

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

OPTIMISATION DES RÉSEAUX SANS FIL ASSISTÉS PAR DES DRONES  
POUR DES COMMUNICATIONS D'URGENCE

MÉMOIRE  
PRÉSENTÉ  
COMME EXIGENCE PARTIELLE  
DE LA MAÎTRISE EN INFORMATIQUE

PAR  
MERIEM HAMMAMI

AOÛT 2023

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.04-2020). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

## REMERCIEMENTS

Je remercie énormément mon directeur de recherche professeur Wessam Ajib et son étudiante en doctorat Cirine Chaieb pour leur excellent encadrement et leur disponibilité. Leur soutien et leur confiance m'ont permis de m'épanouir et de développer mes connaissances dans le domaine de Télécommunication, grâce à leur expertise scientifique, tout au long de la réalisation de mon projet de recherche. Je suis vraiment reconnaissante pour leur effort, leur patience et leur soutien.

J'adresse mes profonds remerciements à ma chère amie Aya Ahmed qui m'a accompagné tout au long de mon parcours universitaire au Canada et ailleurs.

J'aimerais également remercier toute l'équipe de notre laboratoire, en particulier mes amies Aya, Cirine, Manel et Sofia, pour leur aide, leur soutien moral et leurs conseils précieux durant ma maîtrise.

Finalement, j'adresse mes profonds remerciements à ma famille en Tunisie, mon père Naim, ma mère Leila et ma petite sœur Asma, pour leur soutien inconditionnel dans mon parcours académique et leur encouragement à réaliser mes ambitions.

## TABLE DES MATIÈRES

LISTE DES TABLEAUX . . . . .	vi
LISTE DES FIGURES . . . . .	vii
LISTE DES ACRONYMES . . . . .	viii
RÉSUMÉ . . . . .	x
CHAPITRE I INTRODUCTION . . . . .	1
1.1 Mise en contexte . . . . .	1
1.1.1 L'internet tactile . . . . .	1
1.1.2 Les véhicules aériens sans pilotes (UAVs) . . . . .	2
1.1.3 Communications par ondes millimétriques . . . . .	3
1.2 Motivation . . . . .	4
1.3 Problématique . . . . .	5
1.4 Méthodologie et contributions . . . . .	7
1.5 Organisation du mémoire . . . . .	10
1.6 Conclusion . . . . .	10
CHAPITRE II ÉTAT DE L'ART . . . . .	11
2.1 Vue d'ensemble sur les réseaux de communication par UAV . . . . .	12
2.2 Optimisation de l'allocation de ressources et positionnement d'UAVs . . . . .	14
2.3 Allocation de ressources et positionnement d'UAV par algorithme génétique . . . . .	20
2.4 Allocation de ressources et positionnement d'UAV par apprentissage par renforcement profond . . . . .	21
2.5 Conclusion . . . . .	25
CHAPITRE III MODÈLE DU SYSTÈME ET FORMULATION DU PROBLÈME D'OPTIMISATION . . . . .	27

3.1	Modèle du système . . . . .	27
3.1.1	Modèle du canal de propagation . . . . .	30
3.1.2	Modèle du canal de propagation avec une information imparfaite sur l'état du canal . . . . .	32
3.1.3	Modèle de transmission . . . . .	32
3.2	Formulation du problème . . . . .	36
3.3	Analyse de complexité . . . . .	39
3.4	Conclusion . . . . .	40
CHAPITRE IV SOLUTIONS PROPOSÉES . . . . .		42
4.1	Algorithme glouton pour le déploiement des UAVs, association des utili- sateurs et allocation de puissances . . . . .	43
4.1.1	Description de l'algorithme glouton . . . . .	43
4.1.2	Analyse de complexité . . . . .	44
4.2	Algorithme génétique . . . . .	46
4.2.1	Description de la solution . . . . .	46
4.2.2	Analyse de complexité . . . . .	49
4.3	Apprentissage par renforcement profond . . . . .	49
4.3.1	Préliminaires . . . . .	50
4.3.2	Processus de décision de Markov . . . . .	56
4.3.3	Description de l'algorithme DUPRA . . . . .	57
4.3.4	Analyse de complexité . . . . .	58
4.4	Conclusion . . . . .	60
CHAPITRE V RÉSULTATS DE SIMULATION . . . . .		61
5.1	Paramètres de simulation . . . . .	61
5.2	Comparaison des performances des solutions proposées . . . . .	62
5.3	Étude de l'impact des paramètres du système . . . . .	64

5.3.1	Impact du nombre de positions possibles sur les performances de UDPA . . . . .	64
5.3.2	Région de faisabilité du système . . . . .	65
5.3.3	Impact du débit seuil et de la probabilité de coupure seuil . . . .	67
5.3.4	Impact de la puissance maximale des UAVs . . . . .	68
5.3.5	Performance de UDPA avec une méthode de sélection d'utilisateurs dans le cas d'infaisabilité . . . . .	69
5.3.6	Impact du nombre de canaux disponible en réseau d'accès . . . . .	71
5.4	Impact de la connaissance imparfaite de l'information sur les états de canaux sur les performances . . . . .	72
5.5	conclusion . . . . .	73

## LISTE DES TABLEAUX

Tableau	Page
3.1 Liste des symboles et leurs définitions . . . . .	29
5.1 Notations et valeurs des paramètres par défaut . . . . .	62
5.2 Notations et valeurs des paramètres de GA . . . . .	62
5.3 Notations et valeurs des paramètres de DUPRA . . . . .	64
5.4 Évaluation des formules de complexité de algorithmes pour $U = 30$ et $D_{min} = 1$ m . . . . .	64

## LISTE DES FIGURES

Figure	Page
3.1 Modèle du système. . . . .	30
4.1 Étapes de l’algorithme génétique. . . . .	48
4.2 Taxonomie des algorithmes d’apprentissage par renforcement (Dong <i>et al.</i> , 2020). . . . .	53
4.3 Architecture de DDPG. . . . .	56
5.1 Convergence de la solution GA, $\zeta_{fh} = 700$ Mbps, $\hat{O}_{fh} = \hat{O}_{bh} = 5\%$ and $U = 18$ . . . . .	63
5.2 Convergence de DUPRA, pour $U = 14$ . . . . .	63
5.3 Impact de la distance minimale entre les positions possibles des UAVs $D_{min}$ , pour $\zeta_{fh} = 700$ Mbps and $\hat{O}_{fh} = \hat{O}_{bh} = 5\%$ . . . . .	66
5.4 Région de faisabilité pour $\hat{O}_{fh} = \hat{O}_{bh} = 7\%$ . . . . .	67
5.5 Performance de UDPA. . . . .	68
5.6 Impact de la puissance de transmission maximale des UAVs et $B_{bh}$ , pour $\hat{O}_{fh} = \hat{O}_{bh} = 10\%$ et $\zeta_{fh} = 500$ Mbps. . . . .	69
5.7 Performance avec sélection d’utilisateur avec une puissance de transmission maximale limitée des UAVs et $B_{bh}$ , pour $\hat{O}_{fh} = \hat{O}_{bh} = 10\%$ and $\zeta_{fh} = 500$ Mbps. . . . .	70
5.8 Impact du nombre de canaux en réseau d’accès, pour $U = 20$ et $\hat{O}_{fh} = \hat{O}_{bh} = 7\%$ . . . . .	71
5.9 Impact du CSI Imparfait, pour $\zeta_{fh} = 700$ Mbps et $\hat{O}_{fh} = \hat{O}_{bh} = 5\%$ . . . . .	72



## LISTE DES ACRONYMES

AR	Réalité augmentée ( <i>Augmented Reality</i> )
AWGN	Bruit blanc gaussien additif ( <i>Additif White Gaussian Noise</i> )
A2A	Air-Air ( <i>Air to Air</i> )
A2G	Air-sol ( <i>Air to Ground</i> )
BS	Station de base ( <i>Base Station</i> )
CSI	Information sur l'état du canal ( <i>Channel State information</i> )
DCA	algorithme d'optimisation appelé <i>Difference-of-convex algorithm</i>
DDPG	Gradient de politique déterministe et profond ( <i>Deep Deterministic Policy Gradient</i> )
DL	Liaison descendante ( <i>Down link</i> )
DQL	une approche d'apprentissage profond appelée <i>Deep Q-learning</i>
DRL	Apprentissage par renforcement profond ( <i>Deep Reinforcement Learning</i> )
DUPRA	Allocation de ressource et positionnement de UAV basé sur DDPG
FLP	problème de localisation d'installations ( <i>Facility Location Problem</i> )
GS	Station terrestre ( <i>Ground Station</i> )
G2A	Sol-air ( <i>Ground to air</i> )
ITU	Union internationale des télécommunications ( <i>International Telecommunication Union</i> )
LoS	Visibilité directe ( <i>Line of Sight</i> )

MIMO	entrées multiples et sorties multiples ( <i>multiple input multiple output</i> )
mmWave	Ondes millimétriques ( <i>Millimeter-wave</i> )
NLoS	Hors visibilité directe ( <i>Non-Line of Sight</i> )
NOMA	Accès multiple non orthogonal ( <i>Non-Orthogonal Multiple Access</i> )
SINR	Rapport signal sur interférence plus bruit ( <i>Signal-to-Interference-plus-Noise Ratio</i> )
OFDMA	Accès multiple par division de fréquence orthogonale ( <i>orthogonal frequency-division multiple access</i> )
QoS	Qualité de service
RL	Apprentissage par renforcement ( <i>Reinforcement Learning</i> )
TI	Internet tactile ( <i>Tactile Internet</i> )
UAV	Véhicule aérien sans pilote ( <i>Unmanned Aerial Vehicle</i> )
UL	Liaison montante ( <i>Up link</i> )
UPRAP	Problème de positionnement des UAVs et d'allocation de ressources (en anglais <i>UAVs' positioning and resource allocation problem</i> )
URLLC	Communication ultra fiable à faible latence ( <i>Ultra-Reliable Low Latency Communication</i> )
VR	Réalité virtuelle ( <i>Virtual Reality</i> )
3GPP	Projet de partenariat de troisième génération ( <i>Third Generation Partnership Project</i> )

## RÉSUMÉ

Lors des catastrophes naturelles et des incidents imprévus, l'infrastructure de communication traditionnelle, telle que l'ensemble des stations de base, peut devenir temporairement indisponible ou endommagée pour supporter les communications d'urgence comme les téléopérations ou les opérations de sauvetage. Dans de telles situations, la mise en place rapide d'un système de communication d'urgence, qui peut assurer une téléopération d'urgence, est une tâche qui peut s'avérer critique. Considérés comme une solution prometteuse, les véhicules aériens sans pilote (en anglais *Unmanned Aerial Vehicles*, UAVs), communément appelés drones, peuvent être déployés comme de petites stations de base volantes pour offrir et relayer une communication de données entre les médecins distants et les robots d'opération, déployés dans la zone sinistrée, dans les deux directions : montante et descendante, avec des exigences de transmission strictes. Ce projet de maîtrise s'intéresse à un système de communication assisté par des UAVs qui peut assurer une communication de téléopération d'urgence. Nous abordons le problème d'optimisation conjointe de positionnement des UAVs et d'allocation de ressources dans les réseaux sans fil assistés par UAVs afin de minimiser le nombre d'UAVs déployés pour satisfaire tous les utilisateurs sous des contraintes strictes de puissance et de qualité de service définie par la satisfaction de probabilités minimales de coupure de débit dans les liaisons montantes et descendantes. Le problème formulé est un problème non convexe de programmation non linéaire à nombres entiers mixtes et il est prouvé être  $\mathcal{NP}$ -Difficile. En conséquence, nous proposons d'abord deux solutions efficaces et peu complexes : une solution heuristique et une solution métaheuristique basée sur l'algorithme génétique. Nous proposons également un puissant apprentissage par renforcement profond basé sur le gradient de politique déterministe profond où la station de base est modélisée comme un agent central qui apprend les meilleures positions des UAVs et les solutions d'allocation des ressources. Les résultats de simulation montrent que l'algorithme heuristique s'approche de l'algorithme génétique et de l'algorithme de renforcement profond avec une réduction importante de la complexité de calcul. Ils révèlent également l'efficacité des solutions algorithmiques proposées en termes de nombre d'UAVs nécessaires pour servir et satisfaire toutes les exigences de qualité de service de tous les utilisateurs.

**Mots-clés :** UAV, communication à onde millimétrique, téléopération d'urgence, allocation de ressources, algorithme génétique, apprentissage par renforcement profond.

# CHAPITRE I

## INTRODUCTION

### 1.1 Mise en contexte

#### 1.1.1 L'internet tactile

L'internet tactile (en anglais *Tactile Internet*, TI) est l'évolution de l'internet tel que nous le percevons actuellement. La haute disponibilité, la sécurité, les temps de réaction ultrarapides et la fiabilité de l'internet tactile, tel que définis par l'union internationale des télécommunications (en anglais *International Telecommunication Union*, ITU) (ITU, 2014), ajouteront une nouvelle dimension à l'interaction homme-machine en permettant la transmission des sensations tactiles et haptiques en plus des données usuelles.

Grâce à ces changements, les communications sans fil peuvent être la plateforme permettant de contrôler et de diriger des objets réels et virtuels dans de nombreuses situations de notre vie. L'internet tactile promet de créer de nombreuses opportunités pour les marchés technologiques en offrant une grande variété d'applications allant des systèmes de téléopérations et des soins de santé à la sécurité automobile et la réalité augmentée/virtuelle.

Le télédiagnostic, la téléchirurgie et la téléréhabilitation ne sont que quelques-unes des nombreuses applications potentielles de l'internet tactile dans le domaine des soins de santé. Aujourd'hui, l'expertise médicale est liée à l'endroit où se trouve le médecin. Demain, grâce à des outils de télédiagnostic avancés, elle pourra être disponible partout et à tout moment, permettant un examen physique à distance, même par palpation (examen par le toucher). Le médecin pourra commander le mouvement d'un télé-robot à l'endroit où se trouve le patient et recevoir non seulement des informations audiovisuelles, mais aussi un retour haptique critique. La téléopération rend possibles les interventions téléchirurgicales, évitant au patient les déplacements coûteux vers le chirurgien.

Dans le cadre d'une application spécifique telle qu'une téléopération, le principal obstacle qui empêche actuellement l'internet tactile de devenir une réalité tient à son exigence rigoureuse à savoir une latence extrêmement faible. Les conditions de latence pour une réponse haptique dans le corps humain doivent être de l'ordre de la milliseconde pour éviter tout retard perceptible (Promwongsa *et al.*, 2021a). Cela représente un véritable défi pour les futurs réseaux de communication.

### 1.1.2 Les véhicules aériens sans pilotes (UAVs)

Les véhicules aériens sans pilote (en anglais *Unmanned Aerial Vehicle*, UAV), aussi communément appelés drones, sont des avions pilotés par télécommande ou par des programmes informatiques embarqués sans humain à bord. Historiquement, les drones étaient principalement utilisés dans des applications militaires déployées pour la télésurveillance et l'attaque armée, afin de réduire les pertes humaines. Ces dernières années, l'enthousiasme pour l'utilisation des drones dans des applications

civiles et commerciales ne cesse de s'accroître en raison de l'avancement des technologies de fabrication des drones et de la réduction de leur coût, ce qui les rend plus facilement accessibles. Aujourd'hui, les drones sont utilisés dans de nombreuses applications dans une multitude de domaines, tels que l'inspection aérienne, la photographie, l'agriculture de précision, le contrôle du trafic, la recherche et le sauvetage, la livraison de colis et les télécommunications. Tout au long du reste de ce document les mots **drone** et **UAV** sont utilisés interchangeablement.

### 1.1.3 Communications par ondes millimétriques

L'augmentation massive du trafic de données échangées entre les appareils connectés est de plus en plus importante. Afin de prendre en charge l'explosion du trafic de données tout en réduisant la consommation d'énergie et en améliorant la qualité de service (QoS), les réseaux 5G (en anglais *fifth generation*) propose d'utiliser des nouvelles technologies habilitantes. Ces technologies, telles que les réseaux hétérogènes (HetNets), les techniques d'entrées multiples et de sorties multiples (en anglais *multiple input multiple output*, MIMO) et les techniques d'ondes millimétriques (en anglais *millimeter wave*, mmWave), ont été identifiées pour faire de la 5G une réalité.

Les communications par ondes millimétriques, qui fonctionnent entre 30 et 300 GHz, peuvent fournir une plus grande largeur de bande, ce qui permet de transmettre à des débits de données plus élevés. En outre, les communications à ondes millimétriques n'ont aucun effet sur les communications cellulaires traditionnelles en dessous de la bande 6 GHz, car elles fonctionnent à des fréquences différentes. Toutefois, l'une des principales caractéristiques des communications par ondes millimétriques est leur sensibilité aux blocages. Par conséquent, les communications par ondes millimétriques

sont généralement utilisées dans la transmission directionnelle à courte portée avec des antennes directionnelles.

## 1.2 Motivation

La mise en place rapide de systèmes de communication d'urgence, lors des catastrophes naturelles et d'incidents imprévus, est une tâche difficile mais peut être vitale. Dans de telles situations, l'infrastructure de communication traditionnelle, telle que les stations de base, peut devenir temporairement indisponible ou endommagée pour supporter les téléopérations d'urgence. Ces téléopérations visent à exercer des premiers soins d'urgence en permettant aux médecins de contrôler des robots à distance sans devoir rejoindre la zone sinistrée. Considérés comme une solution prometteuse, les véhicules aériens sans pilote (UAVs) peuvent être déployés comme de petites stations de base volantes pour jouer le rôle de relais.

Comme solution envisageable, ce projet de recherche investigate l'intégration des UAVs ainsi que des nouvelles technologies radio telles que les communications par ondes millimétriques dans le réseau afin d'assurer une communication de données entre les médecins et les robots à distance et de garantir les exigences en QoS des applications TI dans les deux directions de liaison montante et descendante.

Le problème d'optimisation conjointe du positionnement des UAVs (ou de la conception de trajectoire des UAVs) et d'allocation de ressources a été abordé dans de nombreux travaux (Zhang *et al.*, 2019) dans la direction de liaison montante ou liaison descendante. Cependant, la prise en compte des communications en liaison montante et descendante simultanément dans le but de minimiser le nombre d'UAVs nécessaires pour satisfaire toutes les exigences strictes du système dans un contexte



de communication d'urgence n'a pas encore été abordée. De plus, la plupart des travaux qui s'intéressent aux applications de l'internet tactile, et plus précisément les téléopérations, proposent des solutions filaires qui se reposent sur l'infrastructure réseau existante (le réseau d'infrastructure avec ses équipements tels que les routeurs). Ce projet considère un réseau mobile sans fil assisté par des UAVs qui peut être utilisé pour les communications de téléopération d'urgence.

### 1.3 Problématique

Dans le but de mettre en place rapidement un système de communication d'urgence dans une zone sinistrée, le déploiement d'un ensemble de UAVs qui servira comme des stations de bases ou des relais pour assurer les communications est une solution prometteuse. Un des cas de communications d'urgence est la téléopération pour permettre à des médecins distants de faire de premiers soins sur des patients en contrôlant des robots. De telles opérations exigent une communication fiable et un échange d'information à très haut débit pour garantir une latence très faible. Parmi les technologies qui permettent d'atteindre ses objectifs, on retrouve les communications par mmWave.

Le déploiement des UAVs comme relais entre les médecins et les robots ainsi que l'utilisation des communications par mmWave semble être la solution pour un déploiement rapide d'un système de communication d'urgence qui peut assurer une téléopération. Toutefois, le choix du nombre nécessaire des UAVs à placer et leurs positions optimales peut s'avérer une tâche compliquée. Ce choix dépend fortement des positions des robots et de l'environnement de communication (les états de canaux de communication entre les UAVs et les robots). L'environnement de communication

joue un rôle primordial dans l'allocation des ressources, notamment l'association des utilisateurs (robots dans ce cas), les canaux de communication et les puissances de transmission. Ainsi, la mise en place des systèmes de communication d'urgence sans fil qui utilisent ces deux technologies nécessite le développement de solutions algorithmiques adaptées de manière à optimiser les ressources utilisées pour garantir la QoS exigée par des applications comme la téléopération.

Dans le cadre de ce projet, nous nous focalisons sur l'étude d'un système assisté par UAVs pour performer une téléopération de premiers soins sur des personnes blessées dans une zone sinistrée. Considérant que l'infrastructure réseau a été démolie dans cette zone, les UAVs jouent le rôle de relais entre une station de base terrestre connectée au réseau d'infrastructure (qui peut être une station de base sur camion), représentant les médecins, et les utilisateurs représentant les robots contrôlés à distance qui vont opérer les personnes. Nous considérons les communications en liaison montante et en liaison descendante pour le réseau d'accès (entre les UAVs et les utilisateurs) et, également, pour le réseau d'amenée (entre la station de base macro et les UAVs). Les communications se font par ondes millimétriques et permettent un haut débit de transmission à courte portée. L'objectif de notre projet est de proposer des solutions algorithmiques qui permettent de minimiser le nombre d'UAVs à déployer dans le réseau tout en assurant la téléopération sous des contraintes sur la QoS de l'ensemble des liaisons de communication, des contraintes de puissances de transmission limitées de l'ensemble des entités du système et des contraintes sur l'association UAV-utilisateur. Pour ce faire, nous optimisons conjointement le positionnement des UAVs à déployer, l'association des utilisateurs aux UAVs et l'allocation des puissances de transmission. La QoS prise en compte est une probabilité de coupure d'un

débit seuil. Cette qualité de service va permettre d'assurer la fiabilité et la continuité du service dans le cadre d'une application de téléopération exigeante.

**Notre objectif est de minimiser le nombre d'UAVs à déployer dans un réseau sans fil assisté par des UAVs par optimisation du positionnement des UAVs, de l'association UAV-utilisateur et de l'allocation de puissances de transmission tout en satisfaisant les exigences de QoS de tous les utilisateurs.**

#### 1.4 Méthodologie et contributions

Pour atteindre l'objectif annoncé dans la section 1.3, nous commençons par définir le modèle mathématique du système afin d'identifier tous les éléments du réseau et de son environnement. Ensuite, nous formulons le problème d'optimisation qui vise à minimiser le nombre d'UAVs à déployer pour servir l'ensemble des utilisateurs du système sous des contraintes sur les associations UAV-utilisateur et des contraintes sur les puissances de transmission limitées des UAVs, des utilisateurs et de la station de base macro tout en garantissant la QoS définie par une probabilité de coupure d'un débit seuil. Notre modèle prend en compte les seuils de probabilité de coupure du débit, car il n'est pas toujours possible de garantir des seuils stricts sur le rapport signal sur interférence plus bruit (en anglais *Signal-to-interference-plus-noise ratio*, SINR) ou de délai (en raison de situations de profond affaiblissement du signal) dans un environnement sans fil. La probabilité de coupure du débit est définie comme la probabilité que le débit de transmission soit inférieur à une valeur de débit prédéfinie seuil. Nous étudions la complexité de notre problème de positionnement d'UAVs et d'allocation de ressource (en anglais *UAVs' positioning and resource allocation pro-*

*blem*, UPRAP) et nous prouvons qu'il est  $\mathcal{NP}$ -Difficile. Enfin, nous proposons trois approches qui permettent de répondre à notre problématique. La première approche est une solution heuristique basée sur un algorithme glouton permettant de déployer les UAVs un à la fois en associant les utilisateurs et en attribuant les puissances minimales satisfaisant les probabilités de coupure seuil. La deuxième approche est une solution basée sur l'algorithme génétique. Les individus, représentant des solutions possibles, ont un ensemble de gènes englobant les variables d'optimisation (la matrice d'association, le vecteur de déploiement des UAVs, les quatre matrices de puissance et la matrice de positions des UAVs). La dernière approche est une solution d'apprentissage par renforcement profond. Les complexités des trois approches sont analysées.

Les principales contributions de notre projet sont résumées comme suit :

- Nous formulons un problème d'optimisation conjointe de positionnement des UAVs et d'allocation des ressources dans les communications en liaison montante et en liaison descendante avec l'objectif de minimiser le nombre des UAVs déployés sous des contraintes de puissances et des exigences de QoS. Les exigences de QoS sont définies par la satisfaction de probabilités minimales de coupure de débit avec des valeurs minimales de débit pour les liaisons montantes et descendantes.
- Nous prouvons la NP-difficulté d'UPRAP en réduisant le problème de localisation d'installations (en anglais *facility location problem*, FLP), qui est un problème connu comme étant NP-difficile, à un cas spécial d'UPRAP.
- Pour résoudre efficacement le problème, nous proposons, tout d'abord, une solution basée sur une approche gloutonne pour sélectionner des positions discrètes pour les UAVs, associer les utilisateurs aux UAVs et allouer des

niveaux de puissance de transmission pour répondre aux exigences de QdS.

- Comme l’algorithme génétique a prouvé sa capacité à converger vers une solution optimale, ou quasi-optimale, après un certain nombre d’itérations dépendant du problème, nous proposons un algorithme métaheuristique basé sur l’approche génétique pour comparer ses performances avec celles obtenues par la solution heuristique.
- Pour tirer profit de l’espace d’action continu de l’algorithme de gradient de politique déterministe profond (en anglais *Deep Deterministic Policy Gradient*, DDPG), nous développons également un algorithme d’apprentissage par renforcement profond (en anglais *deep reinforcement learning*, DRL) centralisé basé sur l’approche DDPG. Le problème considère désormais des positions continues pour les UAVs. Il est modélisé comme un processus de décision de Markov (en anglais *markov decision process*, MDP), où une action est la sélection des positions des UAVs.
- Nous évaluons les performances des algorithmes proposés par le biais de simulations en explorant i) la région de faisabilité du problème, ii) l’impact des paramètres du système (par exemple, le nombre de canaux, le débit seuil requis, la probabilité de coupure requise) sur le nombre d’UAVs à déployer, et iii) l’impact de la connaissance imparfaite de l’information sur l’état des canaux. Les résultats montrent que la solution basée sur le DRL est plus performante que les solutions heuristiques et métaheuristiques et démontrent l’effet de la discrétisation des positions des UAVs sur le nombre d’UAVs à déployer par la solution heuristique.

## 1.5 Organisation du mémoire

Le reste de ce mémoire est organisé comme suit, le chapitre II présente l'état d'art des travaux liés à notre projet de recherche. Le modèle du système et la formulation mathématique sont détaillés dans le chapitre III. Le chapitre IV présente les trois solutions proposées pour résoudre la problématique étudiée et leurs analyses de performances sont énumérées dans le chapitre V. Finalement, nous terminons par une conclusion résumant notre travail.

Un papier de conférence intitulé : "*UAV-Assisted Wireless Networks for Stringent Applications : Resource Allocation and Positioning*" issu de notre projet de mémoire est publié dans les actes de conférence *IEEE Wireless Communications and Networking Conference (WCNC), 26–29 March 2023 in Glasgow, Scotland, UK*.

Un papier de journal intitulé : "*Meeting Stringent QoS Requirements in UAV-Assisted Networks : Resource Allocation and UAVs Positioning Schemes*" a été soumis au journal *IEEE Transactions on Communications*.

## 1.6 Conclusion

Dans ce chapitre, nous avons présenté une mise en contexte pour notre projet de maîtrise. Ensuite nous avons discuté la motivation et la problématique à laquelle nous allons répondre. Enfin, nous avons précisé la méthodologie et les contributions. Dans le prochain chapitre, nous allons présenter les travaux de recherche qui sont liés à notre problématique.



## CHAPITRE II

### ÉTAT DE L'ART

L'internet tactile combinera différentes exigences strictes en termes de latence ultra-faible, de débit élevé, de fiabilité et de sécurité. Les soins de santé à distance, où l'expertise médicale sera disponible partout et à tout moment au lieu d'être limitée à l'endroit où se trouve le médecin, constituent un cas d'utilisation important. Le médecin pourra commander à distance un robot à l'endroit où se trouve le patient, ce qui permettra une intervention physique et un examen à distance avec un retour haptique complet. Par exemple, la téléchirurgie est l'une des applications les plus rigoureuses que l'internet tactile peut permettre (Promwongsa *et al.*, 2021b). Pour réaliser des communications haptiques sur les réseaux mobiles, plusieurs exigences de qualité de service (QoS) doivent être satisfaites, comme un débit minimal et une faible latence. Par conséquent, il est essentiel d'assurer des solutions efficaces d'allocation des ressources pour l'échange d'informations sur la liaison montante (UL) et la liaison descendante (DL) (Aijaz, 2018).

Dans les scénarios de communication d'urgence, les communications par satellite pourraient être une solution potentielle pour transmettre de l'information entre les stations de base distantes et les utilisateurs locaux (Taleb *et al.*, 2011; Wang *et al.*,



2020). Cependant, les communications par satellite sont connues pour leur coût d'introduction, leur délai de propagation élevé et leur débit insuffisant (Watts et Aliu, 2014) qui ne correspondent pas aux exigences des téléopérations (Kapovits *et al.*, 2018). D'autre part, avec le développement des technologies de communication sans fil, plusieurs techniques ont été mises au point pour permettre une capacité élevée, une faible latence et un débit de données élevé, comme par exemple une communication basée sur des véhicules aériens sans pilote (UAVs). Dans le contexte de la téléopération médicale d'urgence où l'infrastructure du réseau terrestre est totalement ou partiellement absente, la communication haptique sans fil assistée par des UAVs est considérée comme une solution prometteuse. Elle peut rompre l'isolement des zones sinistrées et permet aux médecins d'effectuer les premiers soins à distance grâce à leur flexibilité, leur mobilité, leur faible coût et leurs liaisons en visibilité directe (en anglais *Line-of-sight*, LoS).

## 2.1 Vue d'ensemble sur les réseaux de communication par UAV

(Gupta *et al.*, 2016) présente une vue d'ensemble sur les réseaux de communication par UAV et de leurs enjeux. Il commence par une introduction sur les UAVs, leurs domaines d'applications et leurs spécifications matérielles. Cet article identifie quatre défis majeurs pour les réseaux de communication par UAVs. Premièrement, quelques aspects de la conception architecturale de ces réseaux peuvent être non intuitifs à cause de la topologie dynamique et non fixe (des nœuds de la topologie peuvent disparaître comme un UAV peut ne plus être en fonction dans le réseau). Deuxièmement, les protocoles de routage ne peuvent pas être des implémentations simples de protocoles traditionnels, proactifs ou réactifs. Les protocoles de routage réactifs déterminent les routes sous demande. Les protocoles de routage proactifs trouvent

périodiquement des routes en permanence entre deux nœuds quelconques du réseau, indépendamment de la demande. En vue de la nature dynamique de la topologie, les protocoles de routage dans les réseaux de communication par UAV nécessitent une complexité qui gère leur nature. Le troisième défi réside dans le maintien des sessions des utilisateurs transférés d'un UAV, qui n'est plus en service, à un autre UAV actif. Finalement, il s'agit d'un défi de conservation d'énergie des UAVs, connus pour leur énergie limitée, pour étendre la durée de vie du réseau.

Une architecture et un système d'allocation de ressources pour les réseaux de communication par UAVs est proposé par (Kawamoto *et al.*, 2019). L'architecture considérée est composée d'un ensemble de UAV, des stations terrestres (en anglais *Ground Station*, GS) et des stations de base (en anglais *base station*, BS) avec des systèmes de gestion de ressources. Spécifiquement, les UAVs sont connectés aux GSs et leurs envoient une variété d'informations. Les GSs sont liées aux BSs, qui sont des points d'accès à des réseaux existants comme des réseaux cellulaires et des réseaux WiFi. Le système de gestion de ressources est responsable de l'allocation de ressources qui correspondent aux exigences de chaque UAV en fonction des informations collectées sur le réseau. (Kawamoto *et al.*, 2019) considère deux unités séparées pour gérer les ressources. Une unité pour déterminer les ressources à attribuer à partir des exigences de communication de chaque UAV. Cette unité est supposée être incluse dans le système de gestion de ressource. L'autre unité attribue à chaque UAV les ressources comme la fréquence de communication et le temps de communication.

Une vue d'ensemble détaillée de plusieurs aspects des réseaux de communication par UAV est présentée par (Fotouhi *et al.*, 2019). Le premier aspect abordé est la

classification des UAVs et leurs caractéristiques à savoir ; les mécanismes de vol, l'altitude, la vitesse de déplacement, la durée de vol et l'alimentation en énergie. Ensuite, l'article explique les standardisations du groupe de partenariat de troisième génération (en anglais *the third generation partnership project*, 3GPP) qui régissent les activités dans les réseaux de communication par UAV. (Fotouhi *et al.*, 2019) s'est intéressé à deux aspects étudiés dans la standardisation 3GPP. Le premier aspect est la classification des trafics des UAVs et leurs exigences. Ces derniers sont classifiés en : contrôle radio et synchronisation, commande et contrôle, et données d'application. Le deuxième aspect est la modélisation du canal de communication. Ce canal peut être modélisé comme air-sol (en anglais *Air to Ground*, A2G), air-air (en anglais *Air to Air*, A2A) ou sol-air (en anglais *Ground to Air*, G2A). L'article continue par dresser un état d'art d'un ensemble de travaux adressant les défis de déploiement des UAVs comme des stations de bases aériennes dans des réseaux cellulaires. Les défis identifiés sont des défis d'optimisation de placement 3D/2D et de mobilité des UAVs, d'optimisation de la consommation d'énergie des UAVs et d'optimisation du lien de bout en bout des liens de communication quand un UAV en mouvement connecte un utilisateur à l'épine dorsale filaire du réseau.

## 2.2 Optimisation de l'allocation de ressources et positionnement d'UAVs

Plusieurs travaux de recherche ont porté sur l'optimisation de l'allocation des ressources et du positionnement des UAVs dans les réseaux sans fil assistés par UAV. Dans (Tran *et al.*, 2022), les auteurs ont proposé un modèle d'internet des objets (en anglais *Internet of Things*, IoT) assisté par UAV. Un UAV est déployé pour assister l'infrastructure terrestre existante qui est partiellement ou totalement endommagée après un désastre naturel ou dans le cas où la station de base est surchargée du-

rant les heures de pointe. La mission de l’UAV dans cet article est de collecter les informations des dispositifs IoT présents dans le réseau et de les transmettre à une passerelle terrestre. L’objectif de (Tran *et al.*, 2022) est de maximiser le nombre de dispositifs IoT servis par l’UAV sous contraintes de capacité de stockage limitée de l’UAV, des latences maximales permises dans les deux sens de liaison montante (en anglais *Uplink*, UL) et descendante (en anglais *Downlink*, DL) et de la puissance de transmission maximale de l’UAV. Cet objectif est inspiré par l’exigence pratique des mesures de la sécurité humaine qui stipule que plus on collecte des données, mieux on prédit. Ce travail optimise conjointement la bande passante totale, la puissance de transmission, la trajectoire de l’UAV et la capacité de stockage. Comme le problème d’optimisation formulé appartient à la classe des problèmes d’optimisation non linéaire à nombres entiers mixtes (en anglais *mixed-integer non-linear optimization problem*, MINLP) alors un algorithme itératif efficace a été proposé pour obtenir une solution sous-optimale. Premièrement, les variables binaires ont été relaxées en des variables continues. Ensuite, une approche inspirée de la technique d’approximation interne (en anglais *inner approximation*) a été proposée. Les résultats de la simulation montrent l’efficacité de la solution proposée. En effet, ils ont démontré une amélioration significative du nombre d’équipements servis par rapport aux algorithmes de références avec des ressources attribuées fixes ou une trajectoire fixe.

(Kang *et al.*, 2021) considère les communications en UL et en DL dans un système de communication cellulaire basé sur un UAV utilisé comme relais entre une station de base et des utilisateurs fixes. Il propose un algorithme d’allocation de puissance d’UAV et de contrôle de sa trajectoire pour maximiser la somme totale de débit en UL et en DL. Le contrôle de la trajectoire dépend de l’information sur l’état du canal. Cette dernière change à cause de l’évanouissement rapide (en anglais *fast*

*fading*) ce qui entraîne des mouvements irréguliers de l'UAV. Pour résoudre le problème de contrôle de trajectoire, une maximisation ergodique du débit est proposée. L'optimisation de la trajectoire et de l'allocation de ressources est un problème non convexe sous les contraintes d'évitement des collisions, les contraintes de limitation d'altitude et les contraintes de mobilités des utilisateurs. Pour une résolution efficace, la technique d'approximation interne est appliquée de façon itérative. Les résultats de simulation montrent que l'algorithme de maximisation bidirectionnelle de débit est plus performant que les algorithmes considérant une maximisation en UL ou en DL dépendamment de la puissance de transmission de l'UAV.

Dans le cas de déploiement des UAVs comme relais de communication entre une source et une destination, la majorité des travaux considèrent que la durée de la communication des relais n'est pas plus longue que la durée de vol d'un UAV. Cependant, il existe des scénarios qui nécessitent une durée de communication plus longue que la durée de vol d'un UAV. Par exemple, les transmissions de longue durée de flux de données en direct ou les transmissions des données massives exigent une durée de relayage plus longue que le temps de vol d'un UAV. Motivé par la prolongation de la durée de communication de relais dans ces scénarios, (Zhang *et al.*, 2022) propose deux techniques de substitution d'un ensemble d'UAVs qui relaient les données d'une source terrestre à une destination terrestre l'un après l'autre successivement. Le premier schéma est un schéma heuristique de substitution pour prolonger la durée de la communication. Dans le but de maximiser le débit de bout en bout du système, le contrôle de la puissance de transmission et la trajectoire des UAVs sont optimisés conjointement. Comme les transmissions bidirectionnelles simultanées (en anglais *full-duplex*, FD) peuvent être coûteuses à déployer sur un UAV à charge

utile (en anglais *payload*) limitée, alors les transmissions bidirectionnelles alternées (en anglais *half-duplex*, HD) ont été considérées dans ce travail. Pour surmonter le problème de faible efficacité spectrale (la source ne peut envoyer au relais que pendant la demi-durée de communication) causé par ce type de transmission, un deuxième schéma d'efficacité spectrale de substitution d'UAVs a été présenté. Ce schéma permet à un UAV relais de transmettre des données à la destination en même temps qu'un deuxième reçoit les données de la source de façon que la source est en train de transmettre tout le temps. Pour résoudre ce problème d'optimisation non convexe et difficile, un algorithme qui applique la méthode de montée en coordonnées par blocs (en anglais *block coordinate ascent*). Cet algorithme divise le problème en deux sous-problèmes; un pour l'optimisation de trajectoire et l'autre pour le contrôle de puissance de transmission. Puis, une méthode d'approximation convexe successive (en anglais *successive convex approximation*) est appliquée sur les deux sous-problèmes. Les résultats de simulation montrent que les schémas de substitution proposés prolongent effectivement la durée de communication de relais par rapport aux autres schémas qui ne considèrent pas la trajectoire et le contrôle de puissance conjointement.

Le travail de (Guo *et al.*, 2020) a considéré un réseau de relais par UAV multi-utilisateur où un ensemble de paires d'utilisateurs terrestres et un utilisateur cellulaire coexistent dans un réseau. L'utilisateur cellulaire transmet et reçoit de la station de base directement. Les communications entre les émetteurs terrestres et les récepteurs terrestres sont supposées être bloquées à cause des canaux défavorables, ce qui justifie le déploiement de l'UAV comme relais entre eux. Les utilisateurs terrestres et l'utilisateur cellulaire se causent de l'interférence à cause des ressources fréquentielles réutilisées. L'objectif de (Guo *et al.*, 2020) est de maximiser la somme instantanée

des débits de tous les utilisateurs terrestres, tout en assurant l'exigence de débit de l'utilisateur cellulaire. L'optimisation conjointe du positionnement de l'UAV et l'allocation de puissance de transmission a été étudiée. À cause de la nature non convexe du problème, la descente de coordonnées par bloc et l'optimisation successive convexe ont été employées pour diviser le problème en deux sous-problèmes et les résoudre alternativement en améliorant la somme des débits de façon itérative. Les résultats numériques ont montré que l'utilisation des UAVs comme relais a un avantage remarquable comparée aux communications terrestres directes dans le cas des canaux défavorables et des utilisateurs dispersés dans le sol.

Les auteurs de (Shi *et al.*, 2021) ont proposé un système de relais assisté par un UAV pour les transmissions bidirectionnelles simultanées en accès multiple non orthogonal (en anglais *full-duplex non-orthogonal multiple-access*, FD-NOMA) pour améliorer l'efficacité spectrale. Les utilisateurs et la station de base dans le système sont supposés utiliser la technologie de communication FD. Le spectre est supposé être divisé également en plusieurs sous-canaux. À cause de l'énergie limitée de l'UAV, il est déployé pour relayer un flux de données partiel en liaison montante des utilisateurs. D'une autre part, la station de base transmet directement les signaux en liaison descendante vers les utilisateurs. Les signaux transmis en liaison montante et descendante sont superposés sur les mêmes sous-canaux pour augmenter l'efficacité spectrale et énergétique. L'objectif de ce travail est de minimiser l'énergie totale consommée par les communications dans les liaisons montantes et les liaisons descendantes en optimisant conjointement le choix du transmetteur, le jumelage des utilisateurs, l'allocation de puissance et l'assignation de sous-canaux sous des contraintes sur le SINR considérées pour la BS, les utilisateurs et l'UAV. Le problème formulé est

un problème non convexe et pour le résoudre, (Shi *et al.*, 2021) a proposé un schéma d'allocation conjointe de ressources pas à pas en UL et DL qui divise le problème initial en cinq sous-problèmes à savoir ; la détermination du récepteur, le jumelage des utilisateurs, allocation de puissance de transmission, l'attribution des sous-canaux et le contrôle d'accès. Les résultats de simulation ont confirmé l'efficacité spectrale et énergétique du schéma proposé basé sur NOMA par rapport aux méthodes de référence basées sur l'accès multiple par division de fréquence.

Dans (Zeng *et al.*, 2022), un schéma d'accès multiple en mode hybride a été proposé en considérant un système assisté par UAV avec un trafic hétérogène. Les auteurs ont étudié la conception de la trajectoire et l'allocation conjointe de la largeur de bande passante en UL et en DL afin de maximiser le débit moyen minimum. L'accès multiple non orthogonal NOMA a été utilisé pour atteindre les hauts débits de données et l'accès multiple orthogonal OMA a été réservé aux utilisateurs ayant des conditions de canal très mauvaises ou bien une interférence inter-utilisateur très forte pour compenser leurs demandes de débits. Pour résoudre ce problème, (Zeng *et al.*, 2022) a proposé de le diviser en sous-problèmes et de les résoudre d'une manière itérative. L'allocation de bande passante a été résolue avec une trajectoire donnée et en supposant que la puissance de transmission est distribuée également sur la bande passante. Ensuite, le sous-problème d'allocation de puissance a été transformé en une forme convexe. Finalement, la trajectoire de l'UAV a été conçue par une optimisation convexe successive. Les résultats de simulation ont montré que le schéma d'accès multiple hybride proposé est capable de répondre aux demandes de trafic hétérogènes des utilisateurs et d'offrir un débit moyen minimum supérieur, et donc une meilleure équité entre les utilisateurs, en comparant avec les approches par accès multiple



orthogonal ou l'accès multiple non orthogonal séparément.

### 2.3 Allocation de ressources et positionnement d'UAV par algorithme génétique

Les auteurs de (Yang *et al.*, 2019) ont étudié un système de transfert d'énergie sans fil (en anglais *Wireless power tranfert*, WPT) opérationnel par UAV avec plusieurs utilisateurs terrestres. L'objectif de ce travail a été de maximiser l'énergie minimale reçue par tous les utilisateurs en optimisant la trajectoire de l'UAV sous contrainte de la limitation de vitesse. Pour résoudre ce problème intraitable, un schéma de vol et survol successif basé sur l'algorithme génétique a été proposé. Les résultats numériques ont confirmé l'avantage en performance de l'utilisation de l'algorithme génétique en comparaison avec les algorithmes heuristiques proposés dans d'autres travaux similaires et l'algorithme du voyageur de commerce adapté au problème étudié.

(Sawalmeh *et al.*, 2019) s'est intéressé au déploiement dynamique d'un UAV comme station de base aérienne pour fournir une couverture sans fil à une équipe de sauvetage pendant une situation d'urgence. Les utilisateurs mobiles considérés appartiennent à deux types : des utilisateurs à l'extérieur des bâtiments et des utilisateurs à l'intérieur des bâtiments. Les modèles de mouvements considérés diffèrent selon le type de l'utilisateur. Les auteurs de (Sawalmeh *et al.*, 2019) ont proposé trois algorithmes différents pour trouver une trajectoire efficace de l'UAV qui satisfait les exigences de débit minimal des utilisateurs et qui minimise la puissance de transmission totale de l'UAV. La trajectoire optimale a été obtenue en utilisant un algorithme de recherche exhaustive. L'algorithme génétique (GA) et l'algorithme d'optimisation par essaim de particules (en anglais *Particle swarm optimisation*, PSO) ont été pro-

posés comme algorithmes métaheuristiques efficaces pour résoudre le problème. Les deux algorithmes, GA et PSO, ont permis de réduire le temps d'exécution par un facteur de 1/60 et 1/9, respectivement, par rapport à l'approche optimale. La simulation a aussi démontré que le déploiement dynamique de l'UAV est plus performant en termes de puissance de transmission par rapport au déploiement statique de l'UAV.

Un déploiement d'une communication sans fil rapide, efficace et peu coûteuse a été proposé dans (Sambo *et al.*, 2019), où un ensemble de stations de base montées sur des camions est déployé pour un agrégat d'utilisateurs dans une zone sinistrée. Les stations de base n'ont aucun lien de réseau d'amenée avec le réseau d'infrastructure et donc comptent sur un UAV pour venir périodiquement collecter les données et les transférer au réseau d'infrastructure. (Sambo *et al.*, 2019) a proposé une solution basée sur l'algorithme génétique pour trouver la trajectoire qui minimise l'énergie dépensée par l'UAV pour récupérer les données de toutes les stations de base présentes dans le réseau et de retourner au nœud passerelle. La performance de la solution proposée a été comparée à deux autres solutions basées sur l'algorithme génétique qui ont été proposées par d'autres travaux. Les résultats de simulation ont montré que la solution proposée surpasse les deux autres solutions par 14% et 40%.

#### 2.4 Allocation de ressources et positionnement d'UAV par apprentissage par renforcement profond

(Luong *et al.*, 2021) a étudié la liaison descendante d'un réseau cellulaire assisté par des UAVs. Ces UAVs peuvent servir les équipements des utilisateurs coopérativement par des techniques de traitement conjointes. Ils sont contrôlés par un contrôleur central situé dans la station de base macro. Un contrôleur central d'UAVs est res-

ponsable du traitement de tous les signaux en bande de base, de la coordination de l'allocation des ressources et de calcul, ainsi que du transport des données vers les UAVs via des liaisons sans fil en réseau d'accès. L'objectif du travail a été de maximiser le débit global du système pour lequel une optimisation conjointe des positions des UAVs, de l'association UAV-Utilisateur et de la formation des faisceaux (en anglais *Beamforming*) de transmission au niveau des UAVs a été considérée. Le problème d'optimisation a été considéré complexe dû à sa non-convexité et l'indisponibilité d'information sur l'état du canal (CSI) en raison du mouvement des UAVs. Pour résoudre ce problème, (Luong *et al.*, 2021) a proposé un algorithme comprenant deux étapes. Dans la première étape, une méthode d'apprentissage par renforcement distribuée et basée sur un algorithme Q-learning profond (en anglais *Deep Q-learning*, DQL) a été adoptée pour permettre aux UAVs d'apprendre conjointement l'état global du réseau et d'adapter leurs positions selon l'action reçue de la MBS. Dans la deuxième étape, sur la base des positions des UAVs obtenues, un algorithme de différence convexe (en anglais *Difference Convex Algorithm*, DCA) a été proposé pour traiter le problème non convexe résultant de l'association UAV-utilisateur et de la formation de faisceaux de transmission. La sortie de l'algorithme DCA proposée est ensuite utilisée pour calculer la récompense de l'algorithme DQL. Ce processus est itéré jusqu'à la convergence de l'algorithme. Pour améliorer la convergence, les auteurs ont proposé une manière efficace de choisir un état initial de l'algorithme DQL plutôt que d'en choisir un arbitrairement. Les résultats de simulation ont montré que l'algorithme proposé (DQL assisté par DCA) est plus performant qu'un algorithme Q-learning classique en termes de débit global atteint pour le même nombre d'UAVs ainsi qu'il converge plus rapidement.

Les auteurs de (Dai *et al.*, 2022) ont considéré un système de communication basé sur des UAVs en collaboration pour apprendre l'environnement de communication. Étant donné que la puissance de transmission et les ressources énergétiques d'un UAV sont limitées ainsi que l'environnement de communication est complexe et variant dans le temps, il est difficile pour les UAVs de prendre dynamiquement des décisions qui assurent la communication. Une architecture centralisée peut générer un délai de communication et par conséquent affecter l'efficacité de la communication. De ce fait, un algorithme d'allocation de ressources et de placement des UAVs basé sur un apprentissage d'environnement distribué et multi-agent a été proposé dans l'objectif de maximiser l'utilité de l'ensemble des communications des UAVs. Dans cet algorithme, chaque UAV est un agent qui prend la décision de sa position et les ressources attribuées selon une information sur l'état de canal observée partiellement. Pour le réseau global, une méthode d'apprentissage par renforcement multi-agent basée sur l'apprentissage fédéré (en anglais *federated learning*) a été conçue pour assurer l'interaction entre les UAVs. Les résultats de simulation ont démontré que l'algorithme proposé a amélioré l'utilité du réseau global en comparant avec un algorithme d'apprentissage par renforcement multi-agent classique sans interactions entre les agents.

Dans (Wang *et al.*, 2022), les auteurs ont considéré un réseau de communication d'urgence assisté par des UAVs, où un UAV est considéré comme un relais mobile utilisé pour transmettre des informations de données d'une station de base macro (MBS) aux utilisateurs par des transmissions bidirectionnelles alternées. L'objectif de ce travail a été de maximiser la somme de l'efficacité spectrale en optimisant conjointement l'allocation de puissance de la MBS, la sélection de la zone de service de chaque UAV et l'ordonnancement des utilisateurs. Pour ce scénario d'urgence,

(Wang *et al.*, 2022) a proposé une structure de coopération à deux étapes. La première étape était la planification des UAVs dans les zones à couvrir et l'allocation de la puissance de transmission de la MBS et la deuxième étape était l'ordonnancement des utilisateurs. Pour résoudre le problème, les auteurs ont proposé deux solutions, une pour un système qui ignore le délai de temps et l'autre pour un système qui considère le délai de temps. Un algorithme basé sur Q-Network profond (en anglais *Deep Q-Network*, DQN) qui adopte les réseaux de neurones de convolution (en anglais *Convolutional Neural Networks*, CNNs) a été proposé comme solution pour la première étape du système qui ignore le délai de temps. Pour la deuxième étape, le schéma d'ordonnancement des utilisateurs a été modélisé comme un problème d'optimisation binaire divisé en deux sous-problèmes. Dans le cas du système qui considère le délai de temps, l'information sur le canal actuel et la matrice de transition d'états sont inconnues et le DQN estime l'action optimale actuelle en considérant une information obsolète sur le canal. Une solution basée sur le DRL a été proposée pour améliorer l'efficacité du spectre total. Les résultats de simulation ont illustré que la solution de DQN basé sur les CNNs est plus performante que le DQN classique avec les réseaux de neurones profonds (en anglais *Deep Neural Networks*, DNNs) et ceci par la diminution des pas d'apprentissage et l'amélioration des exactitudes de l'apprentissage. La solution proposée a été comparée à un algorithme de recherche exhaustive, un algorithme de Q-learning et un algorithme aléatoire. Cette comparaison a démontré que la solution proposée est plus performante en termes d'efficacité spectrale et est très proche de la performance de la recherche exhaustive avec un temps de calcul moins important.

(Ding *et al.*, 2020) a investigué le problème d’efficacité énergétique et de service de communication pour un système assisté par un UAV. Les auteurs ont formulé une fonction utile d’équité pour exprimer le compromis entre le débit et l’équité entre les utilisateurs. L’objectif a été de maximiser la fonction utile d’équité en optimisant la direction et l’altitude de vol de l’UAV et la bande passante attribuée à chaque utilisateur sous contrainte d’énergie limitée de l’UAV. Pour résoudre le problème, les auteurs de (Ding *et al.*, 2020) ont proposé une solution qui est basée sur un algorithme d’apprentissage par renforcement profond, qui est le gradient de politique déterministe et profond (en anglais *Deep Deterministic Policy Gradient*, DDPG). Les performances de la solution proposée ont été comparées à une solution similaire à celle proposée, mais qui n’optimise pas l’allocation de la bande passante et une autre solution qui calcule l’altitude de vol qui permet au UAV d’arriver à sa destination en épuisant son énergie. Les résultats de simulation ont montré que la solution proposée est de loin plus équitable que les deux solutions de référence. De plus, elle permet de garder le même indice d’équité en augmentant le nombre d’utilisateurs dans le système.

## 2.5 Conclusion

Dans ce chapitre, nous avons présenté les travaux de recherche liés à notre problématique, l’allocation de ressources et de positionnement d’UAV dans les réseaux sans fil, notamment les travaux qui l’optimisent par des méthodes mathématiques, les solutions basées sur l’algorithme génétique et les solutions d’apprentissage par renforcement profond. D’après notre étude des travaux de l’état de l’art, nous avons remarqué que le problème d’optimisation conjointe des position des UAVs et de l’allocation des ressources a été abordé dans plusieurs travaux, soit dans la direction

UL soit dans la direction DL. Cependant, considérer à la fois les communications UL et DL dans le but de minimiser le nombre des UAVs nécessaires pour satisfaire toutes les exigences strictes de tous les utilisateurs dans un contexte de communication d'urgence n'a pas encore été abordé. Dans le chapitre suivant, nous allons définir notre modèle de système et formuler le problème étudié.





## CHAPITRE III

### MODÈLE DU SYSTÈME ET FORMULATION DU PROBLÈME D'OPTIMISATION

Ce chapitre présente le modèle de système utilisé ainsi que la formulation du problème étudié. Nous considérons un réseau sans fil assisté par des UAVs déployés dans une zone de catastrophe pour assurer une communication d'urgence entre des médecins distants et des robots chirurgicaux sur terrain ayant pour tâche d'effectuer les soins de première aide. L'objectif de notre travail est de minimiser le nombre d'UAVs à déployer tout en assurant la communication entre tous les robots et les médecins et en satisfaisant des contraintes strictes sur les probabilités de coupure des débits de transmission de données. Il s'agit d'un problème d'optimisation conjointe d'allocation de puissances de transmission, d'association des utilisateurs et de positionnement des UAVs. Le problème est formulé sous la forme d'un problème non linéaire partiellement en nombres entiers et il est prouvé être un problème  $\mathcal{NP}$ -difficile par restriction.

#### 3.1 Modèle du système

Dans un contexte de désastre naturel, l'infrastructure réseau déjà existante peut être complètement ou partiellement démolie pour pouvoir assurer une communication

d'urgence et sortir la zone en désastre de l'isolement. Nous considérons un réseau sans fil assisté par des UAVs où  $N$  UAV peuvent voler sur une zone couverte par une station de base macro (en anglais *macro base station*, MBS) située au centre pour servir  $U$  utilisateurs qui représentent les robots médicaux contrôlés à distance. Les UAVs sont utilisés comme des relais entre la MBS et les utilisateurs. Soient  $\mathcal{N} = \{1, 2, \dots, N\}$  et  $\mathcal{U} = \{1, 2, \dots, U\}$  qui représentent, respectivement, l'ensemble des UAV disponibles et l'ensemble des utilisateurs. Pour offrir des débits élevés de transmission, nous considérons des communications par onde millimétrique. La bande passante disponible en réseau d'accès pour les communications UL est supposée égale à celle disponible pour les communications DL. Soient  $B_{fh}$  et  $B_{bh}$  les bandes passantes disponibles dans les liaisons du réseau d'accès (c'est-à-dire entre les utilisateurs et les UAVs) et dans les liaisons du réseau d'amenée (c'est-à-dire entre les UAVs et la MBS), respectivement. Sans perte de généralité, nous supposons que le nombre de canaux UL,  $C_{ul}$ , est égal au nombre de canaux DL,  $C_{dl}$ , c'est-à-dire que  $C = C_{ul} = C_{dl}$ . Par conséquent,  $B_{fh}$  est divisé en  $C$  canaux dans chaque direction. Un exemple de ce modèle de système avec deux UAVs, une MBS et des utilisateurs est donné dans la Figure 3.1. Les UAVs déployés sont supposés voler à une altitude fixe  $H_{uav}$ . Soit  $q_n = (x_n, y_n, H_{uav})$ ,  $q_u = (x_u, y_u, 0)$  et  $q_{bs} = (x_{BS}, y_{BS}, H_{bs})$  les positions de l'UAV  $n$ , de l'utilisateur  $u$  et de la MBS respectivement, où  $H_{bs}$  est l'altitude de la station de base. La distance entre l'UAV  $n$  et l'utilisateur  $u$  et la distance entre l'UAV  $n$  et la MBS sont respectivement notées  $d_{n,u} = d_{u,n} = \|q_n - q_u\|$  et  $d_{BS,n} = d_{n,BS} = \|q_{bs} - q_n\|$  où  $\|\cdot\|$  dénote la norme euclidienne. Pour plus de clarté, nous résumons les notations utilisées dans le tableau 3.1.

TABLEAU 3.1 Liste des symboles et leurs définitions

Symboles	Définitions
$\mathcal{N}, N$	Ensemble / Nombre des UAV
$\mathcal{U}, U$	Ensemble / Nombre des utilisateurs
$\mathcal{C}_{ul}, C_{ul}$	Ensemble / Nombre des canaux disponibles en UL
$\mathcal{C}_{dl}, C_{dl}$	Ensemble / Nombre des canaux disponibles en DL
$q_n$	Position de l'UAV $n$
$q_n$	Position de l'utilisateur $u$
$q_{BS}$	Position de MBS
$d_{n,BS} = d_{BS,n}$	Distance entre MBS et UAV $n$
$d_{n,u} = d_{u,n}$	Distance entre utilisateur $u$ et UAV $n$
$h_{BS,n}$	Coefficient du canal DL à petite échelle entre MBS et UAV $n$
$h_{n,BS}$	Coefficient du canal UL à petite échelle entre MBS et UAV $n$
$h_{u,n}$	Coefficient du canal UL à petite échelle entre utilisateur $u$ et UAV $n$
$h_{n,u}$	Coefficient du canal DL à petite échelle entre utilisateur $u$ et UAV $n$
$g_{BS,n}$	Gain de canal DL entre MBS et UAV $n$
$g_{n,BS}$	Gain de canal UL entre MBS et UAV $n$
$g_{u,n}$	Gain de canal UL entre utilisateur $u$ et UAV $n$
$g_{n,u}$	Gain de canal DL entre utilisateur $u$ et UAV $n$
$e_{BS,n}$	Erreur d'estimation du canal DL entre MBS et UAV $n$
$e_{n,BS}$	Erreur d'estimation du canal UL entre MBS et UAV $n$
$e_{u,n}$	Erreur d'estimation du canal UL entre utilisateur $u$ et UAV $n$
$e_{n,u}$	Erreur d'estimation du canal DL entre utilisateur $u$ et UAV $n$
$\sigma_e^2$	Variance de l'erreur de l'estimation du canal
$I_{u,n,c_{ul}}$	Interférence du signal transmis de l'utilisateur $u$ au UAV $n$ sur le canal $c_{ul}$
$I_{n,u,c_{dl}}$	Interférence du signal transmis de l'UAV $n$ à l'utilisateur $u$ sur le canal $c_{dl}$
$\mathbf{b} = [b_n]$	Vecteur d'activation des UAV
$\mathbf{Z} = [Z_{u,n,c_{ul},c_{dl}}]$	Matrice d'association UAV-utilisateur
$p_{u,n}$	Puissance de transmission de l'utilisateur $u$ au UAV $n$
$p_{n,u}$	Puissance de transmission de l'UAV $n$ à l'utilisateur $u$
$p_{BS,n}$	Puissance de transmission du MBS au UAV $n$
$p_{n,BS}$	Puissance de transmission du UAV $n$ au MBS
$R_{n,u,c}, O_{n,u}$	Débit DL entre UAV $n$ et utilisateur $u$ et sa probabilité de coupure
$R_{u,n,c}, O_{u,n}$	Débit UL entre UAV $n$ et utilisateur $u$ et sa probabilité de coupure
$R_{BS,n}, O_{BS,n}$	Débit DL entre UAV $n$ et MBS et sa probabilité de coupure
$R_{n,BS}, O_{n,BS}$	Débit UL entre UAV $n$ et MBS et sa probabilité de coupure

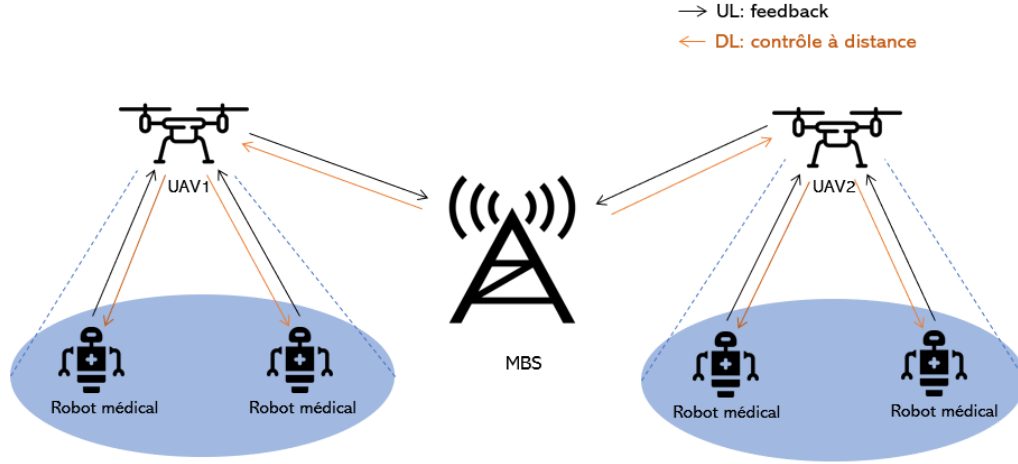


FIGURE 3.1 Modèle du système.

### 3.1.1 Modèle du canal de propagation

Nous considérons un environnement d'évanouissement de *Nakagami-m*, similairement à (Humadi *et al.*, 2021). Nous notons le coefficient d'évanouissement à petite échelle des liens UAV-MBS et UAV-utilisateur par  $h_{\diamond}$ ,  $\diamond \in \{(n, u), (u, n), (n, BS), (BS, n)\}$ , qui suit la distribution de *Nakagami-m*. Le gain de canal à petite échelle des ondes millimétriques  $|h_{\diamond}|^2 \sim \Gamma(m_{\star}, 1/m_{\star})$  est une variable aléatoire *Gamma* indépendante et identiquement distribuée (i.i.d.) avec  $\star \in \{LoS, NLoS\}$  et  $m_{\star}$  sont les paramètres *Nakagami* pour les liaisons en visibilité directe (en anglais *line of sight*, LoS) et les liaisons hors visibilité directe (en anglais *non-line of sight*, NLoS).

Nous supposons que les utilisateurs, les UAVs et la MBS sont équipés de réseaux d'antennes, qui utilisent la technique de formation de faisceaux (en anglais

*beam-forming*), hautement directionnels où un modèle d'antenne sectorielle largement utilisé (Di Renzo, 2015) est considéré. Le motif du réseau d'antennes utilise différentes valeurs  $G_m$  et  $G_s$  pour le lobe principal et les lobes secondaires. Le gain de l'antenne  $G_\diamond$  est modélisé comme :

$$G_\diamond = \begin{cases} G_m & \text{avec probabilité } \theta/2\pi \\ G_s & \text{avec probabilité } (2\pi - \theta)/2\pi \end{cases}, \quad (3.1)$$

où  $\theta$  est l'angle de l'antenne d'émission. Enfin, le gain du canal à ondes millimétriques  $g_\diamond$  est donné par :

$$g_\diamond = G_\diamond f^{-1}(d_\diamond) |h_\diamond|^2, \quad (3.2)$$

où  $f$  est la fonction d'affaiblissement de propagation (en anglais *pathloss*, PL) exprimée comme suit :

$$10 \log_{10}(f(d_\diamond)) = P_{LoS} PL^{LoS}(d_\diamond) + (1 - P_{LoS}) PL^{NLoS}(d_\diamond), \quad (3.3)$$

avec  $P_{LoS}$  est la probabilité d'avoir une liaison LoS (Al-Hourani *et al.*, 2014). Le modèle de propagation utilisé pour calculer l'affaiblissement est donné par (Akdeniz *et al.*, 2014) :

$$PL^*(d_\diamond) = \alpha^* + 10\beta^* \log_{10}(d_\diamond) + \xi [dB], \xi \sim \mathcal{N}(0, \sigma_\xi^{*2}), \quad (3.4)$$

où  $\alpha^*$  et  $\beta^*$  sont les ajustements par les moindres carrés de l'interception flottante, de la pente sur les distances mesurées, respectivement.  $\xi$  est l'ombrage (en anglais *shadowing*) qui suit la loi log-normale  $\mathcal{N}$  de moyenne nulle et de variance  $\sigma_\xi^*$ .

### 3.1.2 Modèle du canal de propagation avec une information imparfaite sur l'état du canal

Afin d'exécuter des algorithmes d'allocation des ressources centralisés, nous supposons que la MBS dispose d'une connaissance parfaite de l'information d'état de canal (en anglais *channel state information*, CSI). Cependant, en raison des fluctuations de canal, du retour limité, de l'estimation de canal et des erreurs de quantification, il est difficile d'obtenir une connaissance parfaite de la CSI en pratique. Nous supposons que la CSI est imparfaitement connue et que le gain de canal estimé à la destination est donné par (Fang *et al.*, 2017) :

$$\tilde{h}_\diamond = h_\diamond + e_\diamond, \quad (3.5)$$

où  $e_\diamond$  est l'erreur d'estimation de canal qui suit une distribution gaussienne complexe centrée à zéro avec une variance  $\sigma_e^2$  (c'est-à-dire  $e_\diamond \sim \mathcal{CN}(0, \sigma_e^2)$ ). Par conséquent, le gain de canal à onde millimétrique avec une CSI imparfaite est  $\tilde{g}_\diamond = G_\diamond f^{-1}(d_\diamond) |\tilde{h}_\diamond|^2$ .

### 3.1.3 Modèle de transmission

Nous considérons la technique d'accès multiple par répartition de fréquences orthogonales (en anglais *orthogonal frequency-division multiple access*, OFDMA) pour gérer les communications simultanées en réseau d'accès dans les deux sens de communication. Nous supposons que les canaux UL et DL sont parfaitement orthogonaux. Pour modéliser le problème d'optimisation, nous introduisons les variables binaires suivantes :

$$b_n = \begin{cases} 1 & \text{si UAV } n \text{ est déployé} \\ 0 & \text{sinon} \end{cases}, \text{ et} \quad (3.6)$$

$$z_{u,n,c_{ul},c_{dl}} = \begin{cases} 1 & \text{si utilisateur } u \text{ est associé à UAV } n \text{ sur les canaux } c_{ul} \text{ et } c_{dl} \\ 0 & \text{sinon} \end{cases}. \quad (3.7)$$

Étant donné que la bande passante en ondes millimétriques est partagée par tous les UAV, les communications sur le même canal seront sujettes à des interférences. Le signal transmis de l'utilisateur  $u$  au UAV  $n$  sur le canal  $c_{ul}$  subirait des interférences exprimées comme suit :

$$I_{u,n,c_{ul}} = \sum_{u' \in \mathcal{U}'} \sum_{n' \in \mathcal{N}'} \sum_{c_{dl} \in \mathcal{C}_{dl}} z_{u',n',c_{ul},c_{dl}} p_{u',n'} g_{u',n}, \quad (3.8)$$

et l'interférence reçue par l'utilisateur  $u$  qui est associé au UAV  $n$  et utilisant le canal  $c_{dl}$  peut être donnée comme suit :

$$I_{n,u,c_{dl}} = \sum_{u' \in \mathcal{U}'} \sum_{n' \in \mathcal{N}'} \sum_{c_{ul} \in \mathcal{C}_{ul}} z_{u',n',c_{ul},c_{dl}} p_{n',u'} g_{n',u}, \quad (3.9)$$

où  $\mathcal{U}' = \mathcal{U} \setminus \{u\}$  et  $\mathcal{N}' = \mathcal{N} \setminus \{n\}$ .

Les valeurs moyennes, notées par  $\mathbb{E}$ , de l'interférence entre l'UAV  $n$  et l'utilisateur  $u$  dans les liaisons UL et DL peuvent être, respectivement, données par :

$$\mathbb{E}(I_{u,n,c_{ul}}) = \sum_{u' \in \mathcal{U}'} \sum_{n' \in \mathcal{N}'} \sum_{c_{dl} \in \mathcal{C}_{dl}} z_{u',n',c_{ul},c_{dl}} p_{u',n'} G_{u',n} f^{-1}(d_{u',n}), \quad (3.10)$$

et

$$\mathbb{E}(I_{n,u,c_{dl}}) = \sum_{u' \in \mathcal{U}'} \sum_{n' \in \mathcal{N}'} \sum_{c_{ul} \in \mathcal{C}_{ul}} z_{u',n',c_{ul},c_{dl}} p_{n',u'} G_{n',u} f^{-1}(d_{u,n'}). \quad (3.11)$$

Finalement, le débit de données UL réalisable de l'utilisateur  $u$  vers l'UAV  $n$  sur le canal  $c_{ul}$  est calculé comme suit :

$$R_{u,n,c_{ul}} = W_{fh} \log_2 \left( 1 + \frac{p_{u,n} |h_{u,n}|^2 f^{-1}(d_{u,n}) G_m}{I_{u,n,c_{ul}} + N_0} \right), \quad (3.12)$$

où  $W_{fh} = B_{fh}/C$  est la bande passante du canal en réseau d'accès,  $N_0 = \sigma^2 W_{fh}$  est la puissance de bruit blanc gaussien additif (AWGN),  $\sigma^2$  est la densité spectrale de puissance du bruit blanc gaussien additif et  $p_{u,n}$  est la puissance de transmission. Le débit de données DL réalisable depuis l'UAV  $n$  vers l'utilisateur  $u$  sur le canal  $c_{dl}$ , noté  $R_{n,u,c_{dl}}$ , peut être calculé en substituant respectivement  $u, n$  et  $c_{ul}$  dans (3.12) par  $n, u$  et  $c_{dl}$ .

Les canaux du réseau d'accès et du réseau d'amenée sont supposés parfaitement orthogonaux. De plus, nous considérons que les UAVs communiquent avec la MBS par une technique d'accès multiple orthogonale. Le débit de données UL réalisable entre l'UAV  $n$  et le MBS peut être donné par :

$$R_{n,BS} = W_{bh} \log_2 \left( 1 + \frac{p_{n,BS} |h_{n,BS}|^2 f^{-1}(d_{n,BS}) G_m}{N'_0} \right), \quad (3.13)$$

où  $W_{bh} = B_{bh} / \sum_{n \in \mathcal{N}} b_n$  est la bande passante du canal en réseau d'amenée,  $p_{n,BS}$  est la puissance de transmission et  $N'_0 = \sigma^2 W_{bh}$ . Le débit de données DL réalisable depuis le MBS vers l'UAV  $n$ , noté  $R_{BS,n}$ , peut être calculé en substituant  $n, BS$  dans (3.13) par  $BS, n$ .

La probabilité de coupure est définie comme la probabilité que le débit de transmission soit inférieur au débit seuil requis. De façon similaire à (Chen *et al.*, 2021),



la probabilité de coupure en UL pour l'utilisateur  $u$  associé à l'UAV  $n$ , (dénotée par  $O_{u,n}$ ) peut être donnée par :

$$\begin{aligned}
O_{u,n} &= \mathbb{P}(R_{u,n,c_{ul}} \leq \zeta_{fh}) \\
&= \mathbb{P}\left(\frac{p_{u,n}|h_{u,n}|^2 f^{-1}(d_{u,n})G_m}{I_{u,n,c_{ul}} + N_0} \leq \gamma_{th,u,n}\right) \\
&= F\left(\frac{\gamma_{th,u,n}(I_{u,n,c_{ul}} + N_0)}{p_{u,n}G_m f^{-1}(d_{u,n})}\right) \\
&\approx \frac{1}{\Gamma(m_\star)}\gamma\left(m_\star, \frac{m_\star\gamma_{th,u,n}(\mathbb{E}(I_{u,n,c_{ul}}) + N_0)}{p_{u,n}G_m f^{-1}(d_{u,n})}\right),
\end{aligned} \tag{3.14}$$

où  $\zeta_{fh}$  est le débit seuil requis pour les transmissions UL et DL dans les liaisons en réseau d'accès,  $\gamma_{th,u,n} = 2^{\zeta_{fh}/W_{fh}} - 1$  est le seuil du rapport signal sur interférence plus bruit (en anglais *signal-to-interference-plus-noise ratio*, SINR) requis,  $F$  est la fonction de distribution cumulative (en anglais *cumulative distribution function*, CDF),  $\Gamma$  est la fonction gamma, et enfin  $\gamma$  est la fonction gamma incomplète inférieure.

De même que (3.14), la probabilité de coupure en DL entre l'UAV  $n$  et l'utilisateur  $u$ , notée par  $O_{n,u}$ , peut être exprimée comme suit :

$$O_{n,u} \approx \frac{1}{\Gamma(m_\star)}\gamma\left(m_\star, \frac{m_\star\gamma_{th,n,u}(\mathbb{E}(I_{n,u,c_{dl}}) + N_0)}{p_{n,u}G_m f^{-1}(d_{u,n})}\right), \tag{3.15}$$

où  $\gamma_{th,n,u} = 2^{\zeta_{fh}/W_{fh}} - 1$  est le seuil SINR.

La probabilité de coupure en UL entre l'UAV  $n$  et la MBS,  $O_{n,BS}$ , peut être exprimée comme suit :

$$\begin{aligned}
O_{n,BS} &= \mathbb{P}(R_{n,BS} \leq \zeta_{bh}) \\
&= \mathbb{P}\left(\frac{p_{n,BS}|h_{n,BS}|^2 f^{-1}(d_{n,BS})G_m}{N'_0} \leq \gamma_{th,n,BS}\right) \\
&= F\left(\frac{\gamma_{th,n,BS}N'_0}{p_{n,BS}G_m f^{-1}(d_{n,BS})}\right) \\
&= \frac{1}{\Gamma(m_\star)}\gamma\left(m_\star, \frac{m_\star\gamma_{th,n,BS}N'_0}{p_{n,BS}G_m f^{-1}(d_{n,BS})}\right),
\end{aligned} \tag{3.16}$$

où  $\zeta_{bh}$  est le débit seuil requis pour les liaisons du réseau d'amenée en UL et DL,  $\gamma_{th,n,BS} = 2^{\zeta_{bh}/W_{bh}} - 1$  est le rapport signal à bruit (en anglais *signal-to-noise ratio*, SNR) seuil requis dans les liaisons UL entre l'UAV  $n$  et la MBS.

De la même manière que (3.16), la probabilité de coupure de DL entre l'UAV  $n$  et la MBS,  $O_{BS,n}$ , peut être exprimée par :

$$O_{BS,n} = \frac{1}{\Gamma(m_\star)}\gamma\left(m_\star, \frac{m_\star\gamma_{th,BS,n}N'_0}{p_{BS,n}G_m f^{-1}(d_{n,BS})}\right), \tag{3.17}$$

où  $\gamma_{th,BS,n} = 2^{\zeta_{bh}/W_{bh}} - 1$  est le seuil SNR.

### 3.2 Formulation du problème

L'objectif de notre travail est d'optimiser les positions des UAVs et les ressources allouées pour minimiser le nombre des UAVs déployés tout en assurant une opération d'urgence sous des contraintes strictes sur les probabilités de coupure et les puissances de transmission. Soit  $\mathbf{b} = [b_n]$  le vecteur binaire d'activation des UAV

de taille  $N$  et  $\mathcal{P} = \{\mathbf{P}_{\text{user}}, \mathbf{P}_{\text{UAV}}^{\text{dl}}, \mathbf{P}_{\text{UAV}}^{\text{ul}}, \mathbf{P}_{\text{BS}}\}$  l'ensemble des matrices des valeurs de puissance. L'ensemble de ces matrices sont définies comme suit : la matrice de puissance des utilisateurs de taille  $U \times N$  est  $\mathbf{P}_{\text{user}} = [p_{u,n}]$ , la matrice de puissance de transmission des UAVs vers les utilisateurs de taille  $N \times U$  est  $\mathbf{P}_{\text{UAV}}^{\text{dl}} = [p_{n,u}]$  et les vecteurs  $\mathbf{p}_{\text{UAV}}^{\text{ul}} = [p_{n,BS}]$  et  $\mathbf{p}_{\text{BS}} = [p_{BS,n}]$  de taille  $N$  représentent, respectivement les puissances de transmission des UAVs vers la MBS et les puissances de transmission de la MBS vers les UAVs. Nous notons également la matrice d'association de taille  $U \times N \times C_{ul} \times C_{dl}$  par  $\mathbf{Z} = [z_{u,n,c_{ul},c_{dl}}]$  et le vecteur de taille  $N$  de la localisation des UAV par  $\mathbf{Q} = [q_n]$ .  $p_{BS}^{\text{max}}$ ,  $p_n^{\text{max}}$  et  $p_u^{\text{max}}$  désignent la puissance de transmission maximale de la MBS, la puissance de transmission du nième UAV et la puissance de transmission maximale de chaque utilisateur, respectivement. Le problème d'optimisation est mathématiquement formulé comme suit :

$$\min_{\mathbf{z}, \mathbf{Q}, \mathbf{b}, \mathcal{P}} \sum_{n \in \mathcal{N}} b_n \quad (\text{P1a})$$

$$\text{s.c} \quad b_n \in \{0, 1\}, \forall n \in \mathcal{N}, \quad (\text{P1b})$$

$$z_{u,n,c_{ul},c_{dl}} \in \{0, 1\}, \forall n \in \mathcal{N}, u \in \mathcal{U}, c_{ul} \in \mathcal{C}_{ul}, c_{dl} \in \mathcal{C}_{dl}, \quad (\text{P1c})$$

$$\sum_{n \in \mathcal{N}} p_{BS,n} b_n \leq p_{BS}^{\text{max}}, \quad (\text{P1d})$$

$$b_n \sum_{u \in \mathcal{U}} \sum_{c_{dl} \in \mathcal{C}_{dl}} \sum_{c_{ul} \in \mathcal{C}_{ul}} p_{n,u} z_{u,n,c_{ul},c_{dl}} + p_{n,BS} \leq p_n^{\text{max}}, \forall n \in \mathcal{N}, \quad (\text{P1e})$$

$$p_{u,n} z_{u,n,c_{ul},c_{dl}} \leq p_u^{\text{max}}, \forall n \in \mathcal{N}, u \in \mathcal{U}, c_{ul} \in \mathcal{C}_{ul}, c_{dl} \in \mathcal{C}_{dl}, \quad (\text{P1f})$$

$$\sum_{n \in \mathcal{N}} \sum_{c_{ul} \in \mathcal{C}_{ul}} \sum_{c_{dl} \in \mathcal{C}_{dl}} z_{u,n,c_{ul},c_{dl}} = 1, \forall u \in \mathcal{U}, \quad (\text{P1g})$$

$$z_{u,n,c_{ul},c_{dl}} O_{n,u} \leq \hat{O}_{fh} \text{ et } z_{u,n,c_{ul},c_{dl}} O_{u,n} \leq \hat{O}_{fh},$$

$$\forall n \in \mathcal{N}, u \in \mathcal{U}, c_{ul} \in \mathcal{C}_{ul}, c_{dl} \in \mathcal{C}_{dl}, \quad (\text{P1h})$$

$$b_n O_{n,BS} \leq \hat{O}_{bh} \text{ et } b_n O_{BS,n} \leq \hat{O}_{bh}, \forall n \in \mathcal{N}, \quad (\text{P1i})$$

$$\sum_{u \in \mathcal{U}} \sum_{c_{dl} \in \mathcal{C}_{dl}} z_{u,n,c_{ul},c_{dl}} \leq 1, \forall n \in \mathcal{N}, c_{ul} \in \mathcal{C}_{ul}, \quad (\text{P1j})$$

$$\sum_{u \in \mathcal{U}} \sum_{c_{ul} \in \mathcal{C}_{ul}} z_{u,n,c_{ul},c_{dl}} \leq 1, \forall n \in \mathcal{N}, c_{dl} \in \mathcal{C}_{dl}, \quad (\text{P1k})$$

$$\sum_{u \in \mathcal{U}} \sum_{c_{ul} \in \mathcal{C}_{ul}} \sum_{c_{dl} \in \mathcal{C}_{dl}} z_{u,n,c_{ul},c_{dl}} \leq b_n C, \forall n \in \mathcal{N}, \quad (\text{P1l})$$

$$b_n - z_{u,n,c_{ul},c_{dl}} \geq 0, \forall n \in \mathcal{N}, u \in \mathcal{U}, c_{ul} \in \mathcal{C}_{ul}, c_{dl} \in \mathcal{C}_{dl}, \quad (\text{P1m})$$

$$q_n b_n \neq q_{n'} b_{n'}, \quad \forall n, n' \in \mathcal{N}, n \neq n'. \quad (\text{P1n})$$

Dans ce problème d'optimisation, les contraintes (P1d) et (P1e) contraignent la puissance de transmission totale de la MBS et des UAVs à ne pas dépasser la puissance maximale disponible. Les contraintes (P1f) garantissent que la puissance de transmission des utilisateurs est limitée. Les contraintes (P1g) garantissent qu'un utilisateur  $u$  peut être associé au maximum à un UAV. Les services fiables et à faible latence sont garantis par les contraintes (P1h) et (P1i) où  $\hat{O}_{bh}$  et  $\hat{O}_{fh}$  sont les seuils de probabilité de coupure pour les liaisons du réseau d'accès et les liaisons du réseau d'amenée, respectivement. Les contraintes (P1j) et (P1k) garantissent que chaque sous-canal UL et DL d'un UAV est utilisé par un seul utilisateur. Les contraintes (P1l) et (P1m) expriment la relation binaire entre le déploiement des UAVs et l'association des utilisateurs. La contrainte (P1n) garantit que deux UAVs déployés ne peuvent pas être placés à la même position.

Le problème (P1) est un problème de programmation non linéaire, non convexe et à nombres entiers mixtes en raison de la fonction logarithmique et des contraintes qui impliquent la matrice de décision  $\mathbf{Z}$  et le vecteur de décision  $\mathbf{b}$ .

### 3.3 Analyse de complexité

Dans ce qui suit, la  $\mathcal{NP}$ -difficulté de (P1) est étudiée en réduisant le problème de localisation d'installation (en anglais *Facility Location Problem*, FLP), qui est un problème connu comme  $\mathcal{NP}$ -difficile, à un cas particulier d'UPRAP. Le lemme suivant prouve ce résultat (Li et Huang, 2008).

*Lemme* : UPRAP formulé dans (P1) est  $\mathcal{NP}$ -difficile.

*Preuve* : Nous prouvons que UPRAP est  $\mathcal{NP}$ -difficile par restriction, c'est-à-dire, nous montrons que UPRAP contient FLP comme cas particulier. Dans FLP, nous avons un ensemble de  $\mathcal{J} = j, \dots, J$  de villes, un ensemble  $\mathcal{F} = i, \dots, F$  de sites potentiels pour les installations. Le coût d'ouverture d'une installation à l'emplacement  $i \in \mathcal{F}$ , si elle est utilisée pour servir un certain nombre de villes, est  $f_i$ . Le coût de connexion d'une ville  $j \in \mathcal{J}$  à l'installation  $i$  est  $c_{ji}$ . De plus, le problème FLP est généralement formulé comme un programme à entiers binaires avec des variables binaires.  $x_{ji}$  indique si la ville  $j$  est connectée à l'installation  $i$  et  $y_i$  indique si l'installation  $i$  est ouverte ou non. Le but principal de FLP est de décider où placer un certain nombre d'installations, afin de couvrir les besoins de service pour un certain nombre de villes, de la manière la plus efficace en minimisant les coûts de placement des différentes installations.

Étant donné une instance de FLP, une instance de UPRAP peut être construite

comme suit :  $F = N$ ,  $J = U$ ,  $c_{ji} = 0$ ,  $f_i = 1$  si l'UAV  $i$  est déployé,  $x_{ji} = z_{j,i,c_{ul},c_{dl}}$  et  $y_i = b_i \forall i \in \mathcal{N}$  et  $j \in \mathcal{U}$ . Nous supposons que il n'y a pas de restriction de puissance (contraintes (P1d), (P1e) et (P1f)), de sorte que toutes les exigences de qualité de service peuvent être satisfaites. Nous supposons également que chaque utilisateur peut être associé à plusieurs UAVs (contraintes (P1g)), chaque canal peut être utilisé par plusieurs utilisateurs (contraintes (P1j) et (P1k)), et enfin les UAVs peuvent prendre la même position mais à des altitudes différentes (contraintes (P1n)). Ces hypothèses impliquent que les contraintes mentionnées ne sont pas applicables. Sous cette restriction, UPRAP équivaut à minimiser le nombre d'UAVs déployés pour satisfaire tous les utilisateurs, sous contraintes de QoS. Par analogie, si les UAVs représentent les installations, les utilisateurs représentent les villes, le coût de déploiement de l'UAV  $i$  est  $f_i = 1$ , et le coût d'association de l'utilisateur  $j$  au UAV  $i$  est  $c_{ji} = 0$ , alors UPRAP devient équivalent à FLP. Puisque FLP est  $\mathcal{NP}$ -difficile, alors UPRAP l'est également. Ceci prouve le lemme.

Résoudre (P1) n'est pas mathématiquement direct et est coûteux en termes de calcul car toutes les combinaisons possibles d'association utilisateur-UAV, d'allocation de puissance et de positions des UAVs doivent être évaluées. Par conséquent, des solutions algorithmiques efficaces sont proposées pour résoudre le problème.

### 3.4 Conclusion

Dans ce chapitre, nous avons défini notre modèle de système, formulé le problème et étudié sa  $\mathcal{NP}$ -difficulté. Dans le chapitre suivant, nous allons présenter les trois solutions proposées. La première suit une approche gloutonne, la deuxième adopte une approche basée sur l'algorithme génétique alors que la troisième est basée sur

une approche d'apprentissage par renforcement profond.





## CHAPITRE IV

### SOLUTIONS PROPOSÉES

Le problème formulé dans la section précédente est un problème  $\mathcal{NP}$ -difficile de nature MINLP non convexe. En raison de la grande complexité de calcul pour résoudre ce type de problème, nous proposons trois différentes approches pour le résoudre avec une complexité réduite.

La première approche est un algorithme glouton qui permet de déployer les UAVs un à un tout en associant les utilisateurs et attribuant les puissances de transmission pour satisfaire les QoS requises. La deuxième approche est un algorithme génétique (GA). Dans cette approche, l'ensemble des individus dans une génération d'une population représentent des solutions faisables pour le problème. À chaque itération (ou génération), les individus subissent des changements/modifications génétiques par croisement et mutation. Après plusieurs générations, les meilleurs individus de la population convergent vers des individus représentant une solution optimale ou sous-optimale. La troisième approche est une solution basée sur l'apprentissage par renforcement profond. Par la suite, nous étudions les complexités de calcul des approches proposées.

#### 4.1 Algorithme glouton pour le déploiement des UAVs, association des utilisateurs et allocation de puissances

Le problème formulé est prouvé être  $\mathcal{NP}$ -difficile qui ne peut pas se résoudre en un temps polynomiale. De ce fait, une solution heuristique peu complexe est proposée.

L'algorithme heuristique résout le problème de façon gloutonne. L'algorithme déploie un UAV à la fois en fonction du nombre d'utilisateurs servis. En effet, tant que les utilisateurs ne sont pas encore tous associés à un UAV, l'algorithme déploie un nouveau UAV. Le déploiement d'un UAV est sujet au respect de plusieurs contraintes sur les puissances de transmissions ainsi que les probabilités de coupure en réseau d'amenée et en réseau d'accès. L'association des utilisateurs aux UAVs doit également respecter des contraintes sur les puissances de transmission et les probabilités de coupure.

##### 4.1.1 Description de l'algorithme glouton

Nous supposons que la MBS dispose de toute l'information sur le nombre d'utilisateurs et les gains du canal. L'algorithme proposé (appelé algorithme de déploiement et de positionnement des UAVs, ou UDPA) déploie les UAVs et choisit leurs positions. Le pseudo-code est donné dans l'Algorithme 1 où les notations suivantes sont utilisées :  $\mathcal{U}_{nas}$  est l'ensemble des utilisateurs pas encore associés,  $\mathcal{Q}$  est l'ensemble des positions possibles,  $p_{BS}$  est la puissance disponible de la MBS et  $p_n$  est la puissance disponible du UAV  $n$ .

L'idée principale de UDPA est de déployer les UAVs un par un tout en choisissant la meilleure position à partir de  $\mathcal{Q}$ , qui maximise le nombre d'utilisateurs associés, jusqu'à ce que toutes les contraintes (P1) soient satisfaites. Puisque tous les utilisateurs doivent être satisfaits, l'algorithme itère sur l'ensemble des utilisateurs non satisfaits, en ligne 3, et ne termine que lorsque cette contrainte est satisfaite. Si cette contrainte n'est pas satisfaite, alors le problème est considéré comme un problème infaisable. A la ligne 1, UDPA initialise le compteur d'UAVs déployés  $n$  à 0. À la ligne 2, tant qu'il y a des utilisateurs non encore assignés, UDPA ajoute un UAV supplémentaire, puis la boucle **pour** (ligne 4) sélectionne sa meilleure position. À la ligne 5, UDPA alloue les niveaux de puissance d'émission minimum aux UAVs déployés qui satisfont les exigences de probabilité de coupure, c'est-à-dire les contraintes (P1d) et (P1i). Avant la boucle **pour** (ligne 9), UDPA trie les utilisateurs non encore associés dans  $\mathcal{U}_{nas}$  selon leurs gains de canal décroissants (stockés dans une matrice). Ensuite, la procédure d'association commence en itérant sur les canaux et les utilisateurs disponibles dans les lignes 9–20. Si les UAVs déployés ne peuvent pas satisfaire toutes les contraintes de (P1), davantage d'UAVs seront déployés. Après chaque association,  $\mathcal{U}_{nas}$ ,  $p_{BS}$  et  $p_n$  sont mis à jour. UDPA s'arrête lorsque tous les utilisateurs sont satisfaits.

#### 4.1.2 Analyse de complexité

UDPA est en exécution tant qu'il reste des utilisateurs non associés (l'ensemble des utilisateurs non associé  $\mathcal{U}_{nas}$  est non vide) ou il a dépassé le nombre maximal d'UAVs disponibles. Pour chacune des itérations, il ajoute un nouveau UAV et cherche la meilleure position possible de l'ensemble de positions disponibles par une boucle **pour** sur  $\mathcal{Q}$ . Pour chaque position, les utilisateurs non associés sont triés par ordre dé-

---

**Algorithme 1** Algorithme UDPA

**Initialiser :**  $\mathbf{P}_{u,n} \leftarrow \mathbf{0}$ ,  $\mathbf{P}_{n,u} \leftarrow \mathbf{0}$ ,  $\mathbf{p}_{n,BS} \leftarrow \mathbf{0}$ ,  $\mathbf{p}_{BS,n} \leftarrow \mathbf{0}$ ,  $\mathbf{Z} \leftarrow \mathbf{0}$ ,  $\mathbf{b} \leftarrow \mathbf{0}$ ,  
 $p_{BS} \leftarrow p_{BS}^{max}$ ,  $p_n \leftarrow p_n^{max}$ ,  $\mathcal{U}_{nas} \leftarrow \mathcal{U}$ .  
1:  $n \leftarrow 0$   
2: **Tant que**  $\mathcal{U}_{nas}$  est non vide et  $n < N$  **faire**  
3:    $n \leftarrow n + 1$   
4:   **Pour** chaque position  $q$  dans l'ensemble de positions  $\mathcal{Q}$  **faire**  
5:     Attribuer  $p_{n,BS}$  et  $p_{BS,n}$  minimales pour satisfaire (P1i).  
6:     **Si** (P1d) **et** (P1i) **alors**  
7:       Mettre à jour  $P_{BS}$  et  $p_n$ .  
8:       Trier  $\mathcal{U}_{nas}$  selon l'ordre décroissant des gains de canaux.  
9:       **Pour**  $u \in \mathcal{U}_{nas}$  **faire**  
10:          **Pour**  $c_{ul} \in \mathcal{C}_{ul}$  **et**  $c_{dl} \in \mathcal{C}_{dl}$  **faire**  
11:           Attribuer  $p_{u,n}$  et  $p_{n,u}$  minimales pour satisfaire (P1h).  
12:           **Si** (P1e) **et** (P1h) **alors**  
13:             Associer l'utilisateur  $u$  avec UAV  $n$ .  
14:              $\mathcal{U}_{nas} \leftarrow \mathcal{U}_{nas} \setminus \{u\}$ .  
15:             **Si**  $\mathcal{U}_{nas} = \emptyset$  **alors**  
16:               **Finir.**  
17:             **Fin Si**  
18:           **Fin Si**  
19:          **Fin Pour**  
20:       **Fin Pour**  
21:       **Si** nombre des utilisateurs associés au UAV  $n$  est  $C$  **alors**  
22:          Sortir de la boucle pour.  
23:       **Fin Si**  
24:     **Fin Si**  
25:   **Fin Pour**  
26: **Fin Tant que**  
27: **Retourner**  $n$

---

croissant des gains de canaux avec une complexité  $\mathcal{O}(U \log(U))$ . Ensuite, pour chaque utilisateur dans  $\mathcal{U}_{nas}$ , parcouru par une boucle **pour**, l'algorithme cherche le canal UL et le canal DL disponibles pour l'associer. La complexité de calcul dans le pire des cas de UDPA est de  $\mathcal{O}(U \times |\mathcal{Q}| \times (U \log U + U \times C^2))$ .

## 4.2 Algorithme génétique

### 4.2.1 Description de la solution

L'algorithme génétique est une technique d'optimisation inspirée des principes de l'évolution biologique. Elle est utilisée pour trouver une solution optimale ou quasi optimale pour des problèmes d'optimisation non convexes et difficiles à résoudre. GA modifie de manière répétée une population d'individus représentant des solutions possibles et crée une nouvelle génération à chaque étape. Au fil des générations successives, la population évolue vers la solution optimale. Dans notre solution, un individu représentant une solution est composé de 7 gènes qui sont les variables d'optimisation :  $(\mathbf{Z}, \mathbf{Q}, \mathcal{P}, \mathbf{b})$ . Tout d'abord, une population initiale aléatoire d'individus, représentant des solutions réalisables, est créée. La fonction de pertinence (en anglais *fitness function*) de chaque individu/solution  $k$  évalué est défini comme suit :

$$fitness[k] = 1 / \sum_{n=1}^N b_n, \quad (4.1)$$

Les étapes de l'algorithme génétique sont détaillées dans la figure 4.1 La génération suivante est formée à partir de la génération actuelle en subissant les opérations génétiques : sélection par élitisme, croisement et mutation.

— **Sélection par élitisme** : Selon un ratio de sélection  $\gamma_e$ , les individus ayant la

meilleure valeur de pertinence sont choisis pour faire partie de la génération suivante.

- **Croisement** : Les parents potentiels pour un croisement sont sélectionnés par la méthode de la roulette (Davis, 1991). Ensuite, en fonction de la probabilité de croisement  $\gamma_c$ , un croisement à un point (en anglais *one-point crossover*) est effectué par paires sur les individus sélectionnés sur les gènes  $\mathbf{Q}$ . Cette opération représente un échange de positions entre deux UAVs dans les solutions représentées par les parents. Les individus nouvellement formés feront partie de la nouvelle génération s'ils respectent les contraintes (P1).
- **Mutation** : L'opération de mutation est contrôlée par une probabilité  $\gamma_m$ . Un individu est muté en changeant une valeur au hasard dans le gène  $\mathbf{Z}$ . Cela signifie que la mutation consiste à changer l'UAV auquel un utilisateur est déjà associé. L'individu muté remplace l'individu original dans la génération suivante s'il respecte les contraintes (P1).

Une méthode de correction est appliquée après chaque opération de croisement et de mutation pour accélérer la convergence et garantir que les contraintes dans l'équation (P1) sont respectées. La correction après le croisement est effectuée comme suit. Tout d'abord, l'algorithme vérifie si les utilisateurs associés peuvent maintenir leurs associations même après que les positions des UAV aient été modifiées. Dans les cas où les contraintes de QdS et de puissance ne peuvent pas être satisfaites, l'algorithme procède à la recherche d'un canal disponible parmi les UAVs déjà déployés pour associer les utilisateurs désassociés en respectant les contraintes de l'équation (P1). Si tous les utilisateurs sont associés et que toutes les contraintes sont respectées, le nouvel individu formé sera intégré à la génération. Sinon, le parent est conservé pour la génération suivante. La correction après la mutation consiste à (i) vérifier si l'association obtenue après le changement aléatoire est réalisable (le nouveau canal choisi

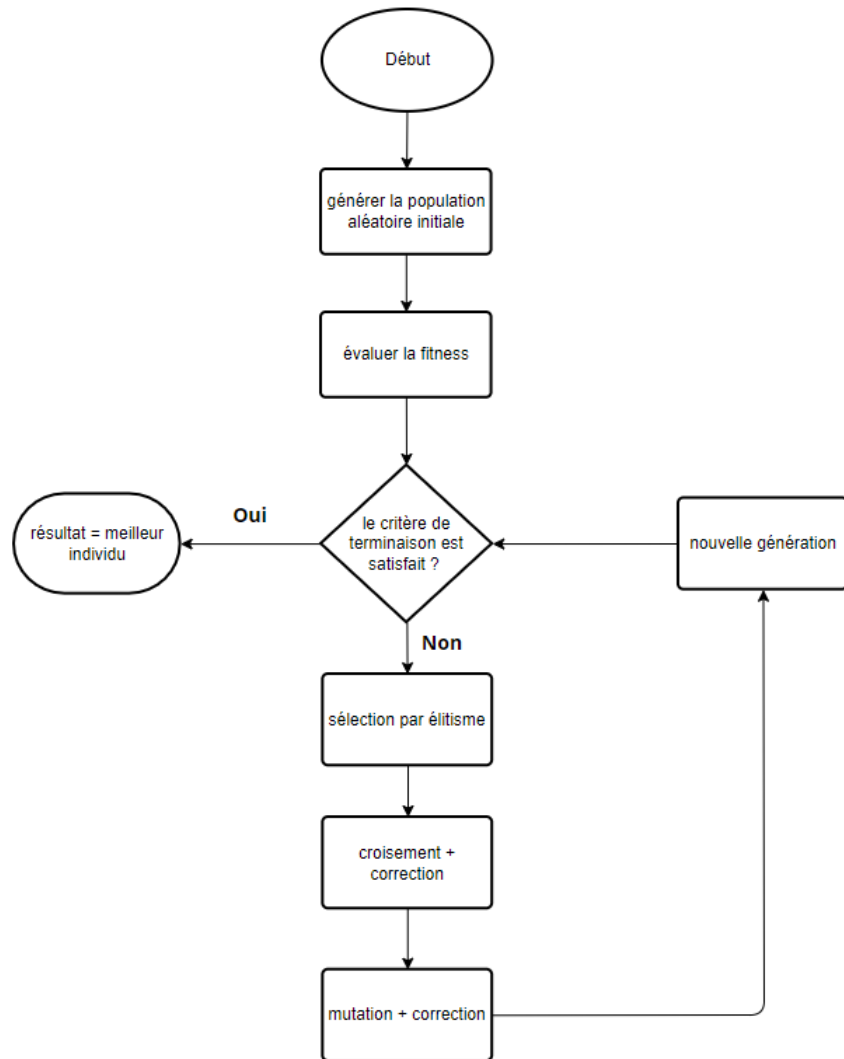


FIGURE 4.1 Étapes de l'algorithme génétique.

aléatoirement est disponible), (ii) ajuster le niveau de puissance alloué en fonction du nouveau gain de canal, et (iii) vérifier si la nouvelle association respecte les seuils de probabilité de défaillance du débit. Si toutes les contraintes de l'équation (P1) sont satisfaites, l'individu muté remplacera l'individu actuel dans la génération suivante.

#### 4.2.2 Analyse de complexité

L'algorithme se termine lorsque le nombre maximum de générations est atteint. La complexité dans le pire des cas de cet algorithme est de  $\mathcal{O}(N_{gen} \times N_{ind} \times (\log N_{ind} + \log N_{ind} + (U + U \times N \times C) + U \times N) + N_{ind} \times O_{pop}^{init})$  avec  $N_{gen}$ , où  $N_{gen}$  est le nombre de générations,  $N_{ind}$  est le nombre d'individus dans une population,  $N_{ind} \times \log N_{ind}$  est la complexité de la sélection par élitisme et est, également, la complexité de la méthode de sélection par roulette,  $N_{ind} \times (U + U \times N \times C)$  est la complexité du croisement et  $N_{ind} \times U \times N$  est la complexité de la mutation. La complexité de la génération d'une population initiale de solutions possibles,  $N_{ind} \times O_{pop}^{init}$ , est aussi prise en considération lors du calcul de la complexité de l'algorithme génétique.

#### 4.3 Apprentissage par renforcement profond

Dans cette section, nous présentons l'algorithme de positionnement d'UAVs et d'allocation des ressources basé sur DDPG (DUPRA). Dans DUPRA, la MBS est considérée comme l'agent. À chaque pas de temps  $t$ , la MBS observe l'état  $s(t)$  et choisit l'action  $a(t)$ . Ensuite, la MBS reçoit la récompense  $r(t)$  et passe à l'état suivant  $s(t+1)$ . Le tuple d'expérience correspondant  $(s(t), a(t), r(t), s(t+1))$  est stocké dans la mémoire de relecture pour entraîner les réseaux.



### 4.3.1 Préliminaires

#### 4.3.1.1 Apprentissage par renforcement

L'apprentissage par renforcement (en anglais *Reinforcement learning*, RL) consiste à apprendre ce qu'il faut faire (comment associer des situations ou des états à des actions) de manière à maximiser un signal de récompense numérique. On ne dit pas à l'apprenant quelles actions il doit entreprendre, mais il doit découvrir les actions les plus gratifiantes en les essayant. Dans les cas les plus intéressants et les plus difficiles, les actions peuvent affecter non seulement la récompense immédiate, mais aussi la situation suivante et, à travers elle, toutes les récompenses ultérieures. Ces deux caractéristiques, à savoir la recherche par essai-erreur et la récompense différée, sont les deux plus importantes caractéristiques de l'apprentissage par renforcement (Sutton et Barto, 2018).

L'un des défis qui se posent dans l'apprentissage par renforcement est le compromis entre exploration et exploitation. Pour obtenir une plus grande récompense, un agent d'apprentissage par renforcement doit privilégier les actions qu'il a essayées dans le passé et qui se sont avérées efficaces pour produire des récompenses. Mais pour découvrir ces actions, il doit essayer des actions qu'il n'a pas sélectionné auparavant. Le dilemme est que ni l'exploration ni l'exploitation ne peuvent être poursuivies exclusivement sans échouer dans la tâche. L'agent doit essayer une variété d'actions et favoriser progressivement celles qui semblent les meilleures.

Au-delà de l'agent et de l'environnement, on peut identifier quatre sous-éléments principaux d'un système d'apprentissage par renforcement : une politique, un signal

de récompense, une fonction de valeur et, éventuellement, un modèle de l'environnement.

- ***Politique*** : Une "politique" définit la manière dont l'agent d'apprentissage se comporte à un moment donné. En gros, une politique est une correspondance entre les états perçus de l'environnement et les actions à entreprendre dans ces états.
- ***Signal de récompense*** : Un "signal de récompense" définit le but d'un problème d'apprentissage par renforcement. À chaque pas de temps, l'environnement envoie à l'agent un nombre unique appelé récompense. Le seul objectif de l'agent est de maximiser la récompense totale qu'il reçoit sur le long terme. Le signal de récompense définit donc les bons et les mauvais événements pour l'agent.
- ***Fonction de valeur*** : Alors que le signal de récompense indique ce qui est bon dans un sens immédiat, une "fonction de valeur" spécifie ce qui est bon à long terme. En effet, la valeur d'un état est le montant total de la récompense qu'un agent peut s'attendre à accumuler dans le futur, à partir de cet état.
- ***Modèle de l'environnement*** : Il s'agit de quelque chose qui imite le comportement de l'environnement ou, plus généralement, qui permet de faire des déductions sur le comportement de l'environnement.

#### 4.3.1.2 Apprentissage par renforcement profond

L'apprentissage par renforcement profond est la combinaison de l'apprentissage profond et de l'apprentissage par renforcement. L'apprentissage par renforcement profond combine des méthodes d'apprentissage de problèmes à haute dimension avec l'apprentissage par renforcement, permettant un apprentissage interactif à haute di-

mension (Plaata, 2020). Le domaine de l'apprentissage profond concerne l'approximation de fonctions dans des problèmes à haute dimension, des problèmes si complexes que les méthodes tabulaires ne peuvent plus résoudre. L'apprentissage profond utilise des réseaux de neurones profonds pour trouver des approximations dans des environnements vastes, complexes et hautement dimensionnels.

#### 4.3.1.3 Taxonomie des algorithmes de RL

Les différents types d'algorithmes de RL sont résumés dans la figure 4.2. Il existe deux principaux types d'algorithmes de RL ; un type qui est basé sur un modèle (en anglais *model-based*) et un autre type qui n'est pas basé sur un modèle, ou sans modèle, (en anglais *model-free*). Dans le "*model-based*", un agent cherche à comprendre son environnement et à créer un modèle en interagissant avec cet environnement. Le modèle permet de faire des inférences sur l'environnement. Par exemple, le modèle peut prédire le prochain état résultant et la prochaine récompense, étant donné un état et une action. Dans un tel système, les préférences ont la priorité sur les conséquences des actions, c'est-à-dire que l'agent essaiera toujours d'effectuer une action qui lui rapportera la récompense maximale, indépendamment des conséquences de cette action.

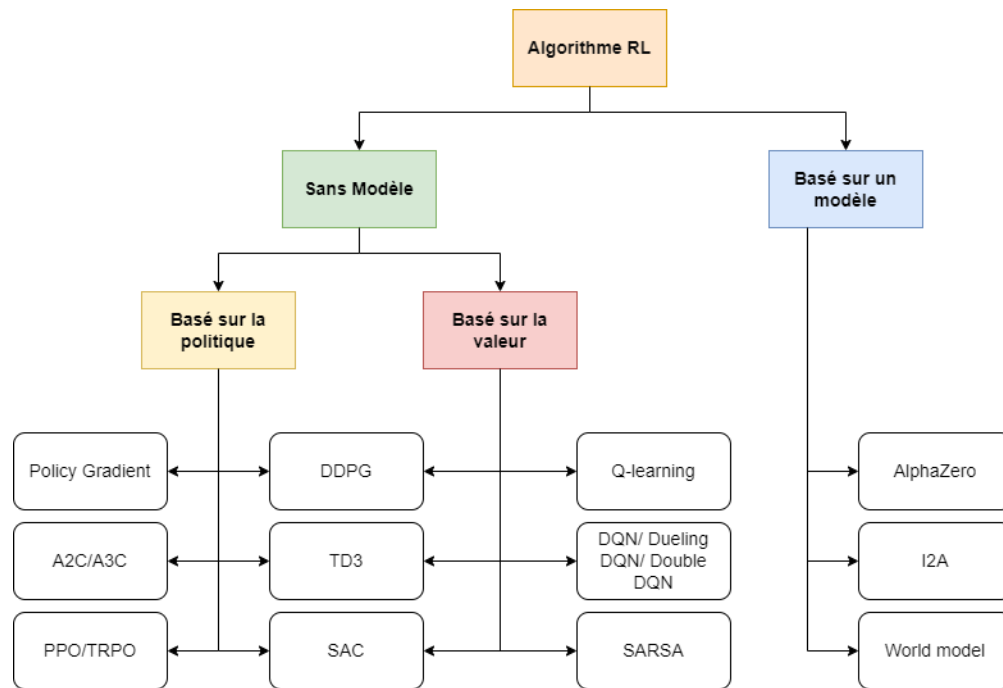


FIGURE 4.2 Taxonomie des algorithmes d'apprentissage par renforcement (Dong *et al.*, 2020).

D'un autre côté, les algorithmes "*model-free*" cherchent à apprendre les conséquences de leurs actions par l'expérience. En d'autres termes, avec un tel algorithme l'agent effectuera une action plusieurs fois et ajustera sa politique pour obtenir des récompenses optimales, en fonction des résultats.

IL existe deux sous-types d'algorithmes "*model-free*" ; des algorithmes basés sur la politique (en anglais *policy-based*) et des algorithmes basés sur la valeur-Q (en anglais *value-based*). Les algorithmes "*value-based*" alternent entre l'estimation de la fonction de valeur dans le cadre de la politique actuelle et l'amélioration (optimisation) de la politique à l'aide de la fonction de valeur estimée. Les algorithmes

"*policy-based*" optimisent directement la politique de l'agent sans estimer ou apprendre une fonction action-valeur. Les valeurs de récompense échantillonnées sont généralement utilisées dans le processus d'optimisation pour améliorer les préférences d'action. Les méthodes basées sur le gradient ou sans gradient sont appliquées dans le processus d'optimisation. Les méthodes basées sur le gradient appliquent toujours le gradient de politique, qui représente la classe d'algorithmes utilisée dans l'apprentissage par renforcement à action continue, bénéficiant de l'extensibilité aux cas de haute dimension. D'autre part, les algorithmes sans gradient ont généralement un processus d'apprentissage plus rapide pour les cas relativement simples de recherche de politique, sans le processus coûteux en calcul des dérivées (Dong *et al.*, 2020).

#### 4.3.1.4 Gradient de la politique déterministe et profond

Nous présentons d'abord brièvement DDPG (Deep Deterministic Policy Gradient) qui est un algorithme d'apprentissage par renforcement profond (en anglais *deep reinforcement learning*, DRL) basé sur l'approche acteur-critique (Konda et Tsitsiklis, 1999). Dans DDPG, le critique  $Q(s, a; \theta^Q)$  évalue la fonction action-valeur sous la politique de l'acteur  $\mu(s; \theta^\mu)$ , où  $\theta^\mu$  et  $\theta^Q$  font référence aux paramètres des réseaux acteurs et des réseaux critiques.

Cependant, un approximateur de fonction non linéaire, par exemple un réseau de neurones profond (en anglais *deep neural network*, DNN), est instable et peut même provoquer une divergence lorsqu'il est appliqué au DRL. Ainsi, deux techniques sont généralement utilisées dans le DRL pour éviter ce problème : la relecture d'expériences et le réseau cible. Le DRL échantillonne un mini lot d'expériences à partir de la mémoire de relecture. Les échantillons aléatoires rompent la corrélation entre les

échantillons successifs et stabilisent le processus d'apprentissage. De plus, les réseaux cibles acteurs et critiques  $\mu'(s; \theta^{\mu'})$  et  $Q'(s, a; \theta^{Q'})$  sont utilisés pour entraîner le réseau critique principal et ont les mêmes architectures que les réseaux principaux et doivent être mis à jour en suivant lentement les réseaux principaux. Plus précisément, le réseau critique peut être entraîné en minimisant la perte suivante :

$$L(\theta^Q) = \frac{1}{B} \sum_t [y_t(t) - Q(s(t), a(t); \theta^Q)]^2, \quad (4.2)$$

où

$$y_t(t) = r(t) + \gamma Q' \left( s(t+1), \mu'(s(t+1); \theta^{\mu'}); \theta^{Q'} \right) \quad (4.3)$$

est la cible de mise à jour,  $\gamma$  est le facteur d'actualisation pour la mise à jour et  $B$  est la taille du lot. En outre, le réseau acteur est entraîné en minimisant la perte de l'acteur :

$$L(\theta^\mu) = \frac{1}{B} \sum_t -Q \left( s(t), \mu(s(t+1); \theta^\mu); \theta^Q \right). \quad (4.4)$$

Les paramètres de mise à jour progressive des réseaux cibles :

$$\theta' \leftarrow \varepsilon \theta + (1 - \varepsilon) \theta', \quad (4.5)$$

avec  $\varepsilon \ll 1$  est le taux de mise à jour (Lillicrap *et al.*, 2015).

L'architecture et les interconnexions entre les différentes composantes de DDPG sont représentées dans la figure 4.3.

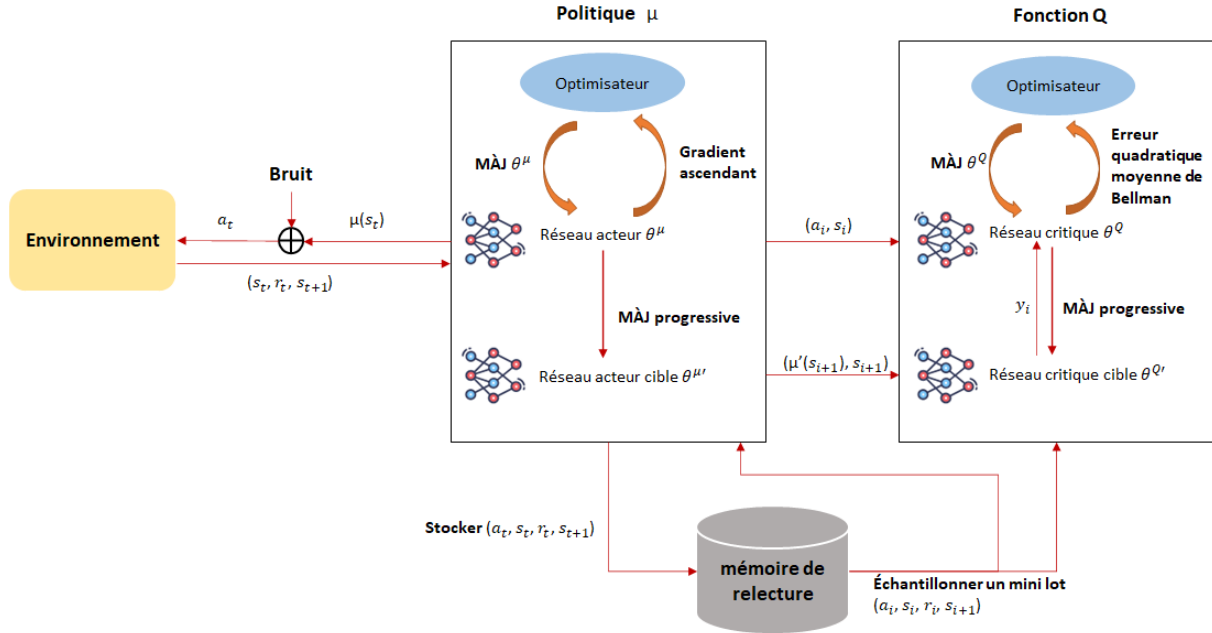


FIGURE 4.3 Architecture de DDPG.

#### 4.3.2 Processus de décision de Markov

Le problème d'optimisation est formulé comme un processus de décision de Markov qui est composé de cinq éléments  $(\mathcal{S}, \mathcal{A}, \mathcal{R}, P, \pi)$ . Comme ce travail vise à minimiser le nombre d'UAVs déployés, une solution multi-agents où les UAVs agissent en tant qu'agents individuels introduit de la complexité. Ainsi, nous considérons une station de base de contrôle centralisée pour déterminer les UAVs déployés. Plus précisément, à chaque intervalle de décision  $t$ , la station de base principale (MBS), qui est considérée comme l'agent central, observe l'état  $s(t)$  dans l'espace des états  $\mathcal{S}$  et choisit une action  $a(t)$  dans l'espace des actions  $\mathcal{A}$  en fonction de sa politique  $\pi$ . L'agent reçoit la récompense immédiate  $r(t)$  en interagissant avec l'environnement. Ensuite, l'agent

passé à l'état suivant  $s(t + 1)$  selon la probabilité  $P$ . Enfin, le tuple d'expérience  $(s(t), a(t), r(t), s(t + 1))$  est stocké dans la mémoire de relecture pour entraîner les réseaux.

- *Espace d'état*  $\mathcal{S}$  : Nous définissons l'état du système à l'instant  $t$  comme  $s(t) \in \mathcal{S}$ , où  $s(t)$  est le nombre d'UAVs déployés à l'instant  $t - 1$ .
- *Espace d'action*  $\mathcal{A}$  : L'action de l'agent consiste à placer des UAVs, où l'action à l'instant  $t$  est  $a(t) = Q(t)$  et  $Q(t)$  représente les emplacements des UAVs à l'instant  $t$ .
- *Récompense*  $\mathcal{R}$  : La récompense est obtenue en fonction de l'action choisie par l'agent. L'objectif est de minimiser le nombre d'UAVs déployés. De plus, pour s'assurer que tous les utilisateurs sont satisfaits par un UAV de service à chaque instant  $t$ , la contrainte de couverture des utilisateurs doit être prise en compte dans la fonction de récompense. Si certains utilisateurs ne sont satisfaits par aucun UAV, une pénalité est encourue par le MBS. La récompense instantanée est calculée comme suit :

$$r(t) = \begin{cases} \frac{1}{\sum_{n=1}^N b_n^{(t)}} & , \text{ si tous les utilisateurs sont satisfaits} \\ c & , \text{ sinon} \end{cases} . \quad (4.6)$$

où  $c < 0$  est une constante négative, représentant une pénalité.

### 4.3.3 Description de l'algorithme DUPRA

L'algorithme de positionnement d'UAVs et d'allocation des ressources basé sur DDPG (DUPRA) est présenté dans l'algorithme 2. Au début de chaque épisode, l'environnement est réinitialisé et aucun UAV n'est déployé dans l'état initial  $s(0)$ .



À chaque pas de temps  $t$ , la MBS choisit une action  $a(t)$  à travers le réseau acteur  $\mu(s(t); \theta^\mu)$ , en ajoutant un bruit d'exploration  $N_e$  selon la ligne 5, afin de promouvoir l'exploration et d'empêcher l'agent de tomber dans une optimisation locale pendant l'entraînement. Nous choisissons un bruit gaussien à moyenne nulle non corrélé, avec un écart-type décroissant au fil du temps. La variance du bruit est initialement fixée à  $N_e^{max}$ . Pour un nombre donné de pas de temps, le bruit d'exploration est annulé et l'agent ne fait plus d'exploration mais sélectionne directement les actions. En fonction de l'action choisie, l'agent déploie les UAVs nécessaires pour servir tous les utilisateurs et attribue les niveaux de puissance de transmission. Si la MBS réussit à satisfaire tous les utilisateurs, elle reçoit une récompense positive telle que décrite ci-dessus et obtient l'état suivant  $s(t + 1)$ . Sinon, l'agent reçoit une récompense négative et l'état suivant est fixé à 0. Ensuite, l'agent stocke le tuple de transition correspondant  $(s(t), a(t), r(t), s(t + 1))$  dans la mémoire de relecture d'expériences  $\mathcal{M}$ .

La mise à jour des réseaux principaux est effectuée aux lignes 11–12. À chaque pas, si la mémoire est pleine, un mini-lot d'expériences est échantillonné à partir de la mémoire d'expérience. Ces échantillons sont utilisés pour entraîner et mettre à jour les réseaux en minimisant les pertes de l'acteur et du critique. Ensuite, les réseaux cibles sont mis à jour en suivant progressivement les réseaux principaux aux lignes 13–15.

#### 4.3.4 Analyse de complexité

Pour analyser la complexité de calcul de DUPRA, nous définissons  $N_l^a$  comme le nombre de neurones dans la  $l$ ème couche du réseau acteur et  $N_k^c$  comme le nombre de neurones dans la  $k$ ème couche du réseau critique. Les réseaux critique et acteur sont

---

**Algorithme 2** Algorithme DUPRA
 

---

- 1: Initialiser aléatoirement le réseau acteur  $\mu(s; \theta^\mu)$ , le réseau acteur cible  $\mu'(s; \theta^{\mu'})$  avec les poids  $\theta^\mu = \theta^{\mu'}$ , le réseau critique  $Q(s, a; \theta^Q)$  et le réseau critique cible  $Q'(s, a; \theta^{Q'})$  avec les poids  $\theta^Q = \theta^{Q'}$ .
  - 2: Initialiser la mémoire de relecture  $\mathcal{M}$ .
  - 3: **Pour chaque** épisode **faire**
  - 4:   **Pour chaque** pas  $t$  **faire**
  - 5:     La MBS choisit l'action  $a(t) = \mu(s(t); \theta^\mu) + N_e$ , où  $N_e$  est le bruit d'exploration.
  - 6:     MBS exécute l'action  $a(t)$ .
  - 7:     Selon  $a(t)$ , les puissances de transmission minimales sont attribuées et les utilisateurs sont associés.
  - 8:     Calculer la récompense  $r(t)$  (4.6) and mettre à jour le nouvel état  $s(t+1)$ .
  - 9:     Stocker le tuple  $(s(t), a(t), s(t+1), r(t))$  dans  $\mathcal{M}$ .
  - 10:    Échantillonner un mini-lot de  $B$  transitions.
  - 11:    Mettre à jour  $\theta^Q$  par minimisation de la perte du réseau critique (4.2).
  - 12:    Mettre à jour  $\theta^\mu$  par minimisation de la perte du réseau acteur (4.4).
  - 13:    Mise à jour progressive pour les réseaux cibles :
  - 14:     $\theta^{Q'} = \varepsilon \theta^Q + (1 - \varepsilon) \theta^{Q'}$ .
  - 15:     $\theta^{\mu'} = \varepsilon \theta^\mu + (1 - \varepsilon) \theta^{\mu'}$ .
  - 16:    **Fin Pour**
  - 17: **Fin Pour**
- 

des réseaux entièrement connectés. La complexité de calcul du réseau d'acteurs est de  $\mathcal{O}(\sum_{l=2}^{L-1} N_{l-1}^a N_l^a + N_l^a N_{l+1}^a)$ , où  $L$  est le nombre de couches des réseaux d'acteurs. De même, la complexité de calcul du réseau critique est de  $\mathcal{O}(\sum_{k=2}^{K-1} N_{k-1}^c N_k^c + N_k^c N_{k+1}^c)$ , où  $K$  est le nombre de couches des réseaux critiques. Les réseaux d'acteurs et de critiques sont formés simultanément et en extrayant  $B$  expériences de la mémoire de relecture pour l'entraînement par rétropropagation. Par conséquent, la complexité de l'algorithme 2 est de  $\mathcal{O}(T \times B \times (\sum_{k=2}^{K-1} N_{k-1}^c N_k^c + N_k^c N_{k+1}^c + \sum_{l=2}^{L-1} N_{l-1}^a N_l^a + N_l^a N_{l+1}^a))$ , où  $T$  est le nombre de pas de temps dans chaque épisode.

#### 4.4 Conclusion

Dans ce chapitre, nous avons présenté les détails de nos trois approches proposées. La première solution est un algorithme glouton qui consiste à déployer les UAVs nécessaires pour assurer la communication dans le réseau et à associer tous les utilisateurs ainsi qu'attribuer les puissances. La deuxième solution est un algorithme basé sur l'algorithme génétique adapté à notre problème. La troisième solution est un algorithme d'apprentissage par renforcement basé sur l'approche du gradient de politique déterministe et profond. Dans le chapitre suivant, nous allons évaluer les performances des trois solutions et discuter des résultats obtenus.



## CHAPITRE V

### RÉSULTATS DE SIMULATION

Ce chapitre présente les résultats de la simulation *Monte-Carlo* pour évaluer les performances de la solution heuristique proposée (UDPA), de la solution génétique (GA) et de la solution basée sur le DRL (DUPRA).

#### 5.1 Paramètres de simulation

Nous considérons une zone circulaire avec un rayon  $R=50\text{m}$ . La MBS est placée au centre de la zone de service, tandis que les utilisateurs (robots télécommandés) sont répartis aléatoirement et uniformément dans la zone. Les valeurs du taux de sélection d'élite, des probabilités de mutation et de croisement sont sélectionnées par simulation. Les hyperparamètres de DDPG (par exemple, le nombre de couches, le nombre de neurones dans les réseaux de l'acteur et du critique, le taux d'apprentissage) sont également choisis par simulations. Le facteur d'actualisation et les valeurs du taux de mise à jour progressive sont choisies de manière similaire à (Lillicrap *et al.*, 2015). Sauf mention contraire, tous les paramètres de simulation énumérés dans les Tableau 5.1, Tableau 5.2 et Tableau 5.3 sont utilisés.

TABLEAU 5.1 Notations et valeurs des paramètres par défaut

Notations	Paramètres	Valeurs
$H_{uav}$	Altitude de vol des UAVs	30 m
$H_{bs}$	Hauteur de l'antenne de la MBS	2 m
$C = C_{ul} = C_{dl}$	Nombre des canaux dans le réseau d'accès	4
$\alpha, \beta, \sigma_\xi$	Paramètres d'affaiblissement de propagation LOS (f=28GHz)	61.4, 2, 5.8
	Paramètres d'affaiblissement de propagation NLOS (f=28GHz)	72, 2.92, 8.7
$m_{LoS}$	Paramètre de <i>Nakagami</i> pour les liens LoS	3
$m_{NLoS}$	Paramètre de <i>Nakagami</i> pour les liens NLoS	2
$\sigma^2$	Densité spectrale de la puissance d'un bruit AWGN	-174 dBm/Hz
$B_{fh}$	Bande passante disponible pour les liens en réseau d'accès	1 GHz
$B_{bh}$	Bande passante disponible pour les liens en réseau d'amenée	10 GHz
$\zeta_{fh}$	Seuil du débit de transmission UL et DL du réseau d'accès	500 Mbps
$\zeta_{bh}$	Seuil du débit de transmission UL et DL du réseau d'amenée	$C \times \zeta_{fh}$
$p_n^{max}$	Puissance de transmission maximale du n-ème UAV	30 dBm
$p_{BS}^{max}$	Puissance de transmission maximale de la MBS	40 dBm
$p_u^{max}$	Puissance de transmission maximale de chaque utilisateur	30 dBm
$\theta$	Largeur de l'angle de l'antenne de transmission	90°
$G_m, G_s$	Gain d'antenne du lobe principale et du lobe secondaire	12 dB, 0 dB

TABLEAU 5.2 Notations et valeurs des paramètres de GA

Notations	Paramètres	Valeurs
$\gamma_e$	Taux de sélection par élitisme	0.1
$\gamma_c$	Probabilité de croisement	0.6
$\gamma_m$	Probabilité de mutation	0.2
$N_{gen}$	Nombre de générations	500
$N_{ind}$	Nombre d'individus dans une génération	50

## 5.2 Comparaison des performances des solutions proposées

La Fig. 5.1 montre la convergence de GA lorsque la population initiale est choisie au hasard et lorsque la solution est obtenue par UDPA. Comme prévu, l'utilisation de l'UDPA comme population initiale permet une convergence plus rapide que l'initialisation aléatoire.

La Fig. 5.2, montre la convergence de DUPRA pour 14 utilisateurs. Il est clair que la récompense moyenne converge rapidement (c'est-à-dire après 200 pas de temps).

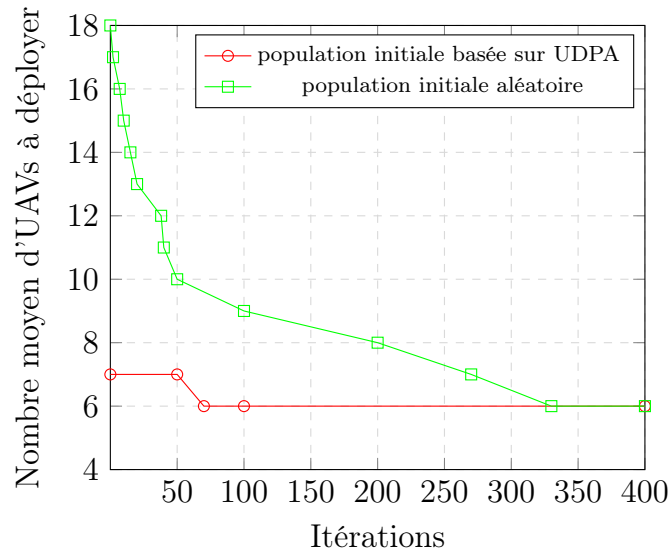


FIGURE 5.1 Convergence de la solution GA,  $\zeta_{fh} = 700$  Mbps,  $\hat{O}_{fh} = \hat{O}_{bh} = 5\%$  and  $U = 18$ .

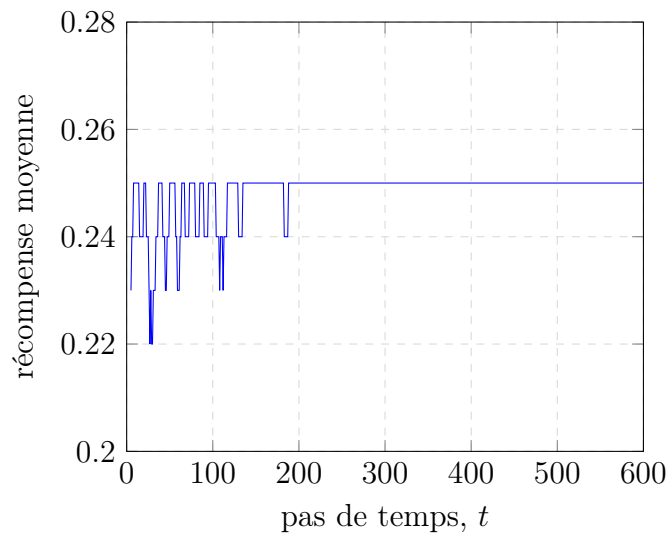


FIGURE 5.2 Convergence de DUPRA, pour  $U = 14$ .

TABLEAU 5.3 Notations et valeurs des paramètres de DUPRA

Notations	Paramètres	Valeurs
$l_{ra}$	Taux d'apprentissage de l'acteur	0.0003
$l_{rc}$	Taux d'apprentissage du critique	0.0003
$\gamma$	Facteur d'actualisation pour la mise à jour	0.99
$\varepsilon$	Taux de mise à jour progressive	0.005
$M$	Taille du tampon de mémoire	1024
$B$	Taille du lot d'échantillons	64
$T$	Pas de temps dans une épisode	600
$(N_1^a, N_2^a, N_3^a)$	Nombre de neurones dans le réseau acteur	(1024,512,256)
$(N_1^c, N_2^c, N_3^c)$	Nombre de neurones dans le réseau critique	(1024,512,256)

### 5.3 Étude de l'impact des paramètres du système

#### 5.3.1 Impact du nombre de positions possibles sur les performances de UDPA

TABLEAU 5.4 Évaluation des formules de complexité de algorithmes pour  $U = 30$  et  $D_{min} = 1$  m

Algorithme	Complexité
UDPA	$4.99 \times 10^7$
GA	$2.6 \times 10^9$
DUPRA	$7.8 \times 10^{10}$

Afin d'évaluer l'impact de la discrétisation de l'espace sur les performances des algorithmes GA et UDPA, la Fig. 5.3 compare ces algorithmes pour différentes valeurs de  $D_{min}$ , la distance minimale entre deux positions discrétisées possibles des UAVs. Il convient de rappeler que DUPRA considère un espace continu. Nous pouvons observer que l'algorithme génétique surpasse l'heuristique (c'est-à-dire UDPA) avec un écart d'au plus 25%, pour  $D_{min} = 2$  m, mais avec une complexité plus élevée, comme indiqué dans le Tableau 5.4. Les résultats de DUPRA sont proches de GA pour un petit nombre d'utilisateurs, mais l'écart de performance devient plus important pour un plus grand nombre d'utilisateurs. Cet écart peut s'expliquer par



le fait que DUPRA sélectionne les positions des UAVs déployés dans un espace de recherche continu, tandis que GA choisit les positions à partir d'un ensemble donné de positions discrètes  $\mathcal{Q}$ . Il peut également être observé qu'une discrétisation plus fine de l'ensemble des positions possibles conduit à de meilleures performances de GA et UDPA. Cela est dû au fait qu'une discrétisation plus fine permet une évaluation plus précise des positions réalisables dans l'espace de recherche, ce qui permet à UDPA et GA de trouver de meilleures solutions au prix d'une complexité computationnelle plus élevée. Nous observons qu'avec  $D_{min} = 1$  m, les UAVs disposent d'un ensemble plus grand de positions parmi lesquelles choisir. Par conséquent, le nombre des UAVs à déployer diminue. Nous pouvons conclure que, en raison de la discrétisation, DUPRA surpasse GA et UDPA. Nous observons que GA, lorsque  $D_{min} = 1$  m, présente des valeurs plus proches de la valeur optimale pour le nombre de générations donné. L'algorithme génétique est plus efficace que UDPA, donc la discrétisation n'a pas un grand impact sur les performances de GA.

### 5.3.2 Région de faisabilité du système

Dans la Fig. 5.4, différentes valeurs de canaux disponibles,  $C$ , sont considérées et la région de faisabilité du système pour  $\hat{O}fh = \hat{O}bh = 7\%$  est présentée. La région de faisabilité est définie comme l'aire sous la courbe qui rassemble l'ensemble des seuils de débit possibles satisfaisant les contraintes du système pour un certain nombre d'utilisateurs. Nous pouvons observer que la région de faisabilité est plus petite lorsque le nombre de canaux disponibles augmente, tant pour UDPA que pour DUPRA. Il est à noter que les régions de faisabilité pour UDPA et GA sont les mêmes car UDPA est utilisé pour générer la population initiale de la solution basée sur GA proposée. Par exemple, en considérant  $C = 2$ , UDPA (et DUPRA) ne peut

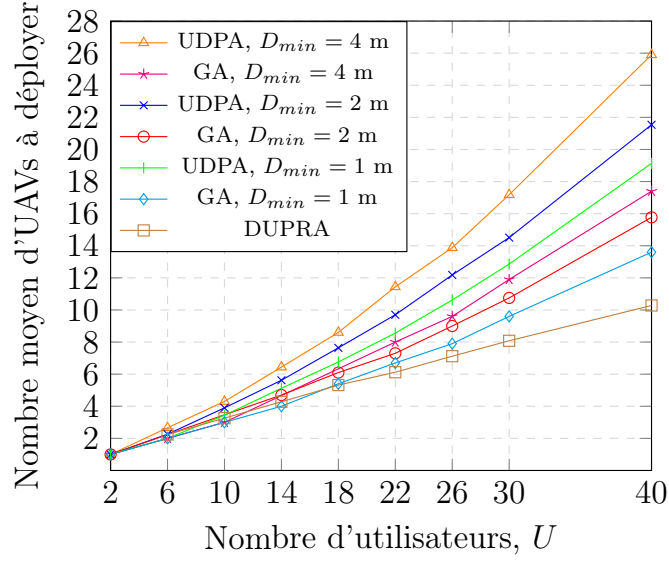


FIGURE 5.3 Impact de la distance minimale entre les positions possibles des UAVs  $D_{min}$ , pour  $\zeta_{fh} = 700$  Mbps and  $\hat{O}_{fh} = \hat{O}_{bh} = 5\%$ .

pas satisfaire  $U = 22$  utilisateurs même si nous déployons un UAV pour chaque utilisateur lorsque le seuil de débit est supérieur à  $\zeta_{fh} = 1.8$  Gbps. D'autre part, même lorsque le nombre de canaux disponibles est plus grand,  $C = 6$ , UDPA ne peut pas satisfaire  $U = 22$  utilisateurs lorsque le seuil de débit est fixé à  $\zeta_{fh} = 0.8$  Gbps, alors que DUPRA peut satisfaire  $U = 22$  utilisateurs si le seuil de débit est inférieur à 0.9 Gbps. Clairement, DUPRA étend la région de faisabilité, comparativement à UDPA. Cela est dû aux positions continues que DUPRA choisit pour les UAVs déployés. En conclusion, l'augmentation du nombre de canaux disponibles conduit à une diminution de la largeur de bande de chaque canal, ce qui conduit à une réduction de la région de faisabilité car il devient plus difficile de satisfaire tous les utilisateurs.

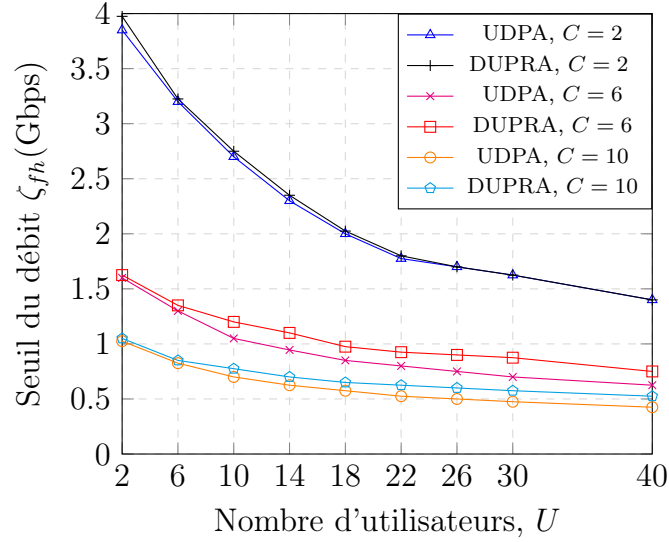


FIGURE 5.4 Région de faisabilité pour  $\hat{O}_{fh} = \hat{O}_{bh} = 7\%$ .

### 5.3.3 Impact du débit seuil et de la probabilité de coupure seuil

La Fig. 5.5 montre l'impact de  $\hat{O}_{fh}$  et  $\zeta_{fh}$  sur les performances du système, où  $\hat{O}_{fh} = \hat{O}_{bh}$ . Il est clair que l'augmentation de  $\zeta_{fh}$  accroît le nombre des UAVs requis. De même, la diminution du seuil de probabilité de coupure augmente le nombre des UAVs à déployer. Par exemple, la diminution de  $\hat{O}_{fh}$  de  $10\%$  à  $7\%$  augmente le nombre minimum des UAVs à déployer d'environ  $10\text{--}15\%$ . Nous pouvons également observer que, pour un grand nombre d'utilisateurs et une probabilité de coupure de  $5\%$ , le système a besoin de déployer en moyenne plus qu'un UAV pour deux utilisateurs afin de garantir un débit de données de  $700\text{ Mbps}$ . Cela est dû aux exigences strictes en termes de qualité de service et aux interférences causées par d'autres transmissions. Cependant, le système doit déployer  $\sim 30\%$  moins d'UAVs pour satisfaire  $\zeta_{fh} = 500\text{ Mbps}$ .

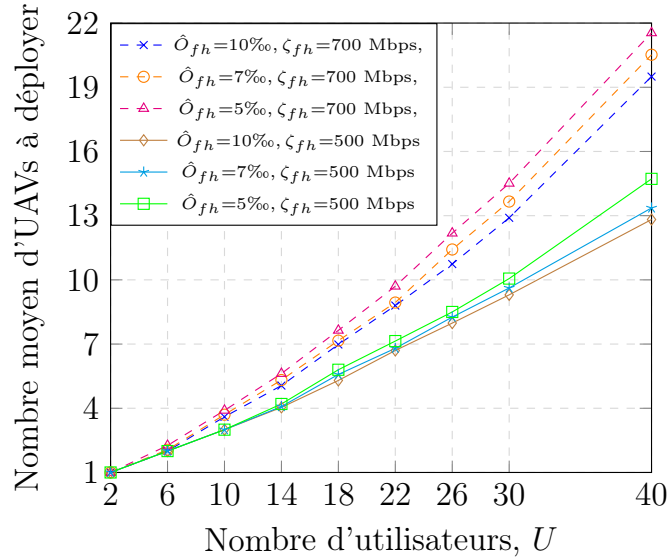


FIGURE 5.5 Performance de UDPA.

### 5.3.4 Impact de la puissance maximale des UAVs

La figure 5.6 montre l'impact de la puissance d'émission disponible au niveau des UAVs,  $p_n^{max}$ , et l'impact de la largeur de bande de la liaison terrestre,  $B_{bh}$ , sur le nombre d'UAVs à déployer. On considère  $\zeta_{fh} = 500$  Mbps et  $\hat{O}_{fh} = \hat{O}_{bh} = 10\%$ . Il est à noter que certaines courbes ne sont pas complétées car elles se situent en dehors des régions de faisabilité. On constate que, lorsque  $p_n^{max}$  est faible (par exemple,  $p_n^{max} = 20$  dBm), le nombre d'UAVs à déployer augmente rapidement avec le nombre d'utilisateurs, alors que la différence de performance est faible lorsque  $p_n^{max}$  est plus élevé. Ces courbes montrent l'impact considérable de la puissance d'émission maximale sur le nombre d'UAVs nécessaires pour satisfaire tous les utilisateurs. En ce qui concerne l'impact des ressources limitées de liaison du réseau d'amenée, nous observons que pour  $p_n^{max}=30$  dBm et  $B_{bh}=2$  GHz, il est impossible de satisfaire les besoins

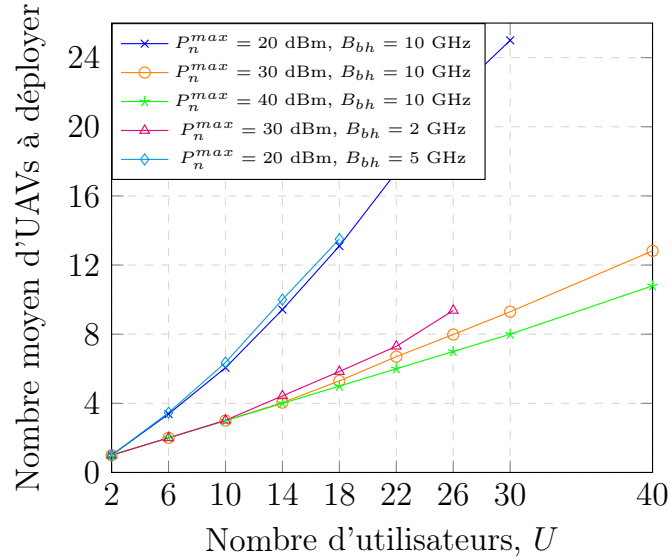


FIGURE 5.6 Impact de la puissance de transmission maximale des UAVs et  $B_{bh}$ , pour  $\hat{O}_{fh} = \hat{O}_{bh} = 10\%$  et  $\zeta_{fh} = 500$  Mbps.

de tous les utilisateurs quand  $U > 26$ . Ce nombre maximal d'utilisateurs pouvant être satisfaits devient 18 lorsque  $p_n^{max} = 20$  dBm et  $B_{bh} = 5$  GHz. Nous pouvons également observer que lorsque le problème est faisable (c'est-à-dire que tous les utilisateurs peuvent être satisfaits), le nombre d'UAVs à déployer augmente légèrement lorsque la largeur de bande passante disponible au niveau de la liaison du réseau d'amenée est limitée.

### 5.3.5 Performance de UDPA avec une méthode de sélection d'utilisateurs dans le cas d'infaisabilité

Lorsque le problème est infaisable, nous proposons d'utiliser une méthode de sélection des utilisateurs et ainsi le problème consiste à satisfaire le nombre maximum d'utilisateurs. La méthode de sélection des utilisateurs consiste à sélectionner les uti-

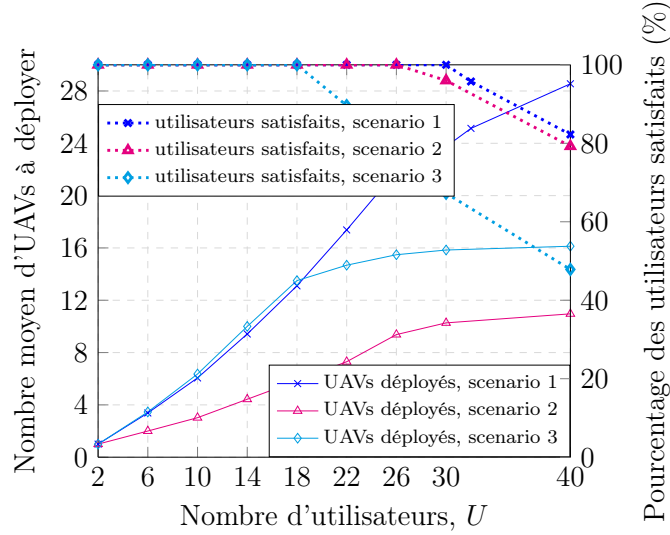


FIGURE 5.7 Performance avec sélection d'utilisateur avec une puissance de transmission maximale limitée des UAVs et  $B_{bh}$ , pour  $\hat{O}_{fh} = \hat{O}_{bh} = 10\%$  and  $\zeta_{fh} = 500$  Mbps.

lisateurs selon leurs gains de canal décroissants. La figure 5.7 montre les performances de la solution heuristique UDPA en utilisant la méthode de sélection des utilisateurs proposée pour satisfaire le nombre maximum d'utilisateurs. Nous considérons trois scénarios avec différentes valeurs de paramètres. Le scénario 1 considère  $P_n^{max} = 20$  dBm et  $B_{bh} = 10$  GHz, le scénario 2 considère  $P_n^{max} = 30$  dBm et  $B_{bh} = 2$  GHz, et le scénario 3 considère  $P_n^{max} = 20$  dBm et  $B_{bh} = 5$  GHz. Les résultats montrent que pour les scénarios 1 et 2, le nombre d'UAVs à déployer atteint sa valeur maximale lorsque le nombre d'utilisateurs dans le système est plus important et que le pourcentage d'utilisateurs satisfaits est plus faible. En outre, l'augmentation de la largeur de bande passantes dans les liaisons du réseau d'amenée mais avec une puissance de transmission limité pour les UAVs (scénario3) augmente le nombre d'UAVs déployés et diminue le nombre d'utilisateurs satisfaits.

### 5.3.6 Impact du nombre de canaux disponible en réseau d'accès

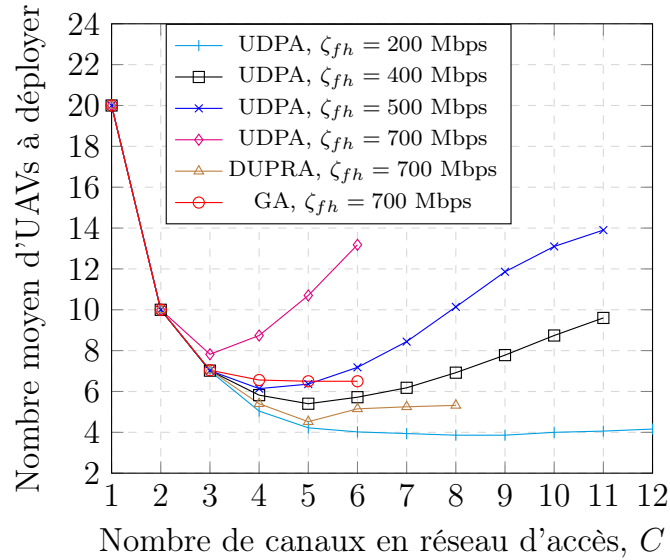


FIGURE 5.8 Impact du nombre de canaux en réseau d'accès, pour  $U = 20$  et  $\hat{O}_{fh} = \hat{O}_{bh} = 7\%$ .

La Fig. 5.8 montre l'impact du nombre de canaux disponibles en réseau d'accès,  $C=C_{ul}=C_{dl}$ , sur le nombre requis d'UAVs en considérant un  $B_{fh}=1\text{GHz}$  fixe. Notez que certaines courbes ne sont pas complétées, car elles atteignent les régions d'infaisabilité. En comparant les performances de l'UDPA pour différentes valeurs de  $C$ , on remarque que lorsque  $C$  augmente, et donc la largeur de bande passante des canaux devient plus étroite, le nombre requis d'UAVs diminue et se stabilise pendant un certain temps avant d'augmenter. Ce comportement peut s'expliquer par le fait qu'un  $C$  plus élevé pourrait diminuer le nombre requis d'UAVs puisqu'un seul UAV peut desservir plus d'utilisateurs, mais à un moment donné, les bandes passantes des canaux deviennent trop étroites. Il devient alors plus difficile de satisfaire les besoins des

utilisateurs et il faut déployer davantage d'UAVs. Il est clair qu'il existe une valeur optimale du nombre de canaux en réseau d'accès à trouver. Dans nos expériences, cette valeur est de 5, 4 et 3 canaux pour  $\zeta_{fh}=300$  Mbps, 500 Mbps et 700 Mbps, respectivement. La même conclusion peut être tirée pour les solutions DUPRA et GA. Lorsque  $\zeta_{fh}=200$ , UDPA atteint le nombre minimum d'UAVs déployés, quel que soit le nombre de canaux disponibles. Cependant, les performances de DDPG et GA sont supérieures à celles de l'UDPA pour  $\zeta_{fh}=700$  Mbps. En fait, le nombre d'UAVs requis pour un déploiement réussi est nettement inférieur lorsqu'on utilise DUPRA et GA. De plus, DDPG étend la faisabilité du problème avec  $C = 8$  en permettant un positionnement continu des UAVs.

#### 5.4 Impact de la connaissance imparfaite de l'information sur les états de canaux sur les performances

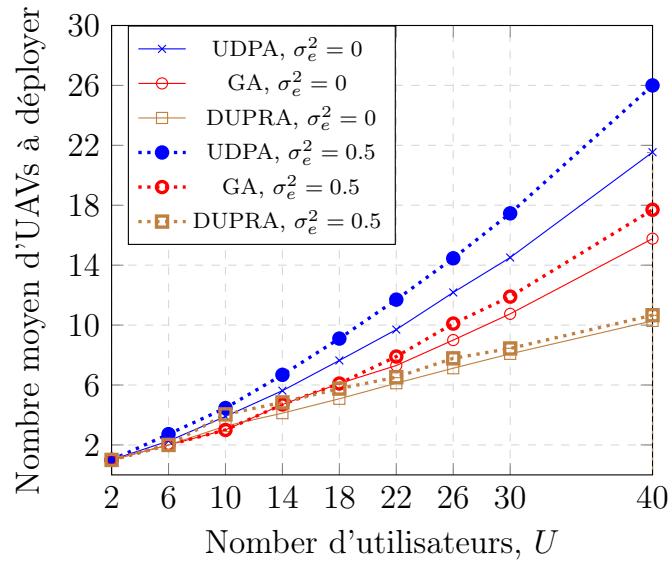


FIGURE 5.9 Impact du CSI Imparfait, pour  $\zeta_{fh} = 700$  Mbps et  $\hat{O}_{fh} = \hat{O}_{bh} = 5\%$ .



La Fig. 5.9 étudie les performances des algorithmes proposés lorsque on considère une connaissance imparfaite de l'information sur l'état du canal. Nous observons, avec une variance d'erreur importante  $\sigma_e^2 = 0.5$ , une détérioration significative des performances de l'algorithme proposé utilisant la solution UDPA ayant pour conséquence l'augmentation du nombre d'UAVs déployés de 20% pour un grand nombre d'utilisateurs. Quant à DUPRA, nous observons que ses performances sont robustes et le nombre d'UAVs à déployer augmente légèrement de 3,5% avec des informations imparfaites sur l'état du canal. GA est également robuste pour un faible nombre d'utilisateurs, mais pour un plus grand nombre d'utilisateurs, il déploie 12% d'UAVs en plus. L'intervalle de confiance n'est pas présenté dans les figures pour des raisons de clarté. Ses limites se situent à 1,5% de la valeur moyenne pour les trois algorithmes.

## 5.5 conclusion

Dans ce chapitre, nous avons évalué et comparé les résultats des trois solutions proposées. Nous avons, également, étudié l'impact des paramètres du système sur les performances des solutions proposées.



## CONCLUSION

Ce projet étudie le problème d'optimisation conjointe du positionnement des UAVs et d'allocation de ressources dans un réseau sans fil assisté par des UAVs qui peut assurer des applications rigoureuses. Ces applications introduisent des exigences strictes pour les communications UL et DL en termes de probabilités de coupure du débit de données qui sont directement liés à la latence. Le problème abordé a été modélisé comme un problème non linéaire à nombres entiers partiels et non convexe dont l'objectif est de minimiser le nombre d'UAVs déployés tout en satisfaisant tous les utilisateurs. Étant donné que le problème d'optimisation considéré a été démontré  $\mathcal{NP}$ -difficile, nous avons proposé une solution heuristique gloutonne, une solution basée sur l'algorithme génétique et une solution basée sur l'apprentissage par renforcement profond, pour le résoudre efficacement. Les résultats de simulation ont montré que les performances de la solution heuristique sont efficaces et se rapprochent de celles de la solution génétique et de la solution par renforcement profond, avec une réduction importante de la complexité de calcul. Ils révèlent également l'efficacité des solutions algorithmiques proposées en termes de nombre d'UAVs déployés pour satisfaire pleinement les exigences de transmission de données de tous les utilisateurs dans les directions de communication montantes et descendantes.

Pour les travaux futurs, nous prévoyons de nous intéresser à des arrivées de trafic à des moments différents dans un réseau sans fil assisté par des UAVs. En effet, nous avons considéré un trafic de données constant et continu dans notre projet.

Cependant, il est plus intéressant de considérer un trafic de données qui n'est pas continu et une arrivée du trafic différente pour chaque utilisateurs dans un réseau sans fil.

## RÉFÉRENCES

- Aijaz, A. (2018). Toward Human-in-the-Loop Mobile Networks : A Radio Resource Allocation Perspective on Haptic Communications. *IEEE Transactions on Wireless Communications*, 17(7), 4493–4508. <http://dx.doi.org/10.1109/TWC.2018.2825985>
- Akdeniz, M. R., Liu, Y., Samimi, M. K., Sun, S., Rangan, S., Rappaport, T. S. et Erkip, E. (2014). Millimeter Wave Channel Modeling and Cellular Capacity Evaluation. *IEEE Journal on Selected Areas in Communications*, 32(6), 1164–1179. <http://dx.doi.org/10.1109/JSAC.2014.2328154>
- Al-Hourani, A., Kandeepan, S. et Lardner, S. (2014). Optimal LAP Altitude for Maximum Coverage. *IEEE Wireless Communications Letters*, 3(6), 569–572. <http://dx.doi.org/10.1109/LWC.2014.2342736>
- Chen, Y.-J., Liao, K.-M. et Chen, Y.-F. (2021). End-to-End Delay Analysis in Aerial-Terrestrial Heterogeneous Networks. *IEEE Transactions on Vehicular Technology*, 70(2), 1793–1806. <http://dx.doi.org/10.1109/TVT.2021.3052250>
- Dai, Z., Zhang, Y., Zhang, W., Luo, X. et He, Z. (2022). A Multi-Agent Collaborative Environment Learning Method for UAV Deployment and Resource Allocation. 8, 120–130. <http://dx.doi.org/10.1109/TSIPN.2022.3150911>
- Davis, L. (1991). Handbook of genetic algorithms.

- Di Renzo, M. (2015). Stochastic Geometry Modeling and Analysis of Multi-Tier Millimeter Wave Cellular Networks. *IEEE Transactions on Wireless Communications*, 14(9), 5038–5057. <http://dx.doi.org/10.1109/TWC.2015.2431689>
- Ding, R., Gao, F. et Shen, X. S. (2020). 3D UAV Trajectory Design and Frequency Band Allocation for Energy-Efficient and Fair Communication : A Deep Reinforcement Learning Approach. *19(12)*, 7796–7809. <http://dx.doi.org/10.1109/TWC.2020.3016024>
- Dong, H., Ding, Z. et Zhang, S. (2020). *Deep Reinforcement Learning : Fundamentals, Research and Applications*. Springer Singapore.
- Fang, F., Zhang, H., Cheng, J., Roy, S. et Leung, V. C. M. (2017). Joint User Scheduling and Power Allocation Optimization for Energy-Efficient NOMA Systems With Imperfect CSI. *35(12)*, 2874–2885. <http://dx.doi.org/10.1109/JSAC.2017.2777672>
- Fotouhi, A., Qiang, H., Ding, M., Hassan, M., Giordano, L. G., Garcia-Rodriguez, A. et Yuan, J. (2019). Survey on UAV Cellular Communications : Practical Aspects, Standardization Advancements, Regulation, and Security Challenges. *IEEE Communications Surveys Tutorials*, 21(4), 3417–3442. <http://dx.doi.org/10.1109/COMST.2019.2906228>
- Guo, Y., Yin, S. et Hao, J. (2020). Joint Placement and Resources Optimization for Multi-User UAV-Relaying Systems With Underlaid Cellular Networks. *IEEE Transactions on Vehicular Technology*, 69(10), 12374–12377. <http://dx.doi.org/10.1109/TVT.2020.3018265>
- Gupta, L., Jain, R. et Vaszkun, G. (2016). Survey of Important Issues in UAV

- Communication Networks. *IEEE Communications Surveys Tutorials*, 18(2), 1123–1152. <http://dx.doi.org/10.1109/COMST.2015.2495297>
- Humadi, K., Trigui, I., Zhu, W.-P. et Ajib, W. (2021). Coverage Analysis of User-Centric Millimeter Wave Networks under Dynamic Base Station Clustering. Dans *ICC 2021 - IEEE International Conference on Communications*, 1–6. <http://dx.doi.org/10.1109/ICC42927.2021.9500654>
- ITU (2014). The Tactile Internet. ITU-T Technology Watch Report. [online]. Récupéré de [https://www.itu.int/dms\\_pub/itu-t/oth/23/01/T23010000230001PDFE.pdf](https://www.itu.int/dms_pub/itu-t/oth/23/01/T23010000230001PDFE.pdf)
- Kang, G.-M., Kim, H. M. et Shin, O.-S. (2021). Bi-directional Power and Trajectory Control for UAV-assisted Cellular Systems. Dans *2021 International Conference on Information and Communication Technology Convergence (ICTC)*, 69–72. <http://dx.doi.org/10.1109/ICTC52510.2021.9621050>
- Kapovits, A., Corici, M.-I., Gheorghe-Pop, I.-D., Gavras, A., Burkhardt, F., Schlichter, T. et Covaci, S. (2018). Satellite communications integration with terrestrial networks. *China Communications*, 15(8), 22–38. <http://dx.doi.org/10.1109/CC.2018.8438271>
- Kawamoto, Y., Nishiyama, H., Kato, N., Ono, F. et Miura, R. (2019). Toward Future Unmanned Aerial Vehicle Networks : Architecture, Resource Allocation and Field Experiments. *IEEE Wireless Communications*, 26(1), 94–99. <http://dx.doi.org/10.1109/MWC.2018.1700368>
- Konda, V. et Tsitsiklis, J. (1999). Actor-critic algorithms. *Advances in neural information processing systems*, 12.

- Li, R. et Huang, H.-C. (2008). A general k-level uncapacitated facility location problem. Dans D.-S. Huang, D. C. Wunsch, D. S. Levine, et K.-H. Jo (dir.). *Advanced Intelligent Computing Theories and Applications. With Aspects of Contemporary Intelligent Computing Techniques*, 76–83., Berlin, Heidelberg. Springer Berlin Heidelberg. [http://dx.doi.org/10.1007/978-3-540-85930-7\\_11](http://dx.doi.org/10.1007/978-3-540-85930-7_11)
- Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D. et Wierstra, D. (2015). Continuous control with deep reinforcement learning. *arXiv preprint arXiv :1509.02971*.
- Luong, P., Gagnon, F., Tran, L.-N. et Labeau, F. (2021). Deep Reinforcement Learning-Based Resource Allocation in Cooperative UAV-Assisted Wireless Networks. *20(11)*, 7610–7625. <http://dx.doi.org/10.1109/TWC.2021.3086503>
- Plaat, A. (2020). *Deep reinforcement learning*. Springer.
- Promwongsa, N., Ebrahimzadeh, A., Naboulsi, D., Kianpisheh, S., Belqasmi, F., Glitho, R., Crespi, N. et Alfandi, O. (2021a). A Comprehensive Survey of the Tactile Internet : State-of-the-Art and Research Directions. *IEEE Communications Surveys Tutorials*, *23(1)*, 472–523. <http://dx.doi.org/10.1109/COMST.2020.3025995>
- Promwongsa, N., Ebrahimzadeh, A., Naboulsi, D., Kianpisheh, S., Belqasmi, F., Glitho, R., Crespi, N. et Alfandi, O. (2021b). A Comprehensive Survey of the Tactile Internet : State-of-the-Art and Research Directions. *IEEE Communications Surveys Tutorials*, *23(1)*, 472–523. <http://dx.doi.org/10.1109/COMST.2020.3025995>
- Sambo, Y. A., Klaine, P. V., Nadas, J. P. B. et Imran, M. A. (2019). Energy Mini-



- mization UAV Trajectory Design for Delay-Tolerant Emergency Communication. Dans *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, 1–6. <http://dx.doi.org/10.1109/ICCW.2019.8757127>
- Sawalmeh, A. H., Othman, N. S., Shakhatreh, H. et Khreishah, A. (2019). Wireless Coverage for Mobile Users in Dynamic Environments Using UAV. *IEEE Access*, 7, 126376–126390. <http://dx.doi.org/10.1109/ACCESS.2019.2938272>
- Shi, W., Sun, Y., Liu, M., Xu, H., Gui, G., Ohtsuki, T., Adebisi, B., Gacanin, H. et Adachi, F. (2021). Joint UL/DL Resource Allocation for UAV-Aided Full-Duplex NOMA Communications. *IEEE Transactions on Communications*, 69(12), 8474–8487. <http://dx.doi.org/10.1109/TCOMM.2021.3110298>
- Sutton, R. S. et Barto, A. G. (2018). *Reinforcement learning : An introduction*. MIT press.
- Taleb, T., Hadjadj-Aoul, Y. et Ahmed, T. (2011). Challenges, opportunities, and solutions for converged satellite and terrestrial networks. *IEEE Wireless Communications*, 18(1), 46–52. <http://dx.doi.org/10.1109/MWC.2011.5714025>
- Tran, D.-H., Nguyen, V.-D., Chatzinotas, S., Vu, T. X. et Ottersten, B. (2022). UAV Relay-Assisted Emergency Communications in IoT Networks : Resource Allocation and Trajectory Optimization. *IEEE Transactions on Wireless Communications*, 21(3), 1621–1637. <http://dx.doi.org/10.1109/TWC.2021.3105821>
- Wang, C., Deng, D., Xu, L. et Wang, W. (2022). Resource Scheduling Based on Deep Reinforcement Learning in UAV Assisted Emergency Communication Networks. *70(6)*, 3834–3848. <http://dx.doi.org/10.1109/TCOMM.2022.3170458>

- Wang, P., Zhang, J., Zhang, X., Yan, Z., Evans, B. G. et Wang, W. (2020). Convergence of Satellite and Terrestrial Networks : A Comprehensive Survey. *IEEE Access*, 8, 5550–5588. <http://dx.doi.org/10.1109/ACCESS.2019.2963223>
- Watts, S. et Aliu, O. G. (2014). 5G resilient backhaul using integrated satellite networks. Dans *2014 7th Advanced Satellite Multimedia Systems Conference and the 13th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, 114–119. <http://dx.doi.org/10.1109/ASMS-SPSC.2014.6934532>
- Yang, T., Hu, Y., Yuan, X. et Mathar, R. (2019). Genetic Algorithm based UAV Trajectory Design in Wireless Power Transfer Systems. Dans *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, 1–6. <http://dx.doi.org/10.1109/WCNC.2019.8885987>
- Zeng, H., Zhu, X., Jiang, Y., Wei, Z., Sun, S. et Xiong, X. (2022). Toward UL-DL Rate Balancing : Joint Resource Allocation and Hybrid-Mode Multiple Access for UAV-BS-Assisted Communication Systems. *IEEE Transactions on Communications*, 70(4), 2757–2771. <http://dx.doi.org/10.1109/TCOMM.2022.3147249>
- Zhang, G., Ou, X., Cui, M., Wu, Q., Ma, S. et Chen, W. (2022). Cooperative UAV Enabled Relaying Systems : Joint Trajectory and Transmit Power Optimization. *IEEE Transactions on Green Communications and Networking*, 6(1), 543–557. <http://dx.doi.org/10.1109/TGCN.2021.3108147>
- Zhang, L., Zhao, H., Hou, S., Zhao, Z., Xu, H., Wu, X., Wu, Q. et Zhang, R. (2019). A Survey on 5G Millimeter Wave Communications for UAV-Assisted Wireless Networks. *IEEE Access*, 7, 117460–117504. <http://dx.doi.org/10.1109/ACCESS.2019.2929241>