal-01271135>
- [50] T. PETRIĆ, M. GOESSENS, L. NUAYMI, A. PELOV, L. TOUTAIN. *Measurements, Performance and Analysis of LoRa FABIAN, a real-world implementation of LPWAN*, June 2016, working paper or preprint, <https://hal-institut-mines-telecom.archives-ouvertes.fr/hal-01331966>

References in notes

-
- [51] M. AGUILERA. *A Pleasant Stroll Through the Land of Infinitely Many Creatures*, in "ACM SIGACT News, Distributed Computing Column", 2004, vol. 35, n^o 2
- [52] F. ANDRÉ. *Exploiting Modern Hardware for Large Scale Nearest Neighbor Search*, Université Rennes 1, November 2016
- [53] D. ANGLUIN. *Local and Global Properties in Networks of Processes*, in "Proc. 12th ACM Symposium on Theory of Computing (STOC'80)", 1980
- [54] K. BIRMAN, M. HAYDEN, O. OZKASAP, Z. XIAO, M. BUDIU, Y. MINSKY. *Bimodal Multicast*, in "ACM Transactions on Computer Systems", May 1999, vol. 17, n^o 2, pp. 41-88
- [55] A. DEMERS, D. GREENE, C. HAUSER, W. IRISH, J. LARSON, S. SHENKER, H. STURGIS, D. SWINEHART, D. TERRY. *Epidemic algorithms for replicated database maintenance*, in "Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC'87)", August 1987
- [56] P. EUGSTER, S. HANDURUKANDE, R. GUERRAOU, A.-M. KERMARREC, P. KOUZNETSOV. *Lightweight Probabilistic Broadcast*, in "ACM Transaction on Computer Systems", November 2003, vol. 21, n^o 4
- [57] D. FREY, R. GUERRAOU, A.-M. KERMARREC, A. RAULT. *Collaborative Filtering Under a Sybil Attack: Analysis of a Privacy Threat*, in "Eighth European Workshop on System Security EuroSec 2015", Bordeaux, France, April 2015 [DOI : 10.1145/2751323.2751328], <https://hal.inria.fr/hal-01158723>
- [58] M. MERRITT, G. TAUBENFELD. *Computing Using Infinitely Many Processes*, in "Proc. 14th Int'l Symposium on Distributed Computing (DISC'00)", 2000
- [59] S. RATNASAMY, P. FRANCIS, M. HANDLEY, R. KARP, S. SHENKER. *A Scalable Content-Addressable Network*, in "Conference of the Special Interest Group on Data Communication (SIGCOMM'01)", 2001
- [60] A. ROWSTRON, P. DRUSCHEL. *Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems*, in "IFIP/ACM Intl. Conf. on Distributed Systems Platforms (Middleware)", 2001
- [61] I. STOICA, R. MORRIS, D. KARGER, F. KAASHOEK, H. BALAKRISHNAN. *Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications*, in "SIGCOMM'01", 2001
- [62] S. VOULGARIS, D. GAVIDIA, M. VAN STEEN. *CYCLON: Inexpensive Membership Management for Unstructured P2P Overlays*, in "Journal of Network and Systems Management", 2005, vol. 13, n^o 2