Activity Report 2011

Project-Team CARAMEL

Cryptology, Arithmetic: Hardware and Software

IN COLLABORATION WITH: Laboratoire lorrain de recherche en informatique et ses applications (LORIA)
# Table of contents

1. Members

2. Overall Objectives
   2.1. Introduction
   2.2. Highlights

3. Scientific Foundations

4. Application Domains
   4.1.1. Cryptography
   4.1.2. Cryptanalysis
   4.1.3. Standardization
      4.1.3.1. Floating-point arithmetic
      4.1.3.2. Curve-based cryptography
      4.1.3.3. Pairing-based cryptography
   4.1.4. Computer algebra systems
      4.1.4.1. Magma
      4.1.4.2. Pari-GP
      4.1.4.3. Sage

5. Software

6. New Results
   6.1. NFS-related results
   6.2. Ballot stuffing in a postal voting system
   6.3. Symmetric cryptanalysis
   6.4. Implementation of cryptographic pairings
   6.5. Multiple-precision arithmetic
   6.6. Proving the complexity of computing endomorphism rings
   6.7. Point counting on curves with real multiplication
   6.8. Computation of isogenies between abelian varieties

7. Partnerships and Cooperations
   7.1. Regional Initiatives
   7.2. National Initiatives
      7.2.1. ANR DEMOTIS (Collaborative Analysis, Evaluation and Modelling of Health Information Technology)
      7.2.2. ANR CHIC (Courbes Hyperelliptiques, Isogénies, Comptage)
   7.3. European Initiatives

8. Dissemination
   8.1. Animation of the scientific community
      8.1.1. Caramel seminar
      8.1.2. Conference organization
   8.2. Committees memberships
   8.3. Vulgarization
   8.4. Invited conferences
   8.5. Teaching
9. Bibliography ................................................................. 14
Project-Team CARAMEL

Keywords: Algorithmic Numbers Theory, Cryptography, Computer Arithmetic, Hardware Accelerators

1. Members

Research Scientists
- Pierrick Gaudry [Team leader, Senior Researcher, CNRS, HdR]
- Jérémy Detrey [Junior Researcher, INRIA]
- Emmanuel Thomé [Junior Researcher, INRIA]
- Paul Zimmermann [Senior Researcher, INRIA, part time, HdR]

Faculty Member
- Marion Videau [Associate Professor, Université Henri Poincaré; on secondment to ANSSI until August 2011]

Technical Staff
- Lionel Müller [ADT grant until October 2011]

PhD Students
- Răzvan Barbulescu [Contrat doctoral, Université Henri Poincaré; started in September 2011]
- Gaëtan Bisson [MESR grant, INPL and Technische Universität Eindhoven; defended on July 14th]
- Romain Cosset [INRIA/DGA grant; until August 31st; defended on November 7th]
- Nicolas Estibals [Contrat doctoral, Université Henri Poincaré; defense planned in 2012]
- Hamza Jeljeli [Contrat doctoral, Université Henri Poincaré; started in September 2011]

Post-Doctoral Fellows
- Sorina Ionica [From November 1st]
- Pascal Molin [From December 1st, 2010 to August 31, 2011]

Administrative Assistant
- Emmanuelle Deschamps [part time]

Others
- Cyril Bouvier [Internship, ENS]
- Benoît Gaudel [Internship, Université de Versailles Saint Quentin; April – July 2011]

2. Overall Objectives

2.1. Introduction

A general keyword that could encompass most of our research objectives is arithmetic. Indeed, in the CARAMEL team, the goal is to push forward the possibilities to compute efficiently with objects having an arithmetic nature. This includes integers, real and complex numbers, polynomials, finite fields, and, last but not least, algebraic curves.

Our main application domains are public-key cryptography and computer algebra systems. Concerning cryptography, we concentrate on the study of the primitives based on the factorization problem or on the discrete-logarithm problem in finite fields or (Jacobians of) algebraic curves. Both the constructive and destructive sides are of interest to CARAMEL. For applications in computer algebra systems, we are mostly interested in arithmetic building blocks for integers, floating-point numbers, polynomials, and finite fields. Also some higher level functionalities like factoring and discrete-logarithm computation are usually desired in computer algebra systems.
Since we develop our expertise at various levels, from most low-level software or hardware implementation of basic building blocks to complicated high-level algorithms like integer factorization or point counting, we have remarked that it is often too simple-minded to separate them: we believe that the interactions between low-level and high-level algorithms are of utmost importance for arithmetic applications, yielding important improvements that would not be possible with a vision restricted to low- or high-level algorithms.

We emphasize three main directions in the CARAMEL team:

- Integer factorization and discrete-logarithm computation in finite fields.
  We are in particular interested in the number field sieve algorithm (NFS) that is the best known algorithm for factoring large RSA-like integers, and for solving discrete logarithms in prime finite fields. A sibling algorithm, the function field sieve (FFS), is the best known algorithm for computing discrete logarithms in finite fields of small characteristic.
  In all these cases, we plan to improve on existing algorithms, with a view towards practical considerations and setting new records.

- Algebraic curves and cryptography.
  Our two main research interests on this topic lie in genus-2 cryptography and in the arithmetic of pairings, mostly on the constructive side in both cases. For genus-2 curves, a key algorithmic tool that we develop is the computation of explicit isogenies; this allows improvements for cryptography-related computations such as point counting in large characteristic, complex-multiplication construction and computation of the ring of endomorphisms.
  For pairings, our principal concern is the optimization of pairing computations, in particular in hardware, or in constrained environments. We plan to develop automatic tools to help in choosing the most suitable (hyper-)elliptic curve and generating efficient hardware for a given security level and set of constraints.

- Arithmetic.
  Integer, finite-field and polynomial arithmetics are ubiquitous to our research. We consider them not only as tools for other algorithms, but as a research theme per se. We are interested in algorithmic advances, in particular for large input sizes where asymptotically fast algorithms become of practical interest. We also keep an important implementation activity, both in hardware and in software.

2.2. Highlights

The highlights for year 2011 in the CARAMEL team are

- the successful organization of the ECC conference, that gathered more than 120 participants;
- the publication of the major result of Robert and Lubicz on explicit isogenies in the prestigious journal Compositio Mathematica.

3. Scientific Foundations

3.1. Cryptography, arithmetic: hardware and software

One of the main topics for our project is public-key cryptography. After 20 years of hegemony, the classical public-key algorithms (whose security is based on integer factorization or discrete logarithm in finite fields) are currently being overtaken by elliptic curves. The fundamental reason for this is that the best-known algorithms for factoring integers or for computing discrete logarithms in finite fields have a subexponential complexity, whereas the best known attack for elliptic-curve discrete logarithms has exponential complexity.

As a consequence, for a given security level $2^n$, the key sizes must grow linearly with $n$ for elliptic curves, whereas they grow like $n^3$ for RSA-like systems. As a consequence, several governmental agencies, like the NSA or the BSI, now recommend to use elliptic-curve cryptosystems for new products that are not bound to RSA for backward compatibility.
Besides RSA and elliptic curves, there are several alternatives currently under study. There is a recent trend to promote alternate solutions that do not rely on number theory, with the objective of building systems that would resist a quantum computer (in contrast, integer factorization and discrete logarithms in finite fields and elliptic curves have a polynomial-time quantum solution). Among them, we find systems based on hard problems in lattices (NTRU is the most famous), those based on coding theory (McEliece system and improved versions), and those based on the difficulty to solve multivariate polynomial equations (HFE, for instance). None of them has yet reached the same level of popularity as RSA or elliptic curves for various reasons, including the presence of unsatisfactory features (like a huge public key), or the non-maturity (system still alternating between being fixed one day and broken the next day).

Returning to number theory, an alternative to RSA and elliptic curves is to use other curves and in particular genus-2 curves. These so-called hyperelliptic cryptosystems have been proposed in 1989 [26], soon after the elliptic ones, but their deployment is by far more difficult. The first problem was the group law. For elliptic curves, the elements of the group are just the points of the curve. In a hyperelliptic cryptosystem, the elements of the group are points on a 2-dimensional variety associated to the genus-2 curve, called the Jacobian variety. Although there exist polynomial-time methods to represent and compute with them, it took some time before getting a group law that could compete with the elliptic one in terms of speed. Another question that is still not yet fully answered is the computation of the group order, which is important for assessing the security of the associated cryptosystem. This amounts to counting the points of the curve that are defined over the base field or over an extension, and therefore this general question is called point-counting. In the past ten years there have been major improvements on the topic, but there are still cases for which no practical solution is known.

Another recent discovery in public-key cryptography is the fact that having an efficient bilinear map that is hard to invert (in a sense that can be made precise) can lead to powerful cryptographic primitives. The only examples we know of such bilinear maps are associated with algebraic curves, and in particular elliptic curves: this is the so-called Weil pairing (or its variant, the Tate pairing). Initially considered as a threat for elliptic-curve cryptography, they have proven to be quite useful from a constructive point of view, and since the beginning of the decade, hundreds of articles have been published, proposing efficient protocols based on pairings. A long-lasting open question, namely the construction of a practical identity-based encryption scheme, has been solved this way. The first standardization of pairing-based cryptography has recently occurred (see ISO/IEC 14888-3 or IEEE P1363.3), and a large deployment is to be expected in the next years.

Despite the raise of elliptic curve cryptography and the variety of more or less mature other alternatives, classical systems (based on factoring or discrete logarithm in finite fields) are still going to be widely used in the next decade, at least, due to resilience: it takes a long time to adopt new standards, and then an even longer time to renew all the software and hardware that is widely deployed.

This context of public-key cryptography motivates us to work on integer factorization, for which we have acquired expertise, both in factoring moderate-sized numbers, using the ECM (Elliptic Curve Method) algorithm, and in factoring large RSA-like numbers, using the number field sieve algorithm. The goal is to follow the transition from RSA to other systems and continuously assess its security to adjust key sizes. We also want to work on the discrete-logarithm problem in finite fields. This second task is not only necessary for assessing the security of classical public-key algorithms, but is also crucial for the security of pairing-based cryptography.

We also plan to investigate and promote the use of pairing-based and genus-2 cryptosystems. For pairings, this is mostly a question of how efficient can such a system be in software, in hardware, and using all the tools from fast implementation to the search for adequate curves. For genus 2, as said earlier, constructing an efficient cryptosystem requires some more fundamental questions to be solved, namely the point-counting problem.

We summarize in the following table the aspects of public-key cryptography that we address in the CARAMEL team.
Another general application for the project is computer algebra systems (CAS), that rely in many places on efficient arithmetic. Nowadays, the objective of a CAS is not only to have more and more features that the user might wish, but also to compute the results fast enough, since in many cases, the CAS are used interactively, and a human is waiting for the computation to complete. To tackle this question, more and more CAS use external libraries, that have been written with speed and reliability as first concern. For instance, most of today’s CAS use the GMP library for their computations with big integers. Many of them will also use some external Basic Linear Algebra Subprograms (BLAS) implementation for their needs in numerical linear algebra.

During a typical CAS session, the libraries are called with objects whose sizes vary a lot; therefore being fast on all sizes is important. This encompasses small-sized data, like elements of the finite fields used in cryptographic applications, and larger structures, for which asymptotically fast algorithms are to be used. For instance, the user might want to study an elliptic curve over the rationals, and as a consequence, check its behaviour when reduced modulo many small primes; and then he can search for large torsion points over an extension field, which will involve computing with high-degree polynomials with large integer coefficients.

Writing efficient software for arithmetic as it is used typically in CAS requires the knowledge of many algorithms with their range of applicability, good programming skills in order to spend time only where it should be spent, and finally good knowledge of the target hardware. Indeed, it makes little sense to disregard the specifics of the possible hardware platforms intended, even more so since in the past years, we have seen a paradigm shift in terms of available hardware: so far, it used to be reasonable to consider that an end-user running a CAS would have access to a single-CPU processor. Nowadays, even a basic laptop computer has a multi-core processor and a powerful graphics card, and a workstation with a reconfigurable coprocessor is no longer science-fiction.

In this context, one of our goals is to investigate and take advantage of these influences and interactions between various available computing resources in order to design better algorithms for basic arithmetic objects. Of course, this is not disconnected from the others goals, since they all rely more or less on integer or polynomial arithmetic.

### 4. Application Domains

#### 4.1. Cryptology

The first application domain for our research is cryptology. This includes cryptography (constructive side) and cryptanalysis (breaking systems). For the cryptanalysis part, although it has practical implications, we do not expect any transfer in the classical sense of the term: it is more directed to governmental agencies and the end-users who build their trust, based on the cryptanalysis effort.

##### 4.1.1. Cryptography

Our cryptographic contributions are related to multiple facets of the large realm of curve-based cryptology. While it is quite clear that a satisfying range of algorithms exists in order to provide cryptographers with elliptic curves having a suitably hard discrete logarithm (as found in cryptographic standards for instance), one must bear in mind that refinements of the requirements and extensions to curves of higher genus raise several interesting problems. Our work contributes to expanding the cryptographer’s capabilities in these areas.
In the context of genus-2 curves, our work aims at two goals. First, improvements on the group law on selected curves yield better speed for the associated cryptosystems. The cryptographic primitives, and then the whole suite of cryptographic protocols built upon such curves would be accelerated. The second goal is the expansion of the set of curves that can be built given a set of desired properties. Using point counting algorithms for arbitrary curves, a curve offering a 128-bit security level, together with nice properties for fast arithmetic, has been computed by CARAMEL[10]. Another natural target for construction of curves for cryptography is also the suitability of curves for pairings. We expect to be able to compute such curves.

Implementations of curve-based cryptography, both in hardware and software, are a necessary step on the way to assessing cryptographic speed. We plan to provide such implementations. In particular, on the hardware side, one of our goals is the design of a complete cryptographic coprocessor, including all the primitives for curve-based and pairing-based cryptography, providing optimized and configurable efficiency vs area trade-off.

4.1.2. Cryptanalysis

Our research on cryptanalysis is important for the cryptographic industry: by detecting weak instances, and setting new records we contribute to the definition of recommended families of systems together with their key sizes. The user’s confidence in a cryptographic primitive is also related to how well the underlying problem is studied by researchers.

In particular, our involvement in computations with “NFS-like” algorithms encompasses of course the task of assessing the computational limits for integer factorization and discrete-logarithm computations. The impact of the former is quite clear as it concerns the RSA algorithm; record-sized computations attract broad interest and determine updates on key-length recommendations. The latter are particularly important for pairing-based cryptography, since, in this context, one naturally encounters discrete-logarithm problems in extension fields of large degree.

4.1.3. Standardization

4.1.3.1. Floating-point arithmetic

The IEEE 754 standard for floating-point arithmetic was revised in 2008. The main new features are some new formats for decimal computations, and the recommendation of correctly rounded transcendental functions. The new decimal formats should not have an impact on our work, since we either use integer-only arithmetic, or arbitrary-precision binary floating-point arithmetic through the GNU MPFR library.

A new standard (P1788) is currently under construction for interval arithmetic. We are not officially involved in this standard, but we follow the discussions, to check in particular that the proposed standard will also cover arbitrary precision (interval) arithmetic.

4.1.3.2. Curve-based cryptography

Elliptic-curve cryptography has been standardized for almost 10 years now, in the IEEE P1363 standard. This standard provides key agreement, signature and encryption schemes, based on integer factorization, discrete logarithm in finite fields and in elliptic curves. There is another standardization effort, called SECG, which is mostly lead by the Certicom company, with the goal to maintain interoperability between different implementations. In particular, the SECG documents give explicit elliptic curves that can be used for cryptography. Similarly, some elliptic curves have been standardized by the US government; the latest version comes from the NSA Suite B that includes only elliptic curves defined over prime fields.

In the long term, those standards are a natural place to promote genus-2 curve cryptography, and by the time we consider that the curves we propose are mature enough, we will look for an industrial partner to help us pushing towards their standardization.

4.1.3.3. Pairing-based cryptography

Despite their very recent discovery, identity-based cryptosystems—and more generally pairing-based cryptosystems—have already spawned several international standardization efforts.
The first standard, part of ISO/IEC 14888-3, was published in 2006. However, it almost exclusively focuses on protocols and therefore is of little interest to us. On the other hand, the IEEE P1363.3 standard, which is still in preparation, is planned to offer more details as to the considered curves and pairings on which the protocols are based.

Although we are not officially involved in the elaboration of this standard, we have already participated in the review process of its first draft.

4.1.4. Computer algebra systems

Some of our software libraries are being used by computer algebra systems. Most of those libraries are free software, with a license that allows proprietary systems to link them. This gives us a maximal visibility, with a large number of users.

4.1.4.1. Magma

Magma is a very large computational algebra package. It provides a mathematically rigorous environment for computing with algebraic, number-theoretic, combinatoric, and geometric objects. It is developed in Sydney, by the team around John Cannon. It is non-commercial (in the sense that its goal is not to make profit), but is not freely distributed and is not open-source.

Several members of the team have visited Sydney — a few years ago — to contribute to the development of Magma, by implementing their algorithms or helping in integrating their software. Our link to Magma exists also via the libraries it uses: it currently links GNU MPFR and MPC for its floating-point calculations, and links GMP-ECM as part of its factorization suite.

4.1.4.2. Pari-GP

Pari/GP is a computational number theory system that is composed of a C library and an interpreter on top of it. It is developed in Bordeaux, where Karim Belabas from the LFANT project-team is the main maintainer. Its license is GPL. Although we do not directly contribute to this package, we have good contact with the developers and in the future, GNU MPFR and MPC could be included.

4.1.4.3. Sage

Sage is a fairly large scale and open-source computer algebra system written in Python. Sage aggregates a large amount of existing free software, aiming at the goal of selecting the fastest free software package for each given task. The motto of Sage is that instead of “reinventing the wheel” all the time, Sage is “building the car”. To date, Sage links GNU MPFR, GMP-ECM, and MPC as optional package since 2010 (this was the result of a huge work done by Philippe Théveny in the MPtools ODL which finished in 2009). Plans exist to link GF2X and CADO-NFS into Sage.

5. Software

5.1. Introduction

A major part of the research done in the CARAMEL team is published within software. On the one hand, this enables everyone to check that the algorithms we develop are really efficient in practice; on the other hand, this gives other researchers — and us of course — basic software components on which they — and we — can build other applications.

5.2. GNU MPFR

Participant: Paul Zimmermann [contact].
GNU MPFR is one of the main pieces of software developed by the CARAMEL team. Since end 2006, with the departure of Vincent Lefèvre to ENS Lyon, it has become a joint project between CARAMEL and the ARÉNAIRE project-team (INRIA Grenoble - Rhône-Alpes). GNU MPFR is a library for computing with arbitrary precision floating-point numbers, together with well-defined semantics, and is distributed under the LGPL license. All arithmetic operations are performed according to a rounding mode provided by the user, and all results are guaranteed correct to the last bit, according to the given rounding mode.

Several software systems use GNU MPFR, for example: the GCC and GFORTRAN compilers; the SAGE computer algebra system; the KDE calculator Abakus by Michael Pyne; CGAL (Computational Geometry Algorithms Library) developed by the Geometrica project-team (INRIA Sophia Antipolis - Méditerranée); Gappa, by Guillaume Melquiond; Sollya, by Sylvain Chevillard, Mioara Joldeș and Christoph Lauter; Genius Math Tool and the GEL language, by Jiri Lebl; Giac/Xcas, a free computer algebra system, by Bernard Parisse; the iRRAM exact arithmetic implementation from Norbert Müller (University of Trier, Germany); the Magma computational algebra system; and the Wcalc calculator by Kyle Wheeler.

The main developments in 2011 are the release of version 3.0.1 in April, and the release of version 3.1.0 (the “canard à l’orange” release) in October. The main changes in GNU MPFR 3.1.0 are the following: thread local storage (TLS) support is now detected automatically, the squaring and division routines got a major speed up thanks to Mulders’ algorithm [20], and a new divide-by-zero exception was introduced. Note that the automatic TLS support did exhibit several compiler bugs (http://www.loria.fr/~zimmerma/software/compilerbugs.html). We had a developers meeting in January 13-14, and in August GNU MPFR was presented at the GNU Hackers Meeting in Paris.

5.3. MPC

Participant: Paul Zimmermann [contact].

MPC is a floating-point library for complex numbers, which is developed on top of the GNU MPFR library, and distributed under the LGPL license. It is co-written with Andreas Enge (LFANT project-team, INRIA Bordeaux - Sud-Ouest). A complex floating-point number is represented by $x + iy$, where $x$ and $y$ are real floating-point numbers, represented using the GNU MPFR library. The MPC library provides correct rounding on both the real part $x$ and the imaginary part $y$ of any result. MPC is used in particular in the TRIP celestial mechanics system developed at IMCCE (Institut de Mécanique Céleste et de Calcul des Éphémérides), and by the Magma computational number theory system.

A new version, MPC 0.9 (Epilobium montanum), was released in February 2011, with new functions, some speed-ups, a few bug fixes, and a logging feature for debugging. Since version 4.5 of GCC, released in May 2010, GCC requires MPC to compute constant complex expressions at compile-time (constant folding), like it requires GNU MPFR since GCC 4.3.

5.4. GMP-ECM

Participants: Cyril Bouvier, Paul Zimmermann [contact].

GMP-ECM is a program to factor integers using the Elliptic Curve Method. Its efficiency comes both from the use of the GMP library, and from the implementation of state-of-the-art algorithms. GMP-ECM contains a library (LIBECM) in addition to the binary program (ECM). The binary program is distributed under GPL, while the library is distributed under LGPL, to allow its integration into other non-GPL software. The Magma computational number theory software and the SAGE computer algebra system both use LIBECM.

During his internship of 4 months in 2011, Cyril Bouvier developed a version of ECM for GPUs. The code was written for NVIDIA GPUs using CUDA. First, the code was written for all NVIDIA cards, and later, it was optimized for the newer Fermi cards. As there is no modular arithmetic library (like GMP) available for GPU, it was necessary to implement a modular arithmetic using array of unsigned integers from scratch, while taking into account constraints of GPU programming. The code was optimized for factoring 1024 bits integers. For now, the code has a throughput roughly four times bigger than GMP-ECM on one core. This code is not yet fully integrated in GMP-ECM but is available in the GMP-ECM svn repository.
The implementation of ECM on GPU uses a different algorithm for scalar multiplication (the binary ladder instead of PRAC) and a different parametrization. This new approach was implemented for CPU in GMP-ECM. It results in a speedup in the execution time of GMP-ECM for finding big factors (more than 20 digits). It will be integrated in the next release of GMP-ECM.

5.5. Finite fields

**Participants:** Pierrick Gaudry, Emmanuel Thomé [contact].

\( \mathbb{F}_q \) is (yet another) library for computing in finite fields. The purpose of \( \mathbb{F}_q \) is not to provide a software layer for accessing finite fields determined at runtime within a computer algebra system like Magma, but rather to give a very efficient, optimized code for computing in finite fields precisely known at compile time. \( \mathbb{F}_q \) is not restricted to a finite field in particular, and can adapt to finite fields of any characteristic and any extension degree. However, one of the targets being the use in cryptology, \( \mathbb{F}_q \) somehow focuses on prime fields and on fields of characteristic two.

\( \mathbb{F}_q \)'s ability to generate specialized code for desired finite fields differentiates this library from its competitors. The performance achieved is far superior. For example, \( \mathbb{F}_q \) can be readily used to assess the throughput of an efficient software implementation of a given cryptosystem. Such an evaluation is the purpose of the “EBats” benchmarking tool. \( \mathbb{F}_q \) entered this trend in 2007, establishing reference marks for fast elliptic curve cryptography: the authors improved over the fastest examples of key-sharing software in genus 1 and 2, both over binary fields and prime fields. These timings are now comparison references for other implementations.

The library’s purpose being the generation of code rather than its execution, the working core of \( \mathbb{F}_q \) consists of roughly 18,000 lines of Perl code, which generate most of the C code. \( \mathbb{F}_q \) is distributed at [http://mpfq.gforge.inria.fr/](http://mpfq.gforge.inria.fr/).

The \( \mathbb{F}_q \) library has undergone no change in 2011.

5.6. gf2x

**Participants:** Pierrick Gaudry, Emmanuel Thomé [contact], Paul Zimmermann.

GF2X is a software library for polynomial multiplication over the binary field, developed together with Richard Brent (Australian National University, Canberra, Australia). It holds state-of-the-art implementation of fast algorithms for this task, employing different algorithms in order to achieve efficiency from small to large operand sizes (Karatsuba and Toom-Cook variants, and eventually Schönhage’s or Cantor’s FFT-like algorithms). GF2X takes advantage of specific processors instruction (SSE, PCLMULQDQ).

The current version of GF2X is 1.0, released in 2010 under the GNU GPL. Since 2009, GF2X can be use as an auxiliary package for the widespread software library NTL, as of version 5.5.

There has been no update of GF2X in 2011, but the software is still maintainted. An LGPL-licensed portion of GF2X is also part of the CADO-NFS software package.

5.7. CADO-NFS

**Participants:** Jérémie Detrey, Pierrick Gaudry, Lionel Muller, Emmanuel Thomé [contact], Paul Zimmermann.

CADO-NFS is a program to factor integers using the Number Field Sieve algorithm (NFS), developped in the context of the ANR-CADO project (November 2006 to January 2010).

NFS is a complex algorithm which contains a large number of sub-algorithms. The implementation of all of them is now complete, but still leaves some places to be improved. Compared to existing implementations, the CADO-NFS implementation is already a reasonable player. Several factorizations have been completed using our implementations.\footnote{http://www.ecrypt.eu.org/ebats/}
Since 2009, the source repository of CADO-NFS is publicly available for download. On October 28, 2011, the 1.1 version of CADO-NFS has been released. Several improvements to the program have been obtained, in practically all areas of the program. In particular, the polynomial selection code described by Thorsten Kleinjung at the CADO workshop in 2008 is now used within CADO-NFS, together with some efficient root-sieve code written by Shi Bai (Australian National University). Overall, CADO-NFS keeps improving its competitiveness over alternative code bases. The lattice siever now supports a sieving region of $2^{31}$ ($I = 16$); its code has been deeply reorganized to allow future improvements that we have in mind but were difficult to implement (proper sieving of powers, sieve according to the parities of the coordinates of the location). The executables in the linear algebra step have been reorganized (now using shared libraries), and now use a code generation mechanism built on top of the MPFQ library for the arithmetic parts. This is in particular meant to ease future accommodation of other base fields that GF(2), which is a requirement for adapting CADO-NFS to discrete logarithm computation. The MPI performance of the linear algebra code has been optimized. Some experimental scripts have been added to execute the sieve on a cluster; these scripts rely on the OAR job scheduler being used, and exploit its “besteffort” mode.

The largest factorizations performed by CADO-NFS in 2011 are a 170-digit integer from aliquot sequence 660 and a 171-digit integer from aliquot sequence 966.

5.8. AVIsogenies

Participants: Gaëtan Bisson [contact], Romain Cosset.

AVISOGENIES (Abelian Varieties and Isogenies) is a Magma package for working with abelian varieties, with a particular emphasis on explicit isogeny computation; it has been publicly released under the LGPLv2+ license in 2010.

Its prominent feature is the computation of $(\ell, \ell)$-isogenies between Jacobian varieties of genus-2 hyperelliptic curves over finite fields; practical runs have involved values of $\ell$ in the hundreds. It also provides procedures for exploring and drawing isogeny graphs, and for computing various complex-multiplication-related structures, such as Shimura’s gothic C group.

In 2011, two incremental versions have been released. They provide the following new features: the characteristic 2 is now supported, and the complete addition laws of [23] have been implemented.

The package can be obtained at http://avisogenies.gforge.inria.fr/.

6. New Results

6.1. NFS-related results

Concerning the number field sieve algorithm for the discrete logarithm problem in prime fields, Răzvan Bărbulescu improved the theoretical complexity of the step called “individual logarithm”, using, at a crucial point, a sequence of ECM steps with well-tuned, increasing parameters. He also proved that an approach similar to Coppersmith’s factoring factory was feasible as well for discrete logarithm, yielding an improved overall complexity if heavy precomputations are allowed [21].

In 2010, Thomas Prest and Paul Zimmermann developed a new algorithm for the polynomial selection in the Number Field Sieve (NFS). This algorithm produces two non-linear polynomials, extending Montgomery’s “two quadratics” method. For degree 3, it gives two skewed polynomials with resultant $O(N^{5/4})$, which improves on Williams $O(N^{4/3})$ recent result. The paper will appear in the Journal of Symbolic Computation [13] and its impact is assessed by the fact that two preprints extending and analyzing the algorithm have already been proposed.
6.2. Ballot stuffing in a postal voting system

In collaboration between many members of the CASSIS and CARAMEL teams, we have studied a postal voting system used by the CNRS for an election involving about 30,000 voters [16]. The structure of the material can be easily understood out of a few samples of voting material (distributed to the voters), without any prior knowledge of the system. Taking advantage of some flaws in the design of the system, we have shown how to perform major ballot stuffing, making possible to change the outcome of the election. Our attack has been tested and confirmed by the CNRS, and the system was quickly fixed for the next elections.

6.3. Symmetric cryptanalysis

Mohamed Ahmed Abdelraheem, Céline Blondeau, María Naya-Plasencia, Marion Videau, and Erik Zenner have proposed an attack against ARMDILLO2, the recommended variant of a multi-purpose cryptographic primitive dedicated to hardware which has been proposed by Badel et al. in 2010. The attack uses a meet-in-the-middle technique that allows us to invert the ARMDILLO2 core function. This makes it possible to perform a key recovery attack when used as a FIL-MAC. A variant of this attack has been applied to the stream cipher derived from the PRNG mode. A (second) preimage attack is also proposed against the hash function mode. All attacks have been validated by implementing cryptanalysis on scaled variants. The experimental results match the theoretical complexities.

The underlying idea of the attacks, the parallel matching algorithm, has also been generalized. The results are presented in the paper [14].

Thomas Fuhr, Henri Gilbert, Jean-René Reinhard, and Marion Videau have studied the security of the two most recent versions of the message authentication code 128-EIA3, which was considered for adoption (and has been adopted) as a third integrity algorithm in the emerging 3GPP standard LTE. An efficient existential forgery attack against the June 2010 version of the algorithm has been presented. This attack allows, given any message and the associated MAC value under an unknown integrity key and an initial vector, to predict the MAC value of a related message under the same key and the same initial vector with a success probability 1/2. The tweaked version of the algorithm that was introduced in January 2011 to circumvent this attack has also been analysed. While this new version offers a provable resistance against similar forgery attacks under the assumption that (key, IV) pairs are never reused by any legitimate sender or receiver, some evidence is given that some of its design features limit its resilience against IV reuse. The results are presented in the paper [18].

6.4. Implementation of cryptographic pairings

The extended version of a work on parallel architectures for the computation of the \( \eta_T \) pairing over supersingular elliptic curves in characteristic 2 and 3, presented at CHES 2009 then accepted at IEEE Transaction on Computers in 2010, was finally published [3]. This paper was the result of a joint effort of Jérémie Detrey and Nicolas Estibals, in collaboration with Jean-Luc Beuchat and Eiji Okamoto (University of Tsukuba, Japan), Francisco Rodríguez-Henríquez (CINVESTAV-IPN, Mexico).

Also, the work on supersingular genus-2 pairings by Diego F. Aranha (UNICAMP, Brazil), Jean-Luc Beuchat (University of Tsukuba, Japan), Jérémie Detrey and Nicolas Estibals was accepted for publication at the Cryptographers’ Track of the RSA Conference (CT-RSA 2012) [15]. Since last year, where only the Eta pairing algorithm was described, several major revisions were undertaken to improve this paper, among which a careful and detailed analysis of the various distortion maps of the considered family of hyperelliptic curves. This study also allowed us to exhibit a somewhat simple distortion map which would enable this curve to benefit from the shorter loop of the Ate pairing algorithm. Exploring this option is currently work in progress, and the results should eventually be submitted to a journal.

6.5. Multiple-precision arithmetic

In [25], Pascal Molin showed that the error function \( \text{erf} \) can be computed very efficiently using a formula involving an integral of a form appropriate for fast evaluation using the trapezoidal scheme. A rigorous
analysis of the scheme in this context allows to get precise bounds on the various errors terms, and therefore
to give a proven complexity result for the multiple-precision evaluation of erf. The good theoretical behaviour
is confirmed by an implementation in Pari.

Together with David Harvey (New York University), P. Zimmermann studied the short division of long
integers, i.e., the division of a $2n$-bit integer by an $n$-bit integer where only the integer quotient is wanted,
or an approximation of it. They gave detailed algorithms with rigorous errors bounds, and implemented them
in GNU MPFR. Using Harvey’s integer middle product code, they obtain a speedup of up to 10% with respect
to the best known implementation [20].

With Guillaume Melquiond (Proval project-team, INRIA Saclay), and Prof. W. Georg Nowak (Institute of
Mathematics, Vienna), P. Zimmermann worked on the numerical approximation of the Masser-Gramain
constant, following some work of Gramain and Weber in 1985. This work disproves a conjecture of Gramain,
and enables one to determine the following approximation of that constant:

$$1.819776 < \delta < 1.819833.$$  

This work has been completed in 2011 [12].
The article “The Great Trinomial Hunt” has been published in the Notices of the AMS [7].

6.6. Proving the complexity of computing endomorphism rings

Subsequent to the work [6] that has been finally published this year, Gaëtan Bisson has been working on
rigorously proving a subexponential running time bound for computing endomorphism rings of ordinary
elliptic curves over finite fields. In the end, the proof rests on only one assumption, namely the extended
Riemann hypothesis (ERH) [4].

In his thesis [1], he has also made substantial advances towards the extension of these algorithms to genus 2
curves.

In studying the above-mentioned algorithms, Gaëtan Bisson, in collaboration with Andrew V. Sutherland, has
designed a low-memory, Pollard-rho type algorithm for finding relations in generic groups [5].

6.7. Point counting on curves with real multiplication

Pierrick Gaudry, David Kohel and Benjamin Smith have designed a new variant of the Schoof algorithm for
point counting on hyperelliptic curves, that can take advantage of the presence of the knowledge of an explicit
and efficient endomorphism coming from real multiplication. In that case, the overall complexity drops from
$O(\log^8 q)$ to $O(\log^5 q)$. Using our algorithm we have computed a 256-bit prime-order Jacobian, suitable for
cryptographic applications, and also the order of a 1024-bit Jacobian. The corresponding paper [19] obtained
the Best Paper Award at the Asiacrypt 2011 conference.

6.8. Computation of isogenies between abelian varieties

Following the work [11] of David Lubicz and Damien Robert (that has just been accepted for publication in
Compositio Mathematica) about the explicit computation of isogenies using theta coordinates, Romain Cosset
and Damien Robert [24] have developed further nice features. In the original paper, only $(\ell^2, \ell^2)$-isogenies
between abelian surfaces were available. It is now possible to handle $(\ell, \ell)$-isogenies between genus 2 curves,
thus providing a more precise tool. Two key elements were necessary: Romain Cosset gave explicit methods
to transfer points between the classical representation with Mumford’s coordinates and the theta functions.
This is a generalisation of the work of Van Wamelen. And Romain Cosset and Damien Robert developed
an explicit algorithm to change the level in the Theta coordinates that are used to represent the geometrical
objects. Many details can be found in Cosset’s thesis [2].
Using the same kind of tools, Christophe Arène and Romain Cosset [23] have constructed the first complete addition law on abelian surfaces. Although they are not yet of any practical use, completeness is a feature that is in principal interesting for cryptographic applications.

The article by Faugère, Lubicz and Robert on computing modular correspondences with Theta constants has finally appeared in Journal of Algebra [9].

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. Function field sieve: implementation and hardware acceleration

Participants: Jérémie Detrey [contact], Pierrick Gaudry, Hamza Jeljeli, Emmanuel Thomé.

The team has obtained for the year 2012 a financial support from the Région Lorraine and INRIA for a project focusing on the hardware implementation and acceleration of the function field sieve (FFS).

The FFS algorithm is currently the best known method to compute discrete logarithms in small-characteristic finite fields, such as may occur in pairing-based cryptosystems. Its study is therefore crucial to accurately assess the key-lengths which such cryptosystems should use. More precisely, this project aims at quantifying how much this algorithm can benefit from recent hardware technologies such as GPUs or CPU-embedded FPGAs, and how this might impact current key length recommendations.

7.2. National Initiatives

7.2.1. ANR DEMOTIS (Collaborative Analysis, Evaluation and Modelling of Health Information Technology)

Participant: Marion Videau.

The project from “programme ARPEGE” involves three INRIA project-teams as a single partner (SMIS, SECRET and CARAMEL) together with colleagues from CECOJI (CNRS) and the company Sopinspace. It has been running from January 2009 and will continue until the beginning of 2012.

The project experiments new methods for the multidisciplinary design of large information systems that have to take into account legal, social and technical constraints. Its main field of application is personal health information systems.

7.2.2. ANR CHIC (Courbes Hyperelliptiques, Isogénies, Comptage)

Participants: Gaëtan Bisson, Romain Cosset, Pierrick Gaudry, Sorina Ionica, Pascal Molin, Emmanuel Thomé [contact].

The team has obtained a financial support from the ANR (“programme blanc”) for a project, common with colleagues from IRMAR (Rennes) and IML (Marseille). The ANR CHIC grant covers the period 09/2009 to 08/2012. The purpose of this ANR project is the study of several aspects of curves in genus 2, with a very strong focus on the computation of explicit isogenies between Jacobians.

This ANR project has been an important source of motivation for both permanent researchers and PhD students, giving notably PhD students the opportunity to meet interested colleagues on a regular basis. In 2011, a server with a huge large of central memory has been bought, to help with CHIC-related experiments. Two PhD thesis were defended (Bisson and Cosset) on the topic.
7.3. European Initiatives

7.3.1. PHC application with EPFL

The team obtained a PHC Germaine de Staël grant in collaboration with the LACAL team from EPFL (Lausanne, Switzerland), in 2011. The grant has been renewed for 2012. This collaboration focuses on integer factorization and discrete logarithms.

8. Dissemination

8.1. Animation of the scientific community

8.1.1. Caramel seminar

Nineteen speakers were invited to our seminar in 2011: Cyril Bouvier, Sorina Ionica, Paul Zimmermann, Diego Aranha, Benoît Gaudel, Cyril Bouvier, Alain Couvreur, Christophe Mouilleron, Marion Videau, Hamza Jeljeli, Benjamin Smith, Xavier Pujol, Răzvan Bărbulescu, Alin Bostan, Martin Albrecht, Bogdan Pasca, Christophe Arène, Frederik Vercauteren, Junfeng Fan.

8.1.2. Conference organization

In 2011, the team organised the ECC 2011 workshop and a Summer School the week before. With more than 120 participants at the workshop and more than 40 participants at the school, it was a great success.

8.2. Committees memberships

- Jérémie Detrey was a member of the hiring committee for ATER positions in computer science (section 27) at Université Henri Poincaré (Nancy I).
- Pierrick Gaudry was referee for the PhD thesis of Mehdi Tibouchi (ENS, Paris 7, Luxembourg), of Vanessa Vitse (Versailles) and of Thomas Izard (Montpellier). He was a member of the PhD thesis jury of Christophe Arène (Marseille), Gaëtan Bisson (Caramel) and Guillaume Batog (Nancy). He participated to the “Comités de selection” in Paris 7, Lille 1 and Nancy 1. He is the principal investigator of a Labex proposal about computer security that have been submitted to the second call. He is deputy director of the LORIA.
- Emmanuel Thomé was a member of the program committee for the WCC2011 (Workshop in Coding and Cryptography) conference. He was elected as a member of the INRIA Evaluation Committee for the period 2011-2015. He was the advisor of Romain Cosset’s PhD thesis, and a member of his PhD jury on nov. 7th.
- Marion Videau is member of the scientific committee of the CCA seminar (Codage, Cryptologie, Algorithmes). She was a member of the PhD jury of Jean-René Reinhard (Versailles-Saint-Quentin-en-Yvelines).
- Paul Zimmermann was a member of the Habilitation thesis of David W. Ritchie (Nancy), of the PhD thesis jury of Christiane Peters (Eindhoven), and of the PhD thesis jury of Julien Cojan (Nancy).

8.3. Vulgarization

P. Gaudry, E. Thomé and P. Zimmermann wrote a vulgarisation article about the RSA-768 integer factorization record for the “Techniques de l’Ingénieur” [22].

8.4. Invited Conferences

P. Gaudry and E. Thomé both gave a one-hour invited talk at the workshop “Elliptic Curve Discrete Logarithm workshop” held in EPFL.

J. Detrey gave a one-hour invited talk at the national workshop “Codage et Cryptographie” held in April 2011 at Saint Pierre d’Oléron.

P. Zimmermann gave an invited talk “GNU MPFR: back to the future” at the MaGiX@LiX 2011 Conference in Palaiseau (France) in September.

E. Thomé gave an invited talk “CADO-NFS: an implementation of the Number Field Sieve” at the MaGiX@LiX 2011 Conference in Palaiseau (France) in September.

8.5. Teaching

- Marion Videau (from September 2011):
  Operating systems: 14 hours (lectures), 14 hours (tutorial sessions), 14 hours (practical sessions), L3, University Henri Poincaré, Nancy 1, France.
  Introduction to cryptography: 15 hours (lectures), 15 hours (tutorial sessions), M1, University Henri Poincaré, Nancy 1, France.

- Jérémie Detrey:
  Introduction to cryptology: 12 hours (lectures), M1, ESIAL, Nancy, France.
  Security of websites: 2 hours (lecture), L1, IUT Charlemagne, Nancy, France.

- Emmanuel Thomé
  Algorithmic Number Theory: 9 hours (lectures), M2, University Paris 7 (Master Parisien de Recherche en Informatique), Paris, France.
  Introduction to cryptography: 3 hours (lecture), L3, École des mines de Nancy, France.
  Factoring algorithms: 6 hours (lectures and practical sessions), M1, École des mines de Nancy, France.

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses


Articles in International Peer-Reviewed Journal


International Conferences with Proceedings


Research Reports


Scientific Popularization


Other Publications


References in notes