

République Algérienne Démocratique et Populaire
Université Abou Bekr Belkaid– Tlemcen
Faculté des Sciences
Département d'Informatique

Mémoire de fin d'études

Pour l'obtention du diplôme de Master en Informatique

Option: Réseaux et Systèmes Distribués (R.S.D)

Thème

Utilisation de l'Apprentissage par Renforcement pour le
Green Networking dans un Réseau de Radio Cognitive

Réalisé par :

- BENSALD Radjaà
- GOUAL Samira

Présenté le 2 juillet 2018 devant le jury composé de :

- LEHSAINI Mohamed (Président)
- AMRAOUI Asma (Examinatrice)
- BENMAMMAR Badr (Encadrant)
- BENDELLA Mohammed Salih (Co-Encadrant)

Année Universitaire 2017-2018

DÉDICACES

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

« وَمَا تَوْفِيقِي إِلَّا بِاللَّهِ عَلَيْهِ تَوَكَّلْتُ وَإِلَيْهِ أُنِيبُ »

Je Dédie ce Mémoire à:

Mes Chers Défunts Grand-Parents

Mes Chers Parents Qui m'Ont Tout Donné

Tous les mots du monde ne sauraient exprimer l'immense amour que je leur porte, ni la profonde gratitude que je leur témoigne pour tous les efforts et les sacrifices que non jamais cessé de consentir pour mon instruction et mon bien-être.

Ma Très Chère et Précieuse Sœur Imane

Tu m'as beaucoup aidé, donné la force, le courage et l'espoir de continuer quand je le perds.

*Mes Chers Seconds Parents ma Tante **Fatiha** et mon Oncle **Abd-El-Hamid***

Ma Petite Sœurette Aicha Katre El-Nada et Frérot Mohamed El-Amine

Mon Très Cher Frère Alaa El-Dine et ma Belle-Sœur Mama

Mes Chers Grands-Parents Fatima et Boucif

Tous Mes Ami(e)s Assia, Khadidja, Fatima, Amel, Amina, Amel, Soumia et Imane et à Tous Ceux Qui Me Sont Chers

Tous Mes Collègues du CEM «Abd El-Moumen Ibnou Ali» et à Tous Mes Élèves

Ma Promotion RSD 2017-2018, Je leur souhaite la réussite et le bonheur.

DÉDICACES

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

« وَمَا تَوْفِيقِي إِلَّا بِاللَّهِ عَلَيْهِ تَوَكَّلْتُ وَإِلَيْهِ أُنِيبُ »

Je Dédie ce Mémoire à:

Mes Très Chers Parents qui m'ont tout donné, et qui ont toujours été là pour moi, et qui m'ont offert un magnifique modèle de labeur et de persévérance.

Mes Espoirs et Ma Force dans la Vie, Mes Très Chers Enfants Zin El Abidine, Adam et Mohamed El Amine je vous souhaite un avenir très brillant. Qu'Allah vous garde toujours.

Mon cher Mari Qui m'A Toujours Soutenu Ce travail a pu voir le jour grâce à ton aide et à ta patience avec moi.

Ma Très Chère et Précieuse Sœur Ouiame et Mes Très Chers Frères pour leurs conseils et leur soutien moral sans oublier ma petite sœur Roaya Ghofran.

REMERCIEMENTS

Ce mémoire a été réalisé, au Laboratoire d'Informatique, **sous la direction conjointe de Messieurs BENMAMMAR Badr et BENDELLA Mohamed Salih, respectivement Maître de Conférences classe A et Doctorant à la Faculté des Sciences de l'Université Abou Bekr Balkaid de Tlemcen.** Qu'il puisse trouver ici l'expression de notre très profonde gratitude pour les nombreuses orientations pertinentes et avisées, qu'il n'a cessées de nous prodiguer tout au long de ce semestre de travail. Qu'il nous soit permis de lui témoigner notre profonde reconnaissance pour avoir bénéficié de ses compétences scientifiques et de son aide matérielle précieuse, sans laquelle nous n'aurions pas pu mener à terme ce travail.

Nous sommes très honorées de compter **Monsieur LEHSAINI Mohamed, Maître de Conférences classe A à la Faculté des Sciences de l'Université Abou Bekr Balkaid de Tlemcen,** parmi les membres du jury de ce mémoire. Qu'il reçoive également le témoignage de notre profonde reconnaissance pour avoir accepté de présider le jury de ce mémoire.

Nous sommes très obligées à **Madame AMRAOUI Asma, Maître de classe B à la Faculté des Sciences de l'Université Abou Bekr Balkaid de Tlemcen,** de nous faire l'honneur en prenant sur son temps pour porter un regard critique sur ce travail et siéger dans le jury en qualité d'examinatrice. À cet égard, nous sommes heureux de lui exprimer nos vifs remerciements et de l'assurer de notre profonde reconnaissance.

Nos remerciements les plus sincères vont à tous ceux qui ont contribué de prêt ou de loin à la mise en forme de ce mémoire.

Merci 

Table des matières

LISTE DES FIGURES	1
LISTE DES TABLEAUX	2
LISTE DES ABRÉVIATIONS	3
INTRODUCTION GÉNÉRALE	4
CHAPITRE I : Réseaux de radio cognitive	5
I.1 Introduction	5
I.2 Radio cognitive	5
I.2.1 Historique	5
I.2.2 Définition de la radio cognitive	5
I.2.3 Relation entre RC et SDR	6
I.2.4 Principe de la radio cognitive	9
I.2.5 Cycle de cognition	9
I.2.6 Fonctionnement de la radio cognitive	12
I.2.7 Composantes de la radio cognitive	13
I.2.8 Domaines d'application de la radio cognitive	15
I.3 Conclusion	16
CHAPITRE II : Green Networking et Apprentissage par renforcement	17
II.1 Introduction	17
II.2 Green Networking	17
II.2.1 Définition	17
II.2.2 Efficacité énergétique dans les réseaux sans fil	18
II.2.3 Modèles de la consommation d'énergie	18
II.2.4 Green Networking dans les réseaux de radio cognitive	22
II.3 Apprentissage par renforcement	25
II.3.1 Apprentissage par renforcement dans les réseaux de radio cognitive	26
II.4 Conclusion	26
CHAPITRE III : Implémentation de l'application et évaluation des résultats	27
III.1 Introduction	27
III.2 Les outils de développement	27
III.2.1 Le langage Java	27
III.2.2 L'IDE NetBeans 8.2	28
III.2.3 L'API JFreeChart 1.0.19	28
III.3 Présentation de l'application	28

III.3.1	L'interface principale	28
III.3.2	Mécanisme de recherche dans la RC classique	32
III.3.3	Apprentissage par renforcement « Q-Learning »	33
III.3.3.1	L'algorithme Q-Learning	34
III.3.4	Résultats obtenus.....	41
III.3.4.1	Comparaison par rapport au temps de recherche	41
III.3.4.2	Comparaison par rapport à la consommation d'énergie.....	43
III.4	Conclusion.....	45
	CONCLUSION GÉNÉRALE	46
	RÉFÉRENCES BIBLIOGRAPHIQUES	47

LISTE DES FIGURES

Figure I. 1 : Relation entre la radio cognitive et la radio logicielle restreinte [4]	7
Figure I. 2 : Architecture de la radio cognitive [5].....	8
Figure I. 3 : Cycle de cognition de Mitola	10
Figure I. 4 : Cycle de cognition simplifié	10
Figure I. 5 : Composantes de la radio cognitive [10]	14
Figure II. 1 : Modélisation de la consommation énergétique dans un réseau de capteur [15]	20
Figure II. 2 : Compromis: largeur de bande, puissance de transmission, distance et débit [27]	23
Figure III. 1 : Page d'accueil d'application	29
Figure III. 2 : Onglet « Jeux de données ».....	30
Figure III. 3 : Choix de l'application utilisé	31
Figure III. 4 : Liste des canaux	32
Figure III. 5 : Partie de l'exécution de la RC classique du 1 ^{er} découpage.....	33
Figure III. 6 : Représentation graphique des pièces.....	35
Figure III. 7 : Matrice de récompense R.....	36
Figure III. 8 : La matrice Q.....	36
Figure III. 9 : Partie du code pour remplir la matrice R	37
Figure III. 10 : Partie d'exécution de la matrice R du premier découpage.....	38
Figure III. 11 : Partie d'exécution de la matrice Q	40
Figure III. 12 : Temps de recherche pour le découpage N°1 (RC vs RC_Q).....	41
Figure III. 13 : Temps de recherche pour le découpage N°2 (RC vs RC_Q).....	42
Figure III. 14 : Temps de recherche pour le découpage N°3 (RC vs RC_Q).....	42
Figure III. 15 : Temps moyen de recherche dans la RC classique.....	43
Figure III. 16 : Consommation d'énergie de la RC et de la RC_Q dans les trois découpages	44
Figure III. 17 : Tableau récapitulatif qui montre l'énergie consommée.....	44
Figure III. 18 : Taux de réduction de l'AR.....	45

LISTE DES TABLEAUX

Tableau III. 1: Les applications utilisées	30
Tableau III. 2 : Les découpages utilisés	31
Tableau III. 3: Exemple sur quelques canaux	38

LISTE DES ABRÉVIATIONS

Acronyme	Signification
AACR	Adaptative Aware Cognitive Radio
BS	Base Station
EMI	ElectroMagnetic Interference
GSM	Global System for Mobile Communication
JRE	Java Runtime Environment
JVM	Java Virtual Machine
LAFS	the Least Active, the First Served
LTM	Long Term Memory
LTE	Long Term Evolution
MPRs	Multi-Point Relays
MAC	Media Access Control
MAFS	the Most Active, the First Served
OLSR	Optimized Link State Routing Protocol
PAA	Power-Aware Alternation
QoI	Quality of Information
RC	Radio Cognitive
RF	Radio Frequency
RNs	Relay Node
SDR	Software Defined Radio
STM	Short Term Memory
SP	Sensory Perception
TIC	Information and Communication Technologies

INTRODUCTION GÉNÉRALE

Introduction générale

Introduction générale

Nous assistons actuellement à la multiplication des normes et des standards de télécommunication vue les progrès récents dans ce domaine. Le nombre croissant de standards normalisés permet d'élargir l'éventail des offres et des services disponibles pour chaque consommateur. D'ailleurs, la plupart des radiofréquences disponibles ont déjà été allouées. Des études ont montré que certaines bandes de fréquences sont partiellement occupées résultant de l'émergence de la radio cognitive (RC) qui est une technologie prometteuse visant à optimiser l'utilisation du domaine spectral tout en augmentant les services de réseau sans fil.

Le Green Networking est un concept général qui désigne les processus utilisés pour optimiser un réseau afin de le rendre plus efficace en consommation énergétique.

A ne pas oublier aussi que les réseaux sans fil ont des impacts négatifs sur notre environnement de vie tel que la radiation électromagnétique et l'émission du gaz carbonique CO₂, une chose qui augmente la pollution de l'environnement.

Dans ce PFE, nous avons utilisé un algorithme d'apprentissage par renforcement (Q-Learning) dans un réseau de radio cognitive. L'objectif visé est de réduire la consommation d'énergie et ainsi satisfaire le concept du « Green Networking ». En effet, il est possible de parfaire des communications "Green" en utilisant la radio cognitive. Dans une transmission, un nœud cognitif est capable de s'adapter aux paramètres dynamiques de la radio et du réseau pour un investissement optimal des ressources radio.

Le reste de ce mémoire est organisé en trois chapitres comme suit:

Dans le premier chapitre, nous allons présenter la radio cognitive avec les différents aspects : principes, architecture, fonction et différents domaines d'application...etc.

Le deuxième chapitre est consacré à la présentation du Green Networking et de l'apprentissage par renforcement que nous allons utiliser pour résoudre le problème de la consommation d'énergie.

Le dernier chapitre est dédié à l'implémentation de la solution proposée (notre contribution dans le cadre de ce PFE) dans les réseaux de radio cognitive. Cette solution consiste à utiliser l'apprentissage par renforcement (algorithme Q-Learning) afin de montrer l'intérêt de la radio cognitive pour le Green Networking. Nous feront une étude comparative par rapport à l'énergie consommée entre les résultats obtenus dans les deux cas : RC classique et RC avec Q-Learning.

CHAPITRE I

Réseaux de Radio Cognitive

I.1 Introduction

Le développement de nouvelles technologies a toujours été dicté par les besoins du moment et la disponibilité de la technique. Ces dernières années ont vu l'apparition de nouvelles technologies et leur utilisation accrue par le grand public, surtout dans la téléphonie mobile. Ce fait, certaines bandes de fréquence et réseaux (GSM, Wifi) sont chères et déjà surchargés. Cependant, l'utilisation du spectre de fréquence n'est pas uniforme : Une bande fréquentielle peut être surchargée pendant qu'une autre reste inutilisée. L'idée a donc naturellement émergé de développer des outils permettant de mieux utiliser le spectre. La solution la plus reconnue est l'utilisation d'un système radio cognitif. C'est le concept qui permet de répondre à ce défi : mieux utiliser le spectre et aussi augmenter les débits. Mais ce développement a conduit à l'émission de gaz carbonique qui a conduit à la pollution. Pour éliminer ce problème, un nouveau concept a émergé « Green Networking » qui pourra révolutionner les processus de communication et les TIC (Technologies de l'Information et de la Communication) de façon général.

Ce chapitre est consacré à la présentation de la radio cognitive dans ses différents aspects principes, architecture, fonction et les différents domaines d'application...etc.

I.2 Radio cognitive

I.2.1 Historique

Le concept de radio intelligente a été introduit en 1999 par Joseph Mitola [1]. Ce dernier travaillait sur la radio logicielle. Il proposa d'utiliser les possibilités de reconfiguration apportées par le logiciel pour transformer les équipements radio statiques en équipements radio dynamiques capables d'adapter leurs paramètres comme la modulation, la puissance d'émission ou les bandes de fréquence utilisées.

L'une des principales applications de la radio intelligente est l'accès opportuniste au spectre (AOS). En effet, le développement de la radio intelligente vivra un essor conséquent à partir de 2005 à la suite de l'article fondateur de Simon Haykin [2] qui propose d'utiliser la radio intelligente pour l'accès opportuniste au spectre.

I.2.2 Définition de la radio cognitive

La radio cognitive est un nouveau paradigme de communication qui attire l'attention des utilisateurs mobiles, des opérateurs réseaux et des fournisseurs de service.

Chapitre I : Réseaux de radio cognitive

Le terme radio cognitive (RC) est utilisé pour décrire un système ayant la capacité de détecter et de reconnaître son cadre d'utilisation, ceci afin de lui permettre d'ajuster ses paramètres de fonctionnement radio de façon dynamique et autonome et d'apprendre des résultats de ses actions et de son cadre environnementale d'exploitation, comme le souligne Mitola.

Un réseau cognitif coordonne les transmissions suivant différentes bandes de fréquences et différentes technologies en exploitant les bandes/technologies disponibles à un instant donné et à un endroit fixé.

Il a besoin d'une station de base capable de travailler sur une large gamme de fréquences afin de reconnaître différents signaux présents dans le réseau et se reconfigurer intelligemment. [3]

I.2.3 Relation entre RC et SDR

Une radio cognitive doit être capable de réaliser trois tâches essentielles. Elle doit surveiller l'environnement (capacité cognitive), analyser les informations collectées (capacité d'auto-organisation) et s'adapter à l'environnement (capacité d'auto-configuration).

La SDR (Radio logicielle restreinte) est capable d'offrir les fonctionnalités de flexibilité, de reconfigurabilité et de portabilité inhérentes à l'aspect d'adaptation de la RC. Par conséquent, cette dernière doit être mise en œuvre autour d'une SDR. En d'autres termes, la SDR est une technologie habilitante pour la radio cognitive [4].

Bien que de nombreux modèles différents soient possibles l'un des plus simples qui décrit la relation entre la RC et la SDR est illustré dans la figure I.1.

Dans ce modèle simple, les éléments de la RC entourent le support SDR. Le moteur cognitif représente la partie chargée de l'optimisation ou du contrôle du module SDR en se basant sur quelque paramètre d'entrée tels que les informations issues de la perception sensorielle ou de l'apprentissage de l'environnement radio, du contexte utilisateur, et de l'état du réseau.

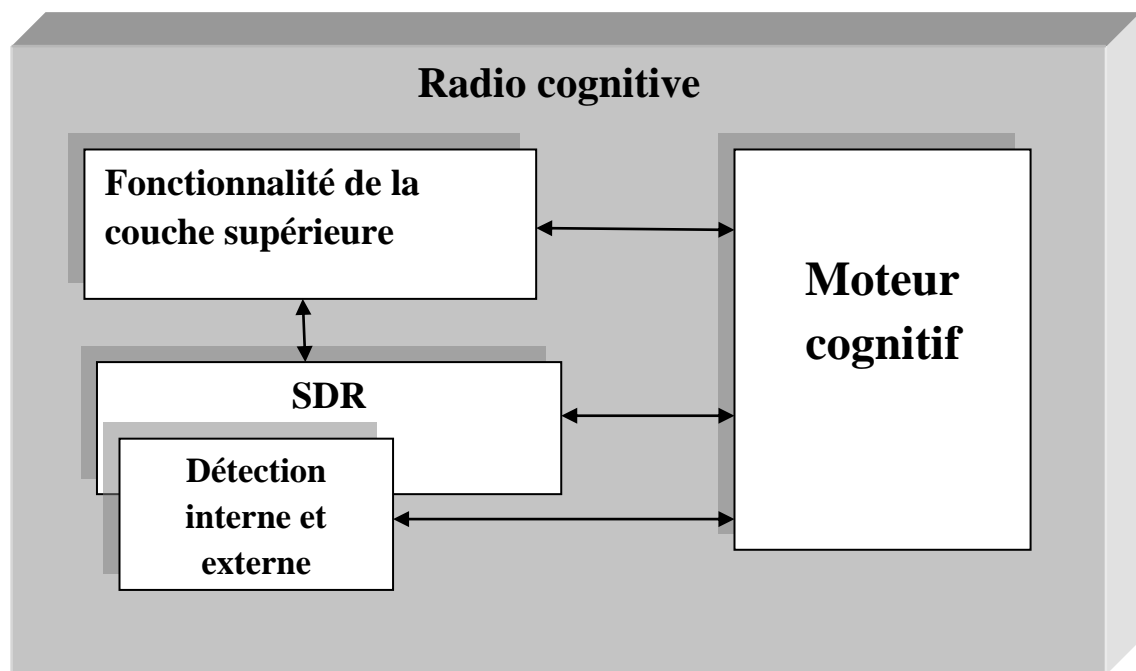


Figure I. 1 : Relation entre la radio cognitive et la radio logicielle restreinte [4]

Selon Mitola, l'architecture de la RC [5] est un ensemble complet et cohérent de règles de conception par lesquelles un ensemble de composants spécifiés réalise un ensemble déterminé de fonctions et de services qui évoluent à travers divers points de conception au cours du temps.

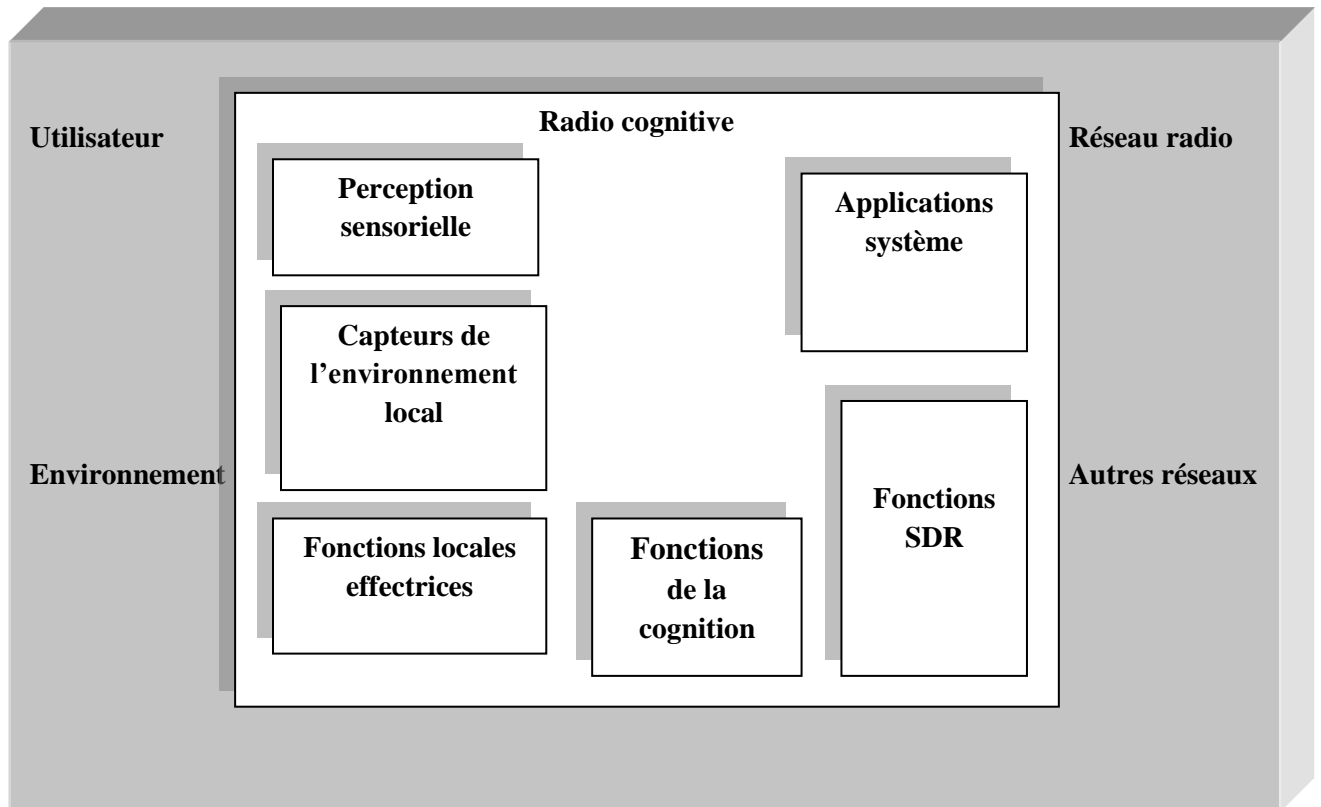


Figure I. 2 : Architecture de la radio cognitive [5]

Les six composantes fonctionnelles dans la radio cognitif sont:

1. La perception sensorielle (SP) de l'utilisateur incluant l'interface haptique (du toucher), acoustique, la vidéo, les fonctions de la détection et de la perception. Les fonctions SP de l'utilisateur peuvent inclure un matériel optimisé, par exemple, pour calculer les vecteurs de flux vidéo en temps réel pour aider la perception d'une scène.
2. Les capteurs de l'environnement local (emplacement, la température, l'accéléromètre, compas, etc.)
3. Les applications systèmes (les services médias indépendants comme un jeu en réseau).
4. Les fonctions SDR (incluant la détection RF et les applications radio de la SDR).
5. Les fonctions de la cognition (pour les systèmes de contrôle, de planification et de l'apprentissage).
6. Les fonctions locales effectrices (synthèse de la parole, du texte, des graphiques et des affiches multimédias).

Chapitre I : Réseaux de radio cognitive

Généralement, les communications vocales avec un annuaire téléphonique, la messagerie texte et l'échange de photos ou de clips vidéo correspondent aux principales applications pour les systèmes SDR. Les applications de l'AACR (Adaptative Aware Cognitive Radio) vont au-delà des services offerts par les systèmes SDR pour une plus grande flexibilité personnelle et le choix de la connectivité sans fil. L'utilisateur pourrait contrôler le passage de l'AACR d'un réseau à un autre en fonction du coût de la connexion.

I.2.4 Principe de la radio cognitive

La radio cognitive est une technologie qui fait appel à l'intelligence des réseaux et des terminaux pour :

- Détecter les besoins de communications des utilisateurs en fonction de l'utilisation.
- Fournir des ressources radio et les services sans fil les plus appropriés à ces besoins. Pour ce faire, la radio cognitive doit disposer de trois principales capacités [6] :
- ✓ **Conscience:** c'est la capacité à prendre conscience de son environnement. Un terminal radio cognitif associera donc environnement spatial, spectral et comportement des usagers pour une meilleure prise en conscience du provisionnement en ressources et un meilleur service.
- ✓ **Adaptation:** c'est la capacité à s'adapter soit à l'environnement (spectral ou technologique), soit à l'utilisateur (besoins et sécurité).
- ✓ **La cognition:** La cognition regroupe les divers processus mentaux allant de l'analyse perceptive de l'environnement à la commande motrice (en passant par la mémorisation, le raisonnement, les émotions, le langage...). Cette définition dépasse donc le seul cadre de la cognition humaine ou animale. La radio cognitive est donc un système qui peut percevoir son environnement, l'analyser, le mémoriser et agir en conséquence.

I.2.5 Cycle de cognition

La composante cognitive de l'architecture de la RC comprend une organisation temporelle, des flux d'inférences et des états de contrôle.

Ce cycle synthétise cette composante de manière évidente. Les stimuli entrent dans la radio cognitive comme des interruptions sensorielles envoyées sur le cycle de la cognition pour une réponse. Une telle radio cognitive observe l'environnement, s'oriente, crée des plans, décide, et puis agit. [7]

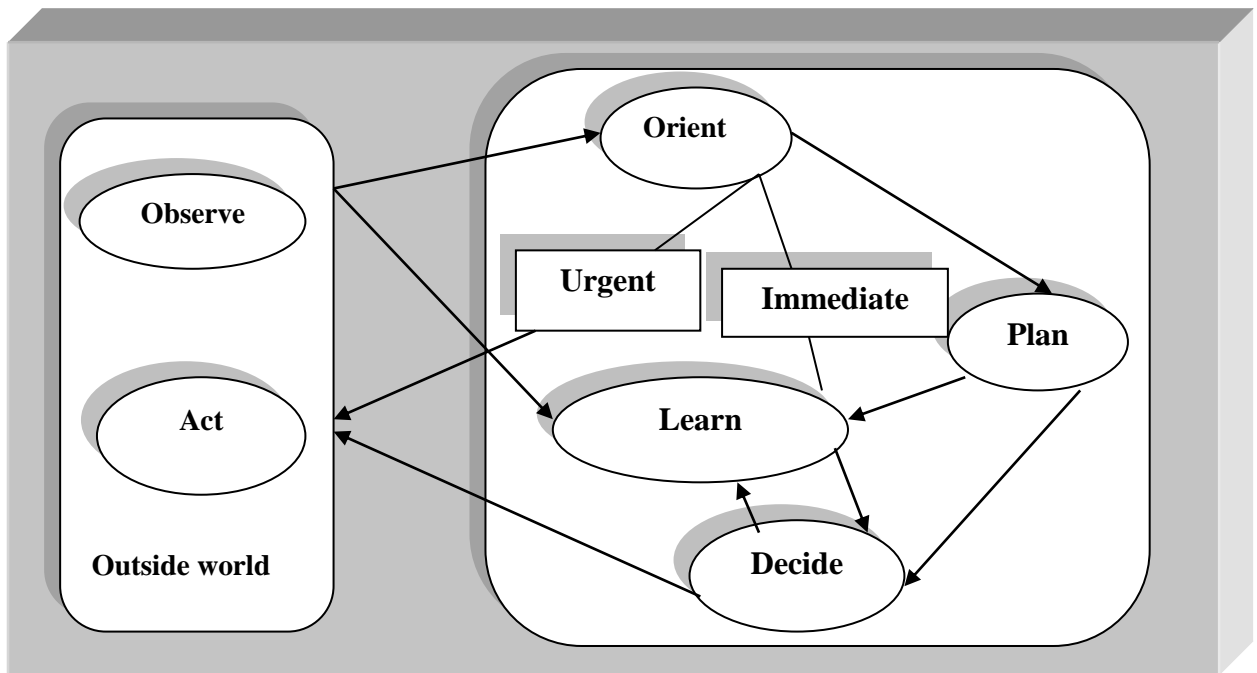


Figure I. 3 : Cycle de cognition de Mitola

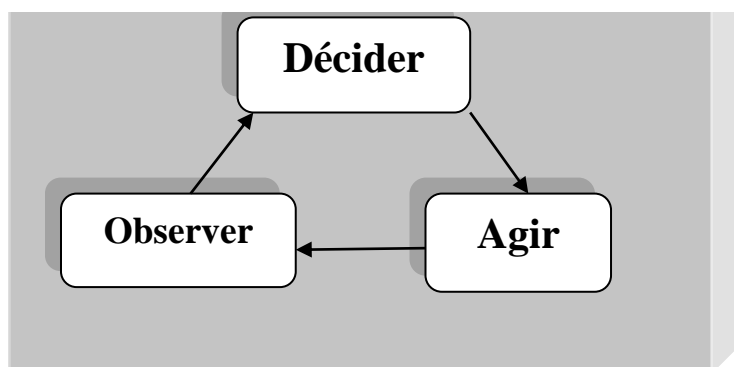


Figure I. 4 : Cycle de cognition simplifié

a. Phase d'observation (détecter et percevoir)

La RC observe son environnement par l'analyse du flux de stimuli entrant. Dans la phase d'observation, la RC associe l'emplacement, la température, le niveau de lumière des capteurs et ainsi de suite pour en déduire le contexte de communication.

Cette phase lie ces stimuli à des expériences antérieures pour discerner les modèles au fil du temps. La RC rassemble les expériences en se souvenant de tout.

Chapitre I : Réseaux de radio cognitive

b. Phase d'orientation

La phase d'orientation détermine l'importance d'une observation en liant à celle-ci une série connue de stimuli. Cette phase fonctionne à l'intérieure des structures de données qui sont analogues à la mémoire à court terme (STM), que les gens emploient pour s'engager dans un dialogue sans forcément se souvenir de tout à la même mesure que dans la mémoire à long terme (LTM). Le milieu naturel fournit la redondance nécessaire pour lancer le transfert de la STM à la LTM. La correspondance entre les stimuli courants et les expériences stockées se fait par reconnaissance des stimuli ou par relieur.

La reconnaissance des stimuli se produit quand il y a une correspondance exacte entre un stimulus courants et une expérience antérieure. La réaction peut être appropriée ou dans l'erreur.

Chaque stimulus est situé dans un contexte plus large, qui inclut d'autres stimuli et les états internes, y compris le temps. Parfois, la phase d'orientation provoque une action qui sera lancé immédiatement comme un comportement réactif "stimulus-réponse".

Une panne d'électricité, par exemple, peut directement invoquer un acte qui sauvegarde les données (le chemin " immédiate " de la phase action sur la figure). Une perte de signal sur un réseau peut invoquer une réaffectation de ressources. Cela peut être accompli via la voie marquée "urgent" dans la figure.

c. Phase de planification

Un message entrant du réseau serait normalement traité par la génération d'un plan (dans la phase de plan, la voie normale). Le plan devrait également inclure la phase de raisonnement dans le temps. Généralement, les réponses réactives sont préprogrammées ou apprises, tandis que d'autres réactions de délibération sont prévues.

d. Phase de décision

La phase de décision sélectionne un plan parmi les plans candidats. La radio peut alerter l'utilisateur d'un message entrant ou reporter l'interruption à plus tard en fonction des niveaux de QoI (Quality of Information) statués dans cette phase.

Chapitre I : Réseaux de radio cognitive

e. Phase d'action

Cette phase lance les processus sélectionnés qui utilisent les effecteurs sélectionnés qui accèdent au monde extérieur ou aux états internes de la RC.

L'accès au monde extérieur consiste principalement à composer des messages qui doivent être envoyés dans l'environnement en audio, ou exprimés dans différents langages appropriés.

Une action Agir sur l'environnement en effectuant, par exemple, des modifications au niveau de la fréquence radio.

f. Phase d'apprentissage

L'apprentissage dépend de la perception, des observations, des décisions et des actions. L'apprentissage initial est réalisé à travers la phase d'observation dans laquelle toutes les perceptions sensorielles sont comparées à l'ensemble de l'expérience antérieure pour compter les évènements et se souvenir du temps écoulé depuis le dernier évènement.

L'apprentissage peut se produire quand un nouveau modèle est créé en réponse à une action [8]. Par exemple, les états internes antérieurs et courants peuvent être comparés avec les attentes pour en apprendre davantage sur l'efficacité d'un mode de communication.

I.2.6 Fonctionnement de la radio cognitive

Les fonctions suivantes sont principales et indispensables pour le bon fonctionnement de la radio cognitive.

Les bandes spectrales inutilisées ont des caractéristiques différentes les unes des autres. Ces caractéristiques sont la fréquence d'opération de la bande spectrale, le débit et le temps. Toutes ces informations changent au cours du temps vu la nature dynamique de l'environnement radio. C'est dans ce contexte que les auteurs [5] ont présenté les nouvelles fonctions requises pour gérer les ressources spectrales dans les réseaux de radio cognitive.

Chapitre I : Réseaux de radio cognitive

a. Gestion de spectre:

- ✓ **L'analyse spectrale:** Elle permet de caractériser les différentes bandes spectrales en termes de fréquence d'opérabilité, de débit, de temps et de l'activité de l'utilisateur primaire.

Cette caractérisation sert à répondre d'une manière précise aux exigences de l'utilisateur équipé d'une radio cognitive. Des paramètres supplémentaires viennent compléter cette caractérisation à savoir le niveau d'interférence, le taux d'erreur du canal, l'évanouissement, le délai et le temps d'occupation de la bande spectrale par un utilisateur à radio cognitive. [9]

- ✓ **Décision spectrale:** Après que toutes les bandes spectrales aient été catégorisées et classifiées, on applique un ensemble de règles décisionnelles pour obtenir la ou les bandes spectrales les plus appropriées à la transmission en cours, tout en prenant compte des exigences de l'utilisateur à radio cognitive.[7]

b. Mobilité de spectre:

La mobilité spectrale est définie comme étant le processus de changement de fréquence d'opérabilité d'un URe. Ce phénomène est enclenché lors de la détérioration de la bande spectrale ou lors de la détection d'un UP. À ce moment là, les deux URC devront évacuer la BS et trouver une autre BS disponible pour continuer la communication.

L'objectif de la mobilité spectrale est d'assurer une transition fluide et rapide lors du changement de la fréquence d'opérabilité.

c. Détection de spectre:

Cette phase a pour objectif la détection de l'état de spectre (libre ou utilisé), afin de le partager avec son utilisateur primaire tout en évitant toutes interférences. Le défi réside dans le fait de mesurer l'interférence au niveau du récepteur primaire causée par les transmissions d'utilisateurs secondaires [5].

I.2.7 Composantes de la radio cognitive

Les différentes composantes d'un émetteur/récepteur radio cognitive qui mettent en œuvre ces fonctionnalités sont présentées dans la figure I.5 [10].

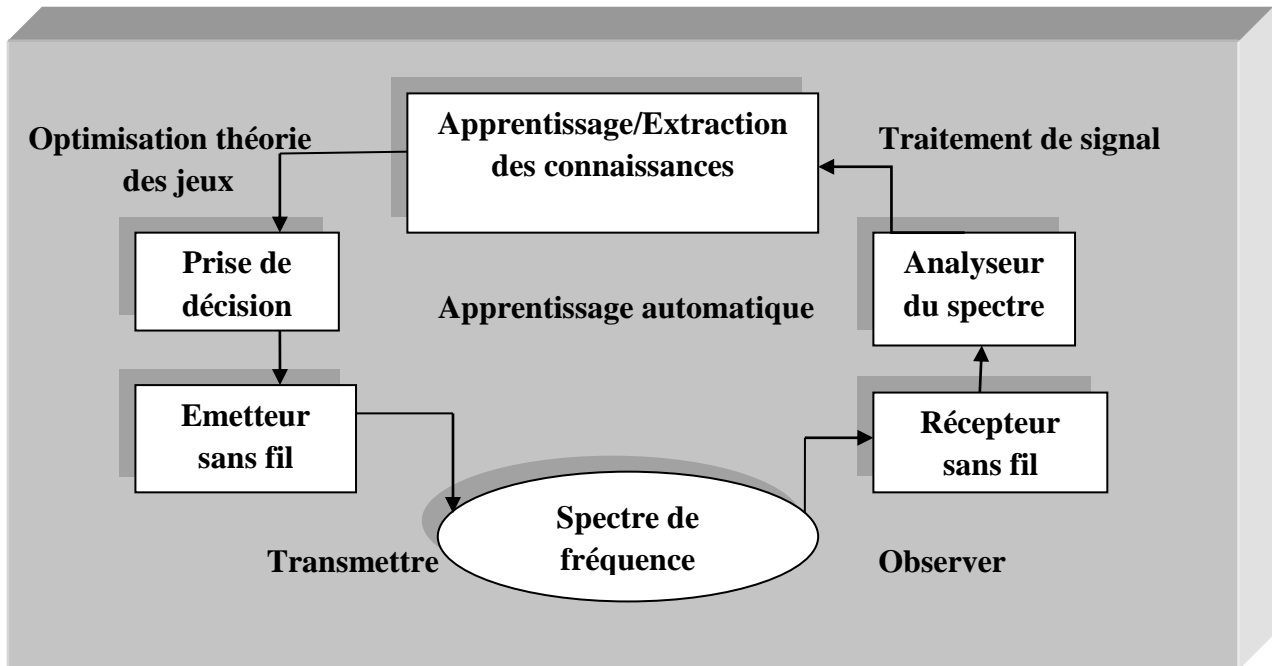


Figure I. 5: Composantes de la radio cognitive [10]

- a. **Emetteur / Récepteur** : c'est le composant majeur avec les fonctions du signal de transmission de données et de réception. En outre, un récepteur sans fil est également utilisé pour observer l'activité sur le spectre de fréquence (détection du spectre). Les paramètres de l'émetteur/récepteur dans le nœud de la RC peuvent être modifiés dynamiquement comme dicté par les protocoles de couches supérieures.
- b. **Analyseur de spectre** : il utilise les signaux mesurés pour analyser l'utilisation du spectre (détecter la signature d'un signal provenant d'un utilisateur primaire et trouver les espaces blancs du spectre pour utilisateurs secondaires). L'analyseur de spectre doit s'assurer que la transmission d'un utilisateur primaire n'est pas perturbée si un utilisateur secondaire décide d'accéder au spectre. Dans ce cas, diverses techniques de traitement du signal peuvent être utilisées pour obtenir des informations sur l'utilisation du spectre.
- c. **Apprentissage et extraction de connaissances** : une base de connaissances de l'environnement d'accès au spectre est construite et entretenue, qui est ensuite utilisée pour optimiser et adapter les paramètres de transmission pour atteindre l'objectif désiré sous diverses contraintes. Les algorithmes d'apprentissage peuvent être appliqués pour l'apprentissage et l'extraction de connaissances.

Chapitre I : Réseaux de radio cognitive

- d. **Prise de décision** : après que la connaissance de l'utilisation du spectre soit disponible, la décision sur l'accès au spectre doit être faite. La décision optimale dépend du milieu ambiant, elle dépend du comportement coopératif des utilisateurs secondaires. Différentes techniques peuvent être utilisées pour obtenir une solution optimale.

Par exemple, la théorie d'optimisation peut être appliquée lorsque le système est modélisé comme une seule entité avec un seul objectif. En revanche, les modèles de la théorie des jeux peuvent être utilisés lorsque le système est composé d'entités multiples, chacun avec son propre objectif. L'optimisation stochastique peut être appliquée lorsque les états du système sont aléatoires.

I.2.8 Domaines d'application de la radio cognitive

On peut appliquer la radio cognitive dans les domaines suivants :

➤ Réseaux de sécurité publique

Les réseaux de sécurité publique peuvent exploiter le réseau de radio cognitive. Ils sont utilisés pour les communications entre la police, les pompiers et le personnel paramédical. De tels réseaux sont également contestés par la quantité limitée de spectre attribué, même avec les extensions récentes des bandes de fréquences de sécurités publiques allouées, le personnel de la sécurité publique ne dispose pas d'une technologie d'opérer dynamiquement à travers les différents segments du spectre.

➤ Réseaux d'urgence:

Les réseaux de sécurité publique et d'urgence peuvent profiter des concepts de la radio cognitive pour fournir la fiabilité et la flexibilité de communication sans fil. Par exemple, dans un scénario où il y a une catastrophe, l'infrastructure de communication standard peut ne pas être disponible, et par conséquent, un système de communication sans fil adaptatif (soit un réseau d'urgence) peut être nécessaire d'être créé pour soutenir la reprise après sinistre. Ce genre de réseau peut utiliser le concept de la radio cognitive pour permettre la transmission sans fil et la réception sur une large gamme du spectre radio.

Chapitre I : Réseaux de radio cognitive

➤ Réseaux militaires

Les communications sécurisées dans les champs de bataille modernes sont devenues une tâche très difficile sachant qu'un réseau de communication du champ de bataille est le seul moyen de communications entre les soldats, les véhicules armés et les autres unités dans le champ de bataille entre eux ainsi qu'avec le siège.

Ceci implique que les réseaux ne nécessitent pas seulement une importante quantité de bande passante, mais aussi une garantie de communication sûre et fiable pour transporter des informations vitales. La radio cognitive est la technologie clé pour la réalisation de ces réseaux densément déployés en se basant sur des stratégies d'accès au spectre opportuniste [11].

➤ Coexistence de différentes technologies sans fil

La radio cognitive est une solution qui fournit la coexistence pour les nouvelles technologies sans fil qui sont en cours d'élaboration afin de réutiliser des fréquences radio allouées à d'autres services sans fil (service TV par exemple).

➤ Services de cyber santé (eHealth services)

Généralement tous les dispositifs médicaux de soins utilisés sont sans fil et sont sensibles aux EMI (interférences électromagnétique). D'autre part, différents dispositifs biomédicaux (équipement et appareil chirurgicaux, de diagnostic et de suivi) utilisent la transmission RF (Radio Frequency). Dans ce cas la radio cognitive peut être appliquée pour l'utilisation du spectre de ces dispositifs mais doit éviter toute interférence [11].

I.3 Conclusion

Dans ce chapitre nous avons donné un aperçu complet sur la radio cognitive en présentant les principaux concepts de cette technologie.

Le prochain chapitre est divisé en deux sections, la première section sera consacrée au concept du Green Networking, la deuxième section présentera l'apprentissage par renforcement.

CHAPITRE II

Green Networking et apprentissage par renforcement

II.1 Introduction

De nos jours, les TIC sont très développées, très avancées et leur utilisation nécessite un coût important de consommation d'énergie fossile d'où l'inquiétude d'un impact néfaste sur l'environnement qui est causé par les émissions de gaz « CO₂ » de ces technologies. La pollution dans le domaine des communications peut être structurée selon ces trois axes principaux : la dissipation d'énergie, la radiation électromagnétique et les déchets. L'impact d'énergie des TIC est évalué à 2% d'émissions de gaz participant à l'effet de serre (équivalentes à celle de toute l'aviation dans le monde) [12]. Des recherches ont permis la naissance d'un nouveau concept « Green Networking » qui pourra révolutionner les processus de communication et les TIC de façon général.

Ce chapitre est consacré à la présentation du Green Networking et de l'apprentissage par renforcement.

II.2 Green Networking

II.2.1 Définition

Le green Networking est un concept récent faisant référence aux processus utilisés pour optimiser un réseau afin de le rendre plus efficace en consommation énergétique. Par conséquent, il couvre les procédés qui permettent de réduire la consommation d'énergie, ce qui va engendrer la réduction du volume de gaz CO₂ émis par les infrastructures de télécommunication participant à l'effet du serre.

Le green Networking est un nouveau paradigme de communication particulièrement important pour toute la communauté scientifique (ingénieurs, universitaires, chercheurs et industriels) travaillant dans le domaine des réseaux. En effet réduire l'impact environnemental des systèmes de communication est un enjeu majeur avec l'augmentation constante du coût de l'énergie et la nécessité de réduire les émissions mondiales de CO₂ pour protéger notre environnement.

Dans ce qui suit, notre étude va se positionner sur les réseaux sans fil.

Et on peut distinguer deux principaux impacts des réseaux sans fil sur l'environnement :

- Radiation électromagnétique : due au rayon émis par le matériel de communication radio et surtout lorsque la portée est grande.

Chapitre II : Green Networking et apprentissage par renforcement

- Les émissions des gaz CO₂ : dues à l'énorme consommation d'énergie par les dispositifs radio.

II.2.2 Efficacité énergétique dans les réseaux sans fil

Un réseau sans fil est un réseau dans lequel au moins deux terminaux sont capables de communiquer entre eux grâce à des signaux radioélectriques. De manière générale, une connexion sans fil est recherchée par commodité (plus besoin de connexion physique), par économie (coût du lien physique) mais également pour s'affranchir des contraintes physiques (implantation dans des milieux dangereux, dans des lieux difficiles d'accès, au sein du corps humain...).

Au sein d'une liaison de données sans fil, une problématique importante est la gestion de l'énergie. En effet, il faut pouvoir alimenter les équipements et transmettre l'information selon une puissance de transmission donnée.

Cette puissance n'est qu'une énergie développée par le dispositif radio pour transmettre ses données d'où l'impact sur la consommation d'énergie puisque si on veut augmenter l'énergie développée (puissance de transmission) on aura besoin d'une énergie de base (grosse charge d'alimentation). Aussi et avec l'essor d'équipements de plus en plus miniaturisés et/ou difficilement rechargeables, il est nécessaire de réduire la consommation énergétique et donc le coût énergétique de ces liaisons qui vont devenir de plus en plus usitées.

II.2.3 Modèles de la consommation d'énergie

Les auteurs dans [13] ont calculé l'énergie nécessaire pour transmettre un paquet dans un réseau sans fil d'un nœud à un autre en utilisant la formule suivante : $E = i * v * t_p$ Joules.

Où i est le courant (en Ampère), v est le voltage (en Volt), et t_p est le temps pris pour transmettre le paquet (en secondes). Dans [14], les auteurs ont démontré que la longévité d'une batterie est inversement proportionnelle.

Par conséquent et en se basant sur ce résultat, les auteurs dans [18] expriment la consommation d'énergie par rapport à la durée de vie d'une batterie comme suit :

Chapitre II : Green Networking et apprentissage par renforcement

$h_{Batt} = \frac{C_{Batt} * V_{Batt}}{P_{tx} + K}$ Heures. Où C_{Batt} est la capacité de la batterie et V_{Batt} est le voltage de la batterie, P_{tx} est la puissance de transmission et K est une constante qui représente la puissance requise pour le traitement des signaux. Une valeur appropriée est $K = -13\text{dBW}$ (decibel watt).

Par ailleurs et dans les réseaux de capteurs sans fil, *Heinzelman et al.* [15] proposent un modèle radio de consommation d'énergie (voir figure II.1). Ainsi, les énergies nécessaires pour émettre $E_{Tx}(s,d)$ et recevoir $E_{Rx}(s)$ des messages sont données comme suit :

- Pour émettre un message de s bits vers un récepteur loin de d mètres, l'émetteur consomme :

$$E_{Tx}(s,d) = E_{Tx} \text{ elec}(s) + E_{Tx} \text{ amp}(s,d)$$

$$E_{Tx}(s,d) = (E_{elec} * s) + (E_{amp} * s * d^2)$$

- Pour recevoir un message de s bits, le récepteur consomme :

$$E_{Rx}(s) = E_{Rx} \text{ elec}(s)$$

$$E_{Rx}(s) = E_{elec} * s$$

E_{elec} et E_{amp} représentent respectivement l'énergie de transmission électronique et d'amplification (voir figure II.1).

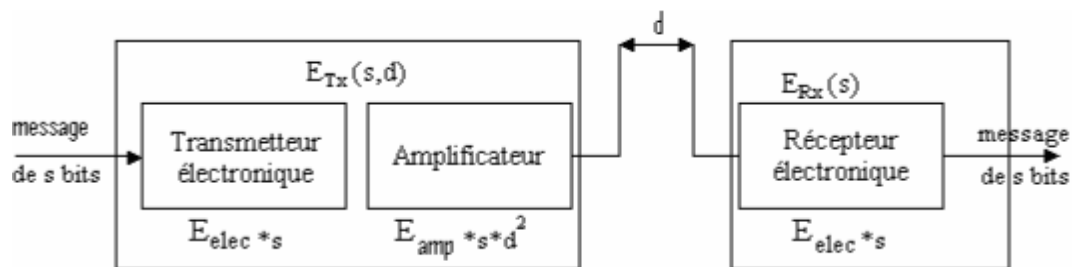


Figure II. 1 : Modélisation de la consommation énergétique dans un réseau de capteur [15]

La question de l'efficacité énergétique dans les réseaux sans fil fut durement étudiée. En réalité, la durée de vie de nombreux équipements réseaux sans fil a été pointée du doigt. Un aperçu non négligeable a été fourni par Jones [16] au sujet de protocoles de réseau sur l'efficacité énergétique pour les réseaux sans fil en 2001, or récemment la ligne de la recherche demeure toujours très active [15].

Chapitre II : Green Networking et apprentissage par renforcement

Les auteurs proposent une démarche, lors d'un travail présenté dans [13], pouvant modifier le protocole de routage OLSR (Optimized Link State Routing Protocol) de façon à ce qu'il soit en mesure d'endurer la notion d'économie d'énergie. Cette amélioration est basée sur l'interaction entre le protocole de routage OLSR et un mécanisme d'amélioration du protocole MAC, appelé Power-Aware Alternation(PAA) se basant sur la suppression de toute activité réseau d'un nœud pendant certaines périodes afin de conserver son énergie. Durant son inactivité, les messages en sa destination seront repris par un ou plusieurs nœuds nommés supporteurs. Ces supporteurs sont choisis à partir de l'ensemble des MPRs (Multi-Point Relays) du protocole OLSR, en tenant compte de la contrainte de leurs niveaux d'énergie. Ainsi, avec cette adaptation, les nœuds ayant une énergie basse sont écartés dans le routage afin de conserver un bon niveau d'énergie pour tous les nœuds mobiles.

Marsan et Meo dans [17] étudient l'économie d'énergie obtenue grâce à la coordination de deux réseaux cellulaires (deux opérateurs différents). Pendant les heures de pointe quotidiennes, les deux réseaux peuvent être allumés, mais l'un d'entre eux pourrait être mis en veille durant la période de bas trafic et les utilisateurs de ce réseau pourront permuter vers l'autre réseau.

Par contre, dans [19] les auteurs ont mis en œuvre deux algorithmes d'ordonnement et de gestion d'endormissement de réseau LTE-Advanced de type relais. Ces algorithmes permettent, d'abord, l'affectation de ressources aux nœuds relais et, ensuite, la gestion de leurs activités selon leurs états (transmission, réception ou endormissement) afin de réduire la consommation énergétique. Cette heuristique permet de calculer un ordonnancement des RNs (Relay Node) et leur endormissement partiel selon deux stratégies différentes en s'appuyant sur le temps d'activation des nœuds. La première stratégie, appelée LAFS (the Least Active, the First Served), commence par ordonner les RNs les moins actifs afin de réduire le nombre de changement d'états et donc minimiser le coût total de réveil. La deuxième stratégie, appelée MAFS (the Most Active, the First Served), commence par ordonner les nœuds les plus actifs dans le but de garantir que ces nœuds peuvent passer en mode endormissement partiel.

Dans un même contexte des systèmes cellulaires, le fait de diminuer la puissance a longtemps été un facteur dans les problèmes des concepteurs et des opérateurs de ces réseaux. Cependant, à mesure que la pression pour des économies d'énergie et des réductions des émissions de gaz participant à l'effet de serre accroit, le concept des réseaux cellulaires verts prend vie.

Chapitre II : Green Networking et apprentissage par renforcement

La plupart des techniques employées pour réduire au minimum la consommation d'énergie dans le cadre des systèmes cellulaires verts essayent d'optimiser la puissance de transmission des stations d'émetteur récepteur [20].

Selon le théorème bien connu de Shannon-Hartley la capacité ou le débit (bits/s) obtenu sur un canal de transmission dépend de la largeur de la bande W (hertz) mais également du rapport signal/bruit (S/N) sur le lien de transmission : $C=W*\log_2 (1+S/N)$.

Cette formule prouve clairement que l'augmentation du débit est linéaire avec la largeur de bande (W) mais elle est logarithmique avec la puissance de transmission. En effet, plus la puissance de transmission est élevée, plus le rapport signal/bruit est meilleur, mais une puissance plus élevée de transmission signifie également une plus grande consommation d'énergie, ce qui aura pour conséquence de limiter la durée de vie de la batterie.

En conséquence, il est possible de limiter la consommation d'énergie en réduisant la puissance de transmission [20].

Finalement, le green Networking dans les réseaux sans fil revient d'une grande partie à gérer le conflit entre la puissance du signal et l'économie d'énergie, c'est-à-dire gérer les paramètres de la radio de manière automatique et adaptatif au besoin. D'où la nécessité d'utiliser les services de la radio cognitive.

II.2.4 Green Networking dans les réseaux de radio cognitive

Le green Networking peut profiter du large paradigme de la radio cognitive. Haykin, dans [2] a défini avec précision l'équipement CR comme un émetteur radioélectrique cognitif qui apprendra de l'environnement et adaptera ses états internes aux variations statistiques des stimuli existants de la radiofréquence. Pour cela, il faut ajuster les paramètres de transmission (par exemple, la bande de fréquence, le mode de modulation et la puissance de transmission) en temps réel. A cela s'ajoute les possibilités cognitives pour détecter le spectre disponible ainsi que la capacité de reconfiguration pour accéder dynamiquement à ce spectre.

Chapitre II : Green Networking et apprentissage par renforcement

Le système de communication intelligent de la radio cognitive augmente l'agilité de spectre et l'efficacité énergétique [21].

Dans la littérature, il y a peu de travaux portant sur la santé de l'utilisateur de la technologie « green » dans la radio cognitive. En particulier, ici, nous souhaitons citer les travaux de Palicot [22 ; 23] qui mettent l'accent sur l'effet nocif de l'utilisation des antennes omnidirectionnelles dans des bornes sans fil. Ceci expose le corps et en particulier la tête aux quantités élevées de rayonnement électromagnétique. Les idées et les suggestions proposées dans ce document pour réduire l'exposition au rayonnement comportent notamment l'utilisation des antennes directionnels.

En effet, un besoin de configuration dynamique des antennes apparaît, ainsi que des puissances de transmission et d'autres paramètres physiques liés aux signaux émis. Cependant, il peut être très difficile de mettre en application de telles adaptations, bien que techniquement faisables avec la radio cognitive dans le cas d'un utilisateur mobile et/ou de propriétés changeantes des canaux de transmission.

En revanche, de nombreux travaux de recherches ont été menés dans le but de réduire la consommation d'énergie, en particulier pendant l'étape de détection. Ainsi dans [24], les auteurs proposent un système de sommeil et de censure combinée pour améliorer l'efficacité énergétique lors de la détection du spectre dans un réseau de capteurs cognitifs. Ils ont analysé la performance de détection de ce système en dérivant théoriquement les probabilités globales de détection et de fausses alarmes. Leur objectif est de minimiser la consommation d'énergie nécessaire à une détection distribuée (distributed sensing), étant donné les contraintes pesant sur les probabilités globales de détection et de fausse alarme, en désignant de manière optimale un taux de sommeil et un seuil de censure. Dans [25] les auteurs ont étudié la conservation de l'énergie et la performance du système d'allocation des ressources décentralisées dans les réseaux radio cognitive fondés sur des utilisateurs secondaires compétitifs. Dans l'approche qu'ils ont proposée, ils appliquent le jeu de la minorité (Minority Game) à la phase la plus importante de l'accès opportuniste au spectre qui est la phase de détection. Les auteurs ont essayé de réaliser une coopération dans un environnement non coopératif sans échange d'informations. Ils ont étudié la solution d'équilibre de Nash pour des stratégies pures et totalement mixtes, en se servant des algorithmes d'apprentissage distribués qui permettent aux utilisateurs cognitifs d'apprendre l'équilibre de Nash. Pour conclure, ils fournissent des résultats numériques pour les approches proposées.

Chapitre II : Green Networking et apprentissage par renforcement

En outre, une minimisation importante de la consommation d'énergie peut être effectuée par une action directe sur les infrastructures de communication. En effet, on peut réduire la consommation d'énergie des équipements en optimisant leur mode de fonctionnement ou même en les arrêtant, cela augmente leur durée de vie et prolonge leur cycle de recharge. En particulier, la radio cognitive offre le contrôle de la puissance de transmission sur chacune des longueurs d'onde élues. Il est clair que plus la puissance de transmission sera faible, plus il en sera de-même pour la consommation d'énergie. Mais si nous diminuons la puissance de transmission, le débit et le rayon de couverture sont impactés, ce qui peut poser des difficultés. En choisissant la bande de fréquence juste, la radio cognitive peut optimiser ces trois paramètres dans le but d'améliorer la consommation d'énergie.

Dans [18], la déduction est spécifique, les auteurs démontrent une relation directe entre la durée de vie de la borne mobile, la largeur de bande et le débit. Par ailleurs, ces chercheurs ont montré que la puissance de transmission d'une station émetteur-récepteur peut avoir un impact négatif sur la consommation d'énergie des bornes sur une cellule voisine. Ainsi, plus la puissance de transmission du point d'accès est élevée, plus l'interférence avec des non-destinataires se produira, et plus les bornes sur les cellules voisines devront augmenter leur puissance de transmission afin de maintenir le même rapport signal/bruit, ce qui va restreindre leur durée de vie. La radio cognitive, en exploitant différentes bandes de fréquence dans les cellules, est capable de détourner cet effet.

Dans le but de minimiser la consommation d'énergie, la radio cognitive doit tenir compte de la relation entre la largeur de bande, la puissance de transmission et la capacité (le débit). Clairement, si la puissance de transmission est réduite, le débit de la communication le sera aussi. Bien que ceci puisse être compensé (légèrement) par l'accroissement de la largeur de bande du canal, sur la majorité des bandes de fréquences utilisées, il n'est pas possible de le faire. En outre, il est mentionné que des transmissions sur des canaux de basse fréquence recouvrent de grandes distances en raison de leur résistance à l'atténuation et à l'interférence. Autrement dit, si deux nœuds placés loin l'un de l'autre ne sont pas à portée de l'autre en utilisant une fréquence donnée, il est possible en employant une fréquence inférieure, et avec la même puissance de transmission, que ces deux nœuds puissent communiquer. Cependant, le débit de cette communication sera certainement plus faible. En effet, en transmettant sur une fréquence inférieure, la largeur de bande est également réduite [20].

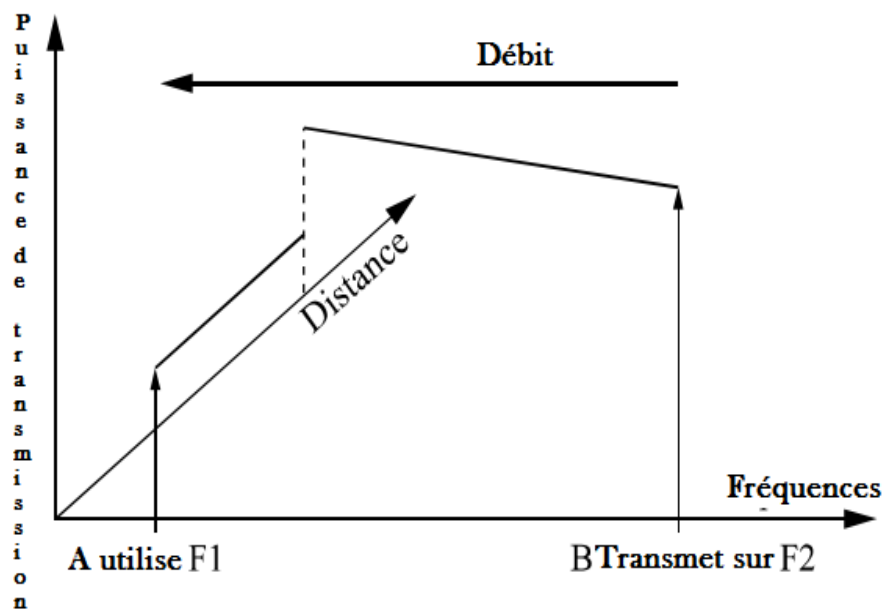


Figure II. 2 : Compromis: largeur de bande, puissance de transmission, distance et débit [20]

En guise de conclusion, il est possible de parfaire des communications "Green" en utilisant la radio cognitive, une technologie apte à adapter dynamiquement tous les paramètres impliqués dans une transmission.

Le schéma montre qu'un nœud radio cognitive A, en choisissant une fréquence basse F1 et avec une faible puissance de transmission, peut effectuer une communication à travers des distances équivalentes à un autre nœud B qui transmet à une puissance plus élevée sur la fréquence F2. Ceci illustre le fait que la radio cognitive, en sélectionnant dynamiquement les bandes inférieures, peut accroître la durée de vie des équipements et optimiser l'efficacité énergétique des réseaux. Cependant, ce gain de durée de vie intervient aux dépens du débit obtenu. Cette contrainte n'est pas un obstacle quand la bande de fréquence plus basse qui a été choisie assure un débit satisfaisant pour la communication.

II.3 Apprentissage par renforcement

L'apprentissage se définit selon [26] comme une notion englobant toute méthode permettant de construire un modèle de la réalité à partir de données.

L'apprentissage par renforcement est un domaine de l'apprentissage automatique (machine Learning en anglais) permettant de résoudre des problèmes de décisions séquentielles dans l'incertain.

On considère un agent (ou encore, sujet, acteur, décideur) et un environnement (ou encore système). L'agent est confronté à une suite de prises de décisions. L'environnement, quant à lui, est caractérisé par un 'Etat' qui évolue dans le temps de manière aléatoire en fonction des décisions prises par l'agent. Une prise de décision consiste en un choix d'action. Ce choix dépend lui-même de l'état actuel de l'environnement. Ainsi, la décision prise à chaque instant induit un changement d'état et, par conséquent, influence la prise de décision suivante. Le déroulement dans le temps de ces événements successifs, qui caractérisent l'interaction entre l'agent et l'environnement, est essentiel [27].

Cette dénomination d'apprentissage par renforcement trouve son origine dans le fait qu'à chaque instant l'agent reçoit des récompenses (ou encore bénéfiques, gains) selon l'état de l'environnement et l'action qu'il a choisie. Ces récompenses le guident dans ses prises de décisions.

Ainsi, le but de l'agent est de choisir ses actions de manière à maximiser les récompenses reçues. En général, l'agent a pour objectif de maximiser, non pas les récompenses reçues à chaque instant, mais celles reçues sur le long terme. L'action permettant de maximiser la récompense immédiate étant donné l'état de l'environnement est a priori différente de celle aboutissant à un grand bénéfice à long terme. En effet, la récompense reçue par l'agent dépendant à la fois de l'action choisie et de l'état de l'environnement, il peut être nécessaire de choisir des actions peu rentables à court terme afin d'atteindre un état dans lequel la récompense est très grande.

II.3.1 Apprentissage par renforcement dans les réseaux de radio cognitive

Le socle de l'apprentissage est sa base de connaissance ou sa base de données conçue pour contenir l'ensemble des informations validées après analyse. L'apprentissage repose sur l'acquisition d'expérience, les connaissances mémorisées qui assurent une réactivité au moteur de raisonnement et de décision dans le temps. Cela suppose donc un entraînement préalable sur des données ou un échantillon valide et suffisant. L'objectif pour la théorie de l'apprentissage est d'apporter un début de réponse aux modèles de prédiction, de planification et d'anticipation très peu examinés dans le domaine des réseaux de radio cognitive.

Dans la littérature, certains travaux proposent des techniques pour améliorer les paramètres de la radio cognitive, pour maximiser les capacités cognitives et rendre adaptable l'accès dynamique au spectre par apprentissage [28].

II.4 Conclusion

Dans ce chapitre nous avons présenté les deux concepts : Green Networking et apprentissage par renforcement. Nous avons présenté en particulier les travaux relatifs à minimiser la consommation d'énergie en utilisant la radio cognitive.

Le prochain et dernier chapitre de ce mémoire sera consacré à la présentation de notre contribution dans le cadre de ce PFE.

CHAPITRE III

*Implémentation de l'application
et évaluation des résultats*

III.1 Introduction

Le progrès des technologies de l'information et de la communication sans fil et leur utilisation dans le monde universitaire ont des conséquences environnementales négatives sur notre planète due à la consommation énergétique qui cause l'émission du gaz carbonique ce qui a conduit à la nécessité de chercher une solution dans les réseaux sans fil.

Dans ce contexte, nous proposons une solution dans les réseaux de radio cognitive (RC) qui consiste à utiliser l'apprentissage par renforcement afin de montrer l'intérêt de la RC pour le Green Networking.

L'objectif est de montrer qu'un algorithme d'apprentissage par renforcement aidera l'utilisateur à choisir le canal optimal pour son application (moins de puissances de signal) contrairement à la RC classique où le même utilisateur pourra utiliser un canal avec une grande puissance de transmission et par conséquent consommer plus d'énergie.

Donc, ce chapitre couvre l'implémentation de l'application et l'évaluation des résultats obtenus.

III.2 Les outils de développement

Dans ce qui suit, nous citons les différents outils de développement que nous avons utilisés pour la réalisation de notre application.

III.2.1 Le langage Java

Java est un langage de programmation orienté objet créé par Sun Microsystems, il était présenté officiellement le 23 mai 1995 au *SunWorld*. La société Sun a été ensuite rachetée en 2009 par la société Oracle qui détient et maintient désormais ce langage.

Java est à la fois un langage de programmation et un environnement d'exécution. Le langage Java a la particularité principale que les logiciels écrits avec ce dernier sont très facilement portables sur plusieurs systèmes d'exploitation. C'est la plate-forme qui garantit la portabilité des applications développées en Java.

Les applications Java peuvent être exécutées sur tous les systèmes d'exploitation pour lesquels a été développée une plate-forme Java, dont le nom technique est JRE (Java Runtime Environment - Environnement d'exécution Java).

Cette dernière est constituée d'une JVM (Java Virtual Machine - Machine Virtuelle Java), le programme qui interprète le code Java et le convertit en code natif. Mais le JRE est surtout constitué d'une bibliothèque standard à partir de laquelle doivent être développés tous les programmes en Java.

III.2.2 L'IDE NetBeans 8.2

NetBeans est un environnement de développement intégré (IDE), placé en open source par Sun en juin 2000. Il supporte une large variété de langages de programmation et d'outils de collaboration.

Compilé en Java, NetBeans est disponible sous Windows, Linux, Solaris (sur x86 et SPARC), Mac OS X ou sous une version indépendante des systèmes d'exploitation (requérant une machine virtuelle Java). Un environnement Java Development Kit JDK est requis pour les développements en Java.

III.2.3 L'API JFreeChart 1.0.19

JFreeChart est une bibliothèque graphique sous Java. Elle est gratuite et permet aux développeurs d'afficher facilement des graphiques de qualité professionnelle dans leurs applications. Le vaste ensemble de fonctionnalités de JFreeChart comprend:

- Une API cohérente et bien documentée, prenant en charge un large éventail de types de graphiques.
- Une conception flexible facile à étendre qui cible à la fois les applications côté serveur et côté client.
- Prise en charge de nombreux types de sortie, y compris les composants Swing et JavaFX, les fichiers image (y compris PNG et JPEG) et les formats de fichiers graphiques vectoriels (y compris PDF, EPS et SVG).
- JFreeChart est un logiciel libre. Il est distribué sous les termes de la LGPL (Lesser General Public License), qui permet l'utilisation dans des applications propriétaires.

III.3 Présentation de l'application

III.3.1 L'interface principale

L'IHM suivante présente la page d'accueil de notre application



Figure III. 1 : Page d'accueil d'application

Le bouton « **Entrer** » lance l'application et le bouton « **Sortir** » la quitte.

En cliquant sur « **Entrer** », nous affichons l'IHM de simulation suivante contenant 4 onglets.

- ✓ **Jeux de données.**
- ✓ **Liste des canaux.**
- ✓ **Temps de recherche.**
- ✓ **Graphes de comparaison.**

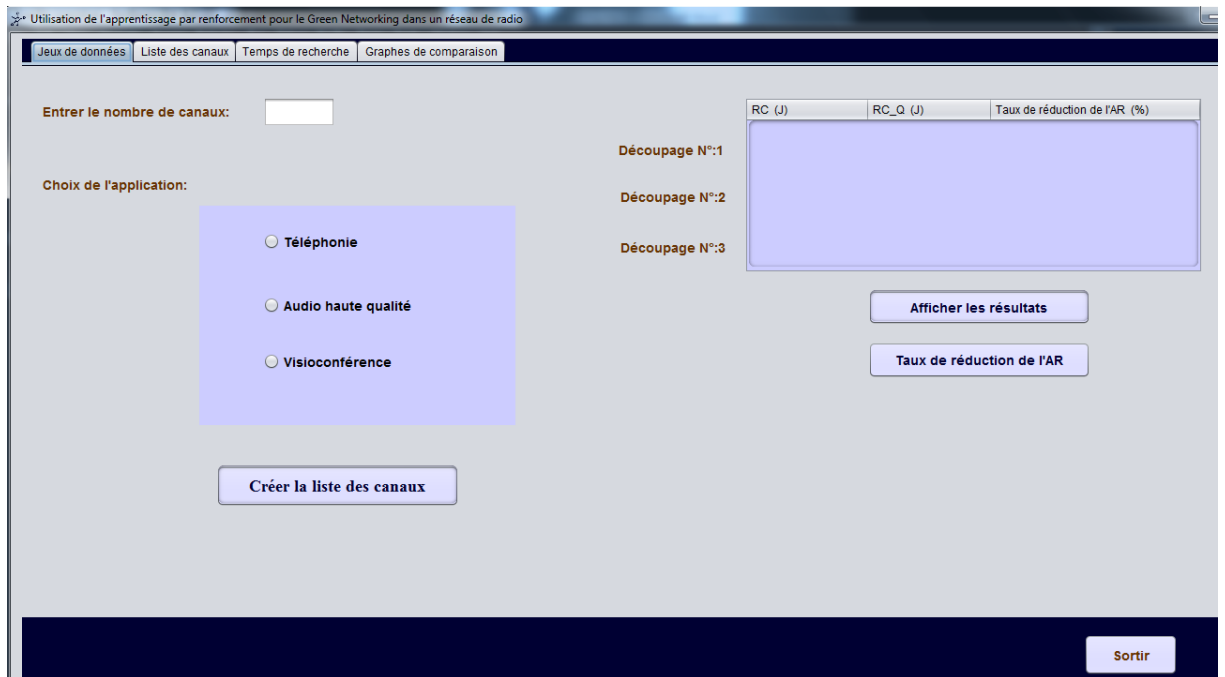


Figure III. 2 : Onglet « Jeux de données »

Commençant par l'onglet « Jeux de données » :

Le tableau ci-dessous représente les trois applications que nous avons utilisées dans ce PFE :

Service	L'application	Débit
Conversationnel (temps réel)	La téléphonie	9.6 kbit/s (GSM)
	La visioconférence	384 kbit/s (3G)
Streaming	Audio haute qualité	128kbit/s (3G)

Tableau III. 1: Les applications utilisées

Dans ce qui suit, nous allons considérer les 3 variables suivantes :

- ✓ **NB_Canaux** : nombre de canaux.
- ✓ **α** : taux de canaux pour lesquels le débit est satisfaisant.
- ✓ **β** : taux de canaux pour lesquels le débit est insatisfaisant.

Par conséquent :

- ✓ Nombre de canaux (débit satisfaisant) = $\alpha * \text{NB_Canaux}$.
- ✓ Nombre de canaux (débit insatisfaisant) = $\beta * \text{NB_Canaux}$.

Tel que : $\alpha + \beta = 1$.

Donc, pour avoir plus de crédibilité dans la tâche de comparaison, nous avons utilisés les 3 découpages suivants :

- ✓ Premier découpage : $\alpha = 0.75$ et $\beta = 0.25$.
- ✓ Deuxième découpage : $\alpha = 0.5$ et $\beta = 0.5$.
- ✓ Troisième découpage : $\alpha = 0.25$ et $\beta = 0.75$.

Le premier découpage favorise les canaux pour lesquels le débit est satisfaisant. Le dernier découpage favorise les canaux pour lesquels le débit est insatisfaisant. Le deuxième découpage est neutre (équitable).

Dans ce qui suit et pour la simulation, nous avons utilisé $NB_Canaux = 80$. Nous avons choisi la **visioconférence** comme application comme indiquer ci-dessous :

Découpage de 80 canaux	Nombre de Canaux (débit insatisfaisant)	Nombre de Canaux (débit satisfaisant)
1 ^{er} découpage	60	20
2 ^{ème} découpage	40	40
3 ^{ème} découpage	20	60

Tableau III. 2 : Les découpages utilisés

The screenshot shows a web form with the following elements:

- A label "Entrer le nombre de canaux:" followed by a text input field containing the value "80".
- A label "Choix de l'application:" followed by a light blue box containing three radio button options:
 - Téléphonie (unselected)
 - Audio haute qualité (unselected)
 - Visioconférence (selected)
- A button at the bottom labeled "Créer la liste des canaux".

Figure III. 3 : Choix de l'application utilisé

En cliquant sur le bouton « **créer la liste des canaux** », nous générons la liste des canaux dans les 3 découpages. L'affichage se fera dans le 2^{ème} onglet : « **Liste des canaux** ».

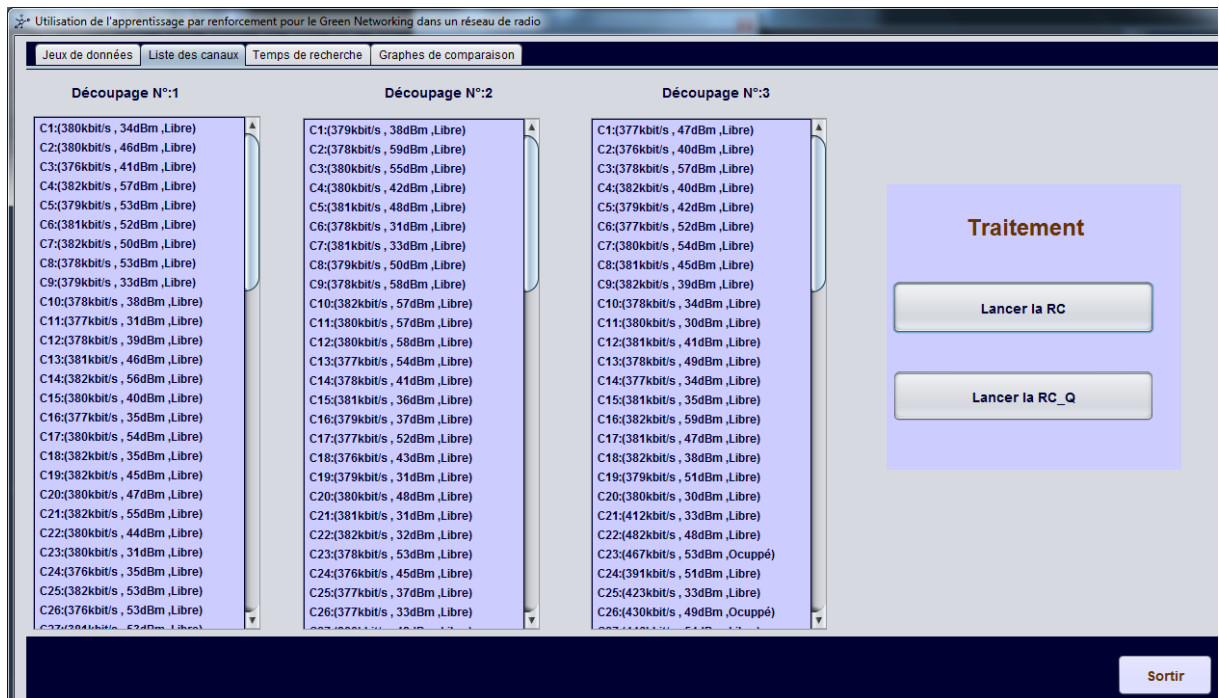


Figure III. 4 : Liste des canaux

Chaque canal est caractérisé par :

- **Le débit** : exprimé en kbit/s.
- **La puissance** : exprimée en dBm.
- **L'état** : libre ou occupé.

Au début de la simulation, tous les canaux sont libres.

Pour les différents résultats obtenus, nous avons considéré la moyenne de **10 simulations** dans chaque découpage.

Le traitement de la RC classique est lancé en cliquant sur le bouton « **Lancer RC** ». Celui de la RC en utilisant le Q-Learning est lancé en cliquant sur le bouton « **Lancer RC-Q** »

III.3.2 Mécanisme de recherche dans la RC classique

La RC classique effectue le sensing pour le choix du canal (trouver un canal avec un débit satisfaisant peu importe la puissance du signal). Une trace d'exécution est montrée dans ce qui suit :

```
****La radio cognitive****  
  
Découpage N°: 1  
  
Simulation N°: 1  
Les différents sauts:  
canal:2  
canal:45  
canal:12  
canal:43  
canal:16  
canal:14  
canal:23  
canal:53  
canal:32  
canal:27  
canal:41  
canal:70  
à partir du canal 1, passer au canal 70 (il devient occupé)  
  
Simulation N°: 2  
Les différents sauts:  
canal:16  
canal:43  
canal:18  
canal:9  
canal:74  
à partir du canal 1, passer au canal 74 (il devient occupé)  
  
Simulation N°: 3  
Les différents sauts:  
canal:68  
à partir du canal 1, passer au canal 68 (il devient occupé)  
  
Simulation N°: 4  
Les différents sauts:  
canal:26  
canal:43  
canal:69  
à partir du canal 1, passer au canal 69 (il devient occupé)  
  
Simulation N°: 5  
Les différents sauts:  
canal:77  
à partir du canal 1, passer au canal 77 (il devient occupé)  
  
Simulation N°: 6  
Les différents sauts:  
canal:37  
canal:2  
canal:70  
canal:19  
canal:6  
canal:2  
canal:77  
canal:19  
canal:72  
à partir du canal 1, passer au canal 72 (il devient occupé)
```

Figure III. 5: Partie de l'exécution de la RC classique du 1^{er} découpage

- Selon la 1^{ère} simulation, nous avons fait 12 sauts pour trouver le canal satisfaisant. Dans la 6^{ème} simulation, c'est uniquement après 3 sauts qu'il a trouvé le canal C70 qui est satisfaisant mais puisque son état est « occupé » (il est déjà pris dans la 1^{ère} simulation), nous élargissons la recherche jusqu'à 6 autres sauts, la même chose pour le canal 77 qui est pris dans la 5^{ème} simulation.
- Chaque fois qu'un canal est sélectionné, il devient occupé.

En cliquant sur le bouton « lancer RC_Q » on lance le traitement de la RC avec QLearning.

L'objectif de cette recherche est d'utiliser le Q-Learning afin de trouver le bon canal pour l'application en un temps réduit.

III.3.3 Apprentissage par renforcement « Q-Learning »

L'Apprentissage par renforcement (AR) permet à un agent d'apprendre à travers une formation sans professeur dans un environnement inconnu visant à maximiser la récompense globale de l'agent vis à vis d'un but à atteindre en considérant uniquement la récompense locale associée à chaque action. Chaque décision prise influant sur les décisions suivantes, l'apprentissage par renforcement peut être vu comme une méthode permettant de résoudre un problème de décision séquentielle.

Les Processus Décisionnels de Markov (PDM) définissent le cadre formel de l'apprentissage par renforcement. L'univers y est fini, stochastique, totalement observable et discret.

Le fait de se placer dans un univers où l'on considère que l'état suivant ne dépend que de l'état courant et de l'action retenue revient à se placer dans un environnement Markovien. Cette hypothèse est communément appelée la "propriété de Markov".

III.3.3.1 L'algorithme Q-Learning

Le Q-Learning est l'un des algorithmes d'apprentissage par renforcement. Il a été proposé comme un moyen d'optimiser les solutions dans les problèmes de processus de décision de Markov. La particularité du Q-Learning est sa capacité à choisir entre des récompenses immédiates et des récompenses différées. A chaque pas de temps, un agent observe le vecteur d'état, puis choisit et applique une action. Lorsque le processus passe à un état, l'agent reçoit un renforcement R. Le but de l'entraînement est de trouver l'ordre séquentiel des actions qui maximise la somme des renforts futurs, conduisant ainsi au chemin le plus court du début à la fin.

La règle de transition du Q-Learning est une formule très simple:

$Q(\text{état}, \text{action}) = Q(\text{état}, \text{action}) + \alpha * (R(\text{état}, \text{action}) + \gamma * \text{Max} [Q(\text{état suivant}, \text{toutes les actions})])$.

Dans ce qui suit, nous allons présenter les variables alpha et gamma ainsi que les deux matrices R et Q.

a. Influence des variables alpha et gamma sur l'algorithme

- ✓ **Vitesse d'apprentissage « alpha »** : détermine à quel point la nouvelle information calculée surpassera l'ancienne. Un facteur de 0 ne ferait rien apprendre à l'agent en question, alors qu'un facteur de 1 ne ferait considérer à l'agent que la dernière information (appelé également **exploration**).
- ✓ **Facteur d'actualisation « gamma »** : détermine l'importance des récompenses futures. Un facteur 0 rendrait l'agent myope en ne considérant seulement les récompenses courantes, alors qu'un facteur proche de 1 ferait intervenir les récompenses les plus lointaines (appelé également **exploitation**).

Dans ce qui suit, nous allons tout d'abord présenter un exemple simple de l'algorithme Q-Learning avant de faire une projection de cet algorithme dans le cadre de notre travail.

b. Exemple simple pour expliquer le fonctionnement du Q-Learning

Supposons que nous ayons 5 chambres dans un bâtiment relié par des portes. Nous numérotons chaque pièce de 0 à 4. L'extérieur du bâtiment peut être considéré comme une grande pièce (5). A noter que les portes 1 et 4 mènent au bâtiment de la pièce 5 (dehors).

Nous allons appeler chaque pièce, y compris à l'extérieur, un «état». Le déplacement de l'agent d'une pièce à une autre est «action». Dans notre diagramme, un "état" est représenté comme un nœud, tandis que "l'action" est représentée par les flèches.

Pour cet exemple, nous supposons que l'agent se trouve dans l'état 2, et à partir de cette pièce, il cherche à se déplacer à l'extérieur du bâtiment (ce sera notre pièce cible). En d'autres termes, la salle visée est celle possédant le numéro 5.

Pour définir cette pièce comme un but, nous allons associer une valeur de récompense à chaque porte. Les portes qui mènent immédiatement au but ont une récompense instantanée de 100. Nous pouvons représenter les pièces sur un graphique, chaque pièce en tant que nœud, et chaque porte en tant que lien.

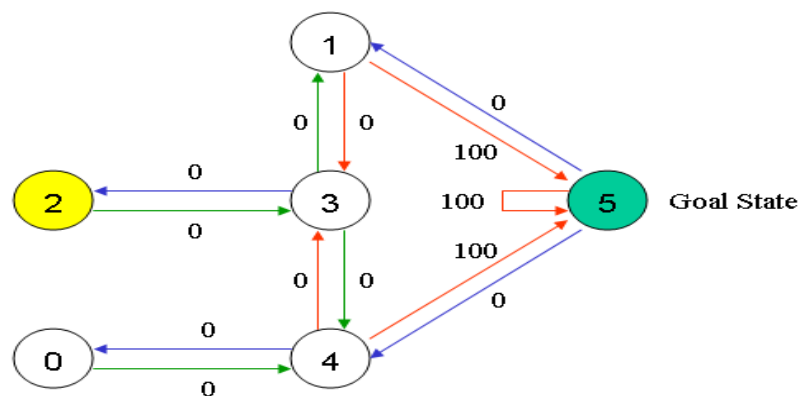


Figure III. 6: Représentation graphique des pièces

Depuis l'état 2, l'agent peut passer à l'état 3 parce que l'état 2 est connecté à 3. De l'état 2, l'agent ne peut pas passer directement à l'état 1 (pas de flèches). A partir de l'état 3, il peut soit aller à l'état 1 ou 4, soit revenir à 2 (voir toutes les flèches à propos de l'état 3). Si l'agent est dans l'état 4, alors les trois actions possibles doivent aller à l'état 0, 5 ou 3. Si l'agent est dans l'état 1, il peut aller soit à l'état 5 ou 3. De l'état 0, il peut seulement aller à l'état 4.

Nous pouvons mettre le diagramme d'état et les valeurs de récompense instantanée dans la table de récompense suivante, "matrice R".

State	Action					
	0	1	2	3	4	5
0	-1	-1	-1	-1	0	-1
1	-1	-1	-1	0	-1	100
2	-1	-1	-1	0	-1	-1
3	-1	0	0	-1	0	-1
4	0	-1	-1	0	-1	100
5	-1	0	-1	-1	0	100

Figure III. 7: Matrice de récompense R

Les -1 dans le tableau représentent des valeurs nulles (c'est-à-dire, où il n'y a pas de lien entre les nœuds (transitions impossibles)). Par exemple, de l'état 0, l'agent ne peut pas aller à l'état 1.

Maintenant, nous allons ajouter une matrice similaire, "Q", à notre agent. Les lignes de la matrice Q représentent l'état actuel de l'agent, et les colonnes représentent les actions possibles menant à l'état suivant (les liens entre les nœuds).

	0	1	2	3	4	5
0	0	0	0	0	80	0
1	0	0	0	64	0	100
2	0	0	0	64	0	0
3	0	80	51	0	80	0
4	64	0	0	64	0	100
5	0	80	0	0	80	100

Figure III. 8: La matrice Q

Une fois que la matrice Q est suffisamment proche de l'état de convergence, nous savons que notre agent a appris les chemins les plus optimaux vers l'état du but. Traçage des meilleures séquences d'états est aussi simple que de suivre les liens avec les valeurs les plus élevées à chaque état.

Par exemple, à partir de l'état initial 2, l'agent peut utiliser la matrice Q comme guide:

A partir de l'état 2, les valeurs maximales de Q suggèrent que l'action passe à l'état 3.

A partir de l'état 3, les valeurs maximales de Q suggèrent deux alternatives: aller à l'état 1 ou 4. Supposons que nous choisissons arbitrairement d'aller à 1.

A partir de l'état 1, les valeurs maximales Q suggèrent que l'action passe à l'état 5.

Ainsi, la séquence est 2 - 3 - 1 - 5.

c. Utilisation du Q-Learning dans notre application

Dans ce PFE, nous avons appliqué le Q-Learning comme suit :

- ✓ **L'état initial** : le canal C1.
- ✓ **L'environnement** : contient 80 canaux.
- ✓ **L'état final** : c'est l'état objectif, c.à.d. trouver un canal ayant un débit satisfaisant pour l'application choisie avec une faible puissance afin de minimiser la consommation d'énergie.

Chaque canal est considéré comme un état, se déplacer d'un canal à un autre est considéré comme une action. L'algorithme aura besoin d'une matrice de récompense R et produira une matrice de quantité Q.

La taille de la matrice R et Q vari selon le nombre de canaux.

1. La matrice de récompense R

Une transition se produit à la suite de l'action à prendre, et un nouvel état est pris en compte. Une récompense est générée en fonction de cette transition. La récompense qui est retournée est alors stockée ou apprise pour cette paire d'actions d'état.

La partie du code suivante montre le remplissage de la matrice R :

```
for(int i=0;i<NB_Canaux;i++){
for (int j=0;j<NB_Canaux;j++){

//si le débit du canal ne satisfait pas le débit de l'application
if(Liste_canaux.get(i).Debit<debit_application){
R[j][i]=penalty; //pénalité = -1
}

//si le débit du canal satisfait le débit de l'application
else {
R[j][i]=0;
//ajouter tous les canaux qui ont un débit satisfaisant à la liste "L_canaux_satisf"
L_canaux_satisf.add(Liste_canaux.get(i));
}
}

//trouver les canaux parmi les canaux de la liste "L_canaux_satisf" qui ont une puissance plus faible
ArrayList<Canal>FP=Faible_Puissance(L_canaux_satisf);
int i=0;
while(FP.get(0).puissance == FP.get(i).puissance && i<FP.size()){
for(int j=0;j<NB_Canaux;j++){
R[j][FP.get(i).Nom]=reward; // récompense= 100
}
i++;
}
```

Figure III. 9: Partie du code pour remplir la matrice R

Chapitre III : Implémentation de l'application et évaluation des résultats

La matrice R contient pour chaque état une ligne d'action remplie comme suit :

- ✓ **100** est la récompense pour l'état final (le canal optimal) celui qui définit une transition possible avec une puissance plus faible par rapport à tous les autres canaux.
- ✓ **0** pour les transitions possibles dans le cas où le débit d'un canal est satisfaisant.
- ✓ **-1** est la pénalité : pour les transitions impossibles dans le cas où le débit du canal est insatisfaisant.

. Les lignes représentent l'état actuel de l'agent, et les colonnes représentent les actions possibles menant à l'état suivant.

La figure suivante présente la matrice de récompense R

Figure III. 10: Partie d'exécution de la matrice R du premier découpage

Par exemple, en choisissant les trois canaux C15, C50 et C71 :

Canal	Débit	Puissance
C15	381 Kbit/s	35 dBm
C45	428 Kbit/s	44 dBm
C67	392 Kbit/s	30 dBm

Tableau III. 3: Exemple sur quelques canaux

- C15 définit une transition impossible parce que son débit est inférieur à celui de la visioconférence (384kbit/s). Donc il va avoir une pénalité -1.

- Les canaux C45 et C67 définissent des transitions possibles parce que le débit de chaque canal est supérieur au débit de la visioconférence (le débit est satisfaisant).
- C67 définit une faible puissance par rapport à C45 donc il va avoir une récompense (+100).

2. La matrice de quantité Q

La lettre 'Q' désigne la fonction qui mesure la qualité d'une action exécutée dans un état donné du système, elle représente la mémoire de ce que l'agent a appris par l'expérience. Avant que l'apprentissage ne débute, la fonction Q est initialisée arbitrairement. Ensuite, après avoir la matrice R, à chaque choix d'action et avec une observation de la récompense et le nouvel état (qui dépend de l'état précédent et de l'action actuelle). Ainsi, est mis à jour. Le cœur de l'algorithme est une mise à jour de la fonction de valeur. La définition de la fonction de valeur est corrigée à chaque étape comme montre l'algorithme suivant :

1. Définissez le paramètre gamma et les récompenses d'environnement dans la matrice R.

2. Initialisez la matrice Q à zéro.

3. Pour chaque cycle d'entraînement

Sélectionnez un état initial aléatoire.

Tant que l'état final n'a pas été atteint

Sélectionnez l'une parmi toutes les actions possibles pour l'état actuel.

En utilisant cette action possible, pensez à passer à l'état suivant.

Obtenez la valeur Q maximale pour cet état suivant en fonction de toutes les actions possibles.

Calculer Q:

$Q(\text{état}, \text{action}) = Q(\text{état}, \text{action}) + \alpha * (R(\text{état}, \text{action}) + \gamma * \text{Max}(\text{état suivant}, \text{toutes les actions}) - Q(\text{état}, \text{action}))$

Définissez l'état suivant comme état actuel.

Fin Tant que

Fin Pour

L'objectif de répéter ce processus est qu'un canal trouve une politique optimale pour chaque état avec une manière récursive. Le fait que l'algorithme Q-Learning converge sans aucune connaissance préalable sur l'environnement le rend très approprié dans un contexte radio cognitive en raison de l'imprévisibilité de ce type de réseau.

La figure suivante présente une partie d'exécution de la matrice Q:

La matrice Q:

à partir du canal 1:	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	16,00	19,20		
à partir du canal 2:	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	19,97	19,20		
à partir du canal 3:	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	19,20	19,84		
à partir du canal 4:	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	19,84	19,97		
à partir du canal 5:	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	19,83	19,20		
à partir du canal 6:	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	19,20		
à partir du canal 7:	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	12,80	16,00		
à partir du canal 8:	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	19,71	15,99		
à partir du canal 9:	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	19,20	19,99		
19,84	19,84	16,00	19,84	19,20	19,20	19,97	16,00	20,00	19,99	0,00	0,00	19,84	19,97	19,97	19,20	19,84	16,00	80,00	19,84	16,00	19,20	19,84	19,20	16,00
19,84	19,84	0,00	16,00	19,20	16,00	0,00	19,20	19,84	19,84	19,84	19,20	19,97	19,20	0,00	16,00	19,20	19,20	96,00	19,20	19,20	19,84	16,00	19,20	19,20
19,20	19,84	19,84	16,00	19,83	19,20	19,20	16,00	19,19	20,00	19,20	16,00	19,19	16,00	0,00	19,20	0,00	19,20	80,00	19,84	19,20	19,97	16,00	19,20	16,00
16,00	0,00	19,97	19,97	19,97	19,84	16,00	16,00	0,00	20,00	19,84	16,00	16,00	19,20	19,84	16,00	19,99	19,97	80,00	19,97	19,99	19,20	0,00	16,08	19,20
16,00	19,20	19,20	19,84	19,97	19,99	19,20	19,20	0,00	19,20	19,20	19,84	16,00	16,00	19,20	19,20	16,00	19,84	96,00	16,00	19,84	16,00	16,00	16,00	19,20
0,00	19,20	19,20	16,00	19,84	19,84	19,97	19,20	19,20	16,00	19,20	19,07	19,84	19,20	0,00	0,00	0,00	19,20	80,00	15,87	19,84	16,00	19,97	19,97	19,20
19,84	19,97	19,20	19,97	19,84	19,99	19,84	19,20	16,00	19,20	19,84	19,20	19,99	19,20	0,00	19,20	0,39	0,00	96,00	19,20	0,00	19,84	0,00	0,00	19,20
16,00	16,00	16,00	19,99	16,00	0,00	16,00	16,00	19,30	16,00	19,07	19,20	19,20	19,84	19,99	19,17	19,84	19,84	80,00	0,00	19,99	16,00	19,20	12,80	19,97
19,20	16,00	19,84	19,97	19,84	19,99	0,00	19,20	19,20	16,00	19,20	0,00	19,20	19,99	16,00	19,84	0,00	16,00	80,00	19,20	0,00	19,97	19,20	19,97	19,84

Figure III. 11: Partie d'exécution de la matrice Q

Nous remarquons que :

- Les valeurs égales à 0 sont les transitions impossibles dans la matrice R.
- Les valeurs qui convergent vers 100 sont les transitions possibles avec une faible puissance dans la matrice R.
- Les autres valeurs sont les transitions possibles.

Dans ce PFE, nous avons pris :

- ✓ **alpha = 0,8** : une probabilité de **80%** qu'une action aléatoire sera choisie (exploration).
- ✓ **gamma = 0,2** : une probabilité de **20%** que la meilleure (ou **max Q** valeur) action sera choisie (exploitation).

Après le remplissage de la matrice Q, le meilleur canal est celui qui a la valeur maximale dans la matrice, l'algorithme suivant explique l'utilisation de la matrice Q

Entrée : (matrice Q, état initial)

1. Définir l'état actuel=état initial
 2. À partir de l'état actuel, rechercher l'action qui produit la valeur Q maximale.
 3. Définir l'état actuel=état suivant.
 4. Aller jusqu'à l'état actuel=état final.
-

III.3.4 Résultats obtenus

III.3.4.1 Comparaison par rapport au temps de recherche

En considérant une unité de temps par saut, le temps de recherche dans les deux cas (RC et RC avec Q-Learning) est égal aux nombres de sauts nécessaires pour trouver le meilleur canal pour l'application.

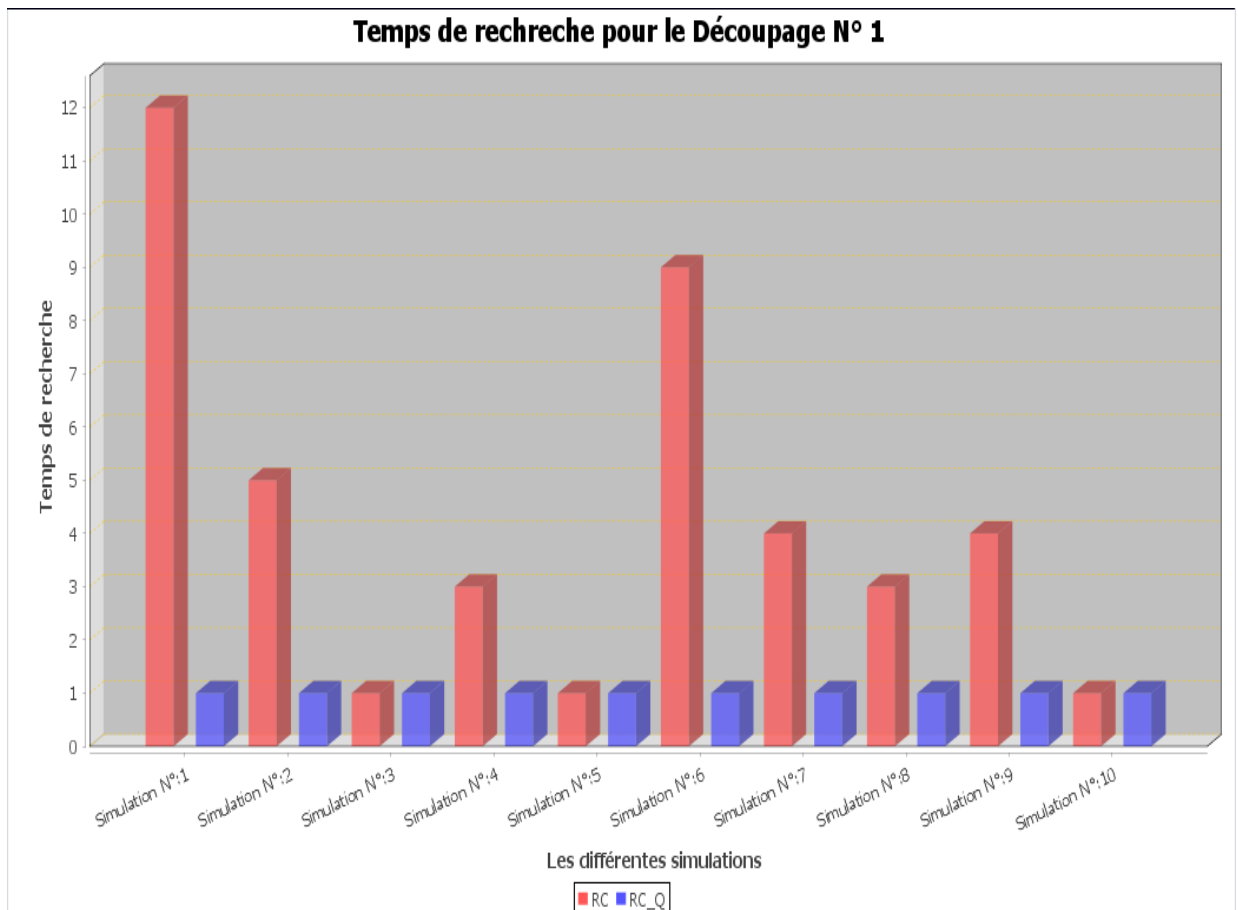


Figure III. 12: Temps de recherche pour le découpage N°1 (RC vs RC_Q)

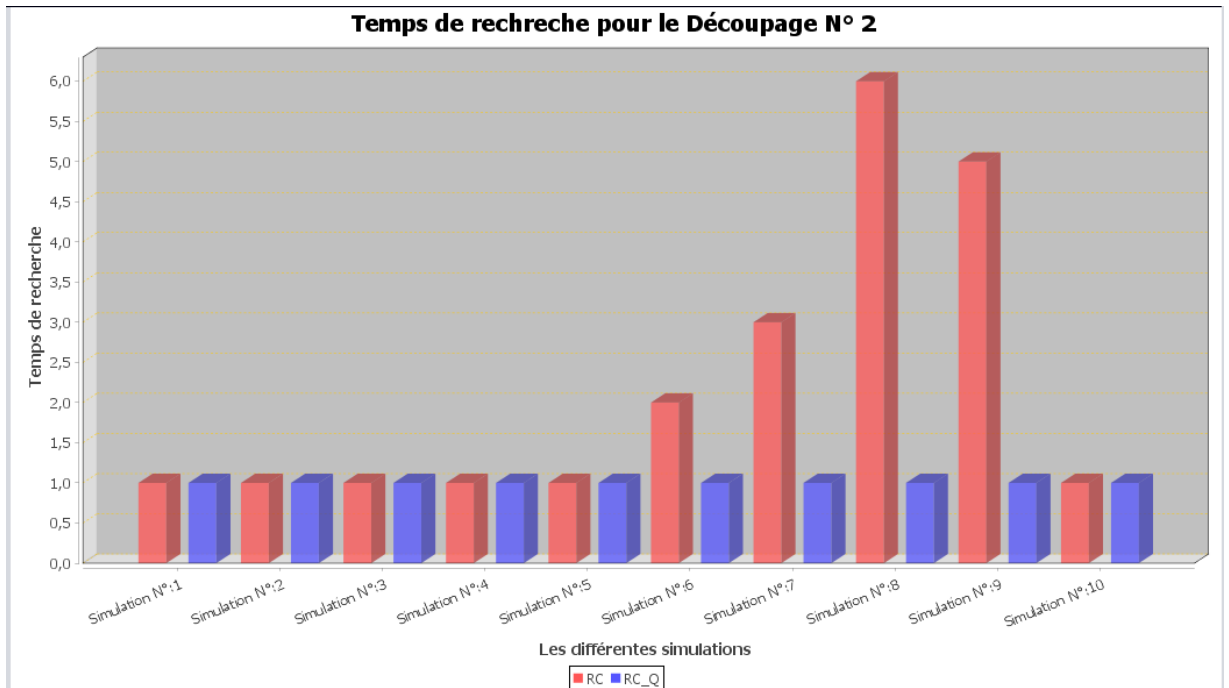


Figure III. 13: Temps de recherche pour le découpage N°2 (RC vs RC_Q)

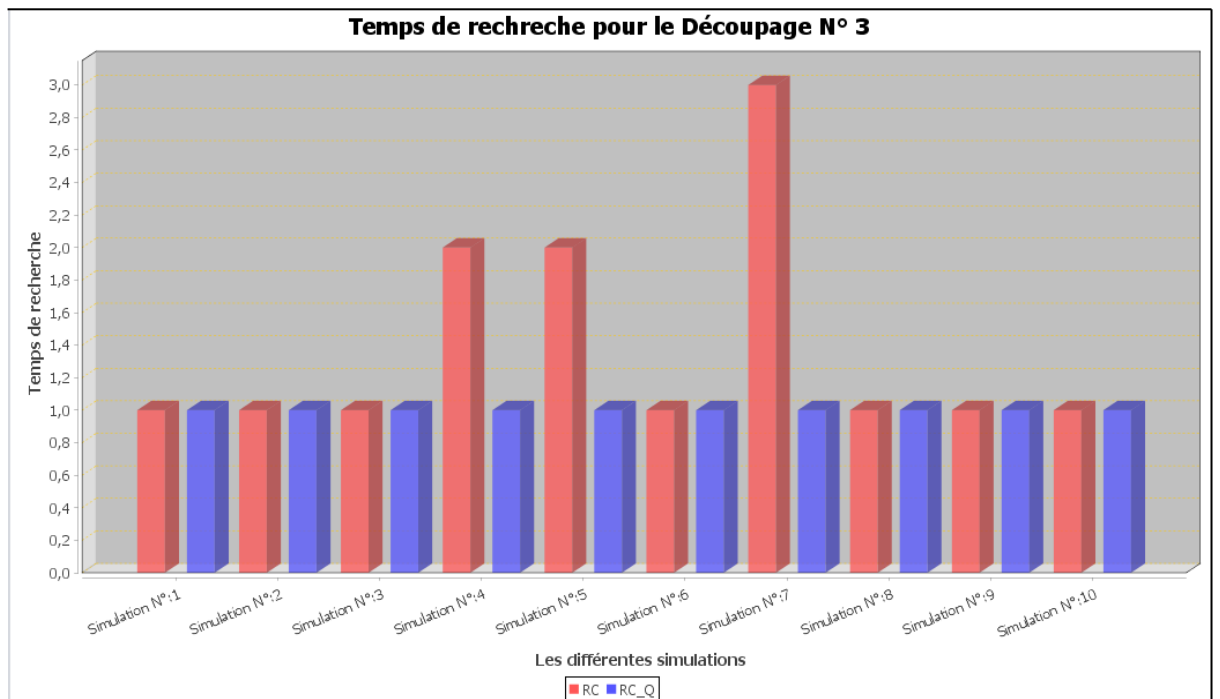


Figure III. 14: Temps de recherche pour le découpage N°3 (RC vs RC_Q)

En analysant les trois figures ci-dessus :

- D'une part, les différents découpages de canaux n'ont aucune influence sur la RC avec Q-Learning, nous avons toujours un seul saut pour trouver le bon canal et donc il consomme une seule unité de temps peu importe le découpage.
- D'une autre part, dans la RC classique, nous remarquons que nous allons faire beaucoup plus de sauts, quand le nombre de canaux qui ont un débit insatisfaisant est plus grand (découpage 1), une chose qui implique plus de temps de recherche. Par contre, avec un nombre important de canaux qui ont un débit satisfaisant (découpage 3) nous avons moins de sauts pour atteindre un bon canal et par conséquent le temps de recherche est réduit.

La figure suivante représente le temps moyen de la recherche dans la RC pour les 3 découpages.

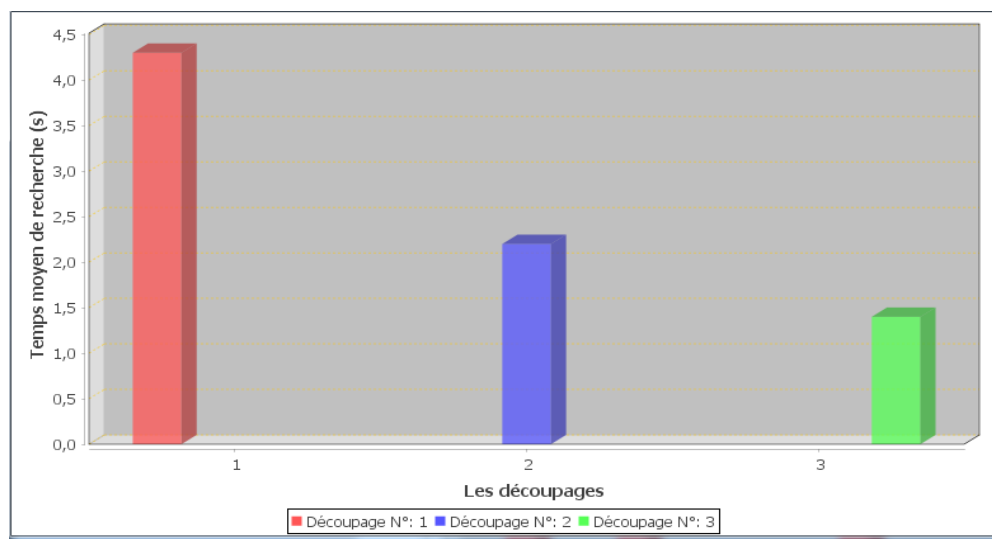


Figure III. 15: Temps moyen de recherche dans la RC classique

III.3.4.2 Comparaison par rapport à la consommation d'énergie

L'énergie est calculée par la formule suivante : $E=i*v*t$.

Tel que :

$i*v$: représente la puissance de transmission.

t = temps de recherche + temps de transmission.

Le temps de recherche : le nombre de saut (en consommant une unité de temps par saut).

Le temps de transmission = $\frac{\text{la taille du paquet}}{\text{débit}}$

Chapitre III : Implémentation de l'application et évaluation des résultats

Dans ce qui suit, nous avons pris une taille de paquet fixe pour toutes les comparaisons : **10kbit**.

La figure ci-dessous présente les graphes de comparaison de l'énergie consommée de la RC classique et de la RC avec Q-Learning dans les différents découpages :

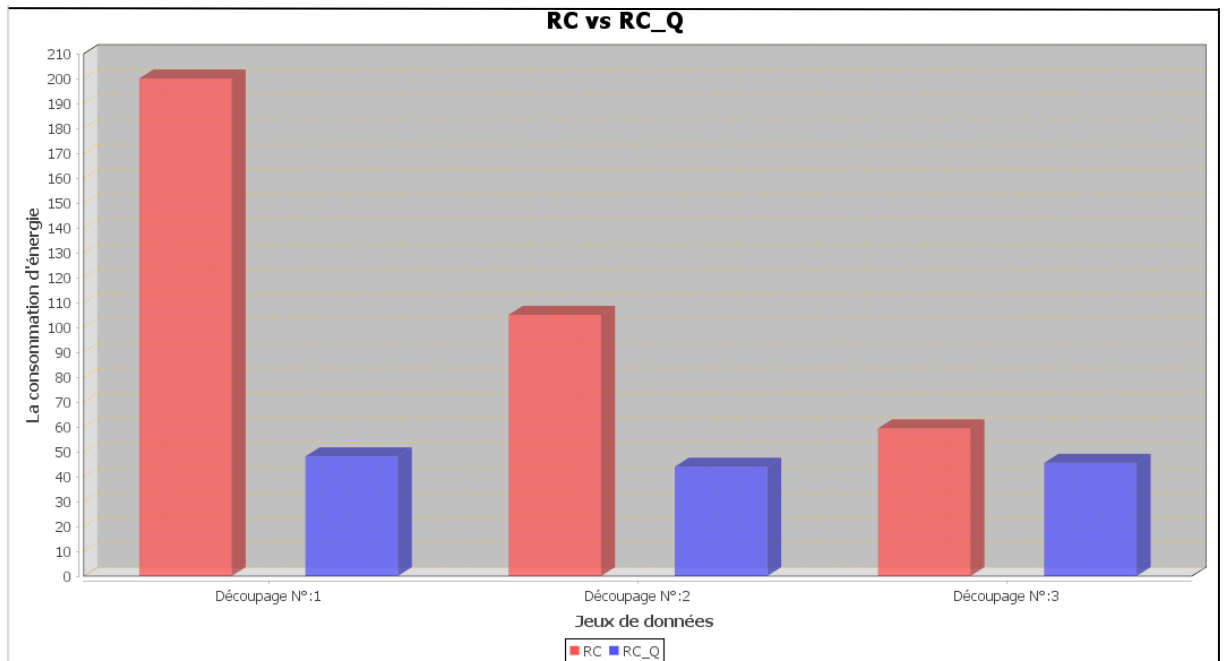


Figure III. 16: Consommation d'énergie de la RC et de la RC_Q dans les trois découpages

La figure suivante résume les résultats obtenus par rapport à l'énergie consommée dans les deux cas.

	RC (J)	RC_Q (J)	Taux de réduction de l'AR (%)
Découpage N°:1	200.0	48.1	75
Découpage N°:2	105.0	44.0	58
Découpage N°:3	59.4	45.5	23

Afficher les résultats

Figure III. 17: Tableau récapitulatif qui montre l'énergie consommée

1. Taux de réduction :

Nous avons calculé le taux de réduction pour montrer le gain obtenu (en termes de pourcentage) en utilisant l'AR dans chaque découpage de la façon suivante : $(RC(J) - RC_Q(J)) / RC(J)$.

L'AR a réduit de 75% l'énergie consommée dans le premier découpage. 58 % dans le deuxième et 23 % dans le troisième. Ce qui est montré dans la figure suivante :

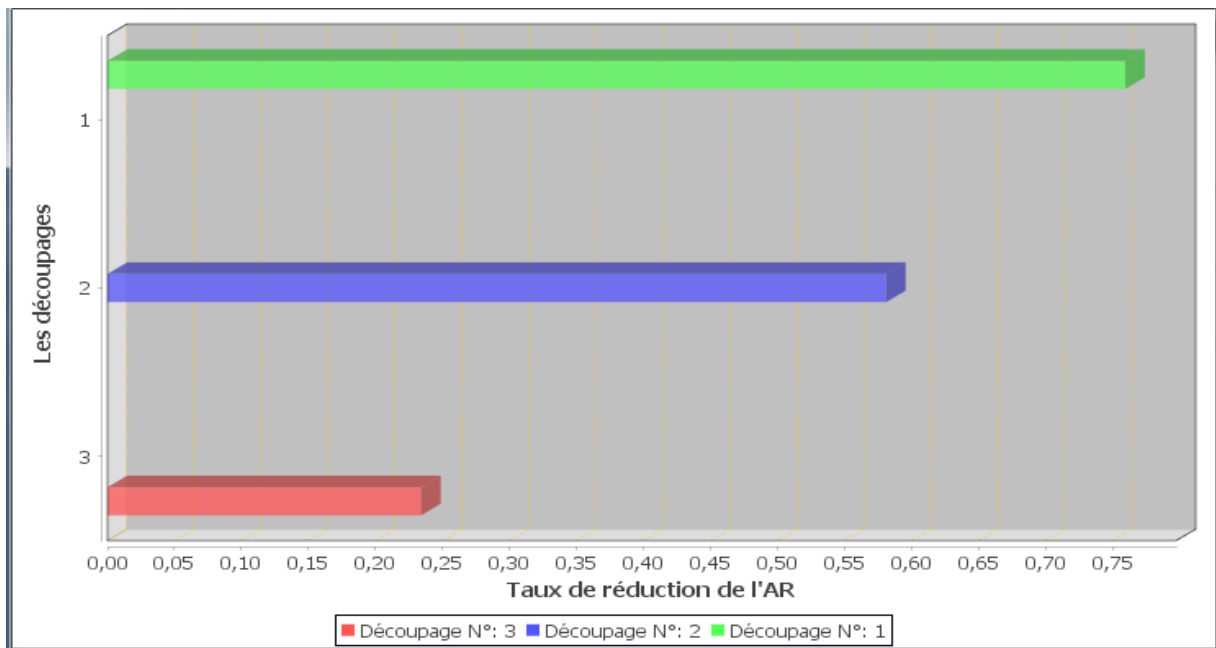


Figure III. 18: Taux de réduction de l'AR

Comme conclusion, nous pouvons affirmer que l'utilisation du Q-Learning a permis au nœud RC de réduire sa consommation d'énergie (peu importe le découpage utilisé). Ce résultat est très intéressant car il montre l'intérêt de ce type d'algorithme dans le contexte du Green Networking

III.4 Conclusion

Nous avons présenté dans ce chapitre notre contribution dans le cadre de ce PFE. En effet, nous avons montré avec des simulations que l'utilisation du Q-Learning dans un contexte radio cognitive donne des résultats très satisfaisants par rapport à la consommation d'énergie. Une chose qui est très importante dans le cadre du Green Networking.

CONCLUSION GÉNÉRALE

Conclusion générale

Dans ce projet de fin d'études, la radio cognitive a été présentée comme un nouveau modèle de communication sans fil, une technologie très prometteuse, visant à réduire le manque de ressources sans fil. Ceci est possible grâce à la distribution dynamique des fréquences. Cela est dû à sa rapidité et sa capacité à adapter intelligemment les standards de communication, ce qui nous permet de réduire la consommation d'énergie et de diriger les réseaux sans fil vers des réseaux «verts».

Nous avons mis en place un mécanisme pour rendre les communications sans fil plus éco énergétiques. Par conséquent, nos solutions utilisent les fonctions de reconfiguration et d'adaptation de la radio cognitive et collectent des informations sur les plages à utiliser. Permettre à ce dernier point de devenir possible grâce à un apprentissage amélioré pour augmenter la durée de vie du réseau et réduire son impact sur l'énergie.

La solution que nous avons proposée dans les réseaux de la radio cognitive qui consiste à utiliser l'apprentissage par renforcement dans le but de déterminer le meilleur canal avec moins de puissance pour une transmission de données quelque soit l'application choisi.

Nous avons travaillé sur les mêmes jeux de données dans les deux cas : RC et RC_Q et nous avons trouvé que dans la RC classique, le temps de recherche pour trouver un canal satisfaisant et qui dépend du nombre de sauts est très élevé en comparant avec la RC_Q qui nécessite seulement une seule unité de temps.

Les résultats obtenus montrent que l'algorithme du Q-Learning aidera l'utilisateur à choisir le canal optimal pour son application qui caractérisé par une faible puissance de signal, contrairement à la RC classique où le même utilisateur pourra utiliser un canal avec une grande puissance de transmission et par conséquent consomme plus d'énergie.

Finalement, l'apprentissage par renforcement à prouver l'intérêt de la radio cognitive pour le Green Networking.

RÉFÉRENCES

BIBLIOGRAPHIQUES

Références bibliographiques

- [1] Mitola J, Maguire GQ. Cognitive radio: making software radios more personal. IEEE personal communications. 1999 Aug;6(4):13-8.
- [2] Haykin S. Cognitive radio: brain-empowered wireless communications. IEEE journal on selected areas in communications. 2005 Feb;23(2):201-20.
- [3] Amraoui A, Benmammar B, Bendimerad FT. Accès Dynamique au Spectre dans le Contexte de la Radio Cognitive. In 2^{ème} édition de la conférence nationale de l'informatique destinée aux étudiants de graduation et de post-graduation 2012 Apr 16.
- [4] Metref A. *Contribution à l'étude du problème de synchronisation de porteuse dans le contexte de la Radio Intelligente* (Doctoral dissertation, Université Rennes 1).
- [5] Amraoui A, Benmammar B, Krief F, Bendimerad FT. Intelligent wireless communication system using cognitive radio. International Journal of Distributed and Parallel Systems. 2012 Mar 1;3(2):91
- [6] Akyildiz IF, Lee WY, Vuran MC, Mohanty S. NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey. Computer networks. 2006 Sep 15;50(13):2127-59
- [7] Ben Dhaou, Ahmed. "Allocation dynamique des bandes spectrales dans les réseaux sans-fil à radio cognitive." (2011).
- [8] I. Ngom et L. Diouf, "la radio cognitive", université Lille 1 USTL, 2008
- [9] Benmammar, B., Amraoui, A., & Krief, F. (2013). A survey on dynamic spectrum access techniques in cognitive radio networks. *International Journal of Communication Networks and Information Security*, 5(2), 68.
- [10] Hossain, E., Niyato, D., & Han, Z. (2009). *Dynamic spectrum access and management in cognitive radio networks*. Cambridge university press.
- [11] Benmammar, Badr, and Asma Amraoui. *Radio resource allocation and dynamic spectrum access*. John Wiley & Sons, 2013.
- [12] <https://www.theclimategroup.org/sites/default/files/archive/files/Smart2020Report.pdf> , Smart report: Enabling the low carbon economy in the information age.
- [13] HEDNA, Saida. *Gestion De L'économie D'énergie Dans Les Réseaux Sans Fil 802.11 Ad Hoc*. Diss. Université de Batna 2, 2011.
- [14] M. Bloem, T. Alpcan, and T. Basar, "A Stackelberg Game for Power Control and Channel Allocation in Cognitive Radio Networks," in *GameComm 2007*. France, Oct. 2007..
- [15] Canming Jiang, Yi Shi, Y. Thomas Hou, and Sastry Kompella. On optimal throughput-energy curve for multi-hop wireless networks. In INFOCOM 2011. The 30th Conference on Computer Communications. IEEE, April 2011.
- [16] Christine E. Jones, Krishna M. Sivalingam, Prathima Agrawal, and Jyh Cheng Chen. A survey of energy efficient network protocols for wireless networks. *Wirel. Netw.*, 7:343-358, 2001..

Références bibliographiques

- [17] Marsan, Marco Ajmone, and Michela Meo. "Energy efficient management of two cellular access networks." *ACM SIGMETRICS Performance Evaluation Review* 37.4 (2010): 69-73..
- [18] GRACE, David, CHEN, Jingxin, JIANG, Tao, et al. Using cognitive radio to deliver 'green' communications. In : 2009 4th International Conference on Cognitive Radio Oriented Wireless Networks and Communications. IEEE, 2009. p. 1-6..
- [19] OUNI, Anis, RIVANO, Hervé, et VALOIS, Fabrice. Réduction de la consommation d'énergie des réseaux cellulaires hétérogènes. In : *15èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications (AlgoTel)*. 2013. p. 1-4..
- [20] KRIEF, Francine (ed.). *Green Networking*. John Wiley & Sons, 2012.
- [21] J. Mitola, "Cognitive Radio," Licentiate proposal, KTH, Stockholm, Sweden
- [22] MICHAEL, Navin, MOY, Christophe, VINOD, Achutavarrier Prasad, *et al.* Area-power trade-offs for flexible filtering in green radios. *Journal of Communications and Networks*, 2010, vol. 12, no 2, p. 158-167.
- [23] PALICOT, Jacques. Cognitive radio: an enabling technology for the green radio communications concept. In : *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*. ACM, 2009. p. 489-494.
- [24] MALEKI, Sina, PANDHARIPANDE, Ashish, et LEUS, Geert. Energy-efficient distributed spectrum sensing for cognitive sensor networks. *IEEE sensors journal*, 2011, vol. 11, no 3, p. 565-573
- [25] IEEE ICC 2014 - Wireless Communications Symposium Green Opportunistic Access for Cognitive Radio Networks: A Minority Game Approach
- [26] CORNUÉJOLS, Antoine et MICLET, Laurent. *Apprentissage artificiel: Concepts et algorithmes*. 2010.
- [27] FILIPPI, Sarah. *Stratégies optimistes en apprentissage par renforcement*. 2010. Thèse de doctorat. Ecole nationale supérieure des télécommunications-ENST.
- [28] OUATTARA, D. *Application de l'apprentissage aux réseaux Radio Cognitive*. Computer Science and Game Theory, 2011.

RÉSUMÉ :

Le nouveau concept « Green Networking » peut profiter du large paradigme de la radio cognitive. L'objectif de notre PFE est de trouver un mécanisme qui permet de minimiser la consommation d'énergie, en l'intégrant au réseau de la radio cognitive. Pour cela, nous avons utilisé l'algorithme de Q-Learning, une technique d'apprentissage par renforcement qui aidera l'utilisateur cognitif à trouver le canal optimal qui a une faible puissance de transmission en garantissant les besoins de son application, et donc une réduction dans la consommation de la batterie.

Mots de clés : Radio Cognitive, Green Networking, Apprentissage par renforcement, Q-Learning.

ABSTRACT:

The new concept "Green Networking" can take advantage of the broad paradigm of cognitive radio. The goal of our final master's project is to find a mechanism that minimizes energy consumption by integrating it into the cognitive radio network. For this, we used the algorithm of Q-Learning, a reinforcement learning technique that will help the cognitive user to find the optimal channel that has a low transmission power by guaranteeing the needs of his application, and therefore a reduction in battery consumption.

Key words: Cognitive Radio, Green Networking, reinforcement learning, Q-Learning.

ملخص:

يمكن لمفهوم "الشبكات الخضراء" الجديد الاستفادة من النموذج الواسع للراديو الإدراكية. الهدف من مشروع نهاية دراستنا هو العثور على آلية تقلل من استهلاك الطاقة من خلال دمجها في الشبكة اللاسلكية الإدراكية. لهذا ، استخدمنا خوارزمية Q-Learning وهي تقنية تعلم معزز ستساعد المستخدم المعرفي على إيجاد القناة المثلى و التي تتمتع بقوة إرسال منخفضة من خلال ضمان احتياجات التطبيق ، وبالتالي انخفاض في استهلاك البطارية.

الكلمات المفتاحية: الراديو الإدراكية، الشبكات الخضراء ، خوارزمية التعلم المعزز ، خوارزمية كمية التعلم.