Rapport scientifique détaillé de l'équipe AMACC

Période 2010-2014

GREYC

À l'attention du lecteur :

Ce document constitue la version longue de la contribution de l'équipe AmacC au rapport d'évaluation du laboratoire GREYC. Il détaille en particulier les aspects scientifiques.

La dernière version de ce document peut être téléchargée à l'URL

`http://clementj01.users.greyc.fr/hceres/amacc-hceres.pdf`

PUBLICATION DU LABORATOIRE GREYC

Version septembre 2015

Table des matières

1. Réalisations

À l'attention du lecteur :

Ce document constitue la version longue de la contribution de l'équipe AmacC au rapport d'évaluation du laboratoire GREYC. Il détaille en particulier les aspects scientifiques.

La dernière version de ce document peut être téléchargée à l'URL
`http://clementj01.users.greyc.fr/hceres/amacc-hceres.pdf`

1 Présentation générale de l'équipe

Au sein du laboratoire généraliste qu'est le GREYC, l'équipe AmacC se caractérise par une identité culturelle forte, celle de l'informatique mathématique (IM). Dans un tel contexte, chacun a un rôle à jouer : grâce à son insertion dans le GREYC, l'équipe AmacC peut rester au contact de problématiques de la vie « réelle » informatique, et elle peut aussi sensibiliser le laboratoire à sa culture propre, et notamment aux problèmes de complexité et d'algorithmique fondamentale.

L'équipe AmacC se retrouve d'abord autour de deux concepts génériques, l'algorithme et la complexité, qui sous-tendent toutes ses activités. Elle y adopte des points de vue complémentaires. Le premier étudie les modèles de calcul et la notion de complexité, dans le pire des cas, via les classes de complexité, tandis que le second travaille dans un cadre aléatoire, avec des modèles probabilistes. L'équipe a aussi des spécialités algorithmiques bien marquées : elle travaille en amont de la Protection de l'Information, notamment en codage et cryptographie, où elle s'appuie sur ses compétences en théorie de l'information et en arithmétique. Elle s'ouvre désormais aussi à d'autres thématiques, liées à la sécurité (internet des objets ou l'environnement embarqué sécurisé) ou à l'algorithmique du web. Elle se structure en trois thèmes :

- Thème 1 : Modèles de calcul et complexité descriptive.
- Thème 2 : Structures aléatoires et analyse d'algorithmes.
- Thème 3 : Protection et traitement de l'information.

L'équipe anime et organise le séminaire hebdomadaire d'Algorithmique [1], depuis maintenant une vingtaine d'années. Ce séminaire est ouvert à toutes les problématiques ayant des aspects algorithmiques ou formels. Parallèlement, et plus récemment, elle a recréé un groupe de travail, qui se réunit environ deux fois par mois, et est un lieu d'expression plus informel pour ses membres, notamment ses doctorants. L'équipe partage l'organisation du séminaire *Cryptographie et Sécurité* [2] (en liaison avec le laboratoire LMNO, l'équipe d'Orange Labs, et l'équipe Monétique et Biométrie). L'équipe participe également activement au groupe de travail « Mots, séquences et entropie(s) » avec le LMNO.

C'est une des équipes provinciales bien visibles parmi les équipes françaises de l'IM. Elle joue fortement le jeu national, en ayant une politique active dans la structure du GdR IM. Depuis

1. `https://clementj01.users.greyc.fr/semalgo/`
2. `https://www.greyc.fr/node/205`

la création de l'ANR, elle s'investit et y réussit bien en portant ou participant à dix projets.

Elle mène une politique ambitieuse de publication en revues, et accroît son ambition sur les conférences internationales de très bon niveau.

1.1 Composition de l'équipe

Au 1er octobre 2015, l'équipe comprend

- 1 collège A (1 PR à l'UFR Sciences) ;
- 9 collèges B (1 CR CNRS, 6 MC à l'UFR Sciences : 1 MC à l'ENSICAEN, 1 MC à l'IUT) ;
- 4 doctorants ;
- 4 membres associés (2 collaborateurs Orange Labs, 1 DR CNRS émérite, 1 MC retraité resté associé).

La responsabilité de l'équipe a été assurée successivement par Brigitte Vallée en 2010, puis par Sylvain Peyronnet de 2011 à mai 2014 et enfin par Julien Clément depuis mai 2014.

L'équipe a connu de multiples changements pendant la période 2010-2015 :

- deux d'entre nous sont partis à la retraite en 2014 (Jean Saquet et Jacques Madelaine) ;
- quatre habilitations ont été soutenues entre 2010 et 2011, et deux de ces nouveaux habilités ont été promus PR à l'extérieur (Fabien Laguillaumie à Lyon et Ayoub Otmani à Rouen) ;
- l'un d'entre nous a rejoint une autre équipe début 2014 (Jean-Marie Le Bars, équipe Monétique et Biométrie) ;
- nous avons recruté en 2012 un professeur qui a demandé en 2014 un détachement pour création d'entreprise (Sylvain Peyronnet) ;
- nous avons recruté un MC en 2013 (Thomas Largillier) ;
- nous accueillons pour 20 mois (janvier 2015 à août 2016) un MC de l'Université de Clermont-Ferrand mis à disposition et nous espérons prolonger cette période d'accueil au-delà (Florent Madelaine) ;
- la DR est devenue émérite en juin 2015 (Brigitte Vallée).

Permanents, chercheurs et enseignants chercheurs

Prénom et nom	Fonction	T1	T2	T3	Date d'entrée/sortie
Ali Akhavi	MC UFR Sciences	✓	✓	✓	2000
Julien Clément	CR CNRS		✓	✓	2005
Etienne Grandjean	PR UFR Sciences	✓			1985
Jerzy Karczmarczuk	MC UFR Sciences	✓			1988
Fabien Laguillaumie	MC UFR Sciences			✓	2006–2012
Thomas Largillier	MC IUT Caen		✓	✓	2013
Jean-Marie Le Bars	MC UFR Sciences			✓	1999-2014
Loïck Lhote	MC ENSICAEN		✓	✓	2007
Florent Madelaine	MC UFR Sciences	✓			2015
Jacques Madelaine	MC UFR Sciences			✓	1985-2014
Ayoub Otmani	MC ENSICAEN			✓	2004-2012
Sylvain Peyronnet	PR UFR Sciences		✓	✓	2012-2014
Ionona Ranaivoson	MC UFR Sciences	✓			1991
Gaétan Richard	MC UFR Sciences	✓		✓	2009
Jean Saquet	MC UFR Sciences			✓	1985-2014
Véronique Terrier	MC UFR Sciences	✓			1991
Brigitte Vallée	DR CNRS		✓	✓	1990–2015

Il y a 4 membres associés : Sébastien Canard et Marc Girault sont ingénieurs de recherche à Orange Labs (Caen) et titulaires de l'HDR. Brigitte Vallée est DR émérite et Jacques Madelaine reste membre associé depuis son départ en retraite en 2014.

Doctorants. Huit thèses ont été soutenues pendant la période. Parmi les docteurs qui ne sont plus post-doctorants, quatre travaillent en entreprise, l'un est enseignant-chercheur et le dernier est professeur de CPGE. En mai 2015, il n'y avait plus que deux doctorants. Depuis, la situation s'est un peu améliorée mais reste préoccupante.

Prénom et nom	Financement	T1	T2	T3	Période	Situation actuelle ou prévue (1/1/2016)
Nicolas Bacquey	CNRS-Région	✓			oct. 2012–oct. 2015	—
Cyril Bazin	Rég.-Mon.			✓	oct. 2005-jan. 2010	Directeur tech. R&D Datexim
Johann Brault-Baron	Minist.	✓			oct. 2009–avril. 2013	Postdoc LIF (Marseille)
Jean Creusefond	Minist.		✓	✓	oct. 2013-oct. 2016	—
Léonard Dallot	CNRS-Région-Orange			✓	oct. 2006–juil. 2010	Ingénieur TazTag (Rennes)
Dimitri Darthenay	Minist.		✓	✓	oct. 2015	
Julien Devigne [3]	CIFRE			✓	oct. 2010–déc. 2013	Ingénieur DGA (Rennes)
Mariya Georgieva	Minist.		✓	✓	oct. 2009–déc. 2013	Ingénieure Gemalto (Paris)
Kanal Hun	Bourse AUF [4]		✓		janv. 2012–déc. 2014	EC au Cambodge
Thu Hien Nguyen Thi	CNRS-Région		✓		oct. 2010–déc. 2014	ATER à Paris 5
Pablo Rotondo	ANR, associé [5]		✓		oct. 2015	
Mathieu Roux	Minist. [6]		✓		oct. 2007–déc. 2011	Professeur de CPGE (Caen)

ATER, Post-Doctorants ou ingénieurs. Nous avons accueilli quinze tels personnels. En plus de nos doctorants (accueillis en ATER en fin de thèse) listés dans la première partie du tableau, nous avons accueilli onze autres personnes ; cinq d'entre elles sont maintenant enseignants-chercheurs, cinq sont ingénieurs, une en recherche d'emploi.

Prénom et nom	Statut	T1	T2	T3	Période au GREYC	Situation actuelle ou prévue (1/1/2016)
Johann Brault-Baron	ATER	✓			sept 2012-août 2013	Postdoc LIF (Marseille)
Léonard Dallot	Postdoc CPER			✓	sept-oct 2010	Ingénieur TazTag (Rennes)
Mariya Georgieva	ATER		✓	✓	sept 2012-août 2013	Ingénieure à Gemalto (Paris)
Thu Hien Ngyuen Thi	ATER		✓		sept 2013-août 2014	ATER Paris 5
Morgan Barbier	Postdoc CPER+ATER			✓	sept 2012-août 2013	MC à l'Ensicaen
Hayat Cheballah	Postdoc ANR		✓	✓	sept 2012-août 2013	Consultante F.Iniciativas
Nicolas Gama	ATER			✓	sept 2009-août 2010	MC à l'U. de Versailles
Valérie Gauthier-Umaña	Postdoc CPER			✓	fév 2012-oct 2103	EC en Colombie
Eleonora Guerrini	ATER+Postdoc CPER			✓	sept 2010-août 2011	MC à Montpellier
Laura Giambruno	ATER		✓		sept 2012-août 2013	Préparation de l'Agrég.
Adrien Laurence	Ingénieur ANR-FUI			✓	jan 2010-jan 2013	Ingénieur Enovea (Rouen)
Jérôme Le Moulec	Ingénieur CPER			✓	jan 2010-fev 2012	Ingénieur W2Next (Caen)
Jérémie Lumbroso	Postdoc CPER		✓		fév 2012-juil 2012	Lecturer à Princeton, USA
Rémy Ménard	Ingénieur FUI			✓	fév 2012-juin 2013	Ingénieur CHU (Caen)
Sylvain Sauvage	Ingénieur FUI			✓	mars 2012-avr 2012	Gérant d'entreprise

1.2 Caractéristiques de l'équipe et faits marquants

Les faits marquants suivants éclairent et caractérisent bien notre équipe.

Résultats scientifiques. Nous avons obtenu durant la période, et de manière régulière, des résultats scientifiques bien reconnus : par exemple, en 2011, 2012, 2013 et 2014, nous avons quatre articles acceptés à STACS [7], sur des thématiques variées et caractéristiques d'AmacC (automates cellulaires, analyse d'algorithmes, combinatoire).

À la mémoire de Philippe Flajolet. Nous avons joué un rôle moteur dans l'organisation d'un

3. Thèse effectuée à Orange Labs.
4. Agence Universitaire de la Francophonie.
5. Inscrit principalement à Paris VII.
6. Partagée à 50% avec le Laboratoire LMNO.
7. La conférence STACS *Symposium on Theoretical Aspects of Computer Science*), se situe parmi les cinq conférences internationales en Informatique Théorique les plus « cotées ». Nous avons également accepté la proposition du *steering committee* de STACS d'accueillir la conférence en 2018 à Caen (présidence du comité de programme et comité d'organisation).

colloque international à la mémoire de Philippe Flajolet[8], et nous faisons partie du comité d'édition de ses œuvres complètes (sept volumes à paraître chez Cambridge Press).

Projets ANR. Depuis la création de l'ANR, l'équipe s'investit pour répondre aux appels à projets et y réussit souvent bien. Sur la période, l'équipe a participé à 10 projets ANR (essentiellement du programme Blanc), 1 fois comme porteur, 6 fois comme partenaire, et 3 fois avec des participations individuelles (voir p. 33 la liste des projets ANR de l'équipe).

GdR IM. Nous sommes fortement impliqués dans le GdR Informatique Mathématique. Prise dans son ensemble, l'équipe participe de manière active à beaucoup de groupes de travail de ce GdR (cinq à six sur une vingtaine en tout). Sur la période, nous avons organisé en particulier les journées annuelles de deux d'entre eux. L'équipe a organisé l'EJCIM 2014 (École Jeunes Chercheurs en Informatique Mathématique) à Caen en 2014. En plus des tâches générales d'organisation, l'équipe a pris en charge trois des cinq cours, et a édité le recueil des cours dans le livre « Une photographie de l'informatique mathématique 2014 » [91]. Enfin, deux d'entre nous sont membres du comité de direction du GdR IM. Nous sommes également responsables du site internet et du serveur de listes du GdR, tous deux hébergés par le GREYC.

Amérique du Sud. Nous tissons des liens resserrés avec l'Amérique du Sud. Le projet CNRS STIC-AMSUD (Argentine, France, Uruguay) que nous portons depuis 2013 s'est récemment renforcé en automne 2014 : le thème ALÉA a intégré le laboratoire international associé (LIA) INFINIS à Buenos Aires (Argentine) en y créant, avec ses partenaires, une nouvelle équipe « Combinatorics and analysis of algorithms in number theory, information theory and cryptography ».

Conseil National du Numérique. L'une d'entre nous est membre du Conseil National du Numérique (proposée par le CNRS) et a notamment participé à la rédaction de trois rapports : Inclusion (automne 2013) [113], Éducation (automne 2014) [119] et Ambition numérique (été 2015) [112].

Formation des enseignants. L'équipe s'implique fortement dans la formation des enseignants de l'informatique dans les lycées (Informatique et Sciences du Numérique). En coordination avec le rectorat, elle a contribué à la mise en place de la formation ; elle anime et participe à la formation elle-même.

Création d'une entreprise. L'un de nous a créé une entreprise, l'entreprise IX-LABS *« Algorithms for the social web, adversarial information retrieval, data science, approximate verification »*.

2 Principaux résultats scientifiques

Nous décrivons ces résultats selon les trois thèmes :

- *Thème 1. Modèles de calcul et complexité descriptive :* Étude des modèles de calcul, automates cellulaires et machines RAM, classes de complexité pour ces modèles et complexité descriptive de ces classes, complexité des problèmes combinatoires ou logiques. Étude du paradigme de la programmation fonctionnelle pure.
- *Thème 2. Structures aléatoires et analyse d'algorithmes :* Analyse probabiliste des structures aléatoires et des algorithmes, avec un volet méthodologique notamment fondé sur la combinatoire analytique et les systèmes dynamiques, et trois volets plus applicatifs, où l'on étudie les mots, les nombres, et le web.
- *Thème 3. Protection et traitement de l'information :* Conception et analyse de la sécurité (via des preuves ou des attaques) de protocoles destinés à protéger l'information (cryptographie, codes correcteurs d'erreurs, fonctions booléennes), avec l'étude des objets et des

8. Philippe Flajolet (1948-2011) est un scientifique de renommée internationale qui a fondé le domaine de la combinatoire analytique [116]. Il a eu une influence déterminante sur l'identité et la construction de l'équipe et joué un rôle scientifique essentiel dans le thème « Structures aléatoires et analyse d'algorithmes ».

algorithmes sous-jacents (issus de l'algèbre discrète ou de la théorie des nombres). Protection de la propriété intellectuelle, internet des objets, sécurité des réseaux. Algorithmique du web.

2.1 Thème 1 : Modèles de calcul et complexité descriptive

Description générale. Ce thème réunit l'équipe autour d'une étude comparative des modèles de calcul centrée d'abord sur leurs algorithmes et leur complexité. On s'intéresse aussi aux modèles de calcul en tant que systèmes dynamiques. Le modèle actuellement le plus étudié dans ce thème est celui des automates cellulaires, d'abord en dimension 1, mais aussi, plus récemment, en dimension 2 ou supérieure. Dans une moindre mesure, on s'intéresse aussi aux modèles de calcul en tant qu'outils de modélisation de phénomènes complexes, en particulier l'évolution des systèmes quantiques.

Résultat marquant. Nous avons développé une famille d'algorithmes permettant de résoudre le problème d'élection de leader sur des configurations périodiques. La généralisation du problème des configurations de dimension 1 [31] aux configurations de dimension 2 [115] a fait émerger à la fois des concepts clés comme celle de racine primitive pour des mots périodiques de dimension 2 [32] et également des outils algorithmiques essentiels comme celui d'horloge spatiale.

2.1.1 Algorithmique et complexité des automates cellulaires

Au travers de problèmes spécifiques, nous développons les outils algorithmiques propres à ce modèle de calcul massivement parallèle.

Automates cellulaires à comportements complexes (N. Bacquey, G. Richard, V. Terrier). Nous cherchons d'abord à construire de tels automates [30]. Notamment, nous avons construit un algorithme pour déterminer la période minimale pour des configurations périodiques en dimension 1 [31] et sa généralisation en dimension 2 [32, 115] qui nécessite d'expliciter de nouvelles abstractions propres à la dimension 2 : « élection de leader », notion de réseau, etc. Récemment, nous avons également mis au point un algorithme qui réalise la synchronisation globale d'un automate, partant de n'importe quelle configuration périodique dont la longueur de la période est un nombre premier.

Nous avons aussi étudié des petites classes de complexité des automates cellulaires et leur puissance de calcul ; en effet, c'est pour ces classes que l'apport du parallélisme est remarquable par rapport au mode séquentiel [87, 104]. Nous avons ainsi proposé un algorithme de parallélisation d'une variante d'automates finis (à plusieurs têtes et mouvement « oblivious ») [37]. Nous nous sommes intéressés aux performances des automates cellulaires vus sous l'aspect fonctionnel et nous avons étudié leurs propriétés [60] ; concernant les limites à la puissance de calcul, nous avons montré l'impossibilité de réaliser l'opération miroir en temps minimal [60]. Enfin, nous avons prouvé l'inclusion de la classe des langages algébriques à croissance linéaire dans la classe temps minimale des automates cellulaires unidirectionnels [79].

Un outil clé qui ressort couramment dans les simulations utilisées est celui de transformations géométriques du diagramme espace/temps. Ceci nous a conduit à une première ébauche de formalisation.

Complexité descriptive des langages d'images reconnaissables et des automates cellulaires (E. Grandjean, G. Richard) Caractériser en logique une classe de complexité permet de mieux comprendre cette classe, d'en justifier l'intérêt et le caractère intrinsèque. Nous avons prouvé dans [20, 62] des caractérisations logiques (logique existentielle du second-ordre, monadique ou non) des classes de langages (de mots ou d'images) suivantes :

(1) la classe des langages reconnaissables d'images ;
(2) la classe des langages reconnus en temps linéaire sur automates cellulaires non déterministes.

Ces caractérisations, valides en toute dimension et donc les plus générales possible, sont basées sur des théorèmes de normalisation logique. Le résultat (1) constitue une généralisation, à toute dimension, de deux résultats classiques bien connus : le Théorème de Büchi-Elgot-Trakhtenbrot (1960) sur les langages rationnels de mots, d'une part ; la caractérisation logique des langages reconnaissables d'images de dimension 2 due à Giammarresi et al. (1996), d'autre part. Le résultat (2) est, lui, totalement nouveau.

Nous nous donnons pour objectif d'établir une caractérisation analogue au résultat (2) pour le cas, plus significatif encore, du temps linéaire des automates cellulaires *déterministes*. Nous utiliserons pour cela des formules de Horn (voir projet).

2.1.2 Algorithmes et complexité des requêtes conjonctives et des CSP

Les requêtes conjonctives sont centrales à deux points de vue : ce sont les requêtes de base des langages de requêtes des bases de données, et elles constituent un formalisme équivalent aux CSP. On sait que la complexité d'une requête conjonctive est déterminée par les propriétés de l'hypergraphe associé à la requête.

Requêtes conjonctives acycliques (J. Brault-Baron, E. Grandjean). Dans la suite des résultats déjà obtenus sur la complexité de l'énumération des requêtes conjonctives acycliques, c'est-à-dire, dont l'hypergraphe est alpha-acyclique, Johann Brault-Baron [39, 94] a construit un algorithme de temps quasi-linéaire pour le calcul de toute requête conjonctive bêta-acyclique (c'est-à-dire, dont l'hypergraphe est bêta-acyclique) : algorithme valide même si la requête contient des négations d'atomes. C'est un résultat optimal, car il a démontré aussi que la condition de bêta-acyclicité est nécessaire pour obtenir un temps quasi-linéaire. Par ailleurs, il a obtenu des caractérisations équivalentes des trois sortes d'acyclicité d'un hypergraphe – alpha, bêta et gamma-acyclicité – qui généralisent et uniformisent celles de la littérature ; ceci, en introduisant une notion de feuille adaptée à chacune des trois acyclicités [94].

Requêtes dans divers fragments de la logique du premier ordre (F. Madelaine). Arrivé au GREYC en décembre 2014, Florent Madelaine étudie aussi la complexité et la complexité descriptive de familles de problèmes génériques modélisant les CSP, les requêtes conjonctives et les problèmes de coloriage. La méthodologie utilisée mêle combinatoire, logique et algèbre [46]. Nous notons ici une synergie forte avec les deux autres équipes du département *Intelligence Artificielle et Algorithmique* du laboratoire : la méthodologie algébrique élaborée est similaire à celle employée par Bruno Zanuttini (de l'équipe MAD) et l'objet d'étude est commun avec Arnaud Lallouet et d'autres chercheurs de CoDaG.

2.1.3 Approches symboliques, algébriques et fonctionnelles

D'autres points de vue sur les modèles de calcul sont également abordés dans l'équipe, notamment des approches symboliques, algébriques ou encore reposant sur des concepts de programmation fonctionnelle.

Problème du tri généralisé ou analyse symbolique des algorithmes (A. Akhavi). Nous définissons dans [114] un cadre pour modéliser les algorithmes « à la façon des systèmes dynamiques symboliques ». Nous considérons les algorithmes déterministes et séquentiels qui finissent toujours et sont réversibles (sans perte d'information durant l'exécution). Un algorithme est alors donné par un ensemble de transformations élémentaires bijectives – donc réversibles – et une stratégie déterministe qui décide, à chaque étape de l'exécution, la transformation élémentaire à appliquer. Sous un ensemble d'hypothèses raisonnables, nous pouvons associer à un algorithme

(identifié à ses traces d'exécutions) un ensemble de systèmes de réécriture sur l'ensemble des mots sur l'alphabet des transformations élémentaires de l'algorithme. Nous caractérisons alors les traces d'exécutions comme les formes normales de ces systèmes de réécriture. Un tel point de vue peut avoir des retombées diverses : description du pire cas d'un algorithme, ou, dans des cas élémentaires comme les tris, obtention d'une relation entre borne inférieure de complexité et géométrie du graphe de Cayley.

Groupes d'automates (A. Akhavi). Dans un transducteur déterministe lettre à lettre, chaque état définit une application sur l'ensemble des mots construits sur l'alphabet du transducteur qui conserve la longueur des mots. Le semi-groupe engendré par tous les états du transducteur (qui devient un groupe lorsque le transducteur est inversible) joue un rôle important historique dans la théorie des (semi)-groupes : il a permis une construction simple des groupes infinis de torsion où il a mis en évidence des premiers exemples de groupes à croissance intermédiaire. Dans ce contexte, la question de la finitude d'un (semi-)groupe est à la fois classique et difficile à résoudre en toute généralité. En collaboration avec Inès Klimann, Sylvain Lombardy, Jean Mairesse et Matthieu Picantin, nous avons considéré le cas particulier du (semi-)groupe associé à un transducteur : est-il fini ou non ? Nous obtenons dans [2] une condition suffisante effective dans le cas des semi-groupes, une condition nécessaire effective dans le cas des groupes et une caractérisation non effective dans les cas des groupes.

Représentation fonctionnelle de l'évolution des systèmes quantiques (J. Karczmarczuk). Il s'agit ici d'étudier le statut informatique des entités quantiques (qui ne sont pas des « données », mais des « calculs » ou des processus), en décrivant la représentation fonctionnelle de l'évolution de ces systèmes, du point de vue de transfert de contrôle. Ainsi, un opérateur qui agit sur un état – ici, un « observable » physique – est assimilable d'un point de vue informatique à une « continuation linéaire » qui appartient aux concepts de la programmation fonctionnelle.

2.2 Thème 2. Structures aléatoires et analyse d'algorithmes

Description générale. Les recherches de ce thème se déclinent selon quatre volets :

- un premier volet qui modélise le concept de source, et les propriétés qui sont essentielles dans l'analyse des structures de données construites sur ces sources ;
- un second volet, dédié à l'analyse probabiliste des algorithmes du texte et de la fouille de données ;
- un troisième volet, dédié à l'analyse probabiliste des algorithmes en arithmétique ;
- un quatrième volet, dédié aux approches probabilistes de l'algorithmique des grands systèmes. C'est une nouvelle activité qui s'est installée suite au recrutement de S. Peyronnet et T. Largillier.

Bien sûr, ces volets ne sont pas étanches. L'analyse dynamique, spécialité locale qui mélange méthodes d'analyse en moyenne et méthodes issues de systèmes dynamiques, et bien décrite dans [83], est un fil conducteur important sous-jacent (mais non omniprésent) dans les trois premiers volets. De plus, le Thème 2, surtout dans ses deux derniers volets, élabore des méthodes et des outils, qui sont souvent appliqués dans le Thème 3. Les méthodes et les outils sont décrits ici, tandis que le versant plus applicatif est décrit dans le Thème 3.

Résultat marquant. L'équipe à élaboré un cadre unificateur pour l'analyse en moyenne des algorithmes classiques de tri et de recherche, qui permet d'évaluer la complexité de ces algorithmes, lorsqu'ils travaillent sur des mots (issus d'une source générale) et que l'opération élémentaire est la comparaison entre symboles. Il y a un double intérêt : la méthode elle-même, et aussi les résultats, qui exhibent des constantes qui expliquent bien l'interaction entre la statégie de l'algorithme et les propriétés probabilistes de la source. Ce travail a donné lieu à trois publications : un

article accepté à STACS [50], suivi de deux articles longs acceptés dans de très bons journaux : CPC [16] et TOCS [15].

2.2.1 Sources et théorie de l'information.

C'est une activité de modélisation autour du concept de source de la théorie de l'information.

Les séries génératrices des sources (J. Clément, K. Hun, T-H. Nguyen-Thi, M. Roux, B. Vallée). L'équipe travaille depuis une dizaine d'années sur un modèle de source très général, et y a associé une série génératrice de probabilités, $\Lambda(s)$ qui n'a pas de pôles pour $\Re s > 1$. La géométrie des pôles de $\Lambda(s)$ dans le demi-plan $\Re s \leq 1$ est essentielle pour décrire l'asymptotique fine des structures de données et des algorithmes travaillant sur les mots de la source. Le travail de l'équipe pendant la période a été déterminant pour élucider cette géométrie, notamment dans le cas où la source est associée à un système dynamique. Dans ce cas, la série $\Lambda(s)$ est engendrée par un opérateur dit « de transfert » qu'il faut donc étudier soigneusement, en étendant à cet opérateur « sécant » les propriétés déjà établies par Dolgopyat pour l'opérateur « tangent ». Une première étape est franchie dans [12] pour la classe de sources dite UNI. La thèse de M. Roux [100] franchit une seconde étape en étudiant la classe de sources dite DIOP : après un premier résultat obtenu dans [57], dans le cas d'une source sans mémoire sur un alphabet fini, nous relions plus généralement [78] la forme de la région sans pôles et les propriétés arithmétiques de la famille de probabilités. Dans la thèse de T-H. Nguyen-Thi [99], et en relation avec l'analyse des algorithmes de sélection, les sources sont aussi étudiées via une autre série génératrice dont il faut également étudier la région sans pôles. Enfin, dans la thèse de K. Hun [98], l'analyse des digital search trees (dst) est reliée aux propriétés de la source « renversée » où on lit les symboles de droite à gauche.

Entropies (L. Lhote). L'équipe développe une collaboration avec les mathématiciens de Caen autour des problématiques sur les mots, les sources ou les entropies. Un premier résultat sur le calcul et l'estimation de taux d'entropie généralisée [14] montre que ces taux prennent au plus quatre valeurs (entropie nulle, entropie infinie positive, taux d'entropie de Shannon ou encore taux d'entropie de Rényi) limitant ainsi leur intérêt pratique. Dans [25], nous étendons les résultats précédents aux taux d'entropie relative généralisée et renormalisée. Nous donnons une valeur explicite des taux d'entropie relative et la renormalisation évite de retrouver les quatre situations triviales précédentes.

Codage de source (J. Clément). Dans une collaboration avec Frédérique Bassino, Gadiel Seroussi et Alfredo Viola, nous étudions les codes préfixes optimaux pour une source sans mémoire, dont les probabilités des symboles suivent une loi géométrique. Cette source produit des mots sur l'alphabet (infini) des entiers et modélise les erreurs résiduelles en compression d'images. On connaît peu de codes préfixes optimaux pour les alphabets infinis, et le plus connu est le code de Golomb. Nous construisons de nouveaux codes, avec de nouvelles propriétés : par exemple, les arbres associés ont une largeur non bornée. Nous avons également montré une propriété de « singularité » des codes obtenus dans le sens qu'un code préfixe qui est optimal pour une certaine valeur du paramètre de la distribution ne peut l'être pour aucune autre valeur. Cela implique que les codes optimaux ne peuvent être décrits de manière compacte [5].

2.2.2 Analyse en algorithmique du texte : recherche de motifs textuels, motifs et fouille de données, algorithmes de tri et de recherche, arbres digitaux.

Recherche de motifs textuels (J. Clément, L. Lhote, B. Vallée). Nous traitons l'analyse probabiliste des motifs textuels de nature diverse.

Ensemble de mots (J. Clément). Avec des méthodes de combinatoire analytique, nous étudions les statistiques d'occurrences d'un motif très général : d'abord, le motif est ici un ensemble fini de mots, où l'on autorise (contrairement au cadre usuel) les mots à être facteurs les uns des

autres ; ensuite, nous comptons toutes les occurrences (chevauchantes ou non) de ce motif. Nous obtenons des expressions précises des constantes qui interviennent dans l'analyse, ce qui permet à la fois un calcul efficace et une meilleure compréhension des phénomènes de corrélation à l'intérieur d'un ensemble de mots [4].

Table des bords (J. Clément, L. Giambruno). Le comportement de nombreux algorithmes de recherche d'un motif n'est pas lié au motif lui-même mais seulement à sa structure d'autocorrélation (qui décrit les recouvrements possibles du motif avec lui-même), elle-même décrite par la table des bords. C'est le cas par exemple pour le célèbre algorithme de Knuth-Morris-Pratt. Dans ce cadre, nous nous intéressons, plutôt qu'à la représentation d'un motif (classiquement sa représentation sur un alphabet), à sa structure combinatoire, en grande partie indépendante de l'alphabet. L'intérêt est double : d'une part, nous souhaitons mieux comprendre la structure des motifs ; d'autre part, nous voulons générer un ensemble minimal de motifs pour valider un algorithme dans une campagne de tests. Dans [49], nous énumérons ces motifs, et nous construisons des algorithmes permettant de passer d'un représentation d'un motif à une autre (table des préfixes, table des bords, ou représentation condensée de la table des bords).

Motifs cachés (L. Lhote). Nous avons étudié dans [73] la loi limite du nombre d'occurrences d'un motif « caché » dans un texte produit par une source dynamique, et mis en évidence une loi limite gaussienne. Même si ce motif caché est le plus simple possible (constitué de deux lettres), la réponse à ce problème « jouet » n'est pas du tout triviale, car la mémoire de la source induit une perte de commutativité des opérateurs de transfert utilisés dans l'analyse.

Mots sturmiens (B. Vallée). Dans le cadre du projet `DynAlco` avec l'Amérique du Sud, et en reprenant la même philosophie que pour les progressions arithmétiques, l'équipe a débuté l'étude des mots sturmiens aléatoires, et en particulier l'étude probabiliste de leur fonction de récurrence, qui mesure les intervalles entre les apparitions d'un même facteur. Il y a un premier résultat accepté [35].

Traverses minimales dans les hypergraphes (L. Lhote). L'existence d'un algorithme « output-polynomial » qui calcule les traverses minimales d'un hypergraphe est encore un problème ouvert. L'article [18] adopte un point de vue probabiliste, introduit des modèles d'hypergraphes aléatoires à la Erdös-Rényi, et calcule le nombre moyen de traverses minimales (et d'ensembles minimaux) dans ces modèles-là. Nous en déduisons des bornes sur la complexité moyenne du calcul de traverses minimales.

Algorithmes de tri et de recherche (J. Clément, T-H. Nguyen-Thi, B. Vallée). C'est le résultat jugé « marquant » du Thème 2. Après l'article fondateur de l'ICALP 2009 qui a ouvert la voie à un nouveau champ de recherche, l'équipe a revisité l'analyse en moyenne des principaux algorithmes de tri et de recherche, lorsque les clés sont des mots produits par une source générale et le coût de la comparaison entre clés est le nombre de symboles utilisés pour comparer les mots. C'est le cadre général de la thèse de T-H. Nguyen Thi [99] où nous étudions toute une classe d'algorithmes classiques avec une méthode unificatrice : les algorithmes de tri dans [16, 50] et les variations de l'algorithme QuickSelect dans [15]. Les constantes qui apparaissent dans l'analyse expliquent bien l'interaction entre la stratégie de l'algorithme et les propriétés probabilistes de la source.

Structures digitales construites sur des sources générales (J. Clément, K. Hun, B. Vallée). Le groupe a analyé finement les deux principales structures digitales "de base" fondées sur un principe de dictionnaire (le trie et le dst) quand elles sont construites sur des mots émis par une source générale.

Structure de trie. L'équipe est revenue sur ses travaux antérieurs sur le trie : J. Clément termine la rédaction d'un chapitre sur ces analyses probabilistes dans un livre en collaboration avec B. Chauvin et D. Gardy. B. Vallée (avec E. Cesaratto) [12] a effectué l'analyse « en distribution » d'un paramètre important du trie, la profondeur moyenne, dans le cas de sources générales dites UNI.

Structure de dst. La structure de digital search tree (dst) est plus complexe à analyser que le trie, car elle dépend de l'ordre d'insertion des mots. Dans sa thèse [98], K. Hun a élaboré un cadre commun à l'analyse de ces deux structures digitales, qui lui a permis de les comparer très finement, et d'expliquer le rôle de la source, via la géométrie des pôles élucidée dans le premier volet. Après l'article de conférence [66], il y a un article long en préparation.

2.2.3 Analyses en arithmétique.

Cet axe de recherche est généralement motivé par la cryptologie, et les analyses qui ont une application plus directe en cryptologie sont décrites dans le Thème 3. C'est un axe original et porteur, car, en général, les cryptologues se préoccupent assez peu d'analyse probabiliste, et les analyseurs en moyenne, assez peu de cryptologie...

Petits quotients (B. Vallée). Dans beaucoup de problèmes arithmétiques (comme celui du paragraphe suivant), les réels (ou les rationnels) dont le développement en fraction continue (DFC) ne contient que des petits quotients jouent un rôle particulier car ils sont « mal approximables » par des rationnels. Nous avons évalué très précisément la probabilité que tous les quotients du DFC d'un rationnel de taille n ait tous ses quotients dans son développement en fraction continue de taille inférieurs à $\alpha(n) \cdot \log n$ pour une fonction α. L'article [11] (avec E. Cesaratto) exhibe une loi 0–1 (si $\alpha(n)$ tend vers 0, la probabilité tend vers 0, et si $\alpha(n)$ tend vers l'infini, la probabilité tend vers 1) ; il affine et généralise des résultats précédents dus à Hensley. Surtout, l'analyse dynamique nous montre un chemin plus naturel que les travaux précédents.

« Pseudo-randomness » d'une progression arithmétique modulaire aléatoire (B. Vallée). Il existe beaucoup de mesures de *pseudo-randomness* d'un générateur pseudo-aléatoire. Le groupe a étudié le cas d'une progression arithmétique modulaire « aléatoire », et l'a relié à une version probabiliste du célèbre théorème des trois distances. Or, si le théorème des trois distances est célèbre, son étude probabiliste n'a pas encore été abordée ; de plus, nous nous intéressons à des modèles probabilistes particuliers (cas où les « entrées » sont rationnelles, par exemple), et nous étudions des mesures diverses de pseudo-randomness, comme la discrépance. Nous comparons aussi la situation d'un rationnel à celle d'un réel, et ce qui se passe si l'entrée réelle est générique ou si elle ne possède que des petits quotients dans son DFC (voir paragraphe précédent). Là encore, les méthodes d'analyse dynamique permettent d'aborder ce problème de manière très naturelle et convaincante. Une version courte [48] (avec E. Cesaratto) est complétée par une version (semi)-longue qui va paraître dans un chapitre de livre [82] et une version longue est en cours de finition (une quarantaine de pages déjà écrites).

Pgcd multiple (J. Creusefond, L. Lhote, B. Vallée). Il y a beaucoup de stratégies pour calculer le pgcd de ℓ entrées (nombres entiers ou polynômes sur un corps fini), et la version la plus naïve (dont l'analyse est cependant proposée par Knuth comme un problème difficile...) calcule les pgcd deux par deux de proche en proche. Cet algorithme se laisse bien analyser en distribution : pour les polynômes, avec des outils de combinatoire analytique ; pour les nombres, avec des outils de combinatoire dynamique. Les analyses et les résultats sont parallèles (même si, comme toujours, l'étude est plus difficile pour les nombres) mais les résultats sont un peu inattendus, car ils exhibent des lois Beta et un phénomène plus « fort » que ce qui était prédit par Knuth. L'article [36] est un article court centré sur les polynômes qui annonce les résultats dans le cas

des nombres, et l'article long [6] qui traite vraiment les deux cadres en parallèle, vient d'être accepté.

Distribution de l'algorithme d'Euclide. (B. Vallée). Nous avions déjà montré avec V. Baladi que beaucoup de coûts dits additifs dans l'algorithme d'Euclide admettaient un théoréme central limite, et nous avions aussi obtenu un théorème local limite (seulement pour des coûts à valeurs dans un réseau), tous deux avec vitesse optimale. Par la suite, V. Baladi a étudié le cas d'un coût général (pas nécessairement à valeurs dans un réseau), et a démontré un theorème local limite, avec des hypothèses supplémentaires assez peu naturelles sur ce coût. Dans [80], nous exhibons une condition sur les coûts sous laquelle on obtient un théorème local limite ; cette condition, à la fois plus générale et plus naturelle, généralise la condition classique du cas d'un système dynamique sans mémoire.

Complexité fine de l'algorithme d'Euclide sur les polynômes (B. Vallée). Nous travaillons ici avec un coût défini non pas par le degré, mais par le nombre de monômes effectivement présents (complexité creuse), et nous utilisons des outils classiques de combinatoire analytique. Dans [7], un schéma d'analyse similaire à ce qui se fait avec le coût classique du degré permet d'obtenir le résultat (en collaboration avec V. Berthé, H. Nakada et R. Natsui).

2.2.4 Approches probabilistes dans l'algorithmique de grands systèmes.

Il s'agit ici de travailler sur de « grands systèmes », que ce soit des réseaux (au sens d'un ensemble d'entités connectées entre elles : ensemble de pages web ou réseau social par exemple) qui apparaissent en algorithmique du web ou des systèmes (probabilistes ou non déterministes) qui apparaissent en vérification : les programmes informatiques donnent des exemples de systèmes probabilistes, tandis que la plupart des protocoles qui sont influencés par un comportement humain ou adversarial (enchères, etc.) donnent des exemples de systèmes non déterministes. Ces approches donnent lieu à des applications dans le Thème 3.

Algorithmique du web (T. Largillier, S. Peyronnet). Nous étudions des réseaux au sens précédent, et cherchons à y détecter rapidement des comportements émergents, afin de mettre en place, pour certaines tâches de base (classement, élection, etc.), des algorithmes qui soient robustes vis-à-vis de ces comportements. Compte-tenu des contraintes (dynamicité, taille des données), seuls des algorithmes probabilistes peuvent être utilisés. Nous avons travaillé en particulier sur la publication robuste dans les plateformes de news en ligne, et avons proposé dans [71] une méthode de filtrage *a priori* du contenu non désirable sur les plateformes d'agrégation de contenu. La méthode repose sur des mécaniques de « Game with a purpose » introduites par Luis Von Ahn, et permet à un sous-ensemble aléatoire d'utilisateurs de valider la publication d'un contenu sur la plateforme. Cette méthode est insensible aux manipulations des utilisateurs tant qu'ils restent minoritaires.

Analyse de grands processus Markoviens (S. Peyronnet). Avec R. Lassaigne, nous avons adopté dans [23] une méthode de double échantillonnage qui permet d'approcher d'aussi près que l'on veut la probabilité, pour un processus Markovien, d'un comportement exprimé par une formule de la logique LTL (Linear Temporal Logic) sous la stratégie optimale d'acquisition de récompense de l'adversaire. L'échantillonnage porte d'abord sur les actions que l'adversaire considère à chaque étape de choix, puis ensuite sur les exécutions restant disponibles suite aux choix de l'adversaire. Ce résultat permet par exemple d'étudier les propriétés temporelles des stratégies d'extraction de ressources. On peut ainsi vérifier si une stratégie optimale de prélèvement d'une ressource naturelle (poisson, pétrole, etc.) amène à la disparition de la ressource.

Vérification approchée et problèmes d'énumérations (S. Peyronnet). Avec Y. Strozecki et M. de Rougemont, nous utilisons dans [77] des méthodes d'échantillonnages « géométriques » pour énumérer les stratégies qui permettent de différencier deux processus Markoviens. Pour cela, nous représentons les stratégies positionnelles d'un processus Markovien par un vecteur de statistiques et nous lui associons un polytope qui nous permet de différencier les processus entre eux, à l'aide notamment de marches aléatoires quasi-isotropes au sein de ce polytope.

2.3 Thème 3. Protection et traitement de l'information

Description générale. Avec le départ de cinq de ses membres (2 promotions de professeur à l'extérieur, 2 départs en retraite, 1 changement d'équipe) et l'arrivée de deux nouveaux membres, le thème a vu son contour scientifique évoluer au cours du quinquennal. C'est pourquoi il a changé de nom. Dans notre projet en 2010, il s'appelait *Protection de l'information, codage et Cryptographie* et il se retrouve maintenant sous le descriptif général *Protection et Traitement de l'information*. Il regroupe des travaux orientés vers les applications, qui trouvent souvent leur méthodologie et leurs outils « en amont » dans le Thème 2. Les travaux sur les procédés cryptographiques et le codage laissent la place, sans disparaître, à des recherches sur la sécurité du web. Les thématiques historiques autour des réseaux euclidiens, de l'internet des objets ou de l'étude en amont des outils issus de la cryptographie restent bien présentes. Une nouvelle thématique se crée autour de l'algorithmique du web. Finalement, les recherches de ce thème s'organisent autour de trois volets :

- le premier volet est dédié à l'amont de la protection de l'information : il étudie l'algorithmique de certains objets qui sont centraux en protection de l'information. En somme ce volet fait vraiment la transition entre les thèmes 2 et 3.
- le deuxième volet est centré sur la protection de l'information, la conception de procédés cryptographiques et leurs analyses.
- le troisième volet regroupe des travaux autour de la représentation et du traitement de l'information, qui est présente dans le web, ou transite plus généralement sur un réseau.

Résultat marquant. Le chiffrement par attributs permet de définir finement une politique de déchiffrement : celle-ci permettra à tous les utilisateurs satisfaisant cette politique de déchiffrer les messages. Ce type de protocole est beaucoup plus pertinent pour gérer des groupes d'utilisateurs que la cryptographie à clé publique, et permet également d'assurer des propriétés d'anonymat. L'équipe a conçu un tel protocole permettant d'implanter des politiques assez expressives, tout en maintenant une taille constante des chiffrés. Ce travail est une première étape fondamentale pour atteindre un protocole qui serait optimal en termes de sécurité, d'expressivité et d'efficacité. Ce résultat, paru à PKC [65], a par la suite été accepté dans le journal TCS [3].

2.3.1 Algorithmique d'objets cryptographiques.

Nous étudions quatre principaux objets, qui sont centraux dans les protocoles cryptographiques, mais que nous étudions ici dans leur rôle amont en protection de l'information : les couplages et les courbes elliptiques, les codes correcteurs d'erreurs, les fonctions booléennes et automates booléens, et enfin les réseaux euclidiens, l'algorithmique de la réduction et leurs modélisations simplifiées.

Couplages et courbes elliptiques (S. Canard, J. Devigne, N. El Mrabet, F. Laguillaumie). Les courbes elliptiques, omniprésentes en cryptographie, sont aussi un objet d'étude en elles-mêmes. Dans [53], nous décrivons l'addition et la multiplication par 2 dans une nouvelle représentation des courbes elliptiques.

Dans [54], nous décrivons une attaque par faute contre l'algorithme de Miller qui utilise les coordonnées d'Edwards. Nous proposons aussi une variante de cet algorithme qui conduit à un

algorithme de calcul de couplage plus efficace dans certains cas [38].

Nous étudions la sécurité et l'efficacité des implémentations basées sur les couplages pour téléphones portables ou smart-cards et nous montrons comment améliorer ces performances [43].

Codes quasi–BCH (M. Barbier). Dans [33], nous décrivons les liens entre trois types de codes : les codes de Reed-Solomon (sur des matrices carrées), les codes entrelacés et les codes quasi–BCH. Nous étendons l'algorithme de Welch-Berlekamp aux codes quasi–BCH et nous montrons que les codes entrelacés offrent une autre approche pour le décodage des codes quasi–BCH.

Nous étudions la complexité des algorithmes de décodage (Welch-Berlekamp et Guruswami-Sudan) quand ils agissent sur les codes de Reed-Solomon généralisés sur des anneaux (commutatifs ou non). Nous en étudions les paramètres et nous proposons des algorithmes pour le décodage par liste [28].

Fonctions booléennes et automates non-déterministes (J.-M. Le Bars, N. Gama). Les fonctions booléennes (ou les automates associés) sont utilisées en cryptographie symétrique.

Fonctions booléennes. Les fonctions booléennes sont utilisables en cryptographie quand elles satisfont certaines propriétés (par exemple, la résilience), et il est donc important de pouvoir générer (aléatoirement) de telles fonctions booléennes. C'est un problème algorithmique difficile, car de telles fonctions sont « rares ». Dans [24], nous proposons un algorithme efficace qui énumère les fonctions 1-résilientes jusqu'à 7 variables. La méthode récursive utilisée s'appuie sur des classes d'équivalence de fonctions booléennes. Dans [10, 45], en améliorant plusieurs aspects de l'algorithme précédent, nous générons les fonctions 1-résilientes à 8 variables. L'explosion combinatoire nous empêche d'aller plus loin.

Automates XOR. Dans [121], nous comparons deux classes d'automates finis non-déterministes, selon qu'ils sont généraux ou implémentés avec des XOR. Nous montrons que, contrairement au cas général, les XOR-automates finis non-déterministes admettent une unique forme minimale, qu'on peut calculer efficacement en temps polynomial. Nous montrons aussi que les opérations efficaces sur les automates finis non-déterministes s'étendent efficacement aux XOR-automates non-déterministes.

Algorithmique des réseaux euclidiens (N. Gama, B. Vallée). Les réseaux euclidiens jouent un rôle central en cryptographie, et le principal algorithme de réduction des réseaux (l'algorithme LLL, du nom de ses inventeurs) est un instrument efficace de cryptanalyse. Suite à la conférence internationale LLL+25 qui a été organisée par l'équipe en 2007 à Caen à l'occasion des 25 ans de l'algorithme, B. Vallée a co-édité (en collaboration avec P. Nguyen) un livre [90] qui regroupe les exposés donnés à la conférence, et qui sont rédigés sous la forme de survey. Ce livre rencontre du succès ; il contient en particulier un chapitre [89] qui décrit les travaux très précis que l'équipe a obtenus sur le comportement de l'algorithme LLL en dimension 2, qui coïncide avec un ancien algorithme, dû Gauss.
Dans [34], nous présentons une variante de l'algorithme de réduction de Gauss, qui a une meilleure borne de complexité que celle fournie par l'algorithme usuel. Nous obtenons ainsi une nouvelle borne (heuristique) de complexité pour la factorisation des entiers de la forme pq^2.
L'énumération de points dans un réseau est une procédure qui est centrale dans la résolution du problème SVP (recherche du plus court vecteur du réseau). Nous revisitons certains algorithmes qui résolvent SVP, et obtenons un gain de complexité exponentiel en utilisant un élagage dit extrême [58].

Modélisation approchée de l'algorithme LLL et des réseaux cryptographiques (A. Akhavi, J. Clément, M. Georgieva, L. Lhote, M. Madritsch, B. Vallée). L'analyse probabiliste de l'algorithme LLL, en particulier dans des modèles d'entrées qui correspondent à la « réalité cryp-

tographique », paraît hors d'atteinte, car le système dynamique sous-jacent est certainement trop complexe pour être analysé exactement. Avec M. Madritsch, post-doctorant au GREYC, et dans le cadre du projet ANR LAREDA, nous avons proposé dans [75] un premier modèle très simplifié de l'algorithme LLL, celui du « tas de sable » (dit modèle M1), où l'on peut déjà exhiber des phénomènes qualitatifs et dégager des heuristiques ; c'est aussi un modèle pédagogique, qui permet à des non spécialistes de comprendre « comment ça marche ». Ce travail s'est prolongé dans le cadre de la thèse de Maryia Georgieva [97], qui a proposé toute une classe de modèles (M2, M3, M4), intermédiaires entre le tas de sable (M1) et le véritable algorithme (M5) : on cherche à « compliquer » progressivement le modèle du tas de sable, pour obtenir d'autres modèles qui restent « simplifiés » tout en devenant cependant plus réalistes. Nous avons commencé par le modèle M2, où nous avons montré que, dans un contexte assez général, le nombre d'itérations suit une loi géométrique, alors qu'un tel résultat n'est pas connu pour l'algorithme LLL. Nous avons débuté l'analyse dans un contexte plus général et obtenons des résultats encourageants. Dans [97], nous décrivons aussi des modèles simplifiés pour les entrées, et montrons qu'il y a essentiellement trois types de réseaux aléatoires qui apparaissent en cryptographie. Les résultats de la thèse ne sont pas encore publiés.

2.3.2 Protocoles cryptographiques.

Il s'agit ici de faire de la cryptographie « en tant que telle », et nous en étudions de multiples facettes, via des protocoles de nature variée (chiffrement, signature, délégation de calcul...) qu'on construit ou qu'on détruit via la cryptanalyse.

Serveurs de re-chiffrement (S. Canard, J. Devigne, E. Guerrini, F. Laguillaumie). Dans le cadre de la thèse CIFRE de Julien Devigne [96], qui s'est déroulée en collaboration avec Orange Labs à Caen, nous étudions les serveurs de re-chiffrement pour le stockage dématérialisé et sécurisé de données personnelles. Nous introduisons un modèle de proxy de re-chiffrement qui généralise tous les modèles connus et nous en étudions ses propriétés [9]. Nous proposons des schémas avec une meilleure sécurité (nous progressons de la sécurité IND-RCCA à la sécurité IND-CCA). Nous mettons aussi en évidence des failles de sécurité qui n'étaient pas prises en compte par certains schémas de serveur de re-chiffrement (basés sur Hash ElGamal) et nous les « corrigeons ». Enfin nous étudions [41] la gestion dynamique des droits d'accès pour l'accès et la gestion des données, et nous introduisons un nouveau modèle de serveur de re-chiffrement dit combiné. Pour la gestion fine des droits d'accès, nous introduisons les serveurs de re-chiffrement sélectifs qui sont compatibles avec les serveurs combinés. Avec ce modèle, il est possible d'obtenir une gestion des droits d'accès plus précise que celles connues [52]. Nous montrons aussi comment réaliser les modèles de serveur proposés.

Délégation de calculs (S. Canard, F. Laguillaumie). Dans [44], nous décrivons un protocole pour la délégation d'un calcul de couplage avec vérifiabilité efficace. Cette méthode est compatible avec d'autres méthodes, ce qui permet d'assurer en plus la confidentialité. Plus généralement, nous proposons une méthode de délégation de calculs cryptographiques qui automatise certains choix et effectue une répartition optimale des calculs entre les systèmes [40]. Même si cette méthode n'est pas complètement nouvelle, elle a l'avantage de s'appliquer à tous les protocoles asymétriques.

Cryptographie à base d'attributs. (F. Laguillaumie). Dans [65], nous introduisons la première politique de chiffrement basée sur les attributs avec deux caractéristiques : la taille des clés secrètes reste d'ordre contant et elle admet une politique de déchiffrement suffisamment expressive. Nous étendons ces résultats dans [3] et [64] où nous décrivons aussi les deux premiers schémas de signatures basées sur les attributs dont la taille reste constante et ne dépend pas du nombre

d'attributs présents dans le prédicat de signature.
Nous étudions les liens entre la sécurité sémantique et l'anonymat des protocoles de chiffrement basés sur l'identité. Nous montrons que l'anonymat sélectif est distinct de l'anonymat adaptatif même avec des schémas sémantiquement sûrs. En revanche, nous prouvons que la sécurité sémantique sélective et l'anonymat adaptatif entraînent la sécurité sémantique adaptative [21].

Protocoles de signature (F. Laguillaumie). Nous proposons un nouveau schéma de signature « indéniable » basé sur les applications bilinéaires. La forge de ces signatures est équivalente à la résolution du problème décisionnel de Diffie Hellman alors que l'anonymat se ramène à un problème décisionnel non standard [22]. En modifiant le schéma, nous introduisons une nouvelle famille de signatures appelées "signatures à durée choisie".
Nous décrivons des constructions génériques de protocoles capables de vérifier si un mot-clé (non chiffré) est présent dans un message chiffré [42]. Ensuite, nous montrons comment ces protocoles peuvent améliorer la sécurité liée à la révocation de signatures de groupe.
Nous introduisons dans [1] le premier protocole en anneau à seuil, basé sur les codes, efficace et prouvé sûr, dont la taille de la clé publique reste raisonnable.

Chiffrement homomorphe (F. Laguillaumie). Dans [47], nous décrivons une approche générique pour concevoir des protocoles homomorphes et prouvés sûrs dans des sous-groupes de points d'une courbe elliptique avec couplage.

Cryptanalyse en cryptographie fondée sur les codes (V. Gauthier-Umaña, A. Otmani). Les versions récentes du cryptosystème de McEliece cherchent à réduire la taille des clés, et s'appuient pour cela sur des codes structurés. Dans [56], nous montrons que la clé privée du cryptosystème de McEliece (ou de certaines variantes fortement structurées) est solution d'un système polynomial de type Vandermonde, et, grâce à des calculs de bases de Groebner (implémentés en MAGMA), nous pouvons attaquer très efficacement plusieurs paramètres cryptographiques.
Nous avons auusi étudié le problème d'indistinguabilité des codes pour des rendements élevés [19, 55]. En utilisant le rang du système polynomial précédent, nous proposons un critère, calculable en temps polynomial, qui (pour des rendements élevés) permet de discriminer complètement un code de Goppa d'un code aléatoire. Ce résultat invalide certaines preuves de sécurité de schémas cryptographiques à rendements élevés.
De fait, le distingueur précédemment construit est fondé sur le « square code » (code généré par le produit des composantes du code public). Avec ce nouveau point de vue, nous proposons de nouvelles attaques qui sont effectives dès lors que les codes sous-jacents sont distinguables des codes aléatoires, ce qui est le cas pour les codes de Reed-Solomon généralisés. Dans [17], nous cassons ainsi complètement deux variantes du cryptosystème de McEliece (le protocole homomorphe de Bogdanov et Lee proposé en 2011 par Baldi et *al.* et une autre variante due à Wieschebrink).
Dans la même veine, nous étudions les protocoles de signature de type KKS qui s'appuient sur des couples de codes correcteurs, et nous proposons une attaque qui retrouve les clés privées dès que les codes choisis sont trop proches l'un de l'autre [76].

2.3.3 Représentation et traitement et de l'information.

Ce volet s'articule autour de recherches sur l'information présente dans l'image (tatouage d'images) dans la structure du web (détection de communautés, filtrage de contenus) et débouche par ailleurs sur deux plateformes, l'une consacrée à l'internet des objets, et la seconde à l'information géographique.

Tatouage de documents structurés (C. Bazin, J-M. le Bars, J. Madelaine). Dans le cadre de la thèse de C. Bazin [93], nous avons étudié le tatouage de documents structurés, qui doit

préserver des contraintes, et nous avons mené en particulier beaucoup d'expérimentations sur les documents géographiques. Cela nous a permis de mieux cerner un domaine encore peu développé et formalisé. Nous avons proposé un schéma général, qui sépare clairement une partie générique commune à tous les documents, d'une partie spécifique à un certain type de document. Cela montre qu'il est indispensable d'avoir une bonne connaissance de la problématique générale du tatouage, ainsi qu'une bonne expertise du type des documents à tatouer.

Travaux sur le web : détection de communautés et filtrage de contenus (J. Creusefond, T. Largillier, S. Peyronnet). Les travaux orientés Web de l'équipe visent à mieux comprendre les comportements sur la toile et à développer des outils qui rendent l'usage de la toile plus agréable aux internautes.

Détection des communautés dans les graphes sociaux. C'est un problème central dans la compréhension des mécanismes de diffusion au sein des réseaux. Dans le cadre de la thèse de Jean Creusefond, nous travaillons au développement d'un algorithme qui détecte le cœur des communautés. Cette méthode s'appuie sur la définition d'une nouvelle métrique, et utilise le parcours de graphe LexDFS. C'est un travail en cours de soumission.

Filtrage de contenus sur le web. Sur le web, il peut y avoir beaucoup de « mauvais » contenu (qui n'intéresse personne), posté par des personnes mal intentionnées soit pour les promouvoir soit pour faire fuir les utilisateurs. Dans [71], nous avons développé une méthode permettant de filtrer *a priori* le mauvais contenu grâce à un mécanisme de vote robuste basé sur des mécaniques de théorie des jeux. Cette méthode a été implémentée avec succès dans un site Web où aucun contenu de mauvaise qualité n'a pu être publié.

Internet des objets et plateforme IoTa (F. Laguillaumie, J. Madelaine). La plateforme IoTa s'est développée dans le cadre d'un projet FUI TACITES, entre 2011-2013, avec beaucoup de partenaires [9] ; elle permet le suivi d'objets équipés de puces RFID par leur lecture à distance (internet des objets). Chaque lecture donne lieu à la création d'un événement (départ usine, mise en palette, etc.) lié au numéro unique stocké dans la puce. Le logiciel EPCGlobal normalise la création de ces événements et leur gestion via un système d'information. La plateforme a permis de valider les couches hautes d'EPCGlobal dans un véritable réseau. Nous avons proposé et expérimenté un nouveau service de découverte décentralisé (distributed DS). Enfin, nous avons sécurisé l'ensemble des communications et proposé un mécanisme de gestion fine des droits d'accès. L'ensemble des logiciels développés est accessible sous licence GPL sur le site : `https://forge.greyc.fr/projects/iota-pub/`.

Plateforme ThemaMap (J. Madelaine, G. Richard). ThemaMap est un outil de cartographie thématique multi plateforme, distribué en tant que logiciel libre (`https://themamap.greyc.fr/`). Il est développé en collaboration par le GREYC, le SAIC-CERTIC (Service d'activités industrielles et commerciales & Centre de ressources technologiques pour les TIC) et le CRH (Centre de Recherche Halieutique Méditerranéenne et Tropicale). L'utilisateur peut combiner plusieurs sources de données (bases de données, fichiers de différents formats), les manipuler et ensuite les visualiser très rapidement sur une carte. Les spécificités du logiciel sont d'offrir, d'une part des opérateurs de calcul utilisant de nombreuses fonctionnalités statistiques et primitives géographiques, d'autre part, en plus des représentations de données statistiques usuelles, celle des boîtes à moustaches et la visualisation de flux.

9. France Telecom Sophia, EADS, Cassidian, LIC, FIME, INVIA, ASK, France Telecom Caen

2. Stratégie et perspectives scientifiques

À l'attention du lecteur :

Ce document constitue la version longue de la contribution de l'équipe AmacC au rapport d'évaluation du laboratoire GREYC. Il détaille en particulier les aspects scientifiques.

La dernière version de ce document peut être téléchargée à l'URL
`http://clementj01.users.greyc.fr/hceres/amacc-hceres.pdf`

1 L'équipe

1.1 Les thématiques

L'équipe garde la même organisation en trois thèmes, avec les mêmes dénominations, même si le Thème 3 évolue dans son contenu.

- Thème 1 : Modèles de calcul et complexité descriptive.
- Thème 2 : Structures aléatoires et analyse d'algorithmes.
- Thème 3 : Protection et traitement de l'information.

1.2 Les membres

1.2.1 Permanents, chercheurs et enseignants chercheurs

Au 1er janvier 2016, il y aura un unique collège A (qui partira en retraite en 2018) et 8 collèges B (dont l'un partira en 2016) ; parmi ces collèges B, il y a deux titulaires de l'HDR dont l'un est CR CNRS.

Prénom et nom	Fonction	T1	T2	T3	Dates d'entrée/sortie
Ali Akhavi	MC UFR Sciences	✓	✓	✓	2000
Julien Clément	CR CNRS– HDR		✓	✓	2005
Etienne Grandjean	PR UFR Sciences	✓			1985 – 2018
Thomas Largillier	MC IUT Caen		✓	✓	2014
Loïck Lhote	MC ENSICAEN		✓	✓	2007
Ionona Ranaivoson	MC UFR Sciences	✓			1991
Gaétan Richard	MC UFR Sciences	✓		✓	2009
Véronique Terrier	MC UFR Sciences – HDR	✓			1991

Jerzy Karczmarczuk prendra sa retraite au 1er septembre 2016 et n'apparaît donc pas dans les perspectives.

1.2.2 Membres associés

Prénom et nom	Statut	T1	T2	T3	Date d'entrée (comme associé)
Marc Girault	Orange Labs			✓	2010
Sébastien Canard	Orange Labs			✓	2010
Florent Madelaine	Mis à disposition		✓		2015
Jacques Madelaine	Collaborateur		✓		2014
Sylvain Peyronnet	PR en disponibilité		✓	✓	2014
Brigitte Vallée	Émérite CNRS		✓	✓	2015

Florent Madelaine est mis à disposition par l'université de Clermont-Ferrand en 2015-2016 ce que nous souhaitons prolonger. Nous l'incluons donc dans le projet de recherche.

1.2.3 Doctorants

Le faible nombre des doctorants est très préoccupant.

Prénom et nom	Financement	T1	T2	T3	Date d'entrée
Nicolas Bacquey	CNRS + Région	✓			oct. 2012
Jean Creusefond	A. Minist.		✓	✓	oct. 2013
Dimitri Darthenay	A. Minist.		✓	✓	oct. 2015
Pablo Rotondo	ANR DynA3S (associé)		✓		oct. 2015

2 Contexte scientifique et stratégie

2.1 Analyse SWOT

Points forts et opportunités

- un rôle de référence en algorithmique et complexité au sein du laboratoire, surtout dans les thématiques du thème « Modèles de calcul et complexité descriptive » ;
- une réelle vie d'équipe ;
- une participation dynamique à la vie du laboratoire, via l'organisation du séminaire Algo (qui se veut plus large qu'un séminaire d'équipe) ;
- des résultats scientifiques bien diffusés et fondés souvent sur des approches originales ;
- une très bonne visibilité nationale, souvent internationale ;
- une excellente insertion dans la communauté nationale (GdR IM et projets ANR) ;
- beaucoup de thèses (8) et d'habilitations (4) soutenues durant la période.
- le recrutement de la paire PR +MC a été une occasion d'ouvrir le profil de l'équipe, et nous voulons conserver cette ouverture malgré le départ du professeur en développant les interactions avec son entreprise ;
- le renforcement des interactions entre thèmes « Modèles de calcul et complexité descriptive » et « Structures aléatoires et analyse d'algorithmes », qui se concrétise aussi par la remise en route du groupe de travail ;
- le rôle des membres associés d'Orange et la complémentarité avec l'équipe « Monétique et Biométrie » en matière de sécurité (organisation du séminaire Cryptographie & Sécurité, interactions dans le Master E-Secure).

Points faibles et risques

- la multiplicité des changements en personnel qui sont intervenus, et leurs répercussions sur les thématiques de l'équipe ;
- le départ des spécialistes de protection de l'information, qui affaiblit le thème « Protection et traitement de l'information » ;
- l'affaiblissement très problématique de l'encadrement, avec actuellement deux collèges A (une DR qui devient émérite en 2015, l'autre collège A qui prend sa retraite en 2018).
- de fortes interactions culturelles entre les thèmes qui, pourtant, ne se concrétisent pas par des publications communes ;

- une adéquation difficile entre le profil de l'équipe et celui des étudiants ; Il s'ensuit un recrutement local de doctorants difficile, pour cette équipe où le bagage des étudiants doit comporter une base importante de mathématiques. L'équipe doit beaucoup compter sur son insertion nationale (ou internationale) pour le recrutement de doctorants.
- une certaine difficulté à diffuser, au sein du laboratoire, les thématiques autour de l'analyse et la complexité en moyenne du thème « Structures aléatoires et analyse d'algorithmes », alors que celles, relatives à la complexité dans le pire des cas, du thème « Modèles de calcul et complexité descriptive » sont mieux comprises ; c'est explicable car le type de mathématiques utilisé dans le thème « Structures aléatoires et analyse d'algorithmes » est plus loin du centre de gravité de la culture du laboratoire.

2.2 Stratégie

Conforter les points forts

- L'équipe AmacC joue un rôle dynamique dans l'animation locale, via l'organisation du séminaire Algo qui est l'une de ses réussites. Elle veut continuer à s'y investir.
- L'équipe est une des équipes de province sur lequel le GdR IM peut s'appuyer, alors que l'informatique mathématique se concentre de plus en plus dans les grands centres. Nous voulons continuer à nous investir dans la vie scientifique et organisationnelle de ce GdR.
- Nous avons eu de très bons succès dans nos dépôts de projets ANR durant la période écoulée, et encore en 2015. Nous comptons montrer le même dynamisme dans la période future dans une conjoncture sans doute plus difficile.
- Nous sommes bien insérés dans la communauté internationale AofA (*Analysis of Algorithms*). Nous sommes particulièrement actifs en Amérique du Sud, où nous développons des collaborations très fructueuses, en particulier au sein du LIA Infinis. Nous continuerons à nous y investir, et nous développerons nos collaborations avec Taïwan.
- Nos publications sont de qualité, et nous voulons continuer cette politique de publications, et accentuer notre effort en publiant plus de résultats intermédiaires dans les grandes conférences internationales du domaine.

Pallier les points faibles et faire face aux risques :

- En 2018, il n'y aura plus de membres collèges A dans l'équipe [1]. Le recrutement d'au moins un PR est indispensable et urgent.
- Le nombre de doctorants (et de post-doctorants) est beaucoup trop faible : il faut profiter de notre bonne insertion aux niveaux national et international pour recruter des étudiants.
- Afin de dynamiser et de renforcer les collaborations et interactions entre thèmes, nous avons remis en route le groupe de travail ; il faut maintenant le pérenniser.
- Nous devons conserver et dynamiser notre participation régionale aux thématiques de sécurité en renforçant les liens avec le groupe d'Orange Labs, l'équipe Monétique et Biométrie et le laboratoire LITIS (au sein de Normastic) ;
- Nous devons conserver notre ouverture en algorithmique plus appliquée notamment via les collaborations avec *ix-Labs* et, plus généralement, via les thématiques du thème 3 « Protection et traitement de l'information ».
- Il faut aussi développer les sujets d'ouverture au sein du département IAA, comme le PEPS HYDRaTA avec l'équipe CODAG ou l'étude en moyenne de protocoles multi-agents avec l'équipe MAD. Cela permettra de renforcer la cohérence du département.

1. Nous pouvons compter par aileurs sur la promotion DR de J. Clément.

3 Projet de recherche de l'équipe AMACC

3.1 Thème 1 : Modèles de calcul et complexité descriptive

Description générale. L'activité du thème s'organise autour de quatre volets :

- un premier volet, qui regroupe des travaux sur l'algorithmique et les classes de complexité des automates cellulaires.
- un second volet, dédié à la « Tight Complexity » (complexité mesurée à un facteur constant près), notamment dans le cadre des automates cellulaires et des problèmes probabilistes.
- un troisième volet, dédié à l'étude des CSP infinis et de fragments de logique monadiques du second ordre.
- un quatrième volet qui définit un cadre qui donne un point de vue symbolique (via des systèmes de réécriture) sur les algorithmes.

3.1.1 Algorithmique et complexité des automates cellulaires

Le groupe étudie leur comportement et leur utilisation, et notamment ce qu'ils permettent de modéliser en tant que systèmes dynamiques. C'est la suite directe des travaux en cours, il s'agit toujours, au travers de problèmes spécifiques, de développer les outils algorithmiques propres à ce modèle de calcul massivement parallèle.

Automates cellulaires à comportement complexe (N. Bacquey, E. Grandjean, G. Richard, V. Terrier). Pour construire de tels automates, une piste sérieuse consiste à partir des résultats obtenus précédemment (voir bilan) sur l'« élection de leader ». Il faut faire un lien avec les agents et, tout particulièrement dans le domaine du *online motion planning*, avec des agents anonymes et sans mémoire mais pouvant marquer l'espace. L'intérêt va bien au delà du transfert de résultat, car cette thématique peut apporter des problèmes intéressants (et potentiellement atteignables) dans le modèle des automates cellulaires.

Petites classes de complexité (N. Bacquey, E. Grandjean, G. Richard, V. Terrier). Là, un de nos objectifs est de préciser les relations entre les langages algébriques et les langages reconnus en temps réel par automates cellulaires unidirectionnels. Partant de notre résultat sur la reconnaissance en temps réel par automates cellulaires unidirectionnels des langages algébriques à croissance linéaire, l'enjeu est maintenant de capturer les langages algébriques à croissance polynomiale. Une étape clé implique de définir un algorithme de synchronisation généralisé. L'ébauche obtenue récemment sur la formalisation des transformations géométriques réalisables sur automate cellulaire en dimension 1 nécessite d'être développée et, autant que possible, étendue dans le cadre des dimensions supérieures.

3.1.2 Algorithmique et « Tight Complexity »

Depuis des années, l'un de nos objectifs constants est de relier, de façon la plus précise possible – *tightly* –, d'une part, l'algorithmique et l'analyse des algorithmes, et d'autre part, la complexité algorithmique : complexité définie donc à une constante multiplicative près, $O(n)$, $O(n \log n)$, $O(n^2)$, etc.

Cela passe bien sûr par l'étude précise des modèles de calcul – RAM, automates cellulaires, essentiellement – et leur algorithmique, mais aussi par la complexité descriptive qui donne une caractérisation logique des classes de complexité, et établit donc leur caractère intrinsèque, indépendant des machines.

Ces dernières années, nous avons étudié la complexité précise des problèmes d'énumération (délai constant ou linéaire sur RAM), puis, plus récemment, la complexité linéaire des automates cellulaires *non déterministes*. Notre objectif maintenant est, d'une part, de caractériser en logique la classe du temps linéaire des automates cellulaires *déterministes*, d'autre part, de mieux cerner

et unifier la « tight complexity » (sur RAM) des problèmes *probabilistes*, intégrant à la fois les aspects Las Vegas et Monte-Carlo.

Automates cellulaires et formules de Horn (E. Grandjean, G. Richard). On sait combien les formules de Horn sont fondamentales en informatique : en programmation logique avec le langage Prolog ; en bases de données avec le langage de requêtes Datalog, etc. On sait aussi que ces formules permettent de caractériser exactement la classe PTIME des problèmes de complexité polynomiale. Dans la suite de ce résultat, nous sommes en train d'exhiber une relation surprenante entre la classe des formules de Horn et la complexité des automates cellulaires. Il s'agit de la « caractérisation exacte » suivante : Un langage de mots ou d'images est reconnu en temps linéaire par un automate cellulaire *déterministe* de dimension d fixée quelconque *si et seulement si* ce langage est défini par une formule de Horn en logique du second ordre avec seulement $d+1$ variables du premier ordre : intuitivement, les d variables correspondent aux d dimensions de l'automate et la $d+1$ ième variable représente le temps. Un tel résultat (dont la preuve reste à confirmer) établit une *symétrie parfaite* (interchangeabilité) entre l'espace et le temps dans le calcul d'un automate cellulaire déterministe, ceci quelle que soit la dimension de l'automate (résultat déjà établi dans le cas non déterministe, voir Bilan).

Une monographie en cours sur la « Tight Complexity » (E. Grandjean). C'est un livre en construction depuis deux ans (environ 80 pages sont déjà écrites, près de la moitié). Il s'agit de décrire les fondements de ce qu'on appelle la « Tight Complexity ». C'est une théorie où la complexité algorithmique est mesurée à un facteur constant près ($O(n)$, $O(n\log n)$, $O(n^2)$, etc.) et qui « s'ajuste exactement » à la complexité intuitive des algorithmes. Celle-ci est définie, à partir d'un modèle de calcul précis, la RAM, qui permet de définir des classes de complexité *robustes*, (invariantes pour de nombreuses variations du modèle) et qui contiennent *exactement* les problèmes de la complexité considérée. Outre son intérêt pour l'analyse rigoureuse des algorithmes, la « Tight Complexity » construit une nouvelle vision, plus précise et différenciée, des classes de problèmes difficiles. Elle établit, grâce à la notion de réduction linéaire, que nombre d'entre eux ont exactement la même complexité : typiquement, parmi les problèmes NP-complets classiques, un grand nombre sont *linéairement* équivalents au problème SAT – c'est-à-dire, lui sont inter-réductibles par des réductions linéaires – donc ont exactement *la même « tight complexité* que SAT. Enfin, la monographie présente les résultats de complexité descriptive concernant la « Tight Complexity ».

Une classe de complexité probabiliste précise et robuste (E. Grandjean, F. Madelaine). En relation avec le thème Aléa, nous étudions les classes de complexité probabilistes et exhibons une classe probabiliste qui nous semble très intéressante la classe des problèmes calculables par algorithmes probabilistes sur RAM, en temps moyen $O(T(n))$ (pour une fonction T fixée) et avec erreur bornée. Cette classe est une variante « tight » de la classe bien connue BPP et nous sommes en train de montrer qu'elle possède de multiples propriétés : elle est précise mais très générale, et aussi très robuste, dans le sens où elle reste inchangée pour beaucoup de variantes de sa définition ; elle intègre tous les problèmes algorithmiques probabilistes usuels, de type Las Vegas ou Monte Carlo, avec la *même complexité*. Nous cherchons aussi une caractérisation logique de cette classe.

3.1.3 Logique et complexité (F. Madelaine)

Feder et Vardi ont introduit une logique – un fragment de MSO – pour capturer les problèmes de contraintes. Cette logique, nommée MMSNP, suit une dichotomie si est seulement si c'est également le cas des problèmes de contraintes. Pour les CSP, la *conjecture de la dichotomie* énoncée par Feder et Vardi stipule que tout problème de contrainte est soit facile (dans P) soit difficile (NP-complet). Par ailleurs, la logique MMSNP est un peu trop expressive et capture en

fait des problèmes de contraintes infinis au sens de Bordisky.

Il se trouve que cette logique est réapparue récemment en bases de données et en représentation des connaissances en correspondance étroite avec la notion de requêtes conjonctives « préfiltrée » par une ontologie, qu'on appellera requête hybride ci-après. En particulier, il serait tout à fait intéressant de pouvoir établir si la logique MMSNP_2 introduite par Florent Madelaine suit une dichotomie entre P et NP-complet (même sous l'hypothèse de la conjecture de la dichotomie) puisque cette logique correspond elle aussi à des requêtes hybrides dans un contexte un peu plus général où l'ontologie est accessible via un langage un peu plus puissant.

Un autre exemple de question concrète concerne la décidabilité / complexité de savoir si une requête hybride peut se reformuler dans un formalisme particulièrement bien connu (typiquement, la logique du premier ordre ou encore DATALOG). Pour approcher ce genre de questions, on se propose d'utiliser d'une part des techniques directes fort de l'expérience acquise sur MMSNP (forme normale, compréhension de la combinatoire sous-jacente) et d'autre part d'utiliser des techniques de l'arsenal de Manuel Bodirsky (théorie des modèles sur les structures homogènes, théorie de Ramsey et amalgamation, algèbre universelle).

3.1.4 Analyse symbolique des algorithmes (A. Akhavi)

Un algorithme est ici vu comme un système dynamique qui à partir de la donnée d'entrée effectue un séquence de transformations élémentaires jusqu'à obtenir le résultat. Cet ensemble de transformations élémentaires est vu d'un point de vu symbolique comme un alphabet et une exécution s'écrit comme un mot sur cet alphabet. Dans notre approche un algorithme est identifié à un système de réécriture qui caractérise les mots qui correspondent à de « véritables » exécutions de l'algorithme.

Dans la suite, nous souhaitons expliciter les règles de réécriture associées à un algorithme donné. Ces règles sont en rapport avec les bases de circuits dans le graphe de Cayley d'un certain groupe d'automorphismes (cela a été déjà fait dans le cas de l'algorithme de Gauss et de quelques algorithmes de tri). Il nous paraît également intéressant de traduire les propriétés classiques des systèmes de réécriture (notamment une mesure de la localité de ces règles) en terme de propriétés des algorithmes associés. Enfin nous cherchons à étudier l'association « algorithmes – systèmes de réécriture » sur plus d'exemples variés.

Un des objectifs de ce travail est également de relier la complexité des algorithmes à la croissance de leurs groupes d'automorphisme.

3.2 Thème 2 : Structures aléatoires et analyse d'algorithmes

Comme dans la période précédente, l'activité du thème s'organise autour de quatre volets :

- un premier volet, méthodologique, qui regroupe des travaux généralistes en combinatoire analytique ;
- un second volet, dédié à l'analyse probabiliste des algorithmes du texte et de la fouille de données ;
- un troisième volet, dédié à l'analyse probabiliste des algorithmes en arithmétique ;
- un quatrième volet dédié aux approches probabilistes de l'algorithmique des grands systèmes.

3.2.1 Méthodes générales en combinatoire analytique

Les thématiques de ce volet sont très présentes dans un nouveau projet ANR auquel l'équipe participe. C'est le projet ANR MetaConc « Méthodes analytiques non-conventionnelles en Combinatoire » de l'appel « Projets de Recherche Collaborative – International » entre France et Taiwan, accepté à la session 2015.

Poisson, Mellin, Rice, Newton (B. Vallée). La combinatoire analytique, surtout quand elle s'applique à des algorithmes ou des structures de texte, adopte souvent une démarche indirecte. Plutôt que de travailler directement dans un modèle où le nombre des entrées est fixé, on considère d'abord que ce nombre suit une loi de Poisson (c'est le modèle de Poisson) ; après avoir fait l'analyse dans ce dernier modèle, on veut revenir au premier, et il y a deux chemins possibles dans la « boite à outils » de l'analyseur d'algorithmes, qui, selon les cas et les goûts, utilise l'un ou l'autre : usage de la transformée de Mellin suivie de la dépoissonisation, ou usage conjoint de deux formules, appelées respectivement Rice et Newton. Il s'agit de comparer ces méthodes en profondeur : sont-elles finalement équivalentes ? nous pensons que la réponse est essentiellement oui ; sinon, il faudrait réfléchir à une stratégie, qui explique quand on peut –ou on doit– utiliser l'une ou l'autre de ces méthodes.

Lois beta (L. Lhote, B. Vallée). Les résultats de l'analyse sur le pgcd multiple [6, 36] font apparaître des lois limites beta pour la distribution du nombre d'itérations. C'est la première fois que de telles lois interviennent dans nos analyses, et elles ne sont guère fréquentes en combinatoire analytique. Nous projetons de mieux comprendre le cadre algorithmique générique où peuvent apparaître de telles lois, en exhibant –comme on le fait pour les lois gaussiennes– des conditions sur les séries génératrices bivariées (et leurs singularités) qui conduisent à de ces lois. Nous voulons étudier les « transitions de phase » entre lois gaussiennes et lois Beta.

Génération aléatoire avec les séries de Dirichlet (J. Clément, B. Vallée). Les séries génératrices construites lors du processus de combinatoire analytique sont aussi utilisées pour engendrer des structures aléatoires de grande taille : cette génération aléatoire, importante en elle-même, est aussi essentielle quand on veut expérimenter sur ces structures. Dans notre cadre original de combinatoire dynamique, nous travaillons naturellement avec des séries génératrices de type Dirichlet, et il faudrait ainsi revisiter toutes les approches classiques de génération aléatoire, qui utilisent des séries génératrices classiques pour les adapter aux séries génératrices de Dirichlet de notre contexte.

3.2.2 Analyse des algorithmes du texte et d'hypergraphes

Entropies (L. Lhote). La collaboration avec les mathématiciens de Caen s'inscrit maintenant dans la durée et le groupe de travail « Mots, Séquences et Entropies » continue à se réunir avec une certaine régularité. Nous souhaitons étendre les calculs de taux d'entropie (généralisé et renormalisé). Jusqu'ici, on a étudié les sources qui satisfont une hypothèse forte dite des quasi-puissances, vérifiée pour beaucoup de sources simples (sources sans mémoire, chaînes de Markov). Ce cadre, où le taux d'entropie de Shannon est non nul, est très propice à l'analyse complexe. Nous projetons de considérer désormais des processus à entropie nulle, où l'analyse va être plus difficile.

Modélisation de sources (L. Lhote, B. Vallée). Les sources dynamiques fournissent un modèle de source avec un taux de corrélation entre symboles consécutifs possiblement fort. Il reste à préciser le statut de ce modèle parmi les autres modèles utilisés. Est-il vrai qu'une source dynamique peut être vue comme la limite d'une suite de chaînes de Markov, quand l'ordre de la chaîne croît ? Inversement, à quelles conditions une suite de chaînes de Markov définit-elle une source dynamique ? Les spécialistes de systèmes dynamiques se sont posés en partie ces questions mais dans un cadre assez différent. Une autre question importante est : Peut-on atteindre par ce modèle le formalisme des chaînes de Markov cachées, très utilisé par exemple dans le contexte de la bio-informatique ? celui des chaînes de Markov à longueur variable (VLMC) ? Ces questions pourraient être reprises dans le groupe de travail caennais sur l'entropie « Mots, Séquences et Entropies ».

Codage de source (J. Clément). La classe des codes bifixes a des propriétés très intéressantes pour la synchronisation, car les mots de ces codes peuvent être décodés instantanément (i.e., dès leur lecture) quel que soit le sens de leur lecture. Depuis qu'elle a été introduite, cette classe se révèle bien plus difficile à étudier que celle des codes préfixes, et donne lieu a des conjectures non élucidées. Dans ce cas, on ne connaît par exemple pas d'analogue du célèbre algorithme de Huffman. La question est difficile et encore l'objet de conjectures. dans un premier temps nous pourrions étudier le cas des codes bifixes pour une source géométrique.

Analyse d'autres algorithmes de recherche et de tri (J. Clément, B. Vallée). Nos travaux récents [15, 16] ont ouvert la voie à un nouveau champ de recherche. Il s'agit de revisiter l'analyse en moyenne des principaux algorithmes de tri et de recherche (ceux qu'on apprend en Master à l'université), lorsque les clés sont des mots produits par une source générale et le coût de la comparaison entre clés est le nombre de symboles utilisés pour comparer les mots. Le tri fusion, le tri par tas ou la recherche dichotomique sont des candidats naturels.

Mots sturmiens (B. Vallée). Les mots sturmiens sont les mots les plus simples qui ne sont pas périodiques. Suite à nos travaux de la période précédente [35], il s'agit d'étudier le comportement d'un mot sturmien aléatoire, quand celui-ci satisfait toutefois à des contraintes spécifiques : il est produit par substitution (et est donc associé à un irrationnel quadratique), ou il est associé à un nombre réel dont les quotients dans le développement en fraction continue sont bornés ; enfin, il est aussi important de comprendre comment un mot périodique « devient » sturmien, en décrivant la transition de phase entre le cas rationnel et le cas réel. Ces études vont être entreprises dans le cadre de la thèse de Pablo Rotondo [2] (co-dirigée par V. Berthé et A. Viola, avec la participation de B. Vallée).

Hypergraphes (L. Lhote). Nous cherchons à raffiner l'analyse des algorithmes de calcul des traverses minimales d'hypergraphes, déjà obtenue dans [18], dans deux directions. Tout d'abord, on peut travailler dans le modèle aléatoire qui généralise le modèle $G(n,p)$ d'Erdös et Rényi à des hypergraphes, et conduit à des hypergraphes quasi-homogènes. Ce modèle simple n'est pas donc pas réaliste. On peut travailler aussi dans d'autres modèles plus adaptés aux applications qui conduisent à des hypergraphes inhomogènes. Enfin, d'autres modèles, qui font « grossir » les hypergraphes, correspondent bien à la stratégie de certains algorithmes d'extraction de traverses. Il serait donc très intéressant d'effectuer l'analyse en moyenne du calcul des traverses minimales dans ces divers modèles. Mais nous songeons aussi à une autre approche, plus méthodologique, fondée sur la combinatoire analytique, qui s'appuierait sur la structure des hypergraphes aléatoires, et serait ainsi plus proche de la structure même des problèmes étudiés.

3.2.3 Analyses en arithmétique

Il s'agit de revisiter l'analyse d'algorithmes classiques de la dimension 1, mais aussi de tenter des analyses en dimension supérieure.

Analyse en distribution d'algorithmes d'Euclide via la méthode des moments (B. Vallée). Ian Morris vient de publier plusieurs travaux qui revisitent les résultats obtenus par l'équipe, en particulier l'analyse en moyenne de l'algorithme Binaire, et l'analyse en distribution de l'algorithme usuel. Morris utilise la méthode des moments à la place de la méthode « à série génératrice bivariée », et son approche donne des résultats plus faibles (pas de vitesse de convergence) mais elle est s'applique aussi plus facilement, sans avoir besoin de résultats forts sur le système dynamique « à la Dolgopyat ». Ce serait intéressant de débuter une collaboration avec Morris, en cherchant à mélanger nos approches.

2. Pablo va être doctorant « associé » au GREYC

Autres algorithmes de pgcd : Plus Moins et SRT (L. Lhote, B. Vallée). L'algorithme Plus-Moins est un algorithme du pgcd très utilisé. C'est une généralisation naturelle « signée » de l'algorithme binaire, dont l'équipe avait obtenu l'analyse en 1998, et qui s'était déjà avérée délicate. Cependant, le signe supplémentaire cause de nouveaux ennuis, et Il faut donc trouver un bon espace fonctionnel, et maintes tentatives ont déjà échoué. C'est le seul algorithme qui reste vraiment rebelle, à ce jour, à l'analyse dynamique...
Appelé ainsi du nom de ses auteurs, l'algorithme SRT est un algorithme de division efficace pour les calculs en « virgule flottante ». Pippenger et McCann ont déjà exhibé le système dynamique et l'ont étudié avec des outils de théorie ergodique. Malgré sa simplicité apparente (c'est un système dont toutes les branches sont affines et de même pente), ce système dynamique est non markovien, et suscite des questions à la fois très naturelles et non triviales.
L'algorithme de Berlekamp-Massey est à la fois un outil de cryptanalyse et une brique essentielle pour le décodage de codes correcteurs. Le groupe a effectué l'analyse en moyenne d'une version de cet algorithme, proche de l'algorithme d'Euclide dit « étendu et interrompu », et il doit maintenant transposer cette analyse au véritable algorithme de Berlekamp-Massey.

Le pgcd multiple sur les entiers (L. Lhote, B. Vallée). En prolongement des travaux de la période précédente [6], on peut chercher à analyser d'autres paramètres (en particulier la complexité en bits), en moyenne ou pourquoi pas en distribution.

Analyse d'algorithmes d'Euclide multi-dimensionnels (L. Lhote, B. Vallée). Il existe aussi beaucoup de généralisations de l'algorithme d'Euclide, de nature variée, qui calculent des approximations rationnelles simultanées (algorithmes de Jacobi-Perron, Brun, Selmer-Euclide...), et sont utilisés en géométrie discrète par exemple. Avec Valérie Berthé, nous souhaitons effectuer l'analyse en moyenne systématique de ces algorithmes afin de les comparer entre eux, notamment en vue de leurs applications géométriques. L'analyse en distribution semble également possible, mais il faudra auparavant étendre les travaux de Dolgopyat-Baladi-Vallée qui ont permis les analyses en distribution de l'algorithme d'Euclide classique, ou bien utiliser l'approche « à la Morris » (voir plus haut).

3.2.4 Approches probabilistes de l'algorithmique des grands systèmes (T. Largillier)

Les recherches que nous avons menées jusqu'ici mettent en place des méthodes qui déclassent l'effet des votes malicieux au sein des sites web communautaires. Nous voulons maintenant mettre en place des méthodes collaboratives de classement, élection, détection du webspam, et cherchons à utiliser des concepts issus de la théorie des jeux. Ces méthodes devront être robustes, c'est-à-dire insensibles à l'action d'une minorité d'utilisateurs malicieux. Il existe déjà quelques travaux embryonnaires, mais surtout une abondante littérature (dans le domaine de la théorie de la décision en économie, et notamment dans la théorie du choix social) qui aborde des problèmes similaires, et nous voudrions l'adapter à notre contexte particulier.

3.3 Thème 3 : Protection et traitement de l'information

Comme dans la période précédente, le thème regroupe les recherches de l'équipe qui sont orientées vers la protection et le traitement de l'information en les organisant comme précédemment en quatre volets.

- le premier volet est dédié à l'étude amont des objets cryptographiques et se concentre sur deux objets principaux : les fonctions booléennes et surtout les réseaux euclidiens. Le recentrage sur ce dernier objet est motivé par le rôle croissant des réseaux euclidiens en cryptographie et par les compétences historiques de l'équipe.
- le deuxième volet est dédié aux protocoles cryptographiques. La cryptographie à base de réseaux euclidiens est en plein essor, et c'est ce qui justifie que nous conservions ce volet,

même si à l'heure présente, aucun membre de l'équipe n'est spécialiste de ce domaine.
- dans le troisième volet, les activités autour de l'internet des objets disparaissent, tandis que les recherches autour du web et de l'information géographique restent bien présentes.
- Dans un quatrième volet, l'équipe souhaite s'ouvrir à deux nouveaux domaines d'application : l'informatique musicale et à la recherche d'informations qui y est associée, et la fouille de données, en liaison avec les hypergraphes du thème 2.

3.3.1 Étude amont des objets cryptographiques

Analyse d'objets cryptographiques Nous voulons d'abord continuer l'analyse de fonctions booléennes, dans un cadre un peu différent ; nous voulons réfléchir aux réductions entre problèmes « pire-cas, cas-moyen » et nous projetons surtout de nous attaquer à deux études autour des réseaux euclidiens et de leur réduction : étude du pire des cas de la dimension 3, modélisations simplifiées des réseaux et de leur réduction.

Diagrammes de décisions binaires (J. Clément). Les diagrammes de décision binaires (BDD) et leurs variantes ont été introduites par Randal Bryant dans les années 80. Il s'agit d'une structure de données parmi les plus populaires et les plus efficaces pour représenter une fonction booléenne en pratique. Si cette structure de données est aujourd'hui bien connue, il reste encore beaucoup de questions ouvertes à son propos. Notamment la structure et les propriétés des fonctions booléennes qui admettent une représentation de taille raisonnable (i.e., non exponentielle) sous cette forme sont mal ou pas connues actuellement. Une application possible serait la génération aléatoire de fonctions booléennes sous la forme de diagrammes de décision binaires qui auraient de « bonnes » propriétés du point de vue des applications cryptographiques.

Réductions « pire-cas, cas-moyen » (B. Vallée, A. Akhavi). L'équipe, avec l'expertise conjointe du thème 1 (réductions entre problèmes) et du thème 2 (distributions probabilistes) est bien armée pour réfléchir sur les réductions « pire-cas moyen-cas » qui sont utilisées dans l'étude de la difficulté de deux problèmes de la « crypto-réseaux » : les problèmes LWE (*Learning With Errors*) et SIS (*Short Integer Solution*).

L'analyse de l'algorithme LLL en dimension $n = 3$ (A. Akhavi, J. Clément). Cette étude pourrait être un test pour la méthode symbolique introduite par A. Akhavi, déjà appliquée avec succès à certains algorithmes de tri et à l'algorithme de réduction de Gauss (LLL pour $n = 2$). Un des objectifs de ce travail est d'identifier précisément le pire cas de l'algorithme LLL pour la dimension 3, la première dimension où ce pire cas n'est pas connu (en collaboration avec Adeline Langlois). Ce travail rejoint avec un autre point de vue l'étude du système dynamique de LLL en dimension 3 du thème 3 (voir page 18).

Modélisations simplifiées des réseaux euclidiens et de leur réduction L'équipe mène quatre études principales ayant trait aux modélisations simplifiées des réseaux euclidiens et de leur réduction.

Modélisation de réseaux cryptographiques (J. Clément, L. Lhote, B. Vallée). Nous voudrions étendre l'approche de la période précédente, que nous avons adoptée lors de la thèse de Mariya Georgieva [] et modéliser d'autres familles de réseaux : ceux qui sont issus de la conception ou de la cryptanalyse de protocoles complètement homomorphes, ceux qui modélisent des problèmes de communications sans fil, ou plus généralement des réseaux structurés (notamment les réseaux dits idéaux), qu'on utilise pour rendre les protocoles plus efficaces.

Analyse des modèles d'exécution de l'algorithme LLL (A. Akhavi, J. Clément, L. Lhote, B. Vallée). Là aussi, nous voudrions poursuivre notre approche de la période précédente. Nous avons proposé trois modèles intermédiaires simplifiés (de simplification décroissante) entre le modèle M1 du tas de sable (exagérément simplifié) et le véritable algorithme LLL (dit modèle M5). Le modèle M2 est un système dynamique déterministe à trou où les coefficients sous–

diagonaux sont supposés fixés (c'est là que réside la simplification), tandis que le modèle M3 est un système probabiliste où ces coefficients sont tirés aléatoirement dans $[-1/2, 1/2]$. Nous avons analysé partiellement le modèle M2, lorsqu'il est gouverné par une stratégie gloutonne ; il reste à terminer l'analyse du modèle M2 et à aborder l'étude du modèle M3. Il faut aussi comparer les résultats théoriques de ces analyses avec les résultats expérimentaux du vrai algorithme LLL, afin d'élaborer des conjectures sur le comportement du vrai algorithme..

Modélisations d'autres algorithmes de réduction de réseaux (A. Akhavi, J. Clément, L. Lhote, B. Vallée). Il existe d'autres algorithmes de réduction de réseaux, plus efficaces que l'algorithme LLL, et effectivement utilisés dans les cryptanalyses (algorithmes BKZ, HKZ, *deep-insertion*, etc.). Une modélisation par tas de sable (modèle M1) a été proposée pour BKZ par une équipe de l'ENS Lyon, et il serait intéressant de proposer des modélisations simplifiées pour ces autres algorithmes.

Réductions de réseaux euclidiens structurés (A. Akhavi, J. Clément, L. Lhote, B. Vallée). Toutes les modélisations simplifiées précédentes (à la fois sur les entrées des algorithmes et leur exécution) permettent de mieux comprendre le comportement de ces algorithmes sur les réseaux véritablement utilisés en cryptographie, et de répondre à la question (au moins dans les modèles simplifiés). Est-il aussi difficile de réduire un réseau structuré qu'un réseau non structuré ?

3.3.2 Protocoles cryptographiques

Ce volet s'est fragilisé à cause de deux départs, suite à des promotions PR. Mais il y a quatre éléments qui plaident pour maintenir ce volet :

- La présence à Caen du groupe de Cryptographie Appliquée d'Orange Labs. C'est un partenaire historique de l'équipe et deux de ses membres sont associés à AmacC et ont co-encadré plusieurs thèses.
- Le Master E-Secure porté pour moitié par l'équipe (avec l'équipe de Monétique-Biométrie).
- Des interactions potentielles avec l'équipe de Monétique-Biométrie.
- Des interactions avec Adeline Langlois, chargée de recherche CNRS, spécialiste de la cryptographie basée sur les réseaux, qui vient d'être nommée à Rennes et avec laquelle nous voulons tisser des collaborations régulières [3].

3.3.3 Traitement et représentation de l'information

Ce volet regroupe des thématiques déjà présentes dans la période précédente : le web et l'information géographique.

Dans le domaine du web (J. Creusefond, T. Largillier, S. Peyronnet). Nous allons continuer à travailler sur la détection de communautés dans les environnements Web et réseaux sociaux, et sur le filtrage de contenu disponible sur le web.

Détection de communautés. Nous voulons étudier les communautés chevauchantes. Les communautés sont souvent vues comme des extensions de la classification des utilisateurs en groupes sémantiquement cohérents. Mais le cloisonnement d'utilisateurs n'est pas toujours la meilleure solution, car chaque utilisateur peut naturellement appartenir à plusieurs communautés, et des méthodes de découpage qui autorisent les communautés à se chevaucher sont donc plus naturelles. Il faut alors regrouper les arêtes du graphe plutôt que les nœuds, et un nœud appartient dès lors à tous les groupes des arêtes auxquelles il est incident.

Filtrage de contenus. L'explosion de contenu disponible sur le Web se produit désormais en continu et il est donc important d'en filtrer le contenu. Il est également important de garantir

3. Nous avons déjà bien réfléchi sur ce projet de recherche, puisqu'Adeline avait placé le GREYC en première position de ses chox d'affectation. Même si l'INS2I n'a pas suivi ce premier choix, ce projet nous tient toujours autant à coeur, de part et d'autre, et nous chercherons à le concrétiser, dans cette situation bien sûr moins favorable.

la neutralité du contenu d'une plateforme afin qu'elle garde la confiance de ses utilisateurs. Il est de plus souhaitable que les utilisateurs soient à la fois impliqués dans le filtrage du contenu, mais aussi déchargés en partie par une procédure automatique. Les outils de théorie de jeux sont fructueux dans le premier cas tandis que les approches de classification (des utilisateurs et du contenu) sont utiles dans le second. Notre objectif final est de lier les deux approches pour n'avoir recours au filtrage utilisateur que dans le cas où la classification automatique échoue.

Informatique géographique et plateforme ThemaMap (J. Madelaine, G. Richard). ThemaMap offre une large palette de représentations visuelles des données statistiques géo-localisées et permet aussi de réaliser de nombreux calculs. Il reste néanmoins à intégrer d'autres opérateurs (calcul de résidus, quelques opérateurs d'analyse spatiale), et à utiliser du parallélisme (*threads*) pour améliorer les performances. De manière plus originale, nous pourrions aussi utiliser des systèmes dynamiques dérivés de réseaux de neurones dans l'analyse qualitative de cartes thématiques. Il sera alors possible de mettre en valeur des situations d'émergence de pôles, ou de phénomènes de frontières dans les thématiques représentées par les utilisateurs.

3.3.4 Nouveaux domaines d'application

Dans ce volet, nous proposons de travailler dans deux nouveaux domaines : la fouille de données et l'information musicale.

Fouille de données et projet HYDRATA (J. Clément, L. Lhote, B. Vallée). Le projet HYDRATA est un projet PEPS (Hypergraphes et Datamining : Algorithmes et Analyses probabilistes), porté par l'équipe et accepté en 2015. Il réunit des membres de l'équipe AmacC (J. Clément, L. Lhote et B. Vallée) et de CoDaG (ainsi que d'autres équipes parisiennes). La structure d'hypergraphe est au cœur du projet, et est étudiée sous des angles complémentaires (combinatoire analytique et fouille de données). L'analyse d'algorithmes permettra sans doute d'expliquer finement les phénomènes qu'on observe en fouille de données sur des problèmes difficiles, dans le pire des cas ou en moyenne. Elle permettra sans doute d'élaborer une meilleure stratégie algorithmique, ou même de concevoir de nouveaux algorithmes. Réciproquement, la fouille de données peut apporter de nouveaux problèmes à l'analyse d'algorithmes, l'incitera à approfondir la structure d'hypergraphe, et à étudier finement des algorithmes très utilisés dans la vie algorithmique « réelle ».

Informatique musicale (J. Clément). Il existe de vastes corpus de fichiers musicaux où la recherche d'informations s'avère difficile, et nous souhaitons initier un projet autour de la recherche d'informations dans les représentations musicales (audio ou symboliques). Il s'agit de développer des outils mathématiques et informatiques pour formaliser des structures musicales, avec des applications qui concernent la transcription automatique, la recherche de motifs musicaux, l'analyse musicale. Des collaborations potentielles sont déjà envisagées : avec des collègues du GREYC (François Rioult, équipe CoDaG), Jean-Baptiste Rolland (Ingénieur de recherche dans l'entreprise Steinberg) et d'autres acteurs locaux (Conservatoire régional de Caen ou « Relais d'Sciences »). À plus long terme, nous pourrions impliquer de manière transversale des compétences variées du laboratoire, allant du traitement du signal à la théorie des langages (dits formels) en passant par la fouille de données et l'algorithmique du texte.

3. Annexe : liste des réalisations

1 Participation à des groupes de recherche, des regroupements thématiques des projets ou des contrats

Régional

- avec le LMNO (laboratoire de Mathématiques) :
  - collaborations entre Loïck Lhote et Valérie Girardin [14],
  - thèse de Mathieu Roux co-encadrée par Brigitte Vallée et Driss Essouabri [100],
  - projet ANR SIMPATIC (Fabien Laguillaumie et John Boxall).
- au sein de la Fédération Norm@Stic :
  - co-animation de l'axe Algorithmique et Combinatoire,
  - participation à l'axe Systèmes Complexes.
  - participation à l'axe transverse Sécurité
- avec Orange Labs : participation conjointe au Master E-Secure ; co-encadrement d'une thèse CIFRE [96].
- avec LEMONLOOP (Normandie Incubation) un contrat d'étude (prestation de service) : « Etude de faisabilité d'un système authentification et de compression entre un client et un serveur LEMONLOOP ». Janvier-Mars 2011 (Jean Saquet, Gaétan Richard).

National

- Groupement de recherche Informatique Mathématique (GdR IM). Prise dans son ensemble, l'équipe participe à 6 groupes de travail (GT) sur la vingtaine au total du GdR.
  - SDA2 (V. Terrier, G. Richard, J. Clément, L. Lhote, B. Vallée),
  - COMATEGE (J. Clément, B. Vallée),
  - ALGA (F. Madelaine, V. Terrier, G. Richard, E. Grandjean, J. Clément, A. Akhavi)
  - ALEA (A. Akhavi, J. Clément, J.-M. Le Bars, L. Lhote, B. Vallée),
  - C2 (A. Akhavi, L. Lhote, F. Laguillaumie, A. Otmani, J.-M. Le Bars, B. Vallée).
  - ARITH (L. Lhote, B. Vallée)
- Liens conservés avec des membres partis dans d'autres laboratoires
  (le LIP, avec Fabien Laguillaumie – le LITIS avec Ayoub Otmani).
- Projets ANR : sur la période, l'équipe a participé à 11 projets ANR (essentiellement du programme Blanc), 1 fois comme porteur, 6 fois comme partenaire, et 4 fois avec des participations individuelles.
  - Projet LAREDA «Lattice Reduction Algorithms : Dynamics, Probabilities, Experiments, Applications», 2007-2011 (GREYC porteur, avec IMB (Dijon), IRISA (Rennes), LIP (Lyon), LIRMM (Montpellier).
  - Projet ENUM « Algorithmes et complexité des problèmes d'énumération », 2007-2011 (ANR Blanche, GREYC partenaire).
  - Projet PACE « Pairings and Advances in Cryptology for E-cash », 2007-2011 (ANR Blanche, GREYC partenaire).
  - Projet WINGS « Widening Interoperability for Networking Global Supply Chains » 2009 - 2011 (partenaire).

  - Projet EMC « Émergence dans les modèles de calcul » 2009-2012 (ANR Blanche, participation individuelle).
  - Projet BOOLE « Quantification des propriétés de structures booléennes avec des méthodes de combinatoire analytique ou des méthodes probabilistes » 2009-2012 (ANR Blanche, GREYC partenaire).
  - Projet MAGNUM « Méthodes algorithmiques de génération aléatoire non uniforme. modèles et applications », 2010-2014 (ANR Blanche, participation individuelle).
  - Projet SIMPATIC « SIM et théorie des couplages pour la sécurité de l'information et des communications », 2012-2015 (GREYC partenaire).
  - Projet DynA3S « DYNamics of gcd algorithms : an Algorithmic, Analytic, Arithmetic, and Symbolic approach », 2013-2016 (ANR Blanche, GREYC partenaire).
  - Projet AGGREG « Algorithmes et complexité des problèmes d'énumération », 2014-2017 (participation individuelle).
  - Projet METACONC « Méthodes analytiques non-conventionnelles en Combinatoire », accepté à la session 2015 (ANR Blanche internationale entre la France et Taiwan, participations individuelles).
- Projet FUI : TACITES « Tag Authentification and Convergence for Internet of Things and Enhanced Security » 2011- 2013 (partenaire).
- Projet PEPS HYDrATA « Hypergraphes et Data-Mining : Algorithmes et Analyses probabilistes» (2015) dans le Programme FaSciDo (Fondements et Applications de la Science des Données) avec le LIPN (Paris Nord).

International

- Projet NSF 2013, avec Manuel Lladser (Boulder, Colorado, USA).
- Projet ECOS Sud (Uruguay), 2008-2011.
- Projet STIC-AMSUD (France, Argentine, Uruguay) 2013-2014.
- Création d'une équipe dans le LIA INFINIS à Buenos Aires (Laboratoire International Associé).
- Invitations de chercheurs : Anahí Gajardo (Universidad de Concepción, Chile), Alfredo Viola (Université de Montevideo, Uruguay), Manuel Lladser (Boulder, Colorado, USA), Eda Cesaratto (Buenos Aires, Argentine), Hsien-Kwei Hwang (Taiwan).

2 Animation scientifique

- Organisation de séminaires locaux :
  - organisation du séminaire ALGO hebdomadaire (existe depuis vingt ans)
  - organisation du groupe de travail AmacC (bimenuel depuis 2014) ;
  - co-organisation du séminaire Sécurité & Cryptographie (avec l'équipe Monétique et Biométrie et le LMNO) ;
  - participation au groupe de travail « Mots, Séquences et Entropies » du LMNO.
- Organisation de colloques et rencontres internationales
  - *« Numbers, Sequences, Lattices : Dynamical Analysis of Algorithms »*, Colloquium for Brigitte Vallée's birthday, Caen, France, 3-4 juin, 2010.
  - *« Workshop on Coding and Cryptography »*, (WCC 2011) (`http://wcc2011.inria.fr`), Paris (France), avril 2011 (participation au comité d'organisation).
  - *« Philippe Flajolet and Analytic Combinatorics »*, Conference in the memory of Philippe Flajolet, Paris-Jussieu, 14-15-16 décembre 2011 (participation au comité d'organisation et au comité de programme).
  - *« EG'60 – Journées en l'honneur de 60 ans d'Etienne Grandjean »*, Caen, les 18 et 19 décembre 2013.
  - *« iSWAG Symposium »*, International Symposium on Web Algorithms, 2-3 juin 2015, Deauville.
- Organisation de colloques et rencontres nationales
  - Journées Frac d'automne 2010 (8ème édition) `https://frac.greyc.fr/frac/2010`.
  - Journées du groupe de tavail SDA2 (2011) `https://jsda2-2011.greyc.fr/`.
  - *« Journées 2013 du GT CMF (Complexité et Modèles Finis) du GdR IM »*, Caen, les 17 et 18 décembre 2013.
  - « L'internet IPv6 opérationnel ?» en 2013 `https://colloque-ipv6.greyc.fr`.
- Edition :
  - Journal of Computer and System Sciences : V. Terrier éditrice depuis fin 2007. Cette revue existe depuis 1967.

  - RAIRO : Brigitte Vallée éditrice sur la période 2011-2013.
- Comités de pilotage : STACS (E. Grandjean de 2007 à 2012), AofA (B. Vallée).
- PC Chair : ISWAG 2015 (S. Peyronnet)
- Participation à des Comités de programme (PC members)
  - AofA : International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (×3),
  - AUTOMATA : International Workshop on Cellular Automata and Discrete Complex Systems (×3),
  - STACS : Symposium on Theoretical Aspects of Computer Science (×1),
  - ANALCO : SIAM Analytic Algorithmics and Combinatorics (colloque satellite de SODA) (×3),
  - IJCAI : International Joint Conference on Artificial Intelligence (×1),
  - ISVC : International Symposium on Visual Computing (×3),
  - CSDM : Complex Systems Design & Management (×1),
  - SCC : International Conference on Symbolic Computation and Cryptography (×2),
  - AFRICACRYPT : International Conference on Cryptology in Africa (×3),
  - PKC : IACR International Conference on Practice and Theory of Public-Key Cryptography (×1),
  - WCC : Workshop on Coding and Cryptography (×1),
  - PROVSEC : International Conference on Provable Security (×1).
- GdR IM :
  - direction du GdR (B. Vallée, 2006-2013)
  - deux membres dans le comité de direction (J. Clément depuis 2012, B. Vallée depuis 2014).
  - responsabilité du site web et de la liste de diffusion (1200 abonnés) (J. Clément, depuis 2012).

3 Visibilité

- Jurys de thèses : participation à 18 thèses extérieures, dont 8 comme rapporteurs.

  École polytechnique (×2), Rennes (×1), Marseille (×1), Orléans (×1), Paris Est (×2), Paris 6 (×2), Cergy (×1), Lyon (×2), Paris 13 (×3), LMNO Caen (×1), Rouen (×1), Barcelone (×1).
- Jurys de HDR : participation à 8 HDR extérieures, dont 3 comme rapporteur.
  - Paris 7 (×2), Graz (×1), Orsay (×1), Lyon (×2), Paris Est (×1), Nancy (×1).
- Conférences invitées à l'international : 7.
  - AofA'11, 22th International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms, Będlewo, Pologne, juin 2011 (J. Clément) [106].
  - CLMPS 2011 : 14th Congress of Logic, Methodology and Philosophy of Science, Nancy, 19-26 juin 2011 (E. Grandjean).
  - LCC 2012 : 13th Workshop on Logic and Computational Complexity, Dubrovnik, Croatie, juin 2012 (E. Grandjean) [107].
  - SMC 2013 : First Workshop on Statistical Model Checking, Rennes, Septembre 2013 (S. Peyronnet).
  - CanaDAM 2013 : 4ième Congrès canadien de mathématiques discrètes et algorithmiques, St. John's, Terre-Neuve, juin 2013 (J. Clément).
- Conférences invitées ou colloquiums en France : 4.
  - Colloquium Jacques Morgenstern (INRIA Sophia-Antipolis), février 2010 (B. Vallée).
  - XXièmes Rencontres Arithmétiques de Caen, juin 2010 (F. Laguillaumie) [108].
  - Séminaire du pôle Algorithmes et Calcul (LIRMM), mars 2014 (J. Clément).
  - Journées Nationales du GdR IM, février 2014 (B. Vallée).
- Invitations à l'étranger : 5.
  - Université de Purdue, Mars 2010 (B. Vallée).
  - Dagstuhl seminar 10061 « Circuits, Logic, and Games » 2010, Allemagne (E. Grandjean).
  - Dagstuhl seminar 14041 "Quantitative Models : Expressiveness, Analysis, and New Applications" 2014, Allemagne, (S.Peyronnet).
  - Dagstuhl seminar 14111 "Combinatorics and Algorithmics of Strings" 2014, Allemagne (J. Clément).
  - Oberwolfach, Novembre 2014 (B. Vallée).
  - Japon (Universités de Keio et Tsukuba), Juillet 2015 (B. Vallée).
- Cours dans des Ecoles de chercheurs et tutoriaux : 10.
  - Semaine Réseaux Euclidiens du mois MathInfo du CIRM, Marseille, février 2010 (deux cours : A. Akhavi [105] et B. Vallée).

  - EJCIM 2010 (Ecole des jeunes chercheus en Informatique Mathématique), Chambéry, mars 2010 (B. Vallée) [111].
  - ADAMA 2012 : École d'automne en Analyse d'Algorithmes et Modèles Aléatoires, Mahdia, Tunisie, octobre 2012 (J. Clément).
  - "Dynamical analytic combinatorics and applications", Universidad Nacional de General Sarmiento, Argentina février 2014 (Trois cours : J. Clément, L. Lhote, B. Vallée).
  - EJCIM 2014 (Ecole Jeunes chercheurs en Informatique Mathématique), Caen, mars 2014 (trois cours pris en charge par l'équipe AmacC. Intervenants : J. Clément, E. Grandjean, T. Largillier, L. Lhote, S. Peyronnet, G. Richard, V. Terrier) [83, 85, 86].
  - "On the Declarative Structure of Quantum Concepts : States and Observables", 16th International Symposium on Principles and Practice of Declarative Programming, PPDP 2014, Canterbury, UK, September 8-10, 2014 (J. Karczmarczuk).
- Communications à destination des industriels : 4.
  - SMX Paris 2014, Search Marketing Expo (S. Peyronnet).
  - SEOCAMPUS 2012, 2013 et 2014, conférence annuelle de l'association professionnelle des référenceurs français (S. Peyronnet).

4 Participation à l'évaluation de structures ou de personnes

- Comités d'évaluation de l'AERES :
  - participation aux comités : LABRI, CEDRIC, LABSTIC (2010), LIRMM (2013)
  - présidence de comité : DI ENS (2012).
- Expertise pour le programme « Retour Post-Doctorants » de l'AERES.
- Comité PEDR : membre du comité.
- Membres extérieurs dans des comités de sélection : 14.
  Paris Créteil ($\times$2), Paris Nord ($\times$4), Nice ($\times$2), Montpellier ($\times$3), Paris 6 ($\times$1), Amiens ($\times$1), Lens ($\times$1).
- Membre nommée au CNU 27 : (B. Vallée, de 2007 à 2010).
- Membre du Jury de thèse Gilles Kahn : (B. Vallée, 2008-2010).
- Concours d'admission CR à l'INS2I du CNRS (B. Vallée, 2010).

5 Responsabilités administratives

- ANR (SIMI2 et comité sectoriel) : B. Vallée :
  - Vice-présidence du comité d'évaluation du programme SIMI2 de l'ANR Blanche, de 2010 à 2013.
  - Membre du Conseil Scientifique Sectoriel STIC de l'ANR, de 2009 à 2013.
  - Membre du comité d'évaluation de projets à mi-parcours (2014, 2015).
- Institut INS2I du CNRS : B. Vallée Directrice adjointe (DAS) mai 2013–Août 2014.
- Alliance Allistène : Animation du groupe 1 (B. Vallée, de 2010 à 2013).
- ACS Section 27 de l'université de Caen : A. Akhavi vice président (élu) depuis 2009.
- Co-responsabilité avec l'équipe Monétique & Biométrie du Master 2 E-Secure, cohabilité avec l'ENSICAEN : J. Saquet, F. Laguillaumie, G. Richard et E. Grandjean ont assumé cette tâche.

6 Interactions avec l'environnement social

B. Vallée est membre du Conseil National du Numérique (CNNUM) depuis janvier 2013, et nommée pour 3 ans. Il y a 30 membres en tout et B. Vallée a été proposée par le CNRS. Elle a participé à la rédaction de trois rapports : Inclusion (automne 2013) [113], Éducation (automne 2014) [119] et Ambition numérique (été 2015) [112].

7 Interaction avec l'enseignement de l'informatique au second degré

G. Richard et J. Saquet ont participé à la mise en place de la formation des enseignants du second degré, pour la spécialité ISN (informatique et sciences du numérique) en classe terminale (depuis

2011).

8 Logiciels

– Internet des objets et plateforme IoTa.
La plateforme IoTa s'est développée dans le cadre d'un projet FUI TACITES, entre 2011-2013, avec beaucoup de partenaires [1] ; elle permet le suivi d'objets équipés de puces RFID par leur lecture à distance (internet des objets). Chaque lecture donne lieu à la création d'un événement (départ usine, mise en palette, etc.) lié au numéro unique stocké dans la puce. L'ensemble des logiciels développés est accessible sous licence GPL sur le site : `https://forge.greyc.fr/projects/iota-pub/`.

– Plateforme ThemaMap.
ThemaMap est un outil de cartographie thématique multi plateforme, distribué en tant que logiciel libre (`https://themamap.greyc.fr/`). Il est développé en collaboration par le GREYC, le SAIC-CERTIC (Service d'activités industrielles et commerciales & Centre de ressources technologiques pour les TIC) et le CRH (Centre de Recherche Halieutique Méditerranéenne et Tropicale).

9 Synthèse

La synthèse est faite sur la période 1er janvier 2010 jusqu'à la rédaction du présent rapport.

Production scientifique

Publications	Nombre total (dont importantes)
Revues Internationales	28 (12)
Revues Nationales	1
Conférences Internationales	51 (12)
Chapitres d'ouvrage	8 (2)
Directions d'ouvrages	3
Thèses	8 (dont 1 cotutelle)
HDR	4
Brevets et logiciels	0+2
Autres	10

Contrats et projets	Nombre
Contrat industriel (prestation, accompagnement de thèse...)	3
Projets FUI	1
Projets ANR	1 (porteur) + 6 (partenaire) + 4 (indiv.)
Projets nationaux (autres)	1
Projets Européens	0
Projets internationaux (hors Europe)	3

Éléments de visibilité

	Nombre
Conférences invitées	7
Comité de rédaction et éditeurs invités	2
Comité de pilotage de conférences	2
Comité de programme de conférences	23
Organisation de conférences, d'écoles, d'ateliers	14
Evaluation et expertise	10
Responsabilités scientifiques	7
Jurys thèse et HDR hors équipe	26

10 Références bibliographiques

Revues Internationales

[1] Carlos Aguilar, Pierre-Louis Cayrel, Philippe Gaborit, and Fabien Laguillaumie. A New Efficient Threshold Ring Signature Scheme Based on Coding Theory. *IEEE Transactions on Information Theory*, pages 4833–4842, hal-01083807, 2011.

1. France Telecom Sophia, EADS, Cassidian, LIC, FIME, INVIA, ASK, France Telecom Caen

[2] Ali Akhavi, Ines Klimann, Sylvain Lombardy, Jean Mairesse, and Matthieu Picantin. On the finiteness problem for automaton (semi)groups. *International Journal of Algebra and Computation*, 22(06) :1250052.1–1250052.26, hal-00695445, Septembre 2012.

[3] Nuttapong Attrapadung, Javier Herranz, Fabien Laguillaumie, Benoît Libert, Elie De Panafieu, and Carla Ràfols. Attribute-Based Encryption Schemes with Constant-Size Ciphertexts. *Theoretical Computer Science*, 422 :15–38, hal-00763158, 2012.

[4] Frédérique Bassino, Julien Clément, and Pierre Nicodème. Counting occurrences for a finite set of words : combinatorial methods. *ACM Transactions on Algorithms*, 8 :31 :1–31 :28, Juillet 2012. hal-00452694, 28 pages.

[5] Frédérique Bassino, Julien Clément, Gadiel Seroussi, and Alfredo Viola. Optimal prefix codes for pairs of geometrically-distributed random variables. *Information Theory, IEEE Transactions on*, 59(4) :2375 – 2395, Avril 2013. hal-00511248, 38 pages.

[6] Valérie Berthé, Loïck Lhote, and Brigitte Vallée. A probabilistic analysis of the plain multiple gcd algorithm. *Journal of Symbolic Computation*, 2015. accepted to the special issue for the ISSAC'14 Conference.

[7] Valérie Berthé, Hitoshi Nakada, Rie Natsui, and Brigitte Vallée. Fine costs for Euclid's algorithm on polynomials and Farey maps. *Advances in Applied Mathematics*, 54 :27–65, hal-01086629, Mars 2014.

[8] Alexandre Borghi, Jérome Darbon, Sylvain Peyronnet, Tony F. Chan, and Stanley Osher. A Simple Compressive Sensing Algorithm for Parallel Many-Core Architectures. *Journal of Signal Processing Systems*, pages 1–20, hal-01086454, Avril 2013.

[9] Sébastien Canard, Julien Devigne, and Fabien Laguillaumie. Improving the Security of an Efficient Unidirectional Proxy Re-Encryption Scheme. *Jounal of Internet Services and Information Security*, pages 140–160, hal-01087041, 2011.

[10] Nicolas Carrasco, Jean-Marie Le Bars, and Alfredo Viola. Enumerative encoding of correlation-immune Boolean functions. *Journal of Theoretical Computer Science (TCS)*, 487 :23–36, hal-01086515, Mai 2013.

[11] Eda Cesaratto and Brigitte Vallée. Small quotients in Euclidean algorithms. *Ramanujan Journal (The)*, 24 :183 – 218, hal-01087902, 2011.

[12] Eda Cesaratto and Brigitte Vallée. Gaussian Distribution of Trie Depth for Strongly Tame Sources. *Combinatorics, Probability and Computing*, 24(01) :54–103, hal-01153052, Janvier 2015.

[13] Hubie Chen, Florent Madelaine, and Barnaby Martin. Quantified Constraints and Containment Problems. *Logical Methods in Computer Science*, 2015. Accepted, to appear.

[14] Gabriela Ciuperca, Valérie Girardin, and Loïck Lhote. Computation and Estimation of Generalized Entropy Rates for Denumerable Markov Chains. *IEEE Transactions on Information Theory*, 57 :4026 – 4034, hal-01082088, 2011.

[15] Julien Clément, James Allen Fill, Thu Hien Nguyen Thi, and Brigitte Vallée. Towards a realistic analysis of the QuickSelect algorithm. *ACM Transactions on Computer Systems*, pages 1–53, hal-01138894, 2015.

[16] Julien Clément, Thu Hien Nguyen Thi, and Brigitte Vallée. Towards a Realistic Analysis of Some Popular Sorting Algorithms. *Combinatorics, Probability and Computing*, 24(01) :104–144, hal-01103998, Janvier 2015.

[17] Alain Couvreur, Philippe Gaborit, Valérie Gautier, Ayoub Otmani, and Jean-Pierre Tillich. Distinguisher-Based Attacks on Public-Key Cryptosystems Using Reed-Solomon Codes. In *International Workshop on Coding and Cryptography - WCC 2013*, pages 181–193, Bergen, Norway, hal-00830594, Avril 2013.

[18] Julien David, Loïck Lhote, Arnaud Mary, and François Rioult. An average study of hypergraphs and their minimal transversals. *Journal of Theoretical Computer Science (TCS)*, 596 :124–141, hal-01086638, Septembre 2015.

[19] Jean-Charles Faugère, Valérie Gauthier-Umana, Ayoub Otmani, Ludovic Perret, and Jean-Pierre Tillich. A Distinguisher for High Rate McEliece Cryptosystems. *IEEE Transactions on Information Theory*, 59(10) :6830–6844, hal-00776068, Juin 2013.

[20] Etienne Grandjean and Frédéric Olive. A logical approach to locality in pictures languages. *Journal of Computer and System Sciences*, pages 1–67, 2015. hal-00786148, 67 pages.

[21] Javier Herranz, Fabien Laguillaumie, and Carla Ràfols. Relations between semantic security and anonymity in identity-based encryption. *Information Processing Letters*, 111 :453 – 460, hal-01084549, 2011.

[22] Fabien Laguillaumie and Damien Vergnaud. Time-Selective Convertible Undeniable Signatures with Compact Conversion Receipts. *Information Sciences*, pages 2458–2475, hal-01087046, 2010.

[23] Richard Lassaigne and Sylvain Peyronnet. Approximate planning and verification for large Markov decision processes. *International Journal on Software Tools for Technology Transfer*, pages 1–11, hal-01149843, 2014.

[24] Jean-Marie Le Bars. Equivalence classes of Boolean functions for first-order correlation. *IEEE Transactions on Information Theory*, pages 1247 – 1261, hal-01083545, Mars 2010.

[25] Loïck Lhote and Valérie Girardin. Rescaling Entropy and Divergence Rates. *IEEE Transactions on Information Theory*, pages 1–16, 2015. hal-01196817, To appear.

[26] Nicolas Ollinger and Gaétan Richard. Four states are enough ! *Journal of Theoretical Computer Science (TCS)*, 412(1-2) :22–32, hal-00469841, 2011.

[27] Ayoub Otmani, Jean-Pierre Tillich, and Léonard Dallot. Cryptanalysis of Two McEliece Cryptosystems Based on Quasi-Cyclic Codes. *Mathematics in Computer Science*, 3(2) :129–140, hal-01083566, 2010.

[28] Guillaume Quintin, Morgan Barbier, and Christophe Chabot. On Generalized Reed-Solomon Codes Over Commutative and Noncommutative Rings. *IEEE Transactions on Information Theory*, 59(9) :5882–5897, hal-00670004, Septembre 2013.

Revues Nationales

[29] Gilles Domalain, Céline Rodriguez, Jacques Madelaine, Christophe Turbout, and Carlota Estrella. THEMAMAP Un outil de cartographie thématique des séries statistiques. *Comité Français de Cartographie*, pages 33–48, 2011. hal-01068735, Article disponible en ligne : http ://www.lecfc.fr/new/articles/207-article-4.pdf.

Conférences Internationales

[30] Nicolas Bacquey. The packing problem : A divide and conquer algorithm on cellular automata. In Enrico FORMENTI, editor, *Automata & JAC 2012*, pages 1–10, Cargese, France, hal-00957117, Septembre 2012.

[31] Nicolas Bacquey. Complexity classes on spatially periodic Cellular Automata. In *STACS 2014*, pages 1–12, Lyon, France, hal-00957130, Mars 2014.

[32] Nicolas Bacquey. Primitive roots of bi-periodic infinite pictures. In *Words 2015*, Words 2015, Local Proceedings, Kiel, Germany, hal-01178256, Septembre 2015.

[33] Morgan Barbier, Clément Pernet, and Guillaume Quintin. On the decoding of quasi-BCH codes. In *WCC-International Workshop on Coding and Cryptography*, Bergen, Norway, hal-00768566, Avril 2013.

[34] Aurore Bernard and Nicolas Gama. Smallest Reduction Matrix of Binary Quadratic Forms. In Guillaume Hanrot, François Morain, and Emmanuel Thomé, editors, *9th International Symposium, ANTS-IX*, volume 6197 of *Algorithmic Number Theory*, pages 32–49, Nancy, France, Juillet 2010. hal-01083360, Lecture Notes in Computer Science.

[35] Valérie Berthé, Eda Cesaratto, Loïck Lhote, Pablo Rotondo, Brigitte Vallée, and Alfredo Viola. Recurrence Function on Sturmian Words : A Probabilistic Study. In *Mathematical Foundations of Computer Science 2015 – 40th International Symposium, MFCS 2015*, volume 9234 of *Lecture Notes in Computer Science*, pages 116–128, Milan, Italy, Août 2015. hal-01199096, Springer.

[36] Valérie Berthé, Jean Creusefond, Loïck Lhote, and Brigitte Vallée. Multiple GCD's. Probabilistic analysis of the plain algorithm. In *38th international symposium on International symposium on symbolic and algebraic computation, ISSAC'13*, Boston, United States, hal-01082245, Juin 2013.

[37] Alex Borello, Gaétan Richard, and Véronique Terrier. A speed-up of oblivious multi-head finite automata by cellular automata. In *Symposium on Theoretical Aspects of Computer Science (STACS2011)*, volume 9, pages 273–283, Dortmund, Germany, hal-00573599, Mars 2011.

[38] John Boxall, Nadia El Mrabet, Fabien Laguillaumie, and Duc-Phong Le. A Variant of Miller's Formula and Algorithm. In *Proceedings of Pairing 2010*, pages 417–434, Ishikawa, Japan, hal-01083368, Décembre 2010.

[39] Johann Brault-Baron. A Negative Conjunctive Query is Easy if and only if it is Beta-Acyclic. In Arnaud Durand Patrick Cegielski, editor, *Computer Science Logic (CSL'12) – 26th International Workshop/21st Annual Conference of the EACSL, CSL 2012, September 3-6, 2012, Fontainebleau, France*, volume 16 of *LIPIcs*, pages 137–151, Fontainebleau, France, hal-00786155, Septembre 2012.

[40] Sébastien Canard, Iwen Coisel, Julien Devigne, Cécilia Gallais, Thomas Peters, and Olivier Sanders. Toward generic method for server-aided cryptography. In *Information and Communications Security - 15th International Conference, ICICS 2013, Beijing, China, November 20-22, 2013. Proceedings*, pages 373–392, 2013.

[41] Sébastien Canard and Julien Devigne. Combined proxy re-encryption. In *Information Security and Cryptology - ICISC 2013 - 16th International Conference, Seoul, Korea, November 27-29, 2013, Revised Selected Papers*, pages 49–66, 2013.

[42] Sébastien Canard, Georg Fuchsbauer, Aline Gouget, and Fabien Laguillaumie. Plaintext-Checkable Encryption. In Orr Dunkelman, editor, *CT-RSA 2012*, volume 7178, pages 332–348, San Francisco, United States, hal-00768305, 2012.

[43] Sébastien Canard, Nicolas Desmoulins, Julien Devigne, and Jacques Traoré. On the implementation of a pairing-based cryptographic protocol in a constrained device. In Michel Abdalla and Tanja Lange, editors, *Pairing-Based Cryptography – Pairing 2012*, volume 7708 of *Lecture Notes in Computer Science*, pages 210–217. Springer Berlin Heidelberg, 2013.

[44] Sébastien Canard, Julien Devigne, and Olivier Sanders. Delegating a pairing can be both secure and efficient. In Ioana Boureanu, Philippe Owesarski, and Serge Vaudenay, editors, *Applied Cryptography and Network Security*, volume 8479 of *Lecture Notes in Computer Science*, pages 549–565. Springer International Publishing, 2014.

[45] Nicolas Carrasco, Jean-Marie Le Bars, and Alfredo Viola. Enumerative encoding of correlation- immune Boolean functions. In *IEEE Information TheoryWorkshop*, Paraty, Brazil, hal-01083891, Octobre 2011.

[46] Catarina Carvalho, Florent Madelaine, and Barnaby Daniel Martin. From complexity to algebra and back : digraph classes, collapsibility and the PGP. In *Logic in Computer Science (LICS)*, pages 462–474, Kyoto, Japan, Juillet 2015. hal-01108627, IEEE.

[47] Guilhem Castagnos and Fabien Laguillaumie. Homomorphic Encryption for Multiplications and Pairing Evaluation. In *Proceedings of SCN 2012*, pages 374–392, Amalfi, Italy, Septembre 2012. hal-01084946, Springer LNCS.

[48] Eda Cesaratto and Brigitte Vallée. Pseudo-randomness of a random Kronecker sequence. In *Latin American Symposium on Theoretical Informatics*, Lecture Notes in Computer Science, pages 157–171, Arequipa, Peru, hal-01084963, Avril 2012.

[49] Julien Clément and Laura Giambruno. On the Number of Prefix and Border Tables. In *LATIN 2014 : Theoretical Informatics – 11th Latin American Symposium*, volume 8392 of *Lecture Notes in Computer Science*, pages 442–453, Montevideo, Uruguay, Mars 2014. hal-00982224, Springer Berlin Heidelberg.

[50] Julien Clément, Thu Hien Nguyen Thi, and Brigitte Vallée. A general framework for the realistic analysis of sorting and searching algorithms. Application to some popular algorithms. In Natacha Portier and Thomas Wilke, editors, *30th International Symposium on Theoretical Aspects of Computer Science (STACS 2013)*, volume 20 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 598–609, Kiel, Germany, 2013. hal-00913309, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[51] Jean Creusefond, Thomas Largillier, and Sylvain Peyronnet. Finding compact communities in large graphs. In *SoMeRis : Social Media and Risk (SoMeRiS2015) workshop of The 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining ASONAM*, Paris, France, Août 2015. hal-01149820, 10 pages, 8 figures.

[52] Julien Devigne, Eleonora Guerrini, and Fabien Laguillaumie. Proxy Re-Encryption Scheme Supporting a Selection of Delegatees. In *7th International Conference on Cryptology AFRICACRYPT 2014*, LNCS, Marrakech, Morocco, Mai 2014. hal-00982549, Springer.

[53] Julien Devigne and Marc Joye. Binary huff curves. In Aggelos Kiayias, editor, *Topics in Cryptology – CT-RSA 2011*, volume 6558 of *Lecture Notes in Computer Science*, pages 340–355. Springer Berlin Heidelberg, 2011.

[54] Nadia El Mrabet. Fault attacks against the Miller algorithm in Edwards Coordinates. In *4th International Conference on Information Security and Assurance, ISA 2010*, pages 72–85, Miyazaki, Japan, hal-01083374, Juin 2010.

[55] Jean-Charles Faugère, Valérie Gauthier-Umana, Ayoub Otmani, Ludovic Perret, and Jean-Pierre Tillich. A distinguisher for high rate McEliece cryptosystems. In *ITW 2011- IEEE Information Theory Workshop*, pages 282–286, Paraty, Brazil, Octobre 2011. hal-01108602, IEEE.

[56] Jean-Charles Faugère, Ayoub Otmani, Ludovic Perret, and Jean-Pierre Tillich. Algebraic Cryptanalysis of McEliece Variants with Compact Keys. In Henri Gilbert, editor, *Eurocrypt 2010 : Proceedings of the 29th International Conference on Cryptology*, volume 6110 of *LNCS – Lecture Notes in Computer Science*, pages 279–298, Monaco, Monaco, Mai 2010. inria-00596632, Springer Verlag.

[57] Philippe Flajolet, Mathieu Roux, and Brigitte Vallée. Digital Trees and Memoryless Sources : from Arithmetics to Analysis. In *DMTCS,Proceedings of AofA'10, DMTCS, proc AM*, pages 231–258, Vienne, Austria, hal-01083405, Juin 2010.

[58] Nicolas Gama, Phong Q. Nguyen, and Oded Regev. Lattice Enumeration Using Extreme Pruning. In *Proceedings of the 29th International Conference on Cryptology – EUROCRYPT 2010*, pages 257–278, Nice, France, hal-01083526, Mai 2010.

[59] Laura Giambruno, Sabrina Mantaci, Jean Néraud, and Carla Selmi. A generalization of Girod's bidirectional decoding method to codes with a finite deciphering delay. In Springer, editor, *Developments in Language Theory 2012*, volume 7410 of *Proceedings. Springer 2012*, pages 471–476, Tapei, Taiwan, hal-00913504, Août 2012.

[60] Anael Grandjean, Gaétan Richard, and Véronique Terrier. Linear functional classes over cellular automata. In Enrico Formenti, editor, *18th international workshop on Cellular Automata and Discrete Complex Systems and 3rd international symposium Journées Automates Cellulaires*, volume 90 of *Electronic Proceedings in Theoretical Computer Science Proceedings*, pages 177–193, Bastia, France, Septembre 2012. hal-01085911, Open Publishing Association.

[61] Etienne Grandjean. Computational Complexity and Logical Definability. In *14th Congress of Logic, Methodology and Philosophy of Science*, Nancy, France, hal-01084455, Juin 2011.

[62] Etienne Grandjean and Frédéric Olive. Descriptive complexity for pictures languages. In Patrick Cégielski and Arnaud Durand, editors, *Computer Science Logic (CSL'12) – 26th International Workshop/21st Annual Conference of the EACSL*, Leibniz International Proceedings in Informatics (LIPIcs), pages 274–288, Fontainebleau, France, Septembre 2012. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[63] Pierre Guillon and Gaétan Richard. Revisiting the Rice Theorem of Cellular Automata. In Jean-Yves Marion and Thomas Schwentick, editors, *27th International Symposium on Theoretical Aspects of Computer Science – STACS 2010*, Proceedings of the 27th Annual Symposium on the Theoretical Aspects of Computer Science, pages 441–452, Nancy, France, Mars 2010. inria-00455736, Inria Nancy Grand Est & Loria.

[64] Javier Herranz, Fabien Laguillaumie, Benoît Libert, and Carla Ràfols. Short Attribute-Based Signatures for Threshold Predicates. In *RSA Conference 2012*, volume 7178 of *Lecture Notes in Computer Science*, pages 51–67, San Francisco, United States, 2012. hal-00611651, Springer.

[65] Javier Herranz, Fabien Laguillaumie, and Carla Ràfols. Constant Size Ciphertexts in Threshold Attribute-Based Encryption. In *13th International Conference on Practice and Theory in Public Key Cryptography 2010*, volume 6056 of *Lecture Notes in Computer Science*, pages 19–34, Paris, France, hal-01083535, Mai 2010.

[66] Kanal Hun and Brigitte Vallée. Typical Depth of a Digital Search Tree built on a general source. In *Proceedings of ANALCO'2014, SIAM Meeting on Analytic Algorithmics and Combinatoric*, pages 1–15, Portland, United States, hal-01087072, Janvier 2014.

[67] Krystian Jobczyk, Maroua Bouzid, Antoni Ligeza, and Jerzy Karczmarczuk. Fuzzy Logic for Preferences expressible by convolutions. In *ECAI 2014 – 21st European Conference on Artificial Intelligence, 18-22 August 2014, Prague, Czech Republic*, Frontiers in Artificial Intelligence and Applications, pages 1041–1042, Prague, Czech Republic, hal-01139634, Août 2014.

[68] Krystian Jobczyk, Maroua Bouzid, Antoni Ligeza, and Jerzy Karczmarczuk. Temporal verbs and adverbs : 'often' and 'many times' and their fuzzy-integral-logic based modelling. In *Logic and Engineering of Natural Language Semantics 11 (LENLS11)*, Tokyo, Japan, hal-01139637, Novembre 2014.

[69] Jerzy Karczmarczuk. Specific "scientific" data structures, and their processing. In *IFIP Working Conference on Domain-Specific Languages DSL 2011*, IFIP Working Conference on Domain-Specific Languages DSL 2011, Bordeaux, France, Septembre 2011. hal-01139633, In Proceedings DSL 2011, arXiv :1109.0323.

[70] Thomas Largillier. Using Subjective Logic to Divide Learners into Groups. In *International Symposium on Web AlGorithms*, Deauville, France, hal-01171300, Juin 2015.

[71] Thomas Largillier, Guillaume Peyronnet, and Sylvain Peyronnet. Rocovo : Robust Communal Publication Scheme. In *Proceedings of the 29th Annual ACM Symposium on Applied Computing ACM*, pages 579–586, Helsinki, Finland, hal-01081279, Juin 2014.

[72] Boris Lesner, Romain Brixtel, Cyril Bazin, and Guillaume Bagan. A Novel Framework to Detect Source Code Plagiarism : Now, Students Have to Work for Real ! In *SAC '10 Proceedings of the 2010 ACM Symposium on Applied Computing*, pages 57–58, Sierre, Switzerland, hal-01067161, Mars 2010.

[73] Loïck Lhote and Manuel E. Lladser. Toward the asymptotic count of bi-modular hidden patterns under probabilistic dynamical sources : a case study. In *DMTCS Proceedings, 23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12)*, Montréal, Canada, hal-01082042, Juin 2012.

[74] Jacques Madelaine, Gaétan Richard, and Gilles Domalain. ThemaMap : a Free Versatile Data Analysis and Visualization Tool. In *International Symposium on Web AlGorithms*, Deauville, France, hal-01171339, Juin 2015.

[75] Manfred Madritsch and Brigitte Vallée. Modelling the LLL Algorithm by Sandpiles. In *Comptes Rendus de la Conférence LATIN 2010*, pages 267–281, Oaxaca, Mexico, hal-01082028, Avril 2010.

[76] Ayoub Otmani and Jean-Pierre Tillich. An Efficient Attack on All Concrete KKS Proposals. In *Post-Quantum Cryptography – PQCrypto 2011*, volume 7071 of *Lecture Notes in Computer Science*, pages 98–116, Taipei, Taiwan, Novembre 2011. hal-00913500, Springer.

[77] Sylvain Peyronnet, Michel De Rougemont, and Yann Strozecki. Approximate Verification and Enumeration Problems. In Springer, editor, *9th International Colloquium on Theoretical Aspects of Computing–ICTAC 2012*, volume 7521 of *Lecture Notes in Computer Science*, pages 228–242, Bangalore, India, Septembre 2012. hal-01149724, Abhik Roychoudhury, Meenakshi D'Souza.

[78] Mathieu Roux and Brigitte Vallée. Information theory : Sources, Dirichlet series, and realistic analyses of data structures. In *Electronic Proceedings of Theoretical Computer Science„Proceedings 8th International Conference Words 2011*, volume 63, pages 199–214, Prague, Czech Republic, Septembre 2011. hal-01082040, EPTCS.

[79] Véronique Terrier. Recognition of linear-slender context-free languages by real time one-way cellular automata. In J. Kari, editor, *AUTOMATA 2015*, volume 9099 of *Lecture Notes in Computer Science*, pages 251–262, Turku, Finland, hal-01149358, Juin 2015.

[80] Brigitte Vallée. The Euclid Algorithm is totally gaussian. In *DMTCS Proceedings, 23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA'12)*, pages 283–302, Montréal, Canada, hal-01086476, Juin 2012.

Ouvrages

[81] Jerzy Karczmarczuk and Helena Kucal. *Wstęp do informatyki kwantowej*. PWN, 2011. hal-01139663, Traduction « Introduction à l'Informatique Quantique » de Michel Le Bellac.

Chapitres d'ouvrages

[82] Eda Cesaratto and Brigitte Vallée. Pseudo-randomness of a random Kronecker sequence. An instance of dynamical analysis. In Michel Rigo. Valérie Berthé, editor, *Combinatorics, Words and Symbolic Dynamics'*, pages 1–40. hal-01153048, 2015.

[83] Julien Clément and Loïck Lhote. Analyse d'algorithmes : calculs de PGCD et algorithmes de tri et de recherche. In *Informatique Mathématique : Une photographie en 2014*. hal-01087191, Mars 2014.

[84] Julien Clément and Mark Daniel Ward. The digital tree process. In Wojciech Szpankowski, editor, *Philippe Flajolet's Collected Papers*, volume III, chapter 3. Cambridge Press, 2016. Volume on Text, Information Theory, and the Mellin Transform). To appear.

[85] Etienne Grandjean, Véronique Terrier, and Gaétan Richard. Automates cellulaires. In *Informatique Mathématique : Une photographie en 2014*. hal-01087506, Mars 2014.

[86] Thomas Largillier and Sylvain Peyronnet. Algorithmique du web : autour du pagerank. In *Informatique Mathématique Une photographie en 2014*. Éditeur Presses Universitaires de Perpignan, hal-01147179, 2014.

[87] Véronique Terrier. Language Recognition by Cellular Automata. In *Handbook of Natural Computing*. Springer Berlin Heidelberg, hal-01086432, 2012.

[88] Brigitte Vallée. On dynamical systems. In Brigitte Vallée, editor, *Philippe Flajolet's Collected Papers*, volume II, chapter 3. Cambridge Press, 2016. Volume on Limit Laws and Dynamical Systems. To appear.

[89] Brigitte Vallée and Antonio Vera. Probabilistic behaviour of lattice reduction algorithms. In *The LLL Algorithm*. hal-01083878, 2010.

Direction d'ouvrages

[90] Phong Q. Nguyen and Brigitte Vallée. *The LLL Algorithm : Survey and Applications*. Information Security and Cryptography. Springer, hal-01141414, 2010.

[91] Sylvain Peyronnet. *Informatique Mathématique : une photographie en 2014*. hal-01146360, 2014.

[92] Brigitte Vallée. *Limit Laws and Dynamical Systems*, volume II. Cambridge Press, 2016. Volume of Philippe Flajolet's Collected Papers. To appear.

Thèses

[93] Cyril Bazin. *Tattoo of geographic data and generalization data to preserve constraints*. Thèse, Université de Caen, tel-01075247, Janvier 2010.

[94] Johann Brault-Baron. *The relevance of the list : propositional logic and complexity of the first order*. Thèse, Université de Caen, tel-01081392, Avril 2013.

[95] Léonard Dallot. *Security of cryptographic protocols based on error correcting codes*. Thèse, Université de Caen, tel-01102440, Juillet 2010.

[96] Julien Devigne. *Re-encryption protocols for the storage of data*. Thèse, Université de Caen, tel-01081377, Décembre 2013.

[97] Mariya Georgieva. *Probabilistic analysis of reduced cryptographic Euclidean networks*. Thèse, Université de Caen, tel-01081679, Décembre 2013.

[98] Kanal Hun. *Analysis of depth of digital trees built on general sources*. Thèse, Université de Caen Basse-Normandie, tel-01136777, Décembre 2014.

[99] Thu Hien Nguyen Thi. *Towards a realistic analysis of sorting and searching algorithms*. Thèse, Université de Caen Basse-Normandie, tel-01134104, Décembre 2014.

[100] Mathieu Roux. *Information theory, Dirichlet series, and analysis of algorithms*. Thèse, Université de caen, tel-01076421, Décembre 2011.

HDR

[101] Julien Clément. *Algorithms, words and random texts*. Habilitation à diriger des recherches, Université de Caen, tel-00913127, Décembre 2011.

[102] Fabien Laguillaumie. *Public-Key Cryptography : Design and Algorithmic*. Habilitation à diriger des recherches, Université de Caen, tel-01083946, Décembre 2011.

[103] Ayoub Otmani. *Contribution to the Cryptanalysis of Code-Based Primitives*. Habilitation à diriger des recherches, Université de Caen Basse Normandie, tel-01138792, Décembre 2011.

[104] Véronique Terrier. *Reconnaissance de langages par automates cellulaires*. Habilitation à diriger des recherches, Université de Caen, tel-01070916, Avril 2011.

Conférences Invitées

[105] Ali Akhavi. Introduction to lattice reduction and to the LLL algorithm. Analysis of the LLL algorithm in the uniform model of the unit ball. In *Math-Info 2010 : Lattice Algorithmics Towards new interactions between mathematics and computer science*, Marseille, France, hal-01083356, Juillet 2010.

[106] Julien Clément. Words occurrences in random texts. In *AofA'11, 22th International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms,*, Będlewo, Poland, hal-01084082, 2011.

[107] Etienne Grandjean. An invitation to Linear Time. In *Logic and Computational Complexity : An International Workshop Series (LCC 12)*, Dubrovnik, Croatia, hal-01087904, Juin 2012.

[108] Fabien Laguillaumie. Factoring pq^2 with quadratic forms and cryptographic applications. In *XXIèmes Rencontres Arithmétiques de Caen*, CAEN, France, hal-01083571, 2010.

[109] Ayoub Otmani. On the Design of Code-Based Signatures. In *Code-based Cryptography Workshop - CBC 2012*, Lyngby, Denmark, hal-00921494, Mai 2012.

[110] Brigitte Vallée. Probabilistic Analysis of lattice reduction algorithms in small dimension. Sandpile model for the LLL algorithm. In *Math-Info 2010 : Lattice Algorithmics Towards new interactions between mathematics and computer science*, Marseille, France, hal-01150557, Juillet 2010.

[111] Brigitte Vallée. Analyse réaliste d'algorithmes de tri et de recherche. In *Ecole des Jeunes Chercheurs en Informatique Mathématique*, Chambéry, France, hal-01083802, 2010.

Autres (poster, autre publication, pré-publication, document de travail, rapport, cours...)

[112] Serge Abiteboul, Nathalie Andrieux, Christine Balagué, Godefroy Beauvallet, Ludovic Blecher, Nathalie Bloch-Pujo, Michel Briand, Virginia Cruz, Pascal Daloz, Marylène Delbourg-Delphis, Stéphane Distinguin, Marie Ekeland, Virginie Fauvel, Cyril Garcia, Audrey Harris, Francis Jutand, Daniel Kaplan, Florence Le Ny, Tristan Nitot, Sophie Pène, Valérie Peugeot, Lara Rouyrès, Jean-Baptiste Rudelle, Cécile Russeil, Bernard Stiegler, Benoît Thieulin, Marc Tessier, and Brigitte Vallée. Ambition numérique, Juin 2015. hal-01180275, Rapport remis au premier ministre.

[113] Serge Abiteboul, Nathalie Andrieux, Christine Balagué, Michel Briand, Cyril Garcia, Audrey Harris, Daniel Kaplan, Florence Le Ny, Sophie Pène, Valérie Peugeot, Benoît Thieulin, and Brigitte Vallée. Citoyens d'une société numérique. Accès, Littératie, Médiations, Pouvoir d'agir, Pour une nouvelle politique d'inclusion, 2013. hal-01144022, Rapport remis à la ministre déléguée chargée des Petites et Moyennes Entreprises, de l'Innovation et de l'Économie numérique.

[114] Ali Akhavi. Generalized sorting problems : a symbolic view on reversible algorithms. Soumis, hal-01196063, Juillet 2015.

[115] Nicolas Bacquey. Leader election on two-dimensional periodic cellular. Soumis, hal-01178250, Mars 2015.

[116] Brigitte Chauvin, Bruno Salvy, Michèle Soria, and Brigitte Vallée. Philippe Flajolet, le fondateur de la combinatoire analytique. *Gazette des Mathématiciens*, 129 :113–114, hal-00680928, Juillet 2011.

[117] Jean Creusefond. A comparison of graph clustering algorithms. In *International Symposium on Web AlGorithms*, Deauville, France, Juin 2015. hal-01171341, Poster.

[118] Jerzy Karczmarczuk. On the Declarative Structure of Quantum Concepts : States and Observables, Septembre 2014. hal-01139638, Tutorial in Principles and Practice of Declarative Programming, (PPDP 2014). Canterbury, Sept. 8 - 10, 2014.

[119] Sophie Pène, Serge Abiteboul, Christine Balagué, Ludovic Blecher, Nathalie Bloch-Pujo, Michel Briand, Cyril Garcia, Francis Jutand, Daniel Kaplan, Pascale Luciani-Boyer, Valérie Peugeot, Bernard Stiegler, and Brigitte Vallée. Jules Ferry 3.0, Bâtir une école créative et juste dans un monde numérique., 2014. hal-01144070, Rapport du conseil national du numérique.

[120] Véronique Terrier. Linear acceleration for one-dimensional cellular automata. In *19th International Workshop on Cellular Automata and Discrete Complex Systems*, pages 97–106, Giessen, Germany, hal-01086617, Septembre 2013.

[121] Jean Vuillemin and Nicolas Gama. Efficient Equivalence and Minimization for Non Deterministic Xor Automata. Research report, inria-00487031, Mai 2010.