

Prise non-coordonnée de décision dans des réseaux intelligents à contraintes

Juan-Antonio Cordero-Fuertes, École Polytechnique

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 des microprocesseurs (loi de Moore), 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 dans la périphérie (« edge » de l'Internet), 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.

Parfois, ces réseaux sont déployés sur des environnements non-traditionnels, relient des dispositifs très limités en termes de capacités de calcul et mémoire (e.g. capteurs), et reposent sur des technologies nouvelles de communication, bien sans fils, bien par courants porteurs en ligne (CPL) ou autres. C'est le cas de ce qu'on appelle, de façon vague, l'Internet des Objets (« Internet of Things », IOT), et qui peut correspondre à des déploiements des réseaux dynamiques de capteurs et de services distribués dans des contextes et avec des ambitions très variés (« bâtiment intelligent », « véhicules connectés », « ville intelligent », « Smart Grid », etc.).

Dans d'autres cas, il s'agit des réseaux filaires, dédiés à la provision de services ou au traitement des informations (« backend »). Dans ces cas, qui est celui des 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, services de « nuage » et de calcul distribué, distribution de contenus, etc.), les contraintes ne sont plus liées à la faible capacité des nœuds, ou aux incertitudes des canaux de communication non-traditionnels, comme dans l'IOT, mais à la magnitude et la variabilité toujours croissante des demandes, des capacités effectives, du trafic et des flux de données à gérer.

Dans tous les cas, on se trouve face à des problématiques très disparates mais qui présentent, d'un point de vue scientifique, des nombreux points en commun – et qui peuvent, donc, être approchées avec des stratégies similaires. Malgré l'hétérogénéité des conditions de communication et d'opération, d'objectifs et de performances, il s'agit toujours de réseaux « intelligents », dans un certain sens, et très dynamiques, composés d'agents autonomes, qui doivent agir et prendre des décisions en temps réel, sans avoir accès à un véritable mécanisme de coordination avec le reste, 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 (nœuds, dispositifs, processus, éléments intermédiaires) d'observer son environnement, apprendre de ces observations locales es partielles, et adapter ses décisions en fonctions des signaux faibles (« weak

signals ») qui peuvent être retenus, de façon que l'évolution globale du réseaux, si elle n'arrive à devenir optimale, puisse au moins évoluer de façon automatique, et non-coordonnée, vers des équilibres acceptables entre les ressources disponibles et les demandes présentes à chaque moment. Dans cette contexte-là, l'« intelligence » des réseaux est donc moins liée à l'optimalité, qu'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.

Dans cette communication, on présente en détail trois exemples différents d'adaptabilité décentralisée dans des réseaux intelligents. Le premier exemple aborde la découverte automatique de chemins (« path-forming ») entre les dispositifs finaux 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), comme Long Range (LoRa) ou Sigfox. Bien que la problématique soit proche au 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 le développement de nouvelles stratégies adaptatives. Le deuxième exemple aborde le problème de la répartition de charges (« load balancing ») dans les datacenters ; là encore, 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 l'exploration de stratégies simples, adaptatives et peu gourmandes en termes de calcul et coordination additionnelle, pour répondre et traiter efficacement les flux de données. Dans le troisième cas, on s'intéresse au problème de la distribution multicast fiable, que ce soit sur des réseaux traditionnels (e.g. dans un centre de données) ou non-traditionnels (c'est-à-dire, dans un déploiement IOT). Le multicast fiable a été longtemps problématique sur Internet, à cause de la rigidité et la complexité imposées sur les routeurs IP impliqués dans l'opération, et la lourdeur des mécanismes de récupération en cas de pertes dans les destinations. Sur la base d'une nouvelle approche qui permet de reformuler cette complexité et alléger cette rigidité dans certaines conditions, on explore les possibilités de récupération décentralisée en cas de pertes, en s'appuyant sur des mécanismes simples de Machine Learning.