

Réactivité, efficacité et apprentissage : problèmes daide à la prise de décision dans des réseaux décentralisés

Juan-Antonio Cordero-Fuertes¹

LIX CNRS, École Polytechnique, Institut Polytechnique de Paris, France

Mots-clés : *communication, intelligence dans les réseaux, routage dynamique, IOT, apprentissage, décentralisation*

1 Introduction

Dans le contexte des réseaux de communication et de l'information et de l'intelligence dans la prise de décision distribuée peut se traduire en termes efficacité pour exploiter les conditions de communication existantes, d'apprentissage pour détecter son variation, et de réactivité pour adapter, sans instruction préalable, son comportement aux conditions observées.

Le développement de l'Internet et des technologies de communication, notamment (mais pas seulement) les technologies de communications sans fils, les progrès dans la miniaturisation et l'augmentation des capacités de calcul, ainsi que l'apparition de nouvelles demandes et nouveaux services qui exploitent ces nouvelles capacités, ont favorisé l'émergence pendant les dernières décennies de réseaux de communication plus hétérogènes, plus complexes et souvent soumises à des contraintes et à des conditions d'opération qui ne correspondent plus à celles des réseaux traditionnels. De plus en plus, la configuration manuelle complète de ces réseaux devient inefficace, trop couteuse et peu pratique, dû à l'extrême variété (statique) et variabilité (dans le temps) des conditions d'opération. Les démarches automatiques ou semi-automatiques (c'est-à-dire, où une définition de politiques ou d'objectifs abstraits de la partie d'opérateurs humains est traduite , implémentée et maintenue en temps réel par des algorithmes adaptatifs autonomes) deviennent donc nécessaires pour gérer avec efficacité une complexité grandissante.

Ces réseaux sont souvent déployés dans la périphérie (edge) de l'Internet. C'est le cas, par exemple, de ce qu'on appelle l'Internet des Objets (Internet of Things, IOT), des réseaux de capteurs et de services distribués dans des contextes et avec des ambitions très variés (bâtiment intelligent , véhicules connectés , ville intelligent , Smart Grid , etc.). Ces réseaux relient souvent des dispositifs très limités en termes de capacités de calcul, communication et mémoire, et qui reposent sur des technologies nouvelles de communication, sans fils ou autres (e.g. par courants porteurs de lignes (CPL)). Dans le contexte des communications mobiles (cellulaires), on peut considérer aussi les infrastructures d'accès des opérateurs, qui relient des stations base entre elles et avec l'Internet. D'autres réseaux périphériques sur Internet (c'est-à-dire, source ou destination de trafic de données) sont des réseaux filaires dans le sens traditionnel, dédiés à la provision de services ou au traitement des informations, comme les grands centres de données (datacenters) qui servent d'infrastructure pour les applications d'IOT et pour d'autres services aux usagers (Big Data et apprentissage automatique, les services de nuage (cloud) et de calcul distribué, distribution de contenus, etc.), ou les réseaux d'entreprise, de plus en plus sophistiquées, intensivement utilisées, et aussi plus vulnérables à des failles, des attaques ou des dysfonctionnements qui ont besoin d'être gérés de plus en plus automatiquement. Dans ces derniers cas, les contraintes ne sont plus liées à la faible capacité des dispositifs, ou aux incertitudes des canaux de communication non-traditionnels, mais à la magnitude et la variabilité toujours croissante des demandes, des capacités effectives, du trafic et des flux de données à gérer.

D'un point de vue scientifique, ces problématiques ont des nombreux points en commun et peuvent être approchées avec des stratégies similaires, même si elles apparaissent dans des contextes très différents et sont sujettes à des objectifs et des restrictions variées. Malgré cette hétérogénéité des conditions de communication, d'opération et de performances, il s'agit toujours de réseaux décentralisés, adaptatifs et très dynamiques, composés d'agents autonomes, qui doivent agir et prendre des décisions en temps réel, avec des mécanismes de coordination entre eux très faibles (voire inexistantes ou trop coûteux, dans certains cas), et qui disposent d'une vue très partielle, et souvent de mauvaise qualité, de la situation globale du réseau, des effets que ses décisions génèrent dans son environnement. Puisque les conditions d'opération (le trafic, les ressources et les capacités, les demandes, la topologie même du réseau) peuvent changer en tout moment, l'objectif premier à achever, c'est l'établissement d'un mécanisme de décision distribué qui permette aux agents impliqués d'observer son environnement, apprendre de ces observations locales, imparfaites et partielles, et adapter ses décisions en fonction des retours qui peuvent être retenus, de façon que l'état du réseau, s'il n'arrive à devenir optimale, puisse au moins évoluer de façon souple, non-disruptive, et non-coordonnée, vers des équilibres et des compromis acceptables entre les ressources disponibles, les restrictions imposées et les demandes présentes à chaque moment. Dans ce contexte-là, l'intelligence des réseaux se mesure à la capacité de s'adapter automatiquement aux changements de l'environnement, et de converger collectivement vers des équilibres satisfaisants en termes de performance, sans entraînement et avec un minimum de complexité et de configuration explicite.

On illustre ces problématiques d'adaptabilité décentralisée dans des réseaux dynamiques, et sa pluralité, avec des exemples qui proviennent de scénarios connectés différents. Le premier aborde la modélisation du comportement transitoire des réseaux de contenus dans des infrastructures intégrées d'accès cellulaire (mobile), aussi que les limites d'une démarche basée seulement en règles. Le deuxième explore la découverte automatique de chemins (path-forming) entre les dispositifs finaux (hosts) et les passerelles d'accès au réseau filaire, dans des réseaux IOT multi-saut, à faible débit et à fortes pertes, basés sur des technologies LPWA (Low-Power Wide Area). Bien que la problématique peut paraître un cas particulier du problème de routage dans des réseaux classiques, les conditions de communications et le type de dispositifs présents dans les réseaux LPWA, interdisent l'usage des techniques et des algorithmes traditionnels de routage, et imposent (ou favorisent) le développement de nouvelles stratégies adaptatives, moins coûteuses et plus simples. Le troisième exemple aborde le problème de la répartition de charges (load balancing) dans les datacenters. La variabilité croissante des demandes et des ressources disponibles dans les nouveaux centres de données (par exemple, avec le développement des techniques de serverless computing dans les infrastructures de nuage), ainsi que l'augmentation du trafic et l'hétérogénéité et complexité des services offerts par ces centres de données, encouragent ici aussi l'exploration de stratégies simples, adaptatives et peu gourmandes en termes de calcul et de coordination additionnelle, pour répondre et traiter efficacement les flux de données. On termine avec une tentative d'abstraction et modélisation des problèmes d'accord dynamique et décentralisé entre agents demandeurs et fournisseurs de contenu (provider-requester model). La motivation pour ce modèle vient du problème de la distribution multicast fiable sur des réseaux génériques, longtemps considéré impraticable sur Internet, et plus particulièrement des mécanismes locaux et adaptatifs de récupération en cas de pertes aux destinations; mais les modèles et la démarche développée pourrait se étendre à d'autres problématiques plus générales, d'adaptabilité de réseaux distribués.

2 Bibliographie

[1] A. R. Ndjiongue, H. C. Ferreira : Power line communications (PLC) technology, more than 20 years of intense research. Transactions on Emerging Telecommunication Technologies, vol. 30, issue 7, 2019. <https://doi.org/10.1002/ett.3575>