IN PARTNERSHIP WITH:
CNRS

Université Versailles Saint-Quentin

Activity Report 2016

Project-Team SMIS

Secured and Mobile Information Systems

IN COLLABORATION WITH: Parallelisme, réseaux, systèmes, modélisation (PRISM)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team SMIS

Creation of the Project-Team: 2004 September 01, end of the Project-Team: 2016 December 31

Keywords:

Computer Science and Digital Science:

1.1.6. - Cloud
1.1.8. - Security of architectures
1.4. - Ubiquitous Systems
3.1.2. - Data management, quering and storage
3.1.3. - Distributed data
3.1.5. - Control access, privacy
3.1.6. - Query optimization
3.1.8. - Big data (production, storage, transfer)
3.1.9. - Database
4.3. - Cryptography
4.7. - Access control
4.8. - Privacy-enhancing technologies

Other Research Topics and Application Domains:

2. - Health
6.4. - Internet of things
6.5. - Information systems
6.6. - Embedded systems
8.2. - Connected city
8.5. - Smart society
9.8. - Privacy
9.10. - Ethics

1. Members

Research Scientists

Nicolas Anciaux [Inria, Researcher, HDR]
Luc Bouganim [Inria, Senior Researcher, HDR]

Faculty Members

Philippe Pucheral [Team leader, Univ. Versailles, Professor, HDR]
Iulian Sandu Popa [Univ. Versailles, Associate Professor]
Guillaume Scerri [Univ. Versailles, Associate Professor, from Sept. 2016]

Engineers

Aydogan Ersoz [Inria]
Oana Manea [Inria, from Nov. 2016]

PhD Students

Athanasia Katsouraki [Inria, until Sept. 2016]
Saliha Lallali [Inria, granted by ANR KISS project, until Jan. 2016]
Paul Tran Van [CozyCloud, CIFRE]

Axel Michel [INSA-CVL, from Sept. 2015]
Julien Loudet [CozyCloud, CIFRE, from April 2016]
Riad Ladjel [IDEX grant, from Oct. 2016]

Visiting Scientist
Benjamin Nguyen [INSA CVL, Professor, HDR]

Administrative Assistant
Emmanuelle Perrot [Inria]

2. Overall Objectives

2.1. Overall Objectives

The research work within the project-team is devoted to the design and analysis of core database techniques dedicated to the definition of secured and mobile information systems.

Ubiquitous computing and ambient intelligence entail embedding data in increasingly light and specialized devices (chips, sensors and electronic appliances for smart buildings, telephony, transportation, health, etc.). These devices exhibit severe hardware constraints to match size, security, power consumption and also production costs requirements. At the same time, they could highly benefit from embedded database functionalities to store data, analyze it, query it and protect it. This raises a first question "Q_1: *How to make powerful data management techniques compatible with highly constrained hardware platforms?*". To tackle this question, SMIS contributes to the design and validation of new storage and indexing models, query execution and optimization techniques, and transaction protocols. The relevance of this research goes beyond embedded databases and may have potential applications for database servers running on advanced hardware.

By making information more accessible and by multiplying –often transparently– the means of acquiring it, ubiquitous computing involves new threats for data privacy. The second question addressed by the project-team is then "Q_2: *How to make smart objects less intrusive?*". New access and usage control models have to be devised to help individuals keep a better control on the acquisition and sharing conditions of their data. This means integrating privacy principles like user's consent, limited collection and limited retention in the access and usage control policy definition. This also means designing appropriate mechanisms to enforce this control and provide accountability with strong security guarantees.

In parallel, thanks to a high degree of decentralization and to the emergence of low cost tamper-resistant hardware, ubiquitous computing contains the seeds for new ways of managing personal/sensitive data. The third question driving the research of the project-team is therefore "Q_3: *How to build privacy-by-design architectures based on trusted smart objects?*". The objective is to capitalize on embedded data management techniques, privacy-preserving mechanisms, trusted devices and cryptographic protocols to define an integrated framework dedicated to the secure management of personal/sensitive data. The expectation is showing that credible alternatives to a systematic centralization of personal/sensitive data on servers can be devised and validating the approach through real case experiments.

3. Research Program

3.1. Embedded Data Management

The challenge tackled is this research action is twofold: (1) to design embedded database techniques matching the hardware constraints of (current and future) smart objects and (2) to set up co-design rules helping hardware manufacturers to calibrate their future platforms to match the requirements of data driven applications. While a large body of work has been conducted on data management techniques for high-end servers (storage, indexation and query optimization models minimizing the I/O bottleneck, parallel DBMS, main memory

DBMS, etc.), less research efforts have been placed on embedded database techniques. Light versions of popular DBMS have been designed for powerful handheld devices; yet DBMS vendors have never addressed the complex problem of embedding database components into chips. Proposals dedicated to databases embedded on chip usually consider small databases, stored in the non-volatile memory of the microcontroller –hundreds of kilobytes– and rely on NOR Flash or EEPROM technologies. Conversely, SMIS is pioneering the combination of microcontrollers and NAND Flash constraints to manage Gigabyte(s) size embedded databases. We present below the positioning of SMIS with respect to international teams conducting research on topics which may be connected to the addressed problem, namely work on electronic stable storage, RAM consumption and specific hardware platforms.

Major database teams are investigating data management issues related to hardware advances (EPFL: A. Ailamaki, CWI: M. Kersten, U. Of Wisconsin: J. M. Patel, Columbia: K. Ross, UCSB: A. El Abbadi, IBM Almaden: C. Mohan, etc.). While there are obvious links with our research on embedded databases, these teams target high-end computers and do not consider highly constrained architectures with non traditional hardware resources balance. At the other extreme, sensors (ultra-light computing devices) are considered by several research teams (e.g., UC Berkeley: D. Culler, ITU: P. Bonnet, Johns Hopkins University: A. Terzis, MIT: S. Madden, etc.). The focus is on the processing of continuous streams of collected data. Although the devices we consider share some hardware constraints with sensors, the objectives of both environments strongly diverge in terms of data cardinality and complexity, query complexity and data confidentiality requirements. Several teams are looking at efficient indexes on flash (HP LABS: G. Graefe, U. Minnesota: B. Debnath, U. Massachusetts: Y. Diao, Microsoft: S. Nath, etc.). Some studies try to minimize the RAM consumption, but the considered RAM/stable storage ratio is quite large compared to the constraints of the embedded context. Finally, a large number of teams have focused on the impact of flash memory on database system design (we presented an exhaustive state of the art in a VLDB tutorial [34]). The work conducted in the SMIS team on bi-modal flash devices takes the opposite direction, proposing to influence the design of flash devices by the expression of database requirements instead of running after the constantly evolving flash device technology.

3.2. Access and Usage Control Models

Access control management has been deeply studied for decades. Different models have been proposed to declare and administer access control policies, like DAC, MAC, RBAC, TMAC, and OrBAC. While access control management is well established, new models are being defined to cope with privacy requirements. Privacy management distinguishes itself from traditional access control is the sense that the data to be protected is personal. Hence, the user's consent must be reflected in the access control policies, as well as the usage of the data, its collection rules and its retention period, which are principles safeguarded by law and must be controlled carefully.

The research community working on privacy models is broad, and involves many teams worldwide including in France ENST-B, LIRIS, Inria LICIT, and LRI, and at the international level IBM Almaden, Purdue Univ., Politecnico di Milano and Univ. of Milano, George Mason Univ., Univ. of Massachusetts, Univ. of Texas and Colorado State Univ. to cite a few. Pioneer attempts towards privacy wary systems include the P3P Platform for Privacy Preservation [36] and Hippocratic databases [29]. In the last years, many other policy languages have been proposed for different application scenarios, including EPAL [40], XACML [39] and WSPL [32]. Hippocratic databases are inspired by the axiom that databases should be responsible for the privacy preservation of the data they manage. The architecture of a Hippocratic database is based on ten guiding principles derived from privacy laws.

The trend worldwide has been to propose enhanced access control policies to capture finer behavior and bridge the gap with privacy policies. To cite a few, Ardagna *et al.* (Univ. Milano) enables actions to be performed after data collection (like notification or removal), purpose binding features have been studied by Lefevre *et al.* (IBM Almaden), and Ni *et al.* (Purdue Univ.) have proposed obligations and have extended the widely used RBAC model to support privacy policies.

The positioning of the SMIS team within this broad area is rather (1) to focus on intuitive or automatic tools helping the individual to control some facets of her privacy (e.g., data retention, minimal collection) instead of increasing the expressiveness but also the complexity of privacy models and (2) to push concrete models enriched by real-case (e.g., medical) scenarios and by a joint work with researchers in Law.

3.3. Tamper-resistant Data Management

Tamper-resistance refers to the capacity of a system to defeat confidentiality and integrity attacks. This problem is complementary to access control management while being (mostly) orthogonal to the way access control policies are defined. Security surveys regularly point out the vulnerability of database servers against external (i.e., by intruders) and internal (i.e., by employees) attacks. Several attempts have been made in commercial DBMSs to strengthen server-based security, e.g., by separating the duty between DBA and DSA (Data Security Administrator), by encrypting the database footprint and by securing the cryptographic material using Hardware Security Modules (HSM) [35]. To face internal attacks, client-based security approaches have been investigated where the data is stored encrypted on the server and is decrypted only on the client side. Several contributions have been made in this direction, notably by U. of California Irvine (S. Mehrotra, Database Service Provider model), IBM Almaden (R. Agrawal, computation on encrypted data), U. of Milano (E. Damiani, encryption schemes), Purdue U. (E. Bertino, XML secure publication), U. of Washington (D. Suciu, provisional access) to cite a few seminal works. An alternative, recently promoted by Stony Brook Univ. (R. Sion), is to augment the security of the server by associating it with a tamper-resistant hardware module in charge of the security aspects. Contrary to traditional HSM, this module takes part in the query computation and performs all data decryption operations. SMIS investigates another direction based on the use of a tamper-resistant hardware module on the client side. Most of our contributions in this area are based on exploiting the tamper-resistance of secure tokens to build new data protection schemes.

While our work on Privacy-Preserving data Publishing (PPDP) is still related to tamper-resistance, a complementary positioning is required for this specific topic. The primary goal of PPDP is to anonymize/sanitize microdata sets before publishing them to serve statistical analysis purposes. PPDP (and privacy in databases in general) is a hot topic since 2000, when it was introduced by IBM Research (IBM Almaden: R. Agrawal, IBM Watson: C.C. Aggarwal), and many teams, mostly north American universities or research centres, study this topic (e.g., PORTIA DB-Privacy project regrouping universities such as Stanford with H. Garcia-Molina). Much effort has been devoted by the scientific community to the definition of privacy models exhibiting better privacy guarantees or better utility or a balance of both (such as differential privacy studied by C. Dwork: Microsoft Research or D. Kifer: Penn-State Univ and J. Gehrke: Cornell Univ) and thorough surveys exist that provide a large overview of existing PPDP models and mechanisms [37]. These works are however orthogonal to our approach in that they make the hypothesis of a trustworthy central server that can execute the anonymization process. In our work, this is not the case. We consider an architecture composed of a large population of tamper-resistant devices weakly connected to an untrusted infrastructure and study how to compute PPDP problems in this context [1]. Hence, our work has some connections with the works done on Privacy Preserving Data Collection (Stevens Institute of Tech. / Rutgers Univ,NJ: R.N.Wright, Univ Austin Texas: V. Shmatikov), on Secure Multi-party Computing for Privacy Preserving Data Mining (Rutgers Univ: J. Vaidya, Purdue Univ: C. Clifton) and on distributed PPDP algorithms (Univ Wisconsin: D. DeWitt, Univ Michigan: K. Lefevre, Rutgers Univ: J. Vaidya, Purdue Univ: C. Clifton) while none of them share the same architectural hypothesis as us.

4. Application Domains

4.1. Application Domains

Our work addresses varied application domains. Typically, data management techniques on chip are required each time data-driven applications have to be embedded in ultra-light computing devices. This situation occurs for example in healthcare applications where medical folders are embedded into smart tokens (e.g.,

smart cards, secured USB keys), in telephony applications where personal data (address book, agenda, etc.) is embedded into cellular phones, in sensor networks where sensors log row measurements and perform local computation on them, in smart-home applications where a collection of smart appliances gather information about the occupants to provide them a personalized service, and more generally in most applications related to ambient intelligence.

Safeguarding data confidentiality has become a primary concern for citizens, administrations and companies, broadening the application domains of our work on access control policies definition and enforcement. The threat on data confidentiality is manifold: external and internal attacks on the data at rest, on the data on transit, on the data hosted in untrusted environments (e.g., Database Service Providers, Web-hosting companies) and subject to illegal usage, insidious gathering of personal data in an ambient intelligence surrounding. Hence, new access control models and security mechanisms are required to accurately declare and safely control who is granted access to which data and for which purpose.

While the application domain mentioned above is rather large, two applications are today more specifically targeted by the SMIS team. The first one deals with privacy preservation in EHR (Electronic Health Record) systems and PCEHR (Personally Controlled EHR) [3]. We are developing technologies tackling this issue and experiment them in the field. The second application area deals with privacy preservation in the context of personal Cloud, that is personal data hosted in dedicated servers staying under the holder's control (e.g., in a personal internet box or in a home automation box).

5. New Software and Platforms

5.1. PLUG-DB ENGINE

FUNCTIONAL DESCRIPTION: PlugDB is a complete platform dedicated to a secure and ubiquitous management of personal data. It aims at providing an alternative to a systematic centralization of personal data. The PlugDB engine is a personal database server capable of storing data (tuples and documents) in tables and BLOBs, indexing them, querying them in SQL, sharing them through assertional access control policies and enforcing transactional properties (atomicity, integrity, durability) [4]. The PlugDB engine is embedded in a tamper-resistant hardware device combining the security of smartcard with the storage capacity of NAND Flash. The personal database is hosted encrypted in NAND Flash and the PlugDB engine code runs in the microcontroller. Complementary modules allow to pre-compile SQL queries for the applications, communicate with the DBMS from a remote Java program, synchronize local data with remote servers (typically used for recovering the database in the case of a broken or lost devices) and participate in distributed computation (e.g., global queries). PlugDB runs both on secure devices provided by Gemalto and on specific secure devices designed by SMIS and assembled by electronic SMEs. Mastering the hardware platform opens up new research and experiment opportunities (e.g., we have recently integrated a Bluetooth module to communicate wirelessly with PlugDB and a fingerprint module to strongly authenticate users) and allows us to engage ourselves in an open-source/open hardware initiative. Open-SW/open-HW contributes to the trust the community of users can put in any privacy preserving solution and is key to enable a diversity of solutions, hence decreasing the risk of class attacks. PlugDB engine has been registered first at APP (Agence de Protection des Programmes) in 2009 - a new version being registered every two years and the hardware datasheets in 2015. PlugDB has been experimented in the field, notably in the healthcare domain. We also recently set up an educational platform on top of PlugDB, named SIPD (Système d'Information privacy-by-Design) and used at ENSIIE, INSA CVL and UVSQ through the Versailles Sciences Lab fablab, to raise students awareness of privacy protection problems and embedded programming. As a conclusion, PlugDB combines several research contributions from the team, at the crossroads of flash data management, embedded data processing and secure distributed computations. It then strongly federates all members of our team (permanent members, PhD students and engineers). It is also a vector of visibility, technological transfer and dissemination and gives us the opportunity to collaborate with researchers from other disciplines around a concrete privacy enhancing platform.

- Participants: Nicolas Anciaux, Luc Bouganim, Philippe Pucheral and Aydogan Ersoz

- Contact: Nicolas Anciaux
- URL: https://project.inria.fr/plugdb/

5.2. Privacy Preserving Mobile Laboratory

FUNCTIONAL DESCRIPTION: We have started to design a privacy preserving mobile laboratory used as an experimental platform for multidisciplinary research launched 'in vivo'. The goal is to conduct reliable surveys and avoid the privacy paradox (what users declare on their privacy behavior is far from what they effectively do). The platform, built on top of PlugDB, includes two android applications, a "server" which takes as input a questionnaire description and broadcast it on demand to the client applications. Users interact with the questionnaire on the client applications, storing the detailed answers in their PlugDB personal server. Then a secure distributed computation takes place (between users' PlugDB servers) and computes non-sensitive global statistics based on potentially sensitive raw answers. A beta-version of this platform was developed during the PhD of Athanasia Katsouraki and was used for a pre-experimentation targeting 140 students. While the experiment was successful, it showed the limitation and complexity of the initial setting (laptops, required Internet access, complexity in the questionnaire deployment). We designed and implement a second platform running on android tablets with a local router and automatic questionnaire deployment. The platform has been demonstrated in several forums and very recently at the Sénat in Paris. This platform represents a backing for two PhD theses on privacy (the first one in economics, the second one in our team) funded in 2016 by the interdisciplinary doctoral program at UPSay (IDI 2016).

- Participants: Nicolas Anciaux, Luc Bouganim, Aydogan Ersoz, Athanasia Katsouraki, Riad Ladjel, Benjamin Nguyen, Remy Pasquion, Paul Tran Van
- Contact: Luc Bouganim
- URL: https://project.inria.fr/plugdb/en/PPML

6. New Results

6.1. Embedded Data Management

Participants: Nicolas Anciaux, Saliha Lallali, Philippe Pucheral, Iulian Sandu Popa [correspondent].

Embedded keyword indexing: In this work, we revisit the traditional problem of information retrieval queries over large collections of files in an embedded context. A file can be any form of document, picture or data stream, associated with a set of terms. A query can be any form of keyword search using a ranking function (e.g., TF-IDF) identifying the top-k most relevant files. The proposed search engine can be used in sensors to search for relevant objects in their surroundings, in cameras to search pictures by using tags, in personal smart dongles to secure the querying of documents and files hosted in an untrusted Cloud, or in a personal cloud securely managed using a tamper resistant smart object. A search engine is usually based on a (large) inverted index and queries are traditionally evaluated by allocating one container in RAM per document to aggregate its score, making the RAM consumption linear with the size of the document corpus. To tackle this issue, we designed a new form of inverted index which can be accessed in a pure pipeline manner to evaluate search queries without materializing any intermediate result. Successive index partitions are written once in Flash and maintained in the background by timely triggering merge operations while files are inserted or deleted from the index. This work was initially published at VLDB'15 [5] and demonstrated at SIGMOD'15 [38]. It constitutes the main contribution of the PhD thesis of Saliha Lallali defended in January 2016. In 2016, we extended this work to demonstrate at EDBT'16 [22] its applicability to set up a secure distributed search engine for the Personal Cloud. We also complemented this work with (1) a thorough analysis of the RAM consumption linked to the main algorithms implementing the solution, (2) the support of conditional top-k queries in a personal Cloud context that we consider as a killer application domain today and (3) new performance measurements with a real dataset (ENRON), representative of this personal Cloud context. These new contributions have been submitted to Information Systems journal.

6.2. Secure Global Computing on Asymmetric Architecture

Participants: Benjamin Nguyen [correspondent], Axel Michel, Philippe Pucheral, Iulian Sandu Popa.

Asymmetric Architecture Computing: This research direction studies the secure execution of various algorithms on data stored in an unstructured network of Trusted Cells (i.e., personal trusted device) so that each user can keep control over her data. The data could be stored locally in a trusted cell or encrypted on some external cloud. Execution takes place on a specific infrastructure called the Asymmetric Architecture (AA): the network of trusted cells, supported by an untrusted cloud supporting IaaS or PaaS. Our objective is to show that many different algorithms and computing paradigms can be executed on AA, thus achieving secure and private computation. Our first contribution in this area was to study the execution of Privacy Preserving Data Publishing algorithms on such an architecture (T. Allard's PhD Thesis). Then we studied general SQL queries in this same execution context. We concentrated on the subset of SQL queries without joins, but including Group By and aggregates, and show how to secure their execution in the presence of honest-but-curious attackers. This work, named SQL-AA and notably published at EDBT'14 [8] and demonstrated at VLDB'15, was part of Quoc-Cuong To's Ph.D defended in 2015. We have extended this framework through a collaboration with INSA Centre Val de Loire, LIFO Lab and University of Paris Nord, LIPN lab and have shown in CoopIS'15 [9] that it is possible to achieve seamless integration of distributed MapReduce processing using trusted cells, while maintaining reasonable performance. In 2016, we added three novel contributions to SQL-AA: (i) an extended privacy analysis in which we consider stronger adversaries with more background knowledge, (ii) an extended threat model in which we consider malicious attacker and propose safety properties to prevent malicious attacks and (iii) we tackled practical issues like exchanging securely shared keys among trusted cells and Querier (GKE protocol) and enforcing access control at query execution time. These new contributions have been published in TODS'16 [15]. In parallel, we are starting a new study in the line of our previous work on Privacy Preserving Data Publishing (PPDP) with the objective to inject individualized privacy requirements in the PPDP protocol. A preliminary contribution has been published at BDA'16 [25] to compute SQL aggregate queries under k-anonymity constraints where each individual contributing to the query may define her own k constraint, thereby letting each one weighting differently the sensitiveness of a given piece of information according to her own situation.

Secure spatio-temporal distributed processing: Mobile participatory sensing could be used in many applications such as vehicular traffic monitoring, pollution tracking, or even health surveying (e.g., to allow measuring in real-time the individual exposure to environmental risk factors or the propagation of an epidemic). However, its success depends on finding a solution for querying a large number of users which protects user location privacy and works in real-time [10]. We addressed these issues and proposed PAMPAS, a privacy-aware mobile distributed system for efficient data aggregation in mobile participatory sensing. In PAMPAS, mobile devices enhanced with secure hardware, called secure probes, perform distributed query processing, while preventing users from accessing other users' data. Secure probes exchange data in encrypted form with help from an untrusted supporting server infrastructure. PAMPAS uses two efficient, parallel, and privacy-aware protocols for location-based aggregation and adaptive spatial partitioning of secure probes. Our experimental results and security analysis demonstrate that these protocols are able to collect, aggregate and share statistics or derived data in real-time, without any privacy leakage. This work is part of Dai Hai Ton That's Ph.D. thesis defended in January 2016, co-supervised by Iulian Sandu Popa. The system implementation was demonstrated in [41], while two papers describing the technical details of the system have been published in 2016 [23], [16].

6.3. Personal Cloud

Participants: Nicolas Anciaux [correspondent], Luc Bouganim, Julien Loudet, Benjamin Nguyen, Philippe Pucheral, Iulian Sandu Popa, Guillaume Scerri, Paul Tran Van.

We are witnessing an exponential increase in the acquisition of personal data about the individuals or produced by them. Today, this information is managed using Web applications, centralizing this data in cloud data servers, under the control of few Web majors [2]. However, it has now become clear that (1) centralizing millions of personal records exposes the data to very sophisticated attacks, linked to a very high potential benefit in case of success (millions of records being revealed), and (2) delegating the management of personal

records without any tangible guarantee for the individuals leads to privacy violations, the data being potentially made accessible to other organizations (e.g., governments, commercial partners) and being subject to lucrative secondary usages (not advertised to the individuals). To face this situation, many recent initiatives push towards the emergence of the Personal Cloud paradigm. A personal cloud can be viewed as a personal server, owned by a given individual, which gives to its owner the ability to store her complete digital environment, synchronize it among various devices and share it with other individuals and applications under control. In the SMIS team, we claim the need of a Secure Personal Cloud, and promote the introduction of a secure (tamper resistant) data engine in the architecture [30]. On this basis, we investigate new data sharing and dissemination models, where usage and access control rules endorsed by the individuals could be enforced and have presented this vision at EDBT'14 [6] and at ADBIS'15 [31]. We have started a cooperation with the startup CozyCloud at the end of 2014. A contract was signed at the end of 2014 to integrate PlugDB in a CozyCloud instance and two CIFRE PhD thesis have been launched so far. Paul Tran Van's PhD thesis explores a new data sharing paradigm dedicated to the personal cloud context. This paradigm, called SWYSWYK (Share What You See with Who You Know), allows to automatically derive intuitive sharing rules from a personal cloud content, to share rules among a community of users and to let each user physically visualize the net effects of these rules on her own Personal Cloud. We propose a reference architecture providing the users with tangible guarantees about the enforcement of SWYSWYK policies and demonstrate through a performance evaluation conducted on a real personal cloud platform that the approach is practical. This work constitutes the core of Paul Tran Van's thesis and is being submitted for publication at VLDB. Preliminary ideas related to this work are presented in ERCIM news'16 [27]. Julien Loudet's PhD thesis is just starting with the objective to explore privacy-preserving distributed computations over personal clouds.

More generally, the personal cloud context gain in importance in our research work. It is even at the heart of our future project-team named PETRUS (PErsonal and TRUSted cloud). PETRUS is expected to take over from the SMIS team beginning of 2017.

6.4. Interdisciplinary study on Privacy-by-Design

Participants: Nicolas Anciaux, Luc Bouganim [correspondent], Athanasia Katsouraki, Benjamin Nguyen, Philippe Pucheral.

The objective of this research action is to study the reciprocal entanglements between economic, legal, societal and technological aspects of the management and exploitation of personal data. Indeed, devising new ways of protecting data privacy cannot be done in isolation; it requires also identifying alternative economic models that are both viable and regulatory compliant. We started an interdisciplinary research work with economists (RITM Lab) and jurists (CERDI and DANTE labs) in the privacy axis of ISN (Institut de la Société Numérique) and plan to pursue it in two projects in preparation: the Convergence Institute I2DRIVE (Interdisciplinary Institute for Data Research: Intelligence, Values and Ethics) and the CNRS Federation SIHS (Sciences Informatiques, Humaines et Sociales) at UVSQ. A first interdisciplinary work conducted in 2016 concerns the design of a privacy preserving platform needed to conduct privacy studies "in vivo". Such platforms are required to validate the effectiveness of privacy preserving solutions, in terms of technical feasibility, lawfulness, acceptability and benefits. To this end, we have designed a privacy preserving mobile lab, derived from the personal cloud platform developed by the team (see 'Software' section). In her PhD thesis, Athanasia Katsouraki developed a beta-version of that platform and used it to perform a pre-experimentation in the context of online form based survey, targeting 140 students. The goal was threefold: (1) to test the effectiveness of the proposed platform, (2) to test the adequation of the questionnaire and experimentation protocol (a result for the experimental economist), and (3) to check the impact of the use of a secure platform on the student's answers. The pre-experimentation showed several improvement axis and led to the actual design of the privacy preserving mobile lab described in the Software section.

Another joint work is related to the design of technical means to help individuals perceive how their personal life is exposed compared to others and to make appropriate protection choices. This work led to the definition of a new principle called Privacy-by-Using [20], that we introduced to try to circumvent the limits of the privacy-by-design principle promoted by the regulator. The confrontation of the Privacy-by-Using principle with Big Data processing [26] has also been studied with jurists and economists.

Finally, we conducted a scientific expertise on behalf of DGCCRF (Direction Générale de la Concurrence, de la Consommation et de la Répression des Fraudes) and of the European Council regarding the draft proposal of "Directive of the European Parliament and of the Council on certain aspects concerning contracts for the online and other distance sales of goods" regulating the payment of numeric goods and services by means of personal data. This led us to a cross-analysis, with researchers in Law and computer scientists, of technical, societal and economic issues linked to the smart disclosure principle, that is, under which conditions and formats individuals can get their data back from service providers [17], [19], [18].

6.5. Formal guarantees

Participant: Guillaume Scerri.

The aim of the action is to investigate the changes required for the PlugDB architecture to be amenable to formal security proofs.

More precisely we started exploring the precise formal guarantees that are desirable for a personal data server. Following work started in Bristol [7], relating to formal guarantees provided by secure hardware, we started studying how one could leverage the low level guarantees provided by secure hardware (PlugDB for example) to cover the more complex operations and guarantees required of a personal data server. The first finding of the action is that a modular architecture is required for formal proofs to be obtainable. This is reflected in the architectural concerns presented in the PETRUS project.

Additionally, we started studying how to leverage secure hardware guarantees in order to perform secure computations on distributed data. A first result in this direction is presented in [33], and submitted to Financial Cryptography 2017.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. Cozy Cloud bilateral contract (Dec 2014 - Nov. 2015)

Partners: Cozy Cloud, Inria-SMIS
SMIS funding: 50k€.
While this initial contract is over, we mention it to explain the increasing relationship being built between Cozy Cloud and our team. Cozy Cloud is a French startup providing a personal Cloud platform. The Cozy product is a software stack that anyone can deploy to run his personal server in order to host his personal data and web services. While centralizing all personal data in the holder's hand is a natural way to reestablish his control on his privacy, this represents an unprecedented threat in case of attacks by an intruder, especially for individuals who are not security experts. The objective of this bilateral contract is to address this issue by integrating the PlugDB solution into the Cozy stack. Roughly speaking, the Cozy data system will be modified in such a way to store only encrypted files and each file access will be intercepted and routed to PlugDB. PlugDB will act as a doorkeeper for the whole individual dataspace by managing the files' metadata, the access control rules defined on these metadata, the decryption keys and the user/application authentication.

7.1.2. Cozy Cloud CIFRE - Tran Van contract (Oct 2014 - Sept 2017)

Partners: Cozy Cloud, Inria-SMIS
SMIS funding: 30k€
In relation with the bilateral contract mentioned above, a CIFRE PhD thesis has been started by Paul Tran Van. The objective is to capitalize on the Cozy-PlugDB platform to devise new access and usage control models to exchange data among devices of the same user (devices may have different levels of trustworthiness) and among different users thanks to a user-friendly sharing model (see the work on the SWYSWYK - Share What You See with Who You Know - model presented above) .

7.1.3. Cozy Cloud CIFRE - Loudet contract (Apr 2016 - Apr 2019

Partners: Cozy Cloud, Inria-SMIS
SMIS funding: 45k€
In relation with the bilateral contract mentioned above, a second CIFRE PhD thesis has been started by Julien Loudet. The objective is to allow for a secure execution of distributed queries on a set of personal clouds associated to users, depending on social links, user's localization or user's profile. The general idea is to build secure indexes, distributed on the users' personal cloud and to devise a secure execution protocol revealing solely the query result to the querier. Such highly distributed secure queries potentially enable new (social) applications fed by user's personal data which could be developed on the Cozy-PlugDB platform.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR PerSoCloud (Jan. 2017 - Jan. 2020)

Partners: Orange Labs (coordinator), Inria-SMIS, Cozy Cloud, Univ. of Versailles.
SMIS funding: 170k€.

The objective of PerSoCloud is to design, implement and validate a fullfledged Privacy-by-Design Personal Cloud Sharing Platform. One of the major difficulties linked to the concept of personal cloud lies in organizing and enforcing the security of the data sharing while the data is no longer under the control of a central server. We identify three dimensions to this problem. Devices-sharing: assuming that the primary copy of user U1's personal data is hosted in a secure place, how to share and synchronize it with U1's multiple (mobile) devices without compromising security? Peers-sharing: how user U1 could exchange a subset of his-her data with an identified user U2 while providing to U1 tangible guarantees about the usage made by U2 of this data? Community-sharing: how user U1 could exchange a subset of his-her data with a large community of users and contribute to personal big data analytics while providing to U1 tangible guarantees about the preservation of his-her anonymity? In addition to tackling these three scientific and technical issues, a legal analysis will guarantee compliance of this platform with the security and privacy French and UE regulation, which firmly promotes the Privacy by Design principle, including the current reforms of personal data regulation.

8.1.2. ANR KISS (Dec. 2011 - Feb. 2016)

Partners: Inria-SMIS (coordinator), Inria-SECRET, LIRIS, Univ. of Versailles, CryptoExperts, Gemalto, Yvelines district.
SMIS funding: 230k€.

The idea promoted in KISS is to embed, in trusted devices, software components capable of acquiring, storing and managing securely various forms of personal data (e.g., salary forms, invoices, banking statements, geolocation data, depending on the applications). These software components form a Personal Data Server which can remain under the holder's control. The scientific challenges include: embedded data management issues tackling regular, streaming and spatio-temporal data (e.g., geolocation data), data provenance-based privacy models, crypto-protected distributed protocols to implement private communications and secure global computations.

8.1.3. PIA - PDP SECSi (May. 2016 - Dec. 2017)

Partners: Cozy Cloud (coordinator), Qwant, Inria-SMIS, FING.
SMIS funding: 149k€.

The objective of this PIA-PDP (Programme Investissement d'Avenir - Protection des Données Personnelles) SECSi project is to build a concrete Personal Cloud platform which can support a large scale deployment of Self Data services. Three major difficulties are identified and will be tackled in this project: (1) how to implement and enforce a fine control of the data flow when personal data are exploited by third party applications, (2) how to protect these same applications when processing is delegated to the personal cloud platform itself and (3) how to implement personalized search on the web without hurting user's privacy.

8.1.4. CAPPRIS Project-Lab (Dec. 2011 - Dec. 2016)

Inria Partners: PRIVATICS (coordinator), SMIS, PLANETE, CIDRE, COMETE.
External partners: Univ. of Namur, Eurecom, LAAS.
Funding: not associated to individual project-teams.

An Inria Project Lab (IPL) is a long-term multi-disciplinary project launched by Inria to sustain large scale risky research actions in line with its own strategic plan. CAPPRIS stands for "Collaborative Action on the Protection of Privacy Rights in the Information Society". The key issues that are addressed are: (1) the identification of existing and future threats to privacy, (2) the definition of formally grounded measures to assess and quantify privacy, (3) the definition of the fundamental principles underlying privacy by design and methods to apply them in concrete situations and (4) The integration of the social and legal dimensions. To assess the relevance and significance of the research results, they are confronted to three classes of case studies CAPPRIS partners are involved in: namely Online Social Networks, Location Based Services and Electronic Health Record Systems.

8.1.5. CityLab@Inria, Inria Project Lab (May 2014 -).

Inria Partners: CLIME, DICE, FUN, MIMOVE, MYRIADS, SMIS, URBANET, WILLOW.
External partners: UC Berkeley.
Funding: not associated to individual project teams.

CityLab@Inria studies ICT solutions toward smart cities that promote both social and environmental sustainability. A strong emphasis of the Lab is on the undertaking of a multi-disciplinary research program through the integration of relevant scientific and technology studies, from sensing up to analytics and advanced applications, so as to actually enact the foreseen smart city Systems of Systems. SMIS contributes to Privacy-by-Design architectures for trusted smart objects so as to ensure privacy to citizens, which is critical for ensuring that urbanscale sensing contributes to social sustainability and does not become a threat. https://citylab.inria.fr/

8.1.6. VALDO (Valorisation et monétisation des données personnelles à l'ère du Big Data), Digital Society Institute (DSI) (May 2015 - Sept. 2016).

Partners: DANTE and SMIS (co-organizers), CERDI, RITM.
SMIS funding: 50K€.

The objective of this project is to study with a multidisciplinary approach (i.e., computer science, law and economics) the impact of putting a certain (e.g., monetary) value on personal data, over the behavior of individuals (that are the rightful owners of the data) and market companies (that make usage of the personal data) in terms of data protection practices and data usage.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- Philippe Pucheral: Co-founder of the bi-annual French Summer School 'Masses de Données Distribuées' and co-organizer of this school in 2016

9.1.1.2. Member of the Organizing Committees

- Benjamin Nguyen: Steering commitee of 'Atelier sur la Protection de la Vie Privée' (APVP) in 2016
- Nicolas Anciaux: co-organizer (with Fabrice Le Guel and Ulysse Roux) of the 'Journée de restitution d'ISN' in 2016

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- Philippe Pucheral: DATA'16, DBKDA'16, BDA'16, EDBT'17
- Luc Bouganim: VLDB'17, EDBT'16, BDA'16 (Demo)
- Benjamin Nguyen: ECML/PKDD'16
- Nicolas Anciaux: SIGMOD'17, EDBT'16, DATA'16, ICT4AWE'16, BDA'16, RESSI'16
- Iulian Sandu Popa: IEEE MobileCloud'16, DATA'16, MobilWare'16, MCSMS'16

9.1.2.2. Reviewer

- Luc Bouganim: EDBT'17
- Iulian Sandu Popa: MDM'16
- Guillaume Scerri : SIGMOD'17, EDBT'17, STACS'17

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Nicolas Anciaux: Associate Editor of the VLDB Journal (since 2015)
- Benjamin Nguyen: Techniques et Sciences Informatiques (TSI) - Revue Française (Rédacteur adjoint, domaine Sécurité et Vie Privée depuis 2016-05)
- Luc Bouganim: Area Editor for PVLDB 2018

9.1.3.2. Reviewer - Reviewing Activities

- Iulian Sandu Popa: IEEE TKDE, ACM Transactions on Storage, International Journal of Digital Earth

9.1.4. Invited Talks

- Invited tutorial: N. Anciaux, L. Bouganim, Tutorial: Decentralized Personal Data Management, using secure devices, BDA summer school: "Masses de données distribuées", Urugne, June 2016 [28].
- Invited talk: N. Anciaux, La valorisation de la donnée par l'usager : approche technique. Centre français de droit comparé, Colloque 'Propriété(s) et données', Paris, 2016 https://linc.cnil.fr/fr/le-centre-francais-de-droit-compare-organise-un-colloque-proprietes-et-donnees
- Invited talk: L. Bouganim, Remy Pasquion, A Secure and Mobile Platform for Experimental Economy, Multidisciplinary Research Seminar at Telecom Paris Tech, Paris, June 2016.

- Invited talk: N. Anciaux, Contrôler ses données personnelles par la décentralisation. Journée Protection de la vie privée : Nouveaux instruments techniques et juridiques, CNRS, November 2016. http://www.univ-orleans.fr/lifo/evenements/CAPPRIS/?page_id=125
- Invited demonstration: L. Bouganim and R. Pasquion, Présentation d'une Plateforme de Questionnaires Respectueux de la Vie Privée, Sommet « Partenariat pour un Gouvernement ouvert », Agora numérique « Pour un Sénat ouvert », December, 8, 2016.
- Invited demonstrations: N. Anciaux, L. Bouganim and P. Tran Van, Partager et disposer de ses données en en gardant la maîtrise, trois démonstrations sur la base de PlugDB: Mobile Platform for Experimental Economy, Secure personal cloud (with Cozy Cloud), Personal Social-Medical Folder. Journée Protection de la vie privée : Nouveaux instruments techniques et juridiques, CNRS, November 2016. http://www.univ-orleans.fr/lifo/evenements/CAPPRIS/?page_id=125
- Invited talk: Iulian Sandu Popa, Distributed Architectures for Privacy-Aware Mobile Participatory Sensing, BIS'16 Workshop, June 2016 https://project.inria.fr/siliconvalley/bis2016-day-2-full-day-working-session-on-smart-cities/
- Invited talk: N. Anciaux, A new approach for secure personal cloud, CITI Talk, Lyon, May 2016 http://www.citi-lab.fr/2016/05/17/citi-talk-a-new-approach-for-secure-personal-cloud-by-nicolas-anciaux-on-26th-may/
- Panel: Nicolas Anciaux, Enjeux et opportunités des modèles ouverts en santé, Paris Open Source Summit, 2016. Trackleader : Olivier de Fresnoye. Panelists : Nicolas Anciaux (Inria), Mobin Yasini(MD, MPH, PhD, mHealth Quality R&D Director), Stéphane Gigandet (Open Food Facts), Clémentine Langlois (Fongwama) http://empoweringopeninnovation.org/enjeux-et-opportunites-des-modeles-ouverts-en-sante/

9.1.5. Scientific Expertise

- N. Anciaux, P. Pucheral : scientific expertise on behalf of DGCCRF (Direction Générale de la Concurrence, de la Consommation et de la Répression des Fraudes) and of the European Council regarding the draft proposal of "Directive of the European Parliament and of the Council on certain aspects concerning contracts for the online and other distance sales of goods" regulating the payment of numeric goods and services by means of personal data
- Benjamin Nguyen: Natural Sciences and Engineering Research Council of Canada (NSERC), Research grants
- Benjamin Nguyen: Membre du comité d'éthique de TeraLab, du Groupe des Ecoles Nationales d'Economie et de Statistique (GENES) dont l'ENSAE.

9.1.6. Research Administration

- Philippe Pucheral: Member of the HDR committee of the STV doctoral school (UVSQ) since 2014
- Philippe Pucheral: Member of the steering committee of the ED STIC doctoral school of University Paris-Saclay, 'Data, Knowledge and Interactions' committee (about 250 PhD students) since 2014
- Philippe Pucheral: Representative of Inria in the 'Comité des Tutelles Formation' of Paris-Saclay University since 2016
- Philippe Pucheral: Member of the 'Commission de Sélection' of UVSQ and of the 'Commission des Appellations' of Télécom SudParis in 2016
- Luc Bouganim: Reviewer for the ANR programs (Evaluation committee CE23), 2016
- Luc Bouganim: Member of the 'Commission de Sélection' of the INSA CVL, 2016
- Nicolas Anciaux: Co-director of the 'Privacy and digital identity' WG at Digital Society Institute (DSI), until June 2016
- Nicolas Anciaux: Reviewer of the CHIST-ERA Call 2015 - SPTIoT
- Nicolas Anciaux: Member of the 'Commission de Sélection' of UVSQ, 2016

- Benjamin Nguyen: Director of LIFO (EA 4022) since July 2016
- Benjamin Nguyen: Responsible of CNRS privacy working group "Sécurité" since september 2016
- Benjamin Nguyen: Director of Digital Affairs (INSA CVL)
- Iulian Sandu Popa: Reviewer for the ANR programs (evaluation committee CE23), 2016
- Iulian Sandu Popa: Member of the 'Commission de Sélection' of the UVSQ, 2016

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master : Philippe Pucheral, co-director of the DataScale master (UPSay), courses in M1 and M2 at UVSQ and at ENSIIE

Master: Nicolas Anciaux, Courses on database internal mechanisms and database security, 80, in Master1 and Master2 (AFTI, Orsay) and in engineering school (ENSTA ParisTech, Telecom Paristech)

Master : Luc Bouganim, Bases de données, 24, niveau M2, CFA AFTI/UVSQ, France

Master : Luc Bouganim, Systèmes d'information Privacy by Design, 54, niveau M2, ENSIIE, INSA CVL, France

Master : Luc Bouganim, Sécurité des bases de données, 10, niveau M2, Télécom ParisTech, France

Licence : Iulian Sandu Popa, Initiation aux bases de données (niveau L2), Bases de données (niveau L3), 96, UVSQ, France

Master : Iulian Sandu Popa, Bases de données relationnelles (niveau M1), Gestion des données spatiotemporelles (niveau M2), Sécurité des bases de données (niveau M2), 96, UVSQ, France

Master : Benjamin Nguyen, Databases, IA, Security, 192, INSA CVL, France

Licence : Guillaume Scerri, Fondements de l'informatique, 36, niveau L1, UVSQ, France

Licence : Guillaume Scerri, Initiation aux bases de données, 43.5, niveau L2, UVSQ, France

Master : Guillaume Scerri, Bases de données, 30, niveau M1, UVSQ France

E-learning

MOOC created in 2016: "Villes intelligentes : défis technologiques et sociétaux", coordinated by V. Issarny and N. Mitton. N. Anciaux, S.Grumbach, V. Issarny, N. Mitton, C. Morin, A. Pathak, H. Rivano. Start date 25 Jan. 2016. In 2016, more than 8000 students have registered to follow the courses. A new edition is planned for 2017. https://www.fun-mooc.fr/courses/inria/41009/session01/about

MOOC created in 2016: Bases de données relationnelles : comprendre pour maîtriser (Serge Abiteboul, Benjamin Nguyen, Philippe Rigaux) https://www.fun-mooc.fr/courses/inria/41008/session01/about

9.2.2. Supervision

PhD: Athanasia Katsouraki, Sharing and Usage Control of Personal Information, UVSQ, September 2016, Luc Bouganim and Benjamin Nguyen

PhD: Saliha Lallali, A Secure Search Engine for the Personal Cloud, UVSQ, January 2016, Nicolas Anciaux, Philippe Pucheral, and Iulian Sandu Popa

PhD: Dai Hai Ton That, Efficient Management and Secure Sharing of Mobility Traces, UVSQ, January 2016, Iulian Sandu Popa and Karine Zeitouni

PhD in progress : Paul Tran Van, Partage de documents sécurisé dans le Cloud Personnel, October 2014, Nicolas Anciaux and Philippe Pucheral

PhD in progress : Axel Michel, Secure Distributed Computations, October 2015, Benjamin Nguyen and Philippe Pucheral

PhD in progress : Julien Loudet, Personal Queries on Personal Clouds, July 2016, Luc Bouganim and Iulian Sandu Popa

PhD in progress : Riad Ladjel, Secure Distributed Computation for the Personal Cloud, October 2016, Nicolas Anciaux and Philippe Pucheral

9.2.3. Juries

Philippe Pucheral: member of the PhD jury of Mahsa Najafzadeh (Paris VI, 22/04/2016)

Benjamin Nguyen: reviewer of the PhD of Tarek SAYAH, Exposition sélective et problème de fuite d'inférence dans le Linked Data, Université Claude Bernard Lyon I, 2016/09/08

Benjamin Nguyen: reviewer of the PhD of Germain JOLLY, Evaluation of payment applications on smart cards, Université de Caen/ENSICAEN, 2016/07/08

Benjamin Nguyen: reviewer of the PhD of Julien LOLIVE, Entrelacement des mécanismes d'identification et de respect de la vie privée pour la protection des contenus externalisés, Télécom Bretagne, 2016/05/13

Benjamin Nguyen: president and reviewer of the PhD jury of Karina SOKOLOVA, Bridging the gap between Privacy by Design and mobile systems by patterns, Université Technologique de Troyes, 2016/04/27

9.3. Popularization

General public (large audience magazines, television, videos):

- "Santé connectée. Une médecine sans médecin ?", interview of N. Anciaux, La Recherche, N°510 Avril 2016. (http://fliphtml5.com/ggwf/ynja p.84)
- "PlugDB, coffre-fort numérique personnel", interview of P.Pucheral, Le Monde Science & Techno, November 28th, 2016.
- "Sur Internet, vos données personnelles valent de l'or", interview of P.Pucheral, Ca m'intéresse, November 2016.
- "Santé connectée, santé partagée", interview of P.Pucheral, Le Point, April 7th 2016.

10. Bibliography

Major publications by the team in recent years

[1] T. ALLARD, B. NGUYEN, P. PUCHERAL. *MetaP: Revisiting Privacy-Preserving Data Publishing using Secure Devices*, in "Distributed and Parallel Databases", June 2014, vol. 32, n° 1, pp. 191-244 [*DOI :* 10.1007/s10619-013-7122-x], https://hal.archives-ouvertes.fr/hal-00934586

[2] N. ANCIAUX, P. BONNET, L. BOUGANIM, B. NGUYEN, P. PUCHERAL, I. SANDU POPA. *Trusted Cells : A Sea Change for Personnal Data Services*, in "CIDR 2013 - 6th Biennal Conference on Innovative Database Research", Asilomar, United States, 2013, 4 p. , http://hal.inria.fr/hal-00768379

[3] N. ANCIAUX, L. BOUGANIM, T. DELOT, S. ILARRI, L. KLOUL, N. MITTON, P. PUCHERAL. *Folk-IS: Opportunistic Data Services in Least Developed Countries*, in "40th International Conference on Very Large Data Bases (VLDB)", Hangzhou, China, Zhejiang University, September 2014, https://hal.inria.fr/hal-00906204

[4] N. ANCIAUX, L. BOUGANIM, P. PUCHERAL, Y. GUO, L. LE FOLGOC, S. YIN. *MILo-DB: a personal, secure and portable database machine*, in "Distributed and Parallel Databases", March 2014, vol. 32, n° 1, pp. 37-63 [*DOI :* 10.1007/s10619-012-7119-x], https://hal.archives-ouvertes.fr/hal-00768355

[5] N. ANCIAUX, S. LALLALI, I. SANDU POPA, P. PUCHERAL. *A Scalable Search Engine for Mass Storage Smart Objects*, in "Proceedings of the 41th International Conference on Very Large Databases (VLDB)", Kohala Coast, Hawaii, United States, August 2015, vol. 8, n° 9, pp. 910-921 [*DOI :* 10.14778/2777598.2777600], https://hal.inria.fr/hal-01176458

[6] N. ANCIAUX, B. NGUYEN, I. SANDU POPA. *Tutorial: Managing Personal Data with Strong Privacy Guarantees*, March 2014, pp. 672-673 [*DOI :* 10.5441/002/EDBT.2014.71], https://hal.inria.fr/hal-01096633

[7] G. SCERRI, B. WARINSCHI, M. BARBOSA, B. PORTELA. *Foundations of Hardware-Based Attested Computation and Application to SGX*, March 2016, pp. 245-260 [*DOI :* 10.1109/EUROSP.2016.28], https://hal.inria.fr/hal-01417137

[8] Q. C. TO, B. NGUYEN, P. PUCHERAL. *Privacy-Preserving Query Execution using a Decentralized Architecture and Tamper Resistant Hardware*, in "17th International Conference on Extending Database Technology (EDBT)", Athens, Greece, March 2014 [*DOI :* 10.5441/002/EDBT.2014.44], https://hal.inria.fr/hal-01096639

[9] C. Q. TO, B. NGUYEN, P. PUCHERAL. *TrustedMR: A Trusted MapReduce System based on Tamper Resistance Hardware* , in "Proceedings of the 23rd International Conference on Cooperative Information Systems (COOPIS)", Rhodes, Greece, October 2015, pp. 38-56 [*DOI :* 10.1007/978-3-319-26148-5_3], https://hal.inria.fr/hal-01254951

[10] D. H. TON THAT, I. SANDU POPA, K. ZEITOUNI. *TRIFL: A Generic Trajectory Index for Flash Storage*, in "ACM Transactions on Algorithms", July 2015, vol. 1, n° 2, 44 p. [*DOI :* 10.1145/2786758], https://hal.inria.fr/hal-01176563

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] A. KATSOURAKI. *Sharing and Usage Control of Personal Information*, Université Paris Saclay, September 2016, https://hal.archives-ouvertes.fr/tel-01425638

[12] S. LALLALI. *A Scalable Search Engine for the Personal Cloud*, Université Paris-Saclay, January 2016, https://hal.inria.fr/tel-01426486

[13] D. H. TON THAT. *Efficient management and secure sharing of mobility traces*, Université Paris-Saclay, January 2016, https://tel.archives-ouvertes.fr/tel-01290834

Articles in International Peer-Reviewed Journals

[14] S. J. PAN, I. SANDU POPA, C. BORCEA. *DIVERT: A Distributed Vehicular Traffic Re-Routing System for Congestion Avoidance*, in "IEEE Transactions on Mobile Computing", January 2017, vol. 16, n° 1, pp. 58-72 [*DOI :* 10.1109/TMC.2016.2538226], https://hal.inria.fr/hal-01426424

[15] C. Q. TO, B. NGUYEN, P. PUCHERAL. *Private and Scalable Execution of SQL Aggregates on a Secure Decentralized Architecture*, in "ACM Transactions on Database Systems", 2016, forthcoming, https://hal.archives-ouvertes.fr/hal-01296432

[16] D. H. TON THAT, I. SANDU POPA, K. ZEITOUNI, C. BORCEA. *Un protocole basé sur des mobiles sécurisés pour une collecte participative de données spatiales en mobilité réellement anonyme*, in "Revue Internationale de Géomatique", August 2016, vol. 26, n° 2, pp. 185-210 [*DOI :* 10.3166/RIG.26.185-210], https://hal.inria.fr/hal-01426358

Articles in National Peer-Reviewed Journals

[17] N. ANCIAUX, P. PUCHERAL, M. BEHAR-TOUCHAIS, V.-L. BENABOU, G. BRUNAUX, A. LEFEVRE, N. MARTIAL-BRAZ, J. ROCHFELD, N. SAUPHANOR-BROUILLAUD, B. SCHULZ, J. SENECHAL, C. ZOLYNSKI. *Dossier Contenus Numériques Revue Contrats, Concurrence, Consommation - Contenus Numériques*, in "Revue Contrats Concurrence Consommation", February 2017, https://hal.archives-ouvertes.fr/hal-01432544

[18] C. BERTHET, C. ZOLYNSKI, N. ANCIAUX, P. PUCHERAL. *" Contenus numériques et récupération des données : un nouvel outil d' 'empouvoirement' du consommateur ? "*, in "Dalloz IP/IT", January 2017, vol. IP IT / 10, https://hal.inria.fr/hal-01429939

[19] P. PUCHERAL, N. ANCIAUX, M. BEHAR-TOUCHAIS, V.-L. BENABOU, N. MARTIAL-BRAZ, J. ROCHFELD, N. SAUPHANOR-BROUILLAUD, B. SCHULZ, J. SENECHAL, C. ZOLYNSKI. *Dossier Contenus Numériques / Données*, in "Contrats concurrence consommation", February 2017, https://hal.inria.fr/hal-01429951

[20] P. PUCHERAL, A. RALLET, F. ROCHELANDET, C. ZOLYNSKI. *La Privacy by design : une fausse bonne solution aux problèmes de protection des données personnelles soulevés par l'Open data et les objets connectés ?*, in "Legicom", January 2016, n° 56, pp. 89-99 [*DOI :* 10.3917/LEGI.056.0089], https://hal.archives-ouvertes.fr/hal-01427983

International Conferences with Proceedings

[21] L. BOUGANIM, A. KATSOURAKI, B. NGUYEN. *DatShA :A Data Sharing Algebra for access control plans*, in "19th International Conference on Extending Database Technology (EDBT 2016)", Bordeaux, France, Proceedings of the 19th International Conference on Extending Database Technology (short paper), March 2016, https://hal.archives-ouvertes.fr/hal-01289023

[22] T. B. T. LE, N. ANCIAUX, S. GILLOTON, S. LALLALI, P. PUCHERAL, I. SANDU POPA, C. CHEN. *Distributed Secure Search in the Personal Cloud*, in "19th International Conference on Extending Database Technology (EDBT 2016)", Bordeaux, France, Proceedings of EDBT'16 (Demo paper), March 2016, pp. 652-655, https://hal.inria.fr/hal-01293409

[23] D. H. TON THAT, I. SANDU POPA, K. ZEITOUNI, C. BORCEA. *PAMPAS: Privacy-Aware Mobile Participatory Sensing Using Secure Probes*, in "International Conference on Scientific and Statistical Database Management (SSDBM '16)", Budapest, Hungary, Proceedings of the 28th International Conference on Scientific and Statistical Database Management, July 2016 [*DOI :* 10.1145/1235], https://hal.inria.fr/hal-01426375

Conferences without Proceedings

[24] A. MICHEL, B. NGUYEN. *Managing distributed queries under anonymity constraints*, in "7e Atelier sur la Protection de la Vie Privée", Toulouse, France, July 2016, https://hal.archives-ouvertes.fr/hal-01424993

[25] A. MICHEL, B. NGUYEN, P. PUCHERAL. *Exécution de requêtes distribuées sous contraintes d'anonymat*, in "32ème Conférence sur la Gestion de Données - Principes, Technologies et Applications", Poitiers, France, November 2016, https://hal.archives-ouvertes.fr/hal-01424989

Scientific Books (or Scientific Book chapters)

[26] P. PUCHERAL, A. RALLET, C. ZOLYNSKI. *Privacy by Design et Big Data*, in "Les big data à découvert", 2016, https://hal.archives-ouvertes.fr/hal-01429075

Scientific Popularization

[27] N. ANCIAUX, B. ANDRÉ, P. PUCHERAL, P. TRAN-VAN. *A Root of Trust for the Personal Cloud*, in "ERCIM News", July 2016, https://hal.archives-ouvertes.fr/hal-01427725

[28] N. ANCIAUX, L. BOUGANIM. *Decentralized Personal Data Management, using secure devices (Tutorial)*, June 2016, BDA summer school: "Masses de données distribuées", Urugne, France, https://hal.archives-ouvertes.fr/hal-01425656

References in notes

[29] R. AGRAWAL, J. KIERNAN, R. SRIKANT, Y. XU. *Hippocratic Databases*, in "Proc. of the Int. Conf. on Very Large Data Bases (VLDB)", 2002

[30] T. ALLARD, N. ANCIAUX, L. BOUGANIM, Y. GUO, L. LE FOLGOC, B. NGUYEN, P. PUCHERAL, I. RAY, I. RAY, S. YIN. *Secure Personal Data Servers: a Vision Paper*, in "Proc. of the 36th Int. Conf. on Very Large Databases (VLDB)", 2010

[31] N. ANCIAUX, B. NGUYEN, I. SANDU POPA. *Towards an Era of Trust in Personal Data Management*, in "Proceedings of the 19th East-European Conference on Advances in Databases and Information Systems (ADBIS '15). Tutorial", Poitiers, France, 2015, https://hal.inria.fr/hal-01176512

[32] A. ANDERSON. *An introduction to the web services policy language (WSPL)*, in "IEEE Computer Society", 2004

[33] R. BAHMANI, M. BARBOSA, F. BRASSER, B. PORTELA, A. SADEGHI, G. SCERRI, B. WARINSCHI. *Secure Multiparty Computation from SGX*, in "IACR Cryptology ePrint Archive", 2016, vol. 2016, 1057 p.

[34] P. BONNET, L. BOUGANIM, I. KOLTSIDAS, S. VIGLAS. *System Co-Design and Data Management for Flash Devices*, in "Very Large Data Bases (Tutorial)", 2011

[35] L. BOUGANIM, Y. GUO. *Database Encryption*, in "Encyclopedia of Cryptography and Security", S. JAJODIA, H. VAN TILBORG (editors), Springer, 2009, pp. 307-312

[36] L. CRANOR. *Web Privacy with P3P*, O'Reilly Media, 2002

[37] B. FUNG, K. WANG, R. CHEN, P. YU. *Privacy-preserving data publishing: A survey of recent developments*, in "ACM Computing Surveys (CSUR)", 2010, vol. 42, n° 4

[38] S. LALLALI, N. ANCIAUX, I. SANDU POPA, P. PUCHERAL. *A Secure Search Engine for the Personal Cloud*, in "Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD '15). Demo paper", Melbourne, Australia, 2015, pp. 1445-1450 [*DOI :* 10.1145/2723372.2735376], https://hal.inria.fr/hal-01176473

[39] T. MOSES. *Extensible access control markup language (XACML) version 2.0*, in "Oasis Standard 200502", 2005

[40] M. SCHUNTER, C. POWERS. *Enterprise privacy authorization language (EPAL 1.1)*, in "IBM", 2003

[41] D. H. TON THAT, I. SANDU POPA, K. ZEITOUNI. *PPTM: Privacy-aware Participatory Traffic Monitoring Using Mobile Secure Probes*, in "Proceedings of the 16th IEEE International Conference on Mobile Data Management (MDM '15). Demo paper", Pittsburgh, United States, 2015, 4 p. , https://hal.inria.fr/hal-01176486