

Résumé :

Avec les récents avantages de la 5G et le développement explosif des réseaux mobiles et celui de l'Internet des objets (IoT), de plus en plus d'applications à forte intensité de calcul et de forte sensibilité à la latence émergent et s'imposent dans les habitudes applicatives quotidiennes et professionnelles des gens. Cependant, en raison des contraintes de taille inhérentes des appareils IoT et des Appareils Mobiles Intelligents (AMI), les capacités limitées de calcul et de la batterie ne permettent pas de satisfaire les exigences accrues de la qualité de service (QoS) de telles applications. On peut citer par exemple les domaines de la réalité augmentée, la reconnaissance faciale, les jeux interactifs et le traitement automatique du langage naturel. Le Mobile Edge Computing (MEC) est une architecture prometteuse qui fournit des ressources de calcul et de communication à proximité des AMI pour décharger les tâches lourdes dont l'exécution locale est coûteuse. Cette opération se fait selon un scénario monotâche ou multitâche. Dans le premier scénario et au moment de la prise de décision, chaque AMI doit décharger une seule tâche ou éventuellement l'exécuter localement. Alors que le scénario multitâche considère que chaque AMI gère une liste de tâches. Toutefois, en raison de la nature dynamique des réseaux MEC (disponibilité de ressources) et de la criticité des ressources de communication et de calcul, la garantie de satisfaire les demandes de tous les AMI d'une station de base est une opération complexe et très délicate.

Dans cette thèse, nous modélisons et résolvons des problèmes de déchargement de calcul avec allocation de ressources où nous visons à minimiser conjointement l'énergie et/ou le temps de traitement globaux ainsi qu'une fonction de pénalité qui modélise les situations critiques des ressources disponibles. Dans ce sens, nous proposons trois contributions. Dans la première, nous adoptons le scénario monotâche en admettant que certains AMI peuvent être prioritaires au service de déchargement de calcul. Nous formulons un problème d'optimisation bi-objectif non linéaire mixte en nombres entiers qui vise à minimiser conjointement la consommation énergétique globale et une fonction de pénalité basée sur la priorité. Ensuite, à l'aide d'une approche d'agrégation pondérée et une décomposition maître-esclaves, nous établissons et étudions une solution efficace du problème résultant. Dans la deuxième contribution et avec le scénario monotâche, nous formulons un problème d'optimisation multiobjectif non linéaire mixte en nombres entiers qui minimise conjointement la consommation énergétique globale, le temps global de traitement et la fonction élaborée de pénalité. Par suite, nous adoptons une approche d'agrégation pondérée avec des poids qui tiennent en compte la pénalité des AMI afin de dériver un problème d'optimisation mono-objectif non linéaire mixte en nombres entiers. En nous basant sur une décomposition maître-esclaves, nous transformons ce dernier problème en un problème d'optimisation combinatoire que nous résolvons par l'approche d'approximation de Markov. Finalement, dans la troisième contribution, nous adoptons le scénario multitâche et nous formulons un problème d'optimisation bi-objectif non linéaire combinatoire qui minimise conjointement la consommation globale d'énergie et une nouvelle fonction de pénalité adaptée au scénario multitâche. En utilisant l'approche d'agrégation pondérée, nous proposons une solution qui se base sur une recherche locale hybride.

Mots clés : Mobile Edge Computing, déchargement de calcul, allocation de ressources, fonction de pénalité, optimisation multiobjectif, décomposition maître-esclaves, approximation de Markov, algorithme Metropolis-Hastings, recherche locale, recherche par force brute, recuit simulé, optimisation discrète par essaim de particules.

MODELING AND SOLVING MULTI-OBJECTIVE OPTIMIZATION PROBLEMS OF COMPUTATION OFFLOADING IN A MOBILE EDGE COMPUTING NETWORK

Abstract :

With the recent 5G advantages and the explosive development of mobile networks and the Internet of Things (IoT), more and more compute-intensive and latency-sensitive applications are emerging and impose themselves in people's daily and professional application habits. However, due to the inherent size constraints in IoT devices and Smart Mobile Devices (SMDs), the limited computing and battery capacities are not sufficient for the increased quality of service (QoS) requirements of such applications. We can cite the fields of augmented reality, facial recognition, interactive games, and automatic natural language processing. Besides, Mobile Edge Computing (MEC) is a promising architecture that provides computing and communication resources close to SMDs to offload heavy tasks that are expensive to execute locally. This operation is done according to a monotask or multitask scenario. In the first scenario and at decision-making time, each SMD must offload a single task or possibly execute it locally. Moreover, multitask scenario considers SMDs where each one manages a list of tasks. However, due to the dynamic nature of MEC networks (resource availability) as well as the criticality of communication and computation resources, satisfying the requests of all SMDs of a base station is a complex and very delicate procedure.

In this thesis, we model and solve computational offloading problems with resource allocation where we aim to jointly minimize the overall energy and/or processing time as well as a penalty function that models critical situations of available resources. In this sense, we propose three contributions. In the first contribution, we adopt the monotask scenario with SMDs that are optionally given priority for an offloading service. We formulate a mixed integer nonlinear bi-objective optimization problem that aims to jointly minimize the overall energy consumption and a priority-based penalty function. Therefore, using a weighted aggregation approach and a master-slave decomposition, we establish and study an efficient solution to the resulting problem. In the second contribution and with the monotask scenario, we formulate a mixed integer nonlinear multiobjective optimization problem that jointly minimize the overall energy consumption, the overall processing time, and the elaborated penalty function. Subsequently, we adopt a weighted aggregation approach using weights that take into account the SMDs' penalty in order to derive a mixed integer nonlinear mono-objective optimization problem. Based on a master-slave decomposition, we transform the last problem into a new combinatorial optimization problem that we solve using a Markov approximation approach. Finally, in the third contribution, we adopt the multitask scenario and formulate a combinatorial nonlinear bi-objective optimization problem that aims to jointly minimize the overall energy consumption beside a new penalty function adapted to the multitask scenario. Using the weighted aggregation approach, we propose and study a hybrid local search-based solution.

Key Words: Mobile Edge Computing, computation offloading, resource allocation, penalty function, multiobjective optimization, master-slave decomposition, Markov approximation, Metropolis-Hastings algorithm, local search, brute force search, simulated annealing, discrete particle swarm optimization.