RESEARCH CENTER
**Lille - Nord Europe**

FIELD

Activity Report 2014

# Section Contracts and Grants with Industry

Edition: 2015-03-24

**ATEAMS Project-Team**

# 6. Bilateral Contracts and Grants with Industry

## 6.1. Bilateral Contracts with Industry

- ING co-financed one PhD position in the context of CWI public-private collaboration program. The goal is to apply domain-specific language technology to revitalize core banking infrastructure.
- AimValley won the CWI research voucher for developing a DSL for state machines in the context of embedded devices. Davy Landman performed the research and development.

**BONSAI Project-Team**

# 7. Bilateral Contracts and Grants with Industry

## 7.1. Bilateral Contracts with Industry

- The PhD thesis of Lea Siegwald is funded by a CIFRE contract with the biotechnology company Gene Diffusion.

**DOLPHIN Project-Team**

# 7. Bilateral Contracts and Grants with Industry

## 7.1. Bilateral Contracts with Industry

- **EDF** (2011-2014): the goal of this bilateral contract is to formulate pricing problems of electrical energy using bilevel mathematical programs.
- **EDF** (2011-2014): this contract models and solves scheduling outages of nuclear plants under uncertainty (Phd of N. Dupin).
- **BeTravel** (2012-2014): this CIFRE project deals with the optimization of group travel plannings (Phd of M. Bue).
- **Tasker** (2011-2014): the goal of this CIFRE project is the multi-objective scheduling of applications in public cloud computing systems (Phd of F. Legillon).
- **Strat-Logic** (2012-2015): the objective of this CIFRE contract is the optimization of economic decisions in a competitive business management simulator (Phd of S. Dufourny).
- **Vekia** (2012-2015): the goal of the CIFRE project is to develop an efficient and generic software for employee scheduling in retail (Phd of M. Gérard).
- **PIXEO** (2014-2015): the objective of this bilateral project is the predictive models and knowledge extraction for insurance web comparator.

**DREAMPAL Team**

# 6. Bilateral Contracts and Grants with Industry

## 6.1. Bilateral Contracts with Industry

Collaboration contract with Nolam Embedded Systems: In conjunction with the CIFRE grant of Venkatasubramanian Viswanathan, a collaboration contract is established with Nolam ES. The objective is to design an innovative embedded computing platform supporting massively parallel dynamically reconfigurable execution model. The use-cases of this platform cover several application domains such as medical, transportation and aerospace.

Collaboration contract with NAVYA: In conjunction with the doctoral grant of Karim Ali, a collaboration contract is established with NAVYA. The objective is to design an innovative embedded system dedicated for dynamic obstacle detection and tracking for autonomous vehicle navigation.

**FUN Project-Team**

# 6. Bilateral Contracts and Grants with Industry

## 6.1. Etineo Partnership

**Participants:** Roudy Dagher, Salvatore Guzzo Bonifacio, Nathalie Mitton [correspondant].

EtiPOPS focuses on portability and flexibility of GOLIATH on several hardwares and in different environments (indoor and outdoor) through the deployment of different applications such as geolocalization. In order to favor the portability, designed solutions in EtiPOPS will respect on-going communication standards which will allow a greater interoperability between heterogeneous hardwares. Publications in 2014 in the framework of EtiPOPS are [32], [11] and software modules.

## 6.2. Traxens partnership

**Participants:** Natale Guzzo, Nathalie Mitton [correspondant].

This collaboration aims to set up a full protocol stack for TRAXENS's guideline.

**LINKS Team**

# 6. Bilateral Contracts and Grants with Industry

## 6.1. Bilateral Contracts with Industry

Innovimax, Cifre and Engineer (2010-2014) The PhD thesis of Tom SEBASTIAN within the QUIXPROC project is supervised byJ.NIEHREN in cooperation with M.ZERGAOUI the head of the INNOVIMAX company. The software development in this context is supported by D. DEBARBIEUX, a senior engineer co-funded by INNOVIMAX and Inria.

**MAGNET Team**

# 7. Bilateral Contracts and Grants with Industry

## 7.1. Bilateral Grants with Industry

### *7.1.1. Cifre SAP (2011-2014)*

**Participants:** Thomas Ricatte, Marc Tommasi, Rémi Gilleron [correspondent].

In business intelligence information systems, one of the first tasks is to acquire and clean internal data and then enrich them with additional sources of informations. This preprocessing step is well planned and specialized for fixed analysis and fixed dashboards. The subject of our collaboration with SAP was included in general objective that (i) specializes this preprocessing task in order to deal with external data coming from networked data like social networks and open relational data, and (ii) simplifies the adaptation of the processing step evolving data analysis tasks. We have focused on the task of merging information acquired from many input data sources represented as graphs, with the final objective of providing a unique graph representation of all data or data models. This research has lead to new graph combination algorithms, but has also raised the need for representing and managing high order relations using graph-like techniques.

RÉMI GILLERON supervises the PhD thesis (Cifre) of Thomas Ricatte together with Yannick Cras from SAP.

### *7.1.2. Cifre Clic and Walk (2013-2016)*

**Participants:** Pauline Wauquier, Marc Tommasi, Mikaela Keller [correspondent].

We start a one to one cooperation with the CLIC AND WALK company that makes marketing surveys by consumers (called clicwalkers). The goal of the company is to understand the community of clicwalkers (40 thousands in one year) and its evolution with two objectives: the first one is to optimize the attribution of surveys to clicwalkers, and the second is to expand company's market to foreign countries. Social data can be obtained from social networks (G+, Facebook, ...) but there is no explicit network to describe the clicwalkers community. But users activity in answering surveys as well as server logs can provide traces of information diffusion, geolocalisation data, temporal data, sponsorship, ...We will study the problem of adaptive graph construction from the clicwalkers network. Node (users) classification and clustering algorithms will be applied. For the problem of survey recommendations, the problem of teams constitution in a bipartite graphs of users and surveys will be studied. Random graph modeling and generative models of random graphs will be one step towards the prediction of the evolution of clicwalkers community.

MIKAELA KELLER and MARC TOMMASI supervise the PhD thesis (Cifre) of PAULINE WAUQUIER on graph-based recommendation together with Guillaume André from CLIC AND WALK.

**MEPHYSTO Team**

# 7. Bilateral Contracts and Grants with Industry

## 7.1. Bilateral Contracts with Industry

The team (C. Chainais and A. Gloria) has had its third bilateral contract with ANDRA (French nuclear waste storage agency) from December 2012 to June 2014. The post-doctoral position of T. Gallouët was funded by this contract.

This collaboration concerned mathematical and numerical issues on a corrosion model, and in particular the identification of steady-states and the design of asymptotic-preserving schemes for a free interface problem.

**MINT Project-Team**

# 7. Bilateral Contracts and Grants with Industry

## 7.1. Bilateral Contracts with Industry

### *7.1.1. Autodesk Research (Feb. 2014-Apr. 2014)*

**Participant:** Fanny Chevalier [correspondant].

The correspondant worked with Autodesk Research as a consultant for the Kitty project. The Inria correspondant, Fanny Chevalier, providied scientific advices on the design and evaluation of the prototype of Kitty [23], a sketch-based tool for authoring dynamic and interactive illustrations.

**MODAL Project-Team**

# 7. Bilateral Contracts and Grants with Industry

## 7.1. Arcelor-Mittal

**Participants:** Christophe Biernacki, Clément Théry.

*Subject:* Supervised and semi-supervised classification on large data bases mixing qualitative and quantitative variables. Arcelor Mittal faced some quality problems in the steel production which lead to supervised and semisupervised classification involving (1) a small number of individuals comparing to the numbers of variables, (2) heterogeneous variables, typically categorical and continous variables and (3) potentially highly correlated variables. A PhD CIFRE grant started on May 2011 on this topic and will finish on 2015.

## 7.2. PGXIS

**Participant:** Christophe Biernacki.

PGXIS is a UK pharmacogenomics company aiming to discover virtual drugs. Its business model relies on its star technology, named Taxonomy3, a ground breaking mathematical method. Applied to Big Genetic Data, it delivers novel drug targets that are biologically confirmed. These drug targets will drive its drug discovery programmes. This six months contract aims at developping mathematical tool for accelerating convergence rate of Taxonomy3. From a scientific point of view, it corresponds to define specific importance sampling methods related to the Monte Carlo process involved in Taxonomy3.

## 7.3. RougeGorge

**Participants:** Christophe Biernacki, Serge Iovleff, Vincent Vandewalle, Vincent Kubicki, Komi Nagbe.

The RougeGorge company sells lingerie item for women. This three months contract aims at defining a new marketing segmentation for customers and also for items. From a scientific point of view, it corresponds to clustering of mixed data, difficulty being provided but the data volume (millions of customers), by the heterogeneity of data (mixed data) and also by many missing data.

## 7.4. Auchan

**Participants:** Christophe Biernacki, Serge Iovleff, Vincent Vandewalle, Vincent Kubicki.

Groupe Auchan SA is a French international retail group and multinational corporation headquartered in Croix. It is one of the world's principal distribution groups with a presence in 12 countries and 269,000 employees. The aim of the two months contracts between Auchan and MODAL is to identify human factors which significantly impact the economical results of the company. From a scientific point of view, it corresponds to regression studies (simple and mixture regression) with missing data and correlated data.

## 7.5. Cap Gemini

**Participants:** Christophe Biernacki, Vincent Vandewalle.

Cap Gemini S.A. is a French multinational corporation headquartered in Paris, with regional activities. It provides IT services and is one of the world's largest consulting, outsourcing and professional services companies with more than 140,000 employees in over 40 countries. The company aims at developping its Big Data ability in regards to its customer needs. A PhD thesis performing specific research to this activity is planned in 2015. In this aim, a preliminary contract has been established since December 2014. It will allow to prepare precisely the research subject.

## 7.6. PIXEO

**Participant:** Christophe Biernacki.

PIXEO is a company allowing online comparisons of insurances. A PhD thesis for optimizing the workflow related to this activity is planned in 2015. In this aim, a preliminary contract has been established since October 2014. It will allow to prepare precisely the research subject. It is a work in collaboration with two members of the Dolphin Inria team (Laetitia Jourdan and Marie-Eléonore Marmion).

## 7.7. AGLAE

**Participants:** Julien Jacques, Cristian Preda, Florence Loingeville.

AGLAE aims to improve analyses, especially chemical and microbiological, of water and other matrices of the environment. In the context of the Ph.D. of Florence Loingeville, we work on ANOVA models for counting data.

## 7.8. Alicante

**Participants:** Julien Jacques, Cristian Preda, Vincent Vandewalle.

ALICANTE develops applications and tools for data coming from health domain. As a participant of the ClinMine ANR project, ALICANTE and GHICL (Groupe Hospitalier de l'Institut Catholique de Lille) provide us well-structured data for clustering hospital stays.

**NON-A Project-Team**

# 7. Bilateral Contracts and Grants with Industry

## 7.1. Bilateral Contracts with Industry

- Agreement with Sick Company for equipment support of the research in the field of the in-door mobile robot navigation.

**RMOD Project-Team**

# 7. Bilateral Contracts and Grants with Industry

## 7.1. SafePython FUI

**Participants:** Damien Cassou [Correspondant], Jean-Baptiste Arnaud, Stephane Ducasse.

**Contracting parties:** CEA, Evitech, Inria, Logilab, Opida, Thales, Wallix.

Beyond embedded computing, there is not so much research and development on the verification of software safety. Recently, some tools have been created for languages such as JAVA, SQL, VB or PHP. Nevertheless, nothing exists for Python even though this language is growing fast. SafePython's goal is to provide code analysis tools applicable to Python programs. This project will define a subset of Python that the developers will have to use to have their programs analyzed.

## 7.2. Sponsoring LAM

**Participants:** Stéphane Ducasse [Correspondant], Marcus Denker.

**Contracting parties:** Inria, LAM Research, Inc.

LAM Research Inc.(http://lamrc.com) is a leading supplier of wafer fabrication equipment and services to the global semiconductor industry. LAM has started to sponsor RMOD in 2014. RMOD used the sponsored funds to pay student internships in 2014.

## 7.3. Resilience FUI

**Participants:** Stéphane Ducasse [Correspondant], Nicolas Petton, Damien Cassou.

**Contracting parties:** Nexedi, Morphom Alcatel-Lucent Bell Labs, Astrium Geo Information, Wallix, XWiki, Alixen, Alterway, Institut Télécom, Université Paris 13, CEA LIST.

Resilience's goal is to protect private data on the cloud, to reduce spying and data loss in case of natural problems. Resilience proposes to develop a decentralized cloud architecture: SafeOS. Safe OS is based on replication of servers. In addition a safe solution for documents should be developed. Sandboxing for Javascript applications should be explored.

We proposed to use WebWorkers as a way to control DOM edition. There is a plethora of research articles describing the deep semantics of JavaScript. Nevertheless, such articles are often difficult to grasp for readers not familiar with formal semantics. We proposed a digest of the semantics of JavaScript centered around security concerns.

## 7.4. Worldline CIFRE

**Participants:** Anne Etien [Correspondant], Nicolas Anquetil, Stéphane Ducasse, Vincent Blondeau.

In the context of a CIFRE PhD we are working on large industrial project characterization. The PhD started in October 2014.

## 7.5. Pharo Consortium

The Pharo Consortium was founded in 2012 and is growing constantly. As of end 2014, it has 14 company members, 10 academic partners and 3 sponsoring companies. Inria supports the consortium with one full time engineer starting in 2011. More at http://consortium.pharo.org.

**SEQUEL Project-Team**

# 7. Bilateral Contracts and Grants with Industry

## 7.1. Bilateral Contracts with Industry

- Deezer, 2013-2014
  **Participants:** Jérémie Mary, Philippe Preux, Romaric Gaudel.

  A research project has started on June 2013 in collaboration with the Deezer company. The goal is to build a system which automatically recommends music to users. That goal is an extension of the bandit setting to the Collaborative Filtering problem.
- Nuukik, 2013-2014
  **Participant:** Jérémie Mary.

  Nuukik is a start-up from Hub Innovation in Lille. It proposes a recommender systems for e-commerce based on matrix factorization. We worked with them specifically on the cold start problem (*i.e* when you have absolutely no data on a product or a customer). This led to promising result and allowed us to close the gap between bandits and matrix factorization. This work led to a patent submission in december 2013.
- Squoring Technologies, 2011-2014
  **Participants:** Boris Baldassari, Philippe Preux.

  Boris Baldassari has been hired by Squoring Technologies (Toulouse) as a PhD student in May 2011. He works on the use of machine learning to improve the quality of the software development process. During his first year as a PhD student, Boris investigated the existing norms and measures of quality of software development process. He also dedicated some time to gather some relevant datasets, which are made of either the sequence of source code releases over a multi-years period, or all the versions stored on an svn repository (svn or alike). Information from mailing-lists (bugs, support, ...) may also be part of these datasets. Tools in machine learning capable of dealing with this sort of data have also been investigated. Goals that may be reached in this endeavor have also been precised.

## 7.2. Bilateral Grants with Industry

- INTEL Corp., 2013 - 2014
  **Participants:** Philippe Preux, Michal Valko, Rémi Munos, Adrien Hoarau.

  This is a research project on Algorithmic Determination of IoT Edge Analytics Requirements. We are attempting to solve the problem of how to automatically predict the system requirements for edge node analytics in the Internet of Things (IoT). We envision that a flexible extensible system of edge analytics can be created for IoT management; however, edge nodes can be very different in terms of the systems requirements around: processing capability, wireless communication, security/cryptography, guaranteed responsiveness, guaranteed quality of service and on-board memory requirements. One of the challenges of managing a heterogeneous Internet of Things is determining the systems requirements at each edge node in the network.

  We suggest exploiting opportunity of being able to automatically customize large scale IoT systems that could comprise heterogeneous edge nodes and allow a flexible and scalable component and firmware SoC systems to be matched to the individual need of enterprise/ government level IoT customers. We propose using large scale sequential decision learning algorithms, particularly contextual bandit modeling to automatically determine the systems requirements for edge analytics. These algorithms have an adaptive property that allows for the addition of new nodes and the re-evaluation of existing nodes under dynamic and potentially adversarial conditions.

**SPIRALS Team**

# 7. Bilateral Contracts and Grants with Industry

## 7.1. ip-label

**Participants:** Nicolas Haderer, Christophe Ribeiro, Romain Rouvoy [correspondant].

A software exploitation licence of the APISENSE® crowdsensing platform has been sold to the ip-label company. They use this platform as a solution to monitor the quality of the GSM signal in the wild. The objective is to provide developers and stakeholders with a feedback on the quality of experience of GSM connection depending on their location.

## 7.2. Microsoft Windows Azure for Research Grant

**Participants:** Nicolas Haderer, Christophe Ribeiro, Romain Rouvoy [correspondant].

The research program associated with this grant consists in leveraging the APISENSE® crowdsensing platform to support the real-time processing of "big" datasets collected in the physical world by a "large" crowd of smartphones. Examples of case studies covered in this area include the automatic inference of roadmaps, the continuous cartography of network coverage quality, or even the detection and the dynamic analysis of earthquakes. However, the unpredictable volume of data to be collected in the wild requires the adoption of elastic computation models and infrastructures to continuously provision the processing capabilities to fit uploads of information reports.

The grant takes the form of virtual credits for accessing the Microsoft Azure cloud computing platform.

## 7.3. Orange Labs

**Participants:** Laurence Duchien [correspondant], Amal Tahri.

This collaboration aims at bridging the gap between home networks and cloud environments for the design, the provisionning and the administration of distributed services. The purpose is to define solutions, essentially software design tools and runtime infrastructures, for the seamless migration of distributed applications and services between home networks and cloud environments. The envisioned approach is based on the research activities that we are conducting in the domain of software product lines.

This collaboration is conducted in the context of the ongoing PhD thesis of Amal Tahri.