PROJET DE THÈSE (2022-2025)
Analyse et vérification de réseaux paramétrés (avec broadcast)

Encadrants de la thèse

— **Christine Tasson**, Professeure d'Informatique à Sorbonne Université, experte en sémantique des applications distribuées.
— **Nathalie Sznajder**, Maîtresse de Conférences en Informatique à Sorbonne Université, experte en théorie des jeux et systèmes paramétrés.
— **Arnaud Sangnier**, Maître de Conférences en Informatique à Université Paris Cité, expert en vérification de systèmes infinis et en réseaux paramétrés.

Contexte

Contexte sociétal et industriel

Les systèmes informatiques modernes sont désormais dans leur grande majorité des systèmes distribués, mettant en jeu des interactions entre des processus géographiquement distincts (que ce soit pour les domaines du commerce électronique, des transports, du grid computing, du cloud computing, des réseaux de robots mobiles, des transports, des réseaux ad-hoc, etc.) selon des architectures fixes ou mobiles. Lorsqu'il s'agit de systèmes critiques, en termes humains ou financiers, il est essentiel de s'assurer de leur fiabilité. Or, l'étude et l'analyse des systèmes distribués soulève des difficultés spécifiques : communications synchrones ou asynchrones, données partagées, données non bornées, nombre de processus non borné, graphes de communications arbitraires, etc.
Les *méthodes formelles* se sont avérées très utiles pour analyser et vérifier des systèmes complexes. Elles comprennent des formalismes de spécification adaptés, des modèles expressifs basés sur les automates (communicants, avec données, paramétrés), des algorithmes pour vérifier si un modèle satisfait sa spécification et des techniques approchées lorsque le problème est indécidable ou trop complexe à résoudre.

Des réseaux avec un nombre non-borné de participants

Nous nous intéresserons dans cette thèse à des systèmes distribués correspondant à des réseaux d'entités communicantes dont le nombre n'est pas fixé a priori et est donc considéré comme un paramètre. En effet, de nombreux algorithmes distribués sont conçus pour fonctionner avec un nombre arbitraire de processus, et les techniques habituelles de vérification automatique ne peuvent s'adapter immédiatement à de tels systèmes. Ils peuvent aussi être utilisés pour représenter des systèmes biologiques comme le modèle proposé dans [BDG+19] où toutes les entités du réseau réagissent à une information commune. Finalement les techniques développées pour leur analyse peuvent être utilisées pour l'amélioration de systèmes industriels dans lesquels les tâches sont assimilées aux entités du réseau et les synchronisations entre tâches sont simulées par les communications ; l'utilisation de méthodes formelles dans ce contexte permet ainsi de mieux comprendre les systèmes, mais aussi d'améliorer leur organisation en fournissant par exemple des méthodes pour planifier les tâches (ce qui correspondrait à la synthèse d'un ordonnanceur paramétré dans le cadre du réseau).

Dans ces réseaux, on suppose que toutes les entités exécutent le même protocole (c'est-à-dire le même programme). Selon le modèle utilisé pour la description du protocole (par exemple, un automate fini, un système à pile ou un automate équipé de choix probabilistes pour representer l'incertitude), ou les moyens de communication entre les entités (passage de message, synchronisation par rendez-vous, broadcast de messages), l'expressivité du modèle change et les résultats sur la vérification de tels systèmes varient (voir par exemple [Esp14] pour une présentation détaillée des différences entre ces modèles).

Les réseaux paramétrés avec broadcast sélectifs

Un cas particulier de tels réseaux paramétrés a été introduit dans [DSZ10], où la communication entre entités se fait par *broadcast sélectif* : seules les entités voisines de l'émetteur peuvent recevoir les messages émis. Ce modèle est particulièrement adapté pour modéliser le comportement de réseaux ad hoc, dans lesquels les entités communiquent par transmission radio et donc seules les entités présentes dans le rayon de transmission de l'émetteur peuvent recevoir le message. Pour ces systèmes, la question de savoir s'il existe un nombre d'agents et une topologie initiale du réseau à partir de laquelle un des processus atteint un état d'erreur n'est décidable que si l'on restreint l'ensemble des topologies de communication possibles. Par contre, si la topologie du réseau peut changer au cours du temps, par exemple car les agents se déplacent dans l'espace, modifiant la portée des ondes radios, ou car on modélise des pertes de messages, de nombreux problèmes de vérification deviennent décidables [DSTZ12, BFS14].

Travaux prévus dans la thèse

Modélisation de la mobilité dans les réseaux paramétrés avec broadcast sélectif

Comme nous l'avons vu précédemment, pour les réseaux avec broadcast sélectif, il a été proposé jusqu'à présent deux hypothèses sur l'évolution de la topologie de communication lors des exécutions, soit elle reste fixe, soit elle peut changer de manière totalement non déterministe. Cette dernière modélisation permet certes d'obtenir des algorithmes de vérification avec de bonnes bornes de complexité (algorithmes en temps polynomial ou en temps non-déterministe polynomial), cependant il s'agit d'une hypothèse qui peut se révéler un peu trop forte en pratique. En effet, en ne faisant aucune hypothèse sur l'évolution du réseau, de très nombreux protocoles vont s'avérer avoir des exécutions erronées, et le risque est de ne certifier "correct" que des systèmes extrêmement simples, voire triviaux. Il peut être cependant plus réaliste de vérifier des systèmes *sous certaines hypothèses* d'évolution du réseau : par un exemple le taux de panne peut être borné (si on modélise la perte de messages), ou la mobilité peut être restreinte par la vitesse de déplacement des entités, engendrant des restrictions sur l'évolution du graphe de communication au cours du temps.
Le premier travail de cette thèse consistera ainsi à proposer de nouvelles manières de modéliser la mobilité dans ces réseaux. Parmi les différentes solutions pour attaquer ce problème de modélisation, il serait possible par exemple d'imposer des règles à chaque étape décrivant la façon dont la topologie peut évoluer, cela pourrait être fait à l'aide d'un système de réécriture de graphes. Une autre solution pourrait être de considérer que la façon dont la topologie de communication évolue est décidée par un environnement incontrôlable. Ceci reviendrait naturellement à modéliser le problème sous la forme d'un jeu à deux joueurs, et restreindre

l'évolution de la topologie du réseau correspondrait à restreindre les actions possibles du joueur modélisant l'environnement. Plusieurs formalismes de jeux à deux joueurs permettant de résoudre des problèmes sur les systèmes paramétrés ont été introduits ces dernières années [BFS14, BLS18] et nous nous y référerons pour développer le modèle le plus pertinent.
Il serait aussi intéressant de pouvoir équiper ces modèles avec une notion de distance et de temps afin de pouvoir modéliser le fait qu'un noeud a besoin d'un certain temps avant d'être hors de portée du rayon transmission d'un autre noeud si auparavant il recevait les messages de ce noeud. Le but ici sera donc d'étudier différentes façons de modéliser cette mobilité et d'analyser l'impact de ces modélisations sur la vérification des systèmes et le cas échéant de proposer des algorithmes efficaces.

Généralisation à des spécifications plus riches

Forts de cette modélisation plus fine de nos problèmes sous différentes formes de jeux à deux joueurs, nous essaierons ensuite d'enrichir le type de propriétés qu'il est possible de vérifier sur ces réseaux mobiles, au delà des questions d'accessibilité d'états de contrôle. Tout d'abord nous essaierons d'intégrer des notions quantitatives dans l'analyse de ces réseaux, permettant de modéliser par exemple la consommation d'énergie provoquée par l'émission d'un message (la consommation d'énergie étant un élément critique des réseaux ad hoc). Nous utiliserons la théorie des jeux afin de rechercher s'il existe des stratégies minimisant la consommation d'énergie tout en maintenant l'objectif à atteindre. Si ces notions de jeux quantitatifs ont été largement étudiés ces dernières années dans le cadre des systèmes finis, nous souhaiterions adapter ces raisonnements au cas des systèmes paramétrés.
D'autre part, dans le cadre de systèmes finis, il est connu qu'il existe une forte relation entre les jeux de parité et le problème de model-checking du μ-calcul (voir par exemple [GTW02]) mais dans le cas de systèmes avec un nombre infinis d'états (dont font partie les réseaux paramétrés pour lesquels le nombre de participants n'est pas borné) une telle relation n'existe pas. Nous souhaitons donc définir un langage logique adapté à la vérification de tels systèmes, qui permettrait de définir des propriétés plus fines que l'on saurait vérifier à l'aide de ce formalisme de jeux.

Expressivité des réseaux paramétrés avec broadcast

Enfin, nous aborderons l'étude des réseaux paramétrés avec broadcast sélectif sous un autre angle : celui de leur expressivité. Nous nous appuierons pour cela sur la connaissance approfondie de ces réseaux acquise durant les premiers travaux de la thèse, ainsi que sur les travaux réalisés sur les protocoles de population. En effet, ceux-ci représentent une autre famille de réseaux paramétrés, introduite dans [AAD+04]. Dans ces réseaux, les entités communiquent par rendez-vous synchrone (deux entités qui se rencontrent peuvent changer d'état selon les règles fixées par le protocole), et il a été montré dans [AAER07] qu'ils permettaient de reconnaître les ensembles semi-linéaires : étant donné un ensemble semi-linéaire sur un ensemble d'étiquettes, il est possible de concevoir un protocole qui amène tous les noeuds dans un état correct si et seulement si les étiquettes des noeuds initiaux appartiennent à l'ensemble semi-linéaire de départ. On s'efforcera dans cette partie d'étudier le pouvoir d'expression des réseaux de broadcast sélectifs en terme de reconnaissance de graphes. Il s'agit de la partie la plus ambitieuse car s'il existe de nombreux formalismes bien établis pour décrire

des mots ou des arbres (en terme d'automates ou de logique), la situation est moins claire pour les graphes et pose de réelles difficultés. Il s'agit cependant d'une piste de recherche extrêmement intéressante aussi bien d'un point de vue théorique qu'en terme d'applications au niveau de la conception de réseaux d'entités mobiles communicant par transmission radio.

Bibliographie

[AAD+04] Dana Angluin, James Aspnes, Zoë Diamadi, Michael J. Fischer, and René Peralta. Computation in networks of passively mobile finite-state sensors. In *Proceedings of the Twenty-Third Annual ACM Symposium on Principles of Distributed Computing, PODC'04*, pages 290–299. ACM, 2004.

[AAER07] Dana Angluin, James Aspnes, David Eisenstat, and Eric Ruppert. The computational power of population protocols. *Distributed Comput.*, 20(4) :279–304, 2007.

[BDG+19] Nathalie Bertrand, Miheer Dewaskar, Blaise Genest, Hugo Gimbert, and Adwait Amit Godbole. Controlling a population. *Log. Methods Comput. Sci.*, 15(3), 2019.

[BFS14] Nathalie Bertrand, Paulin Fournier, and Arnaud Sangnier. Playing with probabilities in reconfigurable broadcast networks. In *Proceedings of the 17th International Conference on Foundations of Software Science and Computation Structures - FOSSACS'14, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS'14*, volume 8412 of *Lecture Notes in Computer Science*, pages 134–148. Springer, 2014.

[BLS18] Benedikt Bollig, Mathieu Lehaut, and Nathalie Sznajder. Round-bounded control of parameterized systems. In *16th International Symposium on Automated Technology for Verification and Analysis, Proceedings of ATVA'18*, volume 11138 of *Lecture Notes in Computer Science*, pages 370–386. Springer, 2018.

[DSTZ12] Giorgio Delzanno, Arnaud Sangnier, Riccardo Traverso, and Gianluigi Zavattaro. On the complexity of parameterized reachability in reconfigurable broadcast networks. In *Proceedings of the IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science - FSTTCS'12*, volume 18 of *LIPIcs*, pages 289–300. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2012.

[DSZ10] Giorgio Delzanno, Arnaud Sangnier, and Gianluigi Zavattaro. Parameterized verification of ad hoc networks. In *Proceedings of the 21th International Conference on Concurrency Theory - CONCUR'10*, volume 6269 of *Lecture Notes in Computer Science*, pages 313–327. Springer, 2010.

[Esp14] Javier Esparza. Keeping a crowd safe : On the complexity of parameterized verification (invited talk). In Ernst W. Mayr and Natacha Portier, editors, *31st International Symposium on Theoretical Aspects of Computer Science (STACS'14)*, volume 25 of *LIPIcs*, pages 1–10. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2014.

[GMT18] Éric Goubault, Samuel Mimram, and Christine Tasson. Geometric and combinatorial views on asynchronous computability. *Distributed Comput.*, 31(4) :289–316, 2018.

[GTW02] Erich Grädel, Wolfgang Thomas, and Thomas Wilke, editors. *Automata, Logics, and Infinite Games : A Guide to Cu rrent Research [outcome of a Dagstuhl seminar, February 2001]*, volume 2500 of *Lecture Notes in Computer Science*. Springer, 2002.