

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE  
UNIVERSITE FARHAT ABBAS – SETIF-1  
(UFAS-1) - (ALGERIE)



Présenté à la Faculté des Sciences  
Département d'Informatique  
Pour l'Obtention du Diplôme de

## **DOCTORAT en SCIENCES**

Par

**Mme. : Nadia ZERGUINE**

### **Thème**

# **Réseaux sans fil et applications temps réel**

Soutenu le : 27/06/2021

devant la commission d'examen :

Pr. ALIOUAT Makhlouf

Prof. à l'UFA- Setif-1

Président

Pr. MOSTEFAI Mohammed

Prof. à l'UFA- Setif-1

Rapporteur

Pr. BENMOHAMED Mohamed

Prof. à l'Université de Constantine

Examineur

Dr. BOUDRIES Abdelmalek

MCA. à l'Université de Bejaia

Examineur

Pr. ALIOUAT Zibouda

Prof. à l'UFA- Setif-1

Invité

*“A ma mère”, “A ma mère”, “A ma mère”*

*“A la mémoire de mon père”*

*“A mon mari”*

*“A mes enfants”*

*“A mes frères et sœurs”*

*“A ma famille”*

*“A tous ceux qui me sont chers”*

---

# Remerciements

---

Tout d'abord, et non des moindres, je voudrais remercier Dieu beaucoup, pour m'avoir donné le pouvoir de terminer cette thèse après toutes ces années.

Je tiens à exprimer mes chaleureux remerciements et toute ma gratitude à mon directeur **Mohammed MOSTEFAI** Professeur à l'Université Ferhat Abbes de Sétif pour m'avoir confié ce travail de recherche et avoir accepté de le diriger, ainsi que pour son aide, ses précieux conseils et surtout sa patience au cours de ces années. Je suis très honorée de l'avoir eu pour directeur.

Je souhaiterais également, exprimer ma gratitude à Madame **Zibouda ALIOUAT** Professeur à l'Université Ferhat Abbes de Sétif, pour la confiance qu'elle m'a témoigné du début à la fin du travail, pour sa grande disponibilité à comprendre et à communiquer, sans oublier ses précieuses intuitions. Je suis ravie d'avoir travaillé en sa compagnie car, outre son appui scientifique, elle a toujours été là pour me soutenir et me conseiller au cours de l'élaboration de cette thèse. Enfin, j'ai été extrêmement sensible à ses qualités humaines d'écoute et de compréhension tout au long de ce travail doctoral. Sincèrement, grâce à elle j'ai pu apprendre beaucoup de choses dont certains forts utiles pour mes travaux académiques bien sûr, mais aussi des choses importantes pour mon développement personnel.

J'aimerais leur adresser tous les deux mes plus vifs remerciements pour tout leur dynamisme, leur tolérance, leur compétences scientifiques que j'ai pu apprécier tout au long de ces années. Ce travail n'aurait jamais pu aboutir sans eux qui ont toujours su me consacrer des moments de leur temps et me témoigner leur soutien et leur confiance. Je souhaite leur transmettre l'expression de ma plus profonde et sincère gratitude.

Un remerciement tout particulier pour **Yacine SLIMANI**, Maître de Conférences à l'Université Ferhat Abbes de Sétif, dont le dynamisme m'a particulièrement encouragé dans la dernière ligne droite. Sa disponibilité pour m'aider à mener à bien la partie expérimentale de ce travail, je le remercie de tout mon cœur.

Je remercie les membres de jury qui m'ont fait l'honneur d'accepter de participer au jury.

Mes remerciements vont également à tous ceux qui ont contribué de près ou de loin à l'élaboration de cette thèse, qu'ils trouvent ici l'expression de ma profonde sympathie.

Une pensée avant de terminer ces remerciements pour toi qui n'a pas vu l'aboutissement de mon travail mais je sais que tu en aurais été très fier de ta fille !!!

Mes derniers remerciements vont à ma famille et notamment à ma mère, mon mari, mes enfants, mes frères, mes sœurs, et tous ceux qui me sont chers, car sans leur croyance en moi, leur constant soutien et leurs encouragements, je n'aurais pas pu mener à bien cette thèse. Cette page serait loin de suffire pour leur exprimer toute ma reconnaissance et mon affection.

Merci infiniment encore à tous.

## ملخص

الشبكات اللاسلكية المتحركة المتخصصة (Mobile Adhoc Network) عبارة عن مجموعة من العقد اللاسلكية التي تشكل شبكة ديناميكية دون أي بنية أساسية موجودة مسبقاً أو بنية مركزية. تستخدم MANETs معيار IEEE 802.11 . للتعامل مع التصادمات في الشبكات اللاسلكية، تستخدم طبقة MAC الخاصة بـ IEEE 802.11 DCF إجراء CSMA / CA للوصول إلى مورد القناة الذي يعتمد على آلية Binary Exponential Backoff (BEB) ، وهذا الأخير يقلل من احتمالية حدوث تصادم ولكن على حساب العديد من مقاييس أداء الشبكة مثل التأخير ، والتدفق ، والإنصاف ، وهي ليست مناسبة للتطبيقات الفورية. لتلبية المتطلبات المختلفة لـ MANET ، يتم توفير وصول فعال للقناة من خلال استنتاج متطور للتحكم في الاصطدامات. مثل هذا الذكاء ممكن بفضل إدخال تقنيات التعلم العميق (DL) في MANETs. في هذه الأطروحة ، تم تطوير نموذج ذكي قائم على DRL لتحسين الوصول إلى قناة طبقة MAC في MANETs تستخدم Q-Learning (QL) وهي إحدى تقنيات DRL المستخدمة لتطوير الآلية الذكية. يأخذ اقتراح MISQ الذكي في الاعتبار بالإضافة إلى المعلومات المتعلقة بحالة مرور أو فشل الإرسال، عدد الحزم التي سيتم إرسالها وعدد التصادمات التي تحدث من قبل كل محطة لتحديد نافذة التنافس المناسبة، من أجل تحديد الأولويات إلى المحطات التي لديها المزيد من الحزم لإرسالها، للنجاح بسرعة في الوصول إلى القناة. تشير النتائج إلى أن الآلية الذكية المقترحة تتعلم بيئات MANETS المختلفة وتحسن الأداء عبر بروتوكول MAC الأساسي. يتم تقييم أداء MISQ في شبكات مختلفة مع التدفق وتأخير الوصول إلى القناة ومعدل تسليم الحزم كمقاييس للأداء.

**الكلمات الرئيسية:** BEB، CSMA / CA، DCF، DRL، MANETs، الوقت الحقيقي، تحسين الأداء، Q-Learning

## Résumé

Un réseau mobile Ad Hoc *MANET* (*Mobile Adhoc Network*) est une collection de stations sans fil formant un réseau dynamique sans infrastructure préexistante ou une architecture centralisée. MANETs utilisent la norme IEEE 802.11. Pour traiter les collisions dans les réseaux sans fil, la couche MAC de l'IEEE 802.11 DCF utilise une procédure CSMA / CA pour accéder à la ressource canal qui est basé sur un mécanisme Binary Exponential Backoff (BEB), ce dernier réduit la probabilité de collision mais au prix de nombreuses mesures des performances réseau telles que le délai, débit, et l'équité, ce qui n'est pas adéquat aux applications temps réel. Pour satisfaire les diverses exigences des MANETs, un accès efficace au canal est prévu grâce à une inférence sophistiquée afin de contrôler les collisions. Une telle intelligence est possible grâce à l'introduction de techniques d'apprentissage profond (DL) dans les MANETs. Dans cette thèse, un paradigme intelligent basé sur DRL est développé pour optimiser l'accès au canal de la couche MAC dans les MANETs. Q-Learning (QL), L'une des techniques DRL, est utilisé pour proposer un mécanisme intelligent d'accès au canal dans les MANETs. La proposition intelligente MISQ prend en compte en plus de l'informations sur l'état de la transmission réussite ou échec, le nombre de paquets à transmettre et les collisions commises par chaque station pour sélectionner la fenêtre de contention appropriée, afin de donner une priorité aux stations qui ont plus de paquets à transmettre, de réussir rapidement à accéder au canal. Les résultats indiquent que le mécanisme intelligent proposé apprend divers environnements MANETS et optimise les performances par rapport au protocole MAC standard. Les performances de MISQ sont évaluées dans divers réseaux avec le débit, le délai d'accès au canal et le taux de livraison des paquets comme mesures de performance.

**Mots clés :** Amélioration des performances, BEB, CSMA/CA, DCF, DRL, MANETs, Temps réel, Q-Learning

## Abstract

An Ad Hoc MANET (Mobile Adhoc NETwork) mobile network is a collection of wireless stations forming a dynamic network without any pre-existing infrastructure or centralized architecture. MANETs use the IEEE 802.11 standard. To deal with collisions in wireless networks, the MAC layer of the IEEE 802.11 DCF uses a CSMA / CA procedure to access the channel resource which is based on a Binary Exponential Backoff (BEB) mechanism, the latter reducing the likelihood of collision but at the cost of many network performance measures such as delay, throughput, and fairness, which is not adequate for real-time applications. To meet the various requirements of MANETs, efficient channel access is provided through sophisticated inference to control collisions. Such intelligence is possible thanks to the introduction of deep learning (DL) techniques in MANETs. In this thesis, an intelligent DRL-based paradigm is developed to optimize MAC layer channel access in MANETs. Q-Learning (QL), one of the DRL techniques, is used to provide an intelligent channel access mechanism in MANETs. The intelligent MISQ proposal takes into account in addition to the information on the state of the transmission pass or fail, the number of packets to be transmitted and the collisions made by each station to select the appropriate contention window, in order to prioritize to stations which have more packets to transmit, to quickly succeed in accessing the channel. The results indicate that the proposed intelligent mechanism learns various MANETS environments and optimizes performance over standard MAC protocol. MISQ performance is evaluated in various networks with throughput, channel access delay, and packet delivery rate as performance measures.

**Keywords:** BEB, CSMA/CA, DCF, DRL, MANETs, Performance improvement, Real Time, Q-learning

# Table des matières

<b>Résumé</b>	
<b>Glossaire Des Acronymes</b> .....	<b>1</b>
<b>Liste Des Symboles</b> .....	<b>3</b>
<b>Liste Des Figures</b> .....	<b>4</b>
<b>Liste Des Tableaux</b> .....	<b>6</b>
<b>Liste Des Algorithmes</b> .....	<b>6</b>
<b>Chapitre 1 : INTRODUCTION</b> .....	<b>7</b>
1.1. Introduction .....	7
1.2. Objectifs et résultats de la thèse .....	9
1.3. Organisation de la thèse .....	9
1.4. Publications Académiques .....	10
<b>Chapitre 2 : RESEAUX SANS FIL AD HOC ET IEEE 802.11</b> .....	<b>12</b>
2.1. Introduction .....	12
2.2. Les réseaux mobiles ad hoc.....	13
2.2.1. Définitions .....	13
2.2.2. Caractéristiques .....	14
2.2.3. Les Applications des MANETs.....	15
2.3. Problèmes de Conception et Contraintes .....	18
2.4. Le standard 802.11 .....	20
2.4.1. Le mode avec infrastructure .....	21
2.4.2. Le mode sans infrastructure (ad hoc) .....	21
2.4.3. Présentation des protocoles MAC IEEE 802.11.....	22
2.5. L'architecture de 802.11 .....	23
2.5.1. La couche physique .....	24
2.5.2. Contrôle d'Accès au Médium.....	25
2.6. Fonctionnement de la couche MAC dans les réseaux ad hoc .....	25

2.6.1. Fonction de Coordination de point PCF .....	25
2.6.2. Fonction de Coordination Distribuée DCF .....	26
2.6.3. L'algorithme du backoff .....	27
2.6.4. Trames MAC .....	33
2.7. Problèmes dans les transmissions sans fil .....	33
2.7.1. Le problème du nœud caché .....	33
2.7.2. Le problème du nœud exposé .....	34
2.7.3. Problème de trou noir .....	35
2.7.4. Problème de l'équité .....	35
2.8. Conclusion .....	36
<b>Chapitre 3 : TRAVAUX ANTERIEURS .....</b>	<b>37</b>
3.1. Introduction .....	37
3.2. Les solutions proposées .....	38
3.2.1. Solutions classiques .....	38
A.Binary Exponential Backoff (BEB) .....	38
B.Linear/Multiplicative Increase and Linear Decrease (LMILD) .....	39
C.Hybrid Backoff Algorithm (HBA) .....	40
D.New Backoff Method (NBM) .....	40
E.Double Increment Double Decrement (DIDD) .....	39
F.Enhanced Fibonacci Backoff (EFB) .....	42
G.Enhanced BEB (E-BEB) .....	42
H.Adaptively Scaled Backoff (ASB) .....	43
I.Improved Backoff Algorithm (I-BEB) .....	41
J.New Binary Exponential Backoff (N-BEB) .....	43
K.Optimized Pessimistic Fibonacci Backoff (PFB) .....	43
3.2.2. Solutions Basées sur l'intelligence .....	44
3.2.2.1. Apprentissage par renforcement profond DRL .....	46
3.2.2.2. Processus de décision de Markov .....	48



3.2.2.3. Algorithmes basés sur la fonction de valeur .....	49
A) Recherche de la valeur optimale .....	49
B) Algorithme RL .....	50
C) Q-Learning .....	51
D) Exploration vs exploitation (Algorithme $\epsilon$ -greedy) .....	52
3.2.2.4. Algorithmes basés sur des politiques .....	53
3.2.2.5. Quelques solutions intelligentes.....	53
A. Cognitive Backoff Mechanism (CB).....	53
B. Centralized Contention Window Optimization with DRL (CCOD) .....	54
C. intelligent QI-based Resource Allocation (iQRA) .....	54
3.3. Conclusion.....	54
<b>Chapitre 4 : O-BEB Un Nouvel Algorithme Adaptatif Et Equitable.....</b>	<b>55</b>
4.1. Introduction .....	55
4.2. Backoff Exponentiel Binaire Opposé (O-BEB) .....	56
4.3. L'implémentation de l'I-BEB et de N-BEB.....	60
4.4. Simulation et analyse des résultats.....	60
a) Débit moyen.....	61
b) Délai moyen de bout en bout .....	61
c) Energie résiduelle moyenne .....	61
d) Taux de paquets délivrés.....	62
e) Equité .....	62
4.5. Conclusion .....	66
<b>Chapitre 5 : M-BEB : Réduction Exponentielle Binaire Améliorée et Equitable .....</b>	<b>68</b>
5.1. Introduction .....	68
5.2. Backoff Exponentiel Binaire modifié (M-BEB) .....	69
5.3. Simulation et analyse des résultats.....	72
a).Débit moyen.....	73
b).Délai moyen de bout en bout .....	73

c).Energie résiduelle moyenne .....	74
d).Taux de paquets délivrés.....	75
e).Indice d'équité.....	75
5.4. Conclusion .....	76
<b>Chapitre 6 :MISQ :Un Mécanisme Intelligent de Sélection de CW basé sur Q-learning</b>	<b>77</b>
6.1. Introduction .....	77
6.2. Processus de décision de Markov et apprentissage par renforcement .....	78
6.3. Estimation de la fenêtre de contention CW.....	79
A. La fonction récompense .....	80
B. L'algorithme MISQ .....	81
B.1. Les Etats .....	81
B.2. Les Actions .....	81
B.3. Algorithme proposé.....	83
6.4. Evaluation des performances .....	85
6.5. Paramètres de simulation .....	89
6.5.1. Débit moyen .....	89
6.5.2. Délai d'accès MAC moyen .....	91
6.5.3. Taux de paquets délivrés .....	92
6.5.4. Indice d'équité .....	94
6.6. Conclusion.....	95
<b>Chapitre 7 : CONCLUSION .....</b>	<b>96</b>
<b>REFERENCES .....</b>	<b>98</b>

# Glossaire Des Acronymes

<b>Acronyme</b>	<b>Signification</b>
<b>ACK</b>	Acknowledgements
<b>AI</b>	Artificial Intelligence
<b>AP</b>	Access Point
<b>ASB</b>	Adaptively Scaled <b>B</b> ackoff
<b>BEB</b>	<b>B</b> inary <b>E</b> xponential <b>B</b> ackoff
<b>BSS</b>	<b>B</b> asic <b>S</b> ervice <b>S</b> et
<b>CB</b>	<b>C</b> ognitive <b>B</b> ackoff <b>M</b> echanism
<b>CCA</b>	<b>C</b> lear <b>C</b> hannel <b>A</b> ssessment
<b>CCOD</b>	<b>C</b> entralized <b>C</b> ontention <b>W</b> indow <b>O</b> ptimization with <b>D</b> RL
<b>CRN</b>	<b>C</b> ognitive <b>R</b> adio <b>N</b> etwork
<b>CRC</b>	<b>C</b> yclic <b>R</b> edundancy <b>C</b> heck
<b>CS</b>	<b>C</b> arrier <b>S</b> ensing
<b>CSMA/CA</b>	<b>C</b> arrier <b>S</b> ense <b>M</b> ultiple <b>A</b> ccess with <b>C</b> ollision <b>A</b> voidance
<b>CW</b>	<b>C</b> ontention <b>W</b> indow
<b>DCF</b>	<b>D</b> istributed <b>C</b> oordination <b>F</b> unction
<b>DIDD</b>	<b>D</b> ouble <b>I</b> ncrement <b>D</b> ouble <b>D</b> ecrement
<b>DL</b>	<b>D</b> eep <b>L</b> earning
<b>DoS</b>	<b>D</b> eny <b>o</b> f <b>S</b> ervice
<b>DSSS</b>	<b>D</b> iscrete <b>S</b> equence <b>S</b> pread <b>S</b> pectrum
<b>DRL</b>	<b>D</b> eep <b>R</b> einforcement <b>L</b> earning
<b>DIFS</b>	<b>D</b> istributed <b>I</b> nter-frame <b>S</b> pace
<b>E-BEB</b>	<b>E</b> nhanced <b>B</b> EB
<b>EFB</b>	<b>E</b> nhanced <b>F</b> ibonacci <b>B</b> ackoff
<b>FDMA</b>	<b>F</b> requency <b>D</b> ivision <b>M</b> ultiple <b>A</b> ccesses
<b>FHSS</b>	<b>F</b> requency <b>H</b> opping <b>S</b> pread <b>S</b> pectrum
<b>HBA</b>	<b>H</b> ybrid <b>B</b> ackoff <b>A</b> lgorithm

<b>HiperLAN 1</b>	<b>H</b> igh <b>P</b> erformance <b>L</b> ocal <b>A</b> rea <b>N</b> etwork type 1
<b>IEEE</b>	<b>I</b> nstitute of <b>E</b> lectrical and <b>E</b> lectronics <b>E</b> ngineers
<b>I-BEB</b>	<b>I</b> mproved <b>B</b> ackoff <b>A</b> lgorithm
<b>IoT</b>	<b>I</b> nternet- <b>o</b> f- <b>T</b> hings
<b>iQRA</b>	<b>i</b> ntelligent <b>Q</b> L–based <b>R</b> esource <b>A</b> llocation
<b>LMILD</b>	<b>L</b> inear/ <b>M</b> ultiplicative <b>I</b> ncrease and <b>L</b> inear <b>D</b> ecrease
<b>LLC</b>	<b>L</b> ogical <b>L</b> ink <b>C</b> ontrol
<b>MAC</b>	<b>M</b> edium <b>A</b> ccess <b>C</b> ontrol
<b>MANET</b>	<b>M</b> obile <b>A</b> d hoc <b>N</b> ETwork
<b>M-BEB</b>	<b>M</b> odified <b>B</b> inary <b>E</b> xponential <b>B</b> ackoff
<b>MDP</b>	<b>M</b> arkov <b>D</b> ecision <b>P</b> rocess
<b>MI</b>	<b>M</b> achine <b>I</b> ntelligence
<b>MIMO</b>	<b>M</b> ultiple- <b>I</b> nput and <b>M</b> ultiple- <b>O</b> utput
<b>MISQ</b>	<b>M</b> écanisme <b>I</b> ntelligent de <b>S</b> élection de <b>C</b> W basé sur <b>Q</b> - <b>L</b> earning
<b>NAV</b>	<b>N</b> etwork <b>A</b> llocation <b>V</b> ector
<b>N-BEB</b>	<b>N</b> ew <b>B</b> inary <b>E</b> xponential <b>B</b> ackoff
<b>NBM</b>	<b>N</b> ew <b>B</b> ackoff <b>M</b> ethod
<b>NS2</b>	<b>N</b> etwork <b>S</b> imulator <b>2</b>
<b>OFDM</b>	<b>O</b> rthogonal <b>F</b> requency <b>D</b> ivision <b>M</b> ultiplexing
<b>O-BEB</b>	<b>O</b> pposite <b>B</b> inary <b>E</b> xponential <b>B</b> ackoff
<b>OSI</b>	<b>O</b> pen <b>S</b> ystems <b>I</b> nterconnection
<b>PCF</b>	<b>P</b> oint <b>C</b> oordination <b>F</b> unction
<b>PFB</b>	<b>O</b> ptimized <b>P</b> essimistic <b>F</b> ibonacci <b>B</b> ackoff
<b>PHY</b>	<b>P</b> hysical <b>L</b> ayer
<b>POMDP</b>	<b>P</b> artially <b>O</b> bserved <b>M</b> DP
<b>SARSA</b>	<b>S</b> tate– <b>A</b> ction– <b>R</b> eward– <b>S</b> tate– <b>A</b> ction
<b>STA</b>	<b>S</b> tation
<b>QL</b>	<b>Q</b> - <b>L</b> earning
<b>QoS</b>	<b>Q</b> uality <b>o</b> f <b>S</b> ervice
<b>RTS/CTS</b>	<b>R</b> equest- <b>T</b> o- <b>S</b> end/ <b>C</b> lear- <b>T</b> o- <b>S</b> end
<b>TCP</b>	<b>T</b> ransmission <b>C</b> ontrol <b>P</b> rotocol
<b>UDP</b>	<b>U</b> ser <b>D</b> atagram <b>P</b> rotocol
<b>Wi-Fi</b>	<b>W</b> ireless <b>F</b> idelity
<b>WLAN</b>	<b>W</b> ireless <b>L</b> ocal <b>A</b> rea <b>N</b> etwork

# Liste Des Symboles

<b>Symbole</b>	<b>Signification</b>
<b>S</b>	Espace d'états
<b>A</b>	Espace d'actions
<b>T</b>	Probabilité de transition
<b>V(s)</b>	Valeur estimée de l'état $s$
<b>V*</b>	Politique optimale
<b>Q(s, a)</b>	Valeur estimée de l'action $a$ à l'état $s$
<b>Q*(s, a)</b>	Fonction de valeur optimale
<b>P(a ss')</b>	Probabilité de passer de l'état $s$ à $s'$ sachant $a$
<b>r(s)</b>	Récompense pour l'état $s$
<b>CW<sub>min</sub></b>	Valeur minimale de CW
<b>CW<sub>max</sub></b>	Valeur maximale de CW
<b>a<sub>i</sub></b>	Une action
<b>s<sub>i</sub></b>	Un état
<b><math>\pi</math></b>	Politique
<b><math>\alpha</math></b>	Taux d'apprentissage entre [0, 1]
<b><math>\gamma</math></b>	Facteur d'actualisation entre [0, 1]
<b><math>\epsilon</math></b>	Probabilité d'exploration/exploitation $\in [0, 1]$
<b><math>\Delta</math></b>	Différence temporelle

# Liste Des Figures

<b>FIGURE 2.1</b> : UN RESEAU AD HOC .....	14
<b>FIGURE 2.2</b> : CHANGEMENT DE TOPOLOGIE DANS UN RESEAU AD HOC.....	15
<b>FIGURE 2.3</b> : RESEAU 802.11 EN MODE INFRASTRUCTURE.....	21
<b>FIGURE 2.4</b> : RÉSEAU 802.11 AD-HOC 1 .....	21
<b>FIGURE 2.5</b> : LA COUCHE PHY ET LA COUCHE LIAISON DES DONNEES DANS IEEE 802.11.....	24
<b>FIGURE 2.6</b> : EXEMPLE D'AUGMENTATION EXPONENTIELLE DE CW .....	28
<b>FIGURE 2.7</b> : METHODE D'ACCES AU MEDIUM DCF .....	29
<b>FIGURE 2.8</b> : ORGANIGRAMME DE BEB .....	30
<b>FIGURE 2.9</b> : METHODE D'ACCES AU MEDIUM AVEC RTS/CTS.....	32
<b>FIGURE 2.10</b> : DIFFERENCE ENTRE DUREES IFS.....	33
<b>FIGURE 2.11</b> : LE PROBLEME DE LA STATION CACHEE.....	34
<b>FIGURE 2.12</b> : LE PROBLEME DE LA STATION EXPOSE .....	35
<b>FIGURE 3.1</b> : QUELQUES SOLUTIONS BEB DANS LA LITTERATURE.....	38
<b>FIGURE 3.2</b> : ORGANIGRAMME DE L'I-BEB « CAS DE SUCCES DE TRANSMISSION ».....	41
<b>FIGURE 3.3</b> : ORGANIGRAMME DE N-BEB.....	43
<b>FIGURE 3.4</b> : TYPES D'APPRENTISSAGE AUTOMATIQUE .....	45
<b>FIGURE 3.6</b> : DIAGRAMME DE L'APPRENTISSAGE PAR RENFORCEMENT.....	47
<b>FIGURE 4.1</b> : ORGANIGRAMME DE L'O-BEB .....	59
<b>FIGURE 4.2</b> : DEBIT MOYEN .....	63
<b>FIGURE 4.3</b> : DELAI MOYEN.....	63
<b>FIGURE 4.4</b> : ENERGIE RESIDUELLE MOYENNE.....	64
<b>FIGURE 4.5</b> : INDICE D'EQUITE.....	65
<b>FIGURE 4.6</b> : TAUX DE PAQUETS DELIVRES.....	65
<b>FIGURE 4.7</b> : TAUX DE PERTE.....	66

<b>FIGURE 5.1 : ORGANIGRAMME DE M-BEB</b> .....	71
<b>FIGURE 5.2 : DEBIT MOYEN</b> .....	73
<b>FIGURE 5.3 : DELAI MOYEN</b> .....	73
<b>FIGURE 5.4 : ENERGIE RESIDUELLE MOYENNE</b> .....	74
<b>FIGURE 5.5 : TAUX DE PAQUETS DELIVRES</b> .....	75
<b>FIGURE 5.6 : INDICE D'EQUITE</b> .....	76
<b>FIGURE 6.1 : DIAGRAMME ETATS TRANSITIONS DE LA PROPOSITION</b> .....	82
<b>FIGURE 6.2: PROTOCOLE MAC BASE SUR Q-LEARNING</b> .....	83
<b>FIGURE 6.3 : COMPARAISON DE DEBIT DANS UN RESEAU DE 25 NŒUDS DE TAILLE FILE</b> D'ATTENTE =10 AVEC (A) E =0.3 (B) E = 0.6 ET (C) E =0.8.....	87
<b>FIGURE 6.4 : COMPARAISON DE DEBIT DANS UN RESEAU DE 25 NŒUDS DE TAILLE FILE</b> D'ATTENTE =40 AVEC (A) E =0.3 (B) E = 0.6 ET (C) E =0.8.....	89
<b>FIGURE 6.5: DEBIT MOYEN DANS (A) TAILLE FILE D'ATTENTE =10, (B) TAILLE FILE D'ATTENTE =20, (C) TAILLE FILE D'ATTENTE =40</b> .....	90
<b>FIGURE 6.6: DELAI D'ACCES MAC POUR (A): TAILLE FILE D'ATTENTE =10, (B) TAILLE FILE D'ATTENTE =20, (C) TAILLE FILE D'ATTENTE =40</b> .....	92
<b>FIGURE 6.7: PACKET DELIVERY RATIO POUR (A): TAILLE FILE D'ATTENTE =10, (B) TAILLE FILE D'ATTENTE =20, (C) TAILLE FILE D'ATTENTE =50, (D): MOYENNE DE TRANSMISSION REUSSIE</b> .....	93
<b>FIGURE 6.8: INDICE D'EQUITE DE JAIN, (A) TAILLE FILE D'ATTENTE =20, (B) TAILLE FILE D'ATTENTE = 40</b> .....	94

# Liste Des Tableaux

<b>TABLEAU 2.1 : APPLICATIONS DES RESEAUX AD HOC MOBILES.</b> .....	16
<b>TABLEAU 2.2 : LES DIFFERENTES EXTENSIONS DU 802.11</b> .....	22
<b>TABLEAU 4.1 : PARAMETRES DE SIMULATION « O-BEB »</b> .....	60
<b>TABLEAU 4.2 : INDEX D'EQUITE</b> .....	62
<b>TABLEAU 5.1 : PARAMETRES DE SIMULATION » « M-BEB »</b> .....	72
<b>TABLEAU 6.1 : PARAMETRES DE SIMULATION « MISQ »</b> .....	86

# Liste Des Algorithmes

<b>ALGORITHME 3.1: Q-LEARNING</b> .....	53
<b>FONCTION 4.1 : RESET_LOG_CW()</b> .....	57
<b>FONCTION 4.2: INCR_FIBO_CW()</b> .....	58
<b>FONCTION 5.1 : RESET_CW()</b> .....	70
<b>FONCTION 5.2 : INCREM_CW()</b> .....	70
<b>ALGORITHME 6.1 : CW OPTIMISATION UTILISANT MISQ</b> .....	85



# INTRODUCTION

## 1.1. Introduction

Les réseaux mobiles Ad Hoc MANET (**M**obile **A**d hoc **N**ETworks) sont devenus une option très courante dans les techniques de déploiement des réseaux de l'industrie et des communautés au cours de la dernière décennie. Certaines des raisons qui peuvent être citées pour une utilisation aussi répandue des réseaux locaux sans fil WLAN (**W**ireless **L**ocal **A**rea **N**etworks) sont le faible coût d'infrastructure, la facilité de développement, le support pour la communication mobile, le déploiement sans câblage et la facilité d'ajouter de nouveaux utilisateurs au réseau [8]. Cela dit, l'utilisation de réseaux Ad Hoc à une telle échelle ne prouve pas l'absence des failles dans ce type de réseaux, en fait, elle éclaire de nombreux domaines où les réseaux Ad Hoc ne se comportent pas comme attendu, et par conséquent, il souffre de graves problèmes de performance. Les principales raisons de ces problèmes de performance résident dans le mécanisme d'accès aux médias partagés et dans les protocoles de routage. Le but de ce dernier est d'avoir un établissement de routes efficace entre une paire de nœuds, afin que les messages puissent être livrés en temps opportun. Les contraintes de bande passante et de puissance sont les facteurs importants à prendre en compte dans les réseaux sans fil en général et dans les réseaux Ad Hoc en spécifique car le réseau sans fil Ad Hoc multi-sauts repose sur chaque nœud du réseau pour agir en tant que routeur et en tant que transmetteur de paquets. Cette dépendance place la demande de la bande passante et du calcul dans l'hôte mobile à être prise en compte lors du choix ou de la conception du protocole. Le protocole de contrôle d'accès au médium MAC (**M**edium **A**ccess **C**ontrol) est un autre défi dans un réseau Ad Hoc mobile. Puisque le Médium est une ressource partagée et rare dans un tel réseau, contrôler efficacement l'accès à ce médium partagé est-elle même une tâche compliquée.

Il existe de nombreuses normes pour les réseaux locaux sans fil parmi lesquelles : HiperLAN<sup>1</sup> 1 (**H**igh **P**erformance **L**ocal **A**rea **N**etwork type 1), HiperLAN 2<sup>2</sup>, Bluetooth, et IEEE 802.11. La norme IEEE 802.11, qui est maintenant utilisée dans la plupart des travaux de la communauté ad hoc. Ce standard a été conçu pour permettre un accès équitable à toutes les stations concurrentes. Il couvre les deux couches du modèle OSI (**O**pen **S**ystems **I**nterconnection), la couche de MAC et la couche physique (Phy). Le mécanisme d'accès au canal distribué DCF (**D**istributed **C**oordination **F**unction) est le mécanisme d'accès de base défini par la couche MAC, et est exécuté au niveau de chaque station.

La principale motivation du travail de cette thèse est de mettre en évidence les problèmes et les défis du protocole MAC IEEE802.11 pour les réseaux mobiles ad hoc d'une part. Et d'une autre part, comment augmenter l'efficacité et la robustesse de la couche MAC pour fournir des garanties sur le délai, sur les taux de pertes ou encore sur la bande passante pour les applications temps réel, dans les réseaux sans fil sans infrastructure (sans point d'accès). Ainsi de mettre l'accent sur les problèmes d'équité, donc, de garantir une certaine qualité de service QoS (Quality of Service).

Dans les dernières années, le domaine de l'apprentissage en profondeur DL (**D**eep **L**earning) s'est développé afin de permettre des capacités d'intelligence artificielle dans les technologies de communication sans fil. Les chercheurs pensent que les WLAN peuvent optimiser les performances en introduisant DL dans l'allocation des ressources de la couche MAC. L'apprentissage par renforcement profond DRL (**D**eep **R**einforcement **L**earning) est une technique de DL qui est motivée par la sensibilité comportementale et l'idiologie de contrôle, où un agent peut atteindre un objectif en interagissant avec son environnement [10, 15, 74]. DRL utilise des modèles d'apprentissage spécifiques, tels que le processus de décision de Markov (MDP), le MDP partiellement observé (POMDP) et le Q-Learning (QL) [53].

Notre travail de thèse consiste à s'intéresser aux moyens d'améliorer la couche MAC 802.11 en particulier la fonction DCF, afin d'augmenter les paramètres de qualité de service (QoS), ainsi de supporter des applications temps réel. Dans cette thèse, de nouvelles approches sont proposées pour améliorer et optimiser les performances de l'algorithme Binary Exponential Backoff utilisé par la fonction DCF. Ainsi, des nouvelles approches sont proposées, les deux premières solutions apportent des modifications de calcul de la CW dans les deux cas transmission réussie et collision, alors que la troisième approche proposée, basée sur l'un des modèles de DRL, le Q-Learning est utilisé pour calculer la valeur de la CW adéquate pour chaque station sans fil dans le réseau. Le protocole MAC intelligent proposé, optimise les performances des réseaux mobiles sans fil, en particulier dans les environnements de réseaux

---

<sup>1</sup> Homologue européen de l'IEEE 802.11; défini par ETSI (l'Institut européen des normes de télécommunications).  
Type 1 : Radio LAN, 5,15 GHz et 17,1 GHz, débit max : 23,5 Mbps.

<sup>2</sup> Type 2 : Accès sans fil à courte portée à IP, 5 GHz, débits : 6 Mbps à 54 Mbps.

ayant un nombre important de stations.

## 1.2. Objectifs et résultats de la thèse

Cette thèse apporte plusieurs contributions pour résoudre le problème de sélection des fenêtres de contentions et développer une stratégie et une méthode pour apprendre comment évolue l'accès au canal sans fil dans les MANETs. A ces fins, trois objectifs spécifiques sont identifiés, qui sont décrits ci-après :

1. Le premier objectif de cette thèse est de surligner un état de l'art des recherches sur la qualité de service dans les réseaux ad hoc. Ces réseaux qui ne disposent pas concrètement de protocoles de qualité de service standards adaptés à leurs spécificités, car la topologie variable de ce type de réseaux et le manque d'une infrastructure fixe posent problème pour mettre en place de politiques de gestion de la qualité de service (publié dans la Publication IV).
2. Le second objectif de cette thèse est d'étudier et de quantifier les enjeux et les défis du standard IEEE802.11 dans les MANETs, plus spécifiquement dans différents réseaux sans fil. On constate que l'augmentation de la densité des périphériques réseau est directement proportionnelle à la collision dans le réseau et entraîne une dégradation des performances du réseau. Le mécanisme d'accès au canal CSMA / CA traditionnel est évalué à cette fin. L'algorithme de réduction exponentielle binaire (BEB) de CSMA / CA gère aveuglément l'évitement de collision dans les réseaux sans fil (pour une discussion détaillée, veuillez-vous reporter à la section 2.7). Deux algorithmes O-BEB et M-BEB basés sur l'incrémentaire linéaire de la fenêtre de contention CW (Contention Window), et qui améliorent le problème de cécité du BEB, sont proposés dans la publication II et la publication III.
3. Le troisième objectif de la thèse est de développer une méthode intelligente pour accéder au canal sans fil. Dans cette thèse, les potentiels du paradigme d'apprentissage par renforcement profond sont étudiés pour l'optimisation des performances du protocole MAC dans un réseau mobile sans fil. Un mécanisme intelligent de sélection de CW basé sur Q-Learning (MISQ) est proposé à cet effet, où Q-Learning est l'un des techniques dominantes de DRL. Le mécanisme MISQ est publié dans la Publication I.

## 1.3. Organisation de la thèse

La présente thèse est organisée comme suit :

Le chapitre II introduit les réseaux ad hoc où nous donnons des notions de base, des généralités, et les domaines d'application d'une part. Et d'une autre part, étudier la norme 802.11 utilisée pour la communication sans fil dans les MANETs, où nous citons les différents

amendements, l'architecture du IEEE802.11, les différents modes de fonctionnement, le mécanisme CSMA/CA...etc.

Dans le chapitre III, nous donnons un panorama synthétique des travaux antérieurs consacrés à l'optimisation de protocole MAC 802.11 dans les réseaux sans fil où nous classifions et étudions les approches proposées dans la littérature. Parmi ces travaux, ceux qui utilisent des techniques de l'apprentissage par renforcement, nous introduisons ainsi, les DL, le DRL, les processus décisionnels de Markov (MDP), ainsi que la technique DRL, utilisée dans notre travail, le Q-Learning (QL).

Les trois chapitres suivants exposent nos différentes contributions, et sont organisés comme suit :

Le chapitre IV décrit notre première contribution O-BEB au problème d'optimisation du standard 802.11 dans les réseaux mobiles Ad hoc, qui se porte sur une méthode inversée de calcul de la fenêtre de contention, contrairement au BEB classique, les performances d'O-BEB sont évaluées en utilisant plusieurs métriques de performances.

Le chapitre V est consacré à la deuxième contribution M-BEB au problème posé, où nous avons proposé une approche linéaire de calcul de CW en cas de succès et une diminution de la taille de CW en cas de collision. Les mêmes métriques de performances utilisées pour l'O-BEB sont utilisées pour l'évaluation de M-BEB.

Dans le chapitre VI, nous proposons une approche de sélection de CW intelligente basé sur une technique de DRL qui est Q-Learning, l'évaluation de cette approche est réalisée sous différentes simulations avec différentes métriques.

Enfin, nous terminerons notre thèse par une conclusion générale, et nous énoncerons des perspectives de recherche.

#### 1.4. Publications Académiques

- I. **Journal:** Zerguine, N., Mostefai, M., Aliouat, Z., Slimani, Y. (2020). Intelligent CW selection mechanism based on Q-learning (MISQ). *Ingénierie des Systèmes d'Information*, Vol. 25, No. 6, pp. 803-811. <https://doi.org/10.18280/isi.250610>
- II. **Conference:** N. Zerguine, Z. Aliouat, M. Mostefai and S. Harous, (2020). M-BEB: Enhanced and Fair Binary Exponential Backoff, *14th International Conference on Innovations in Information Technology (IIT)*, Al Ain, 2020, pp. 142-147, doi: 10.1109/IIT50501.2020.9299014.

- III. **Conference:** N. Zerguine, Z. Aliouat, M. Mostefai, (2019). Enhanced Collisions rate by a new Adaptive and Fair Binary Exponential Backoff, The 1st International Conference on Innovative Trends in Computer Science (CITCS'2019), Guelma, Algérie.
- IV. **Conference:** N. Zerguine, M. Mostefai, (2006). Un état de l'art sur : Les réseaux ad hoc et la qualité de service, Conférence Internationale sur l'Informatique et ses Applications (CIIA06), Saida, Algérie.



# RESEAUX SANS FIL AD HOC ET IEEE 802.11

## 2.1. Introduction

Les réseaux sans fil ont connu un progrès remarquable et une large utilisation dans notre vie quotidienne ces dernières années, grâce à leur simplicité, flexibilité, ainsi qu'à leur coût compétitif et à leur haut débit d'une part. D'une autre part, la technologie des communications sans fil qui a joué un rôle important au sein des réseaux informatiques. Elle a offert des solutions pour permettre la mobilité ainsi que des services cruciaux, essentiellement lorsque l'installation d'infrastructures est impossible. Les réseaux sans fil sont une technologie émergente qui permet aux utilisateurs d'accéder aux services et à l'information, indépendamment de leur position géographique.

Les réseaux sans fil peuvent être classés en deux types : Réseaux sans fil avec infrastructure, qui utilisent une station centrale et réseaux sans fil sans infrastructure ou ad hoc. Un réseau ad hoc mobile est constitué d'un grand nombre d'entités qui se déplacent librement dans un environnement, utilisant des interfaces radio sans fil pour communiquer entre elles.

Le présent travail est dans le domaine des réseaux mobiles ad hoc, dans ce qui suit nous présentons les MANETs, leurs caractéristiques, leurs problèmes. La technologie permettant l'établissement des WLAN sera l'objectif de ce deuxième chapitre, où, nous détaillons la norme 802.11, ses amendements, son architecture, ses modes de fonctionnement ainsi que les problèmes rencontrés.

## 2.2. Les réseaux mobiles ad hoc

### 2.2.1. Définitions

Le terme « Ad Hoc » est une expression latine qui signifie pour ce ou plus spécifiquement pour cet objet seulement. Ce terme a été appliqué en particulier aux réseaux sans fil constitués d'entités communicantes qui appartiennent à un réseau uniquement pendant une session de communication et sont à portée radio avec d'autres entités. Un réseau ad hoc est une collection de stations (STA)<sup>3</sup> mobiles, interconnectées par une technologie sans fil formant un réseau dynamique et temporaire sans l'aide de toute infrastructure préexistante ou administration centralisée [17] (Figure 2.1). Il n'y a pas de préarrangement concernant le rôle spécifique que chaque station devrait assumer. Au lieu de cela, chaque STA prend sa décision indépendamment, en fonction de la situation du réseau, sans utiliser une infrastructure réseau préexistante [1]. Par exemple, deux ordinateurs équipés de cartes adaptatrices sans fil peuvent établir un réseau indépendant lorsqu'ils se trouvent à portée l'un de l'autre. Dans un réseau Ad Hoc, les connexions entre entités sont généralement temporaires, car les STAs peuvent être ajoutées ou supprimées, par l'entrée ou la sortie du réseau respectivement. En général, les réseaux Ad Hoc fonctionnent de manière autonome, mais peuvent être connectés à un autre réseau tel que l'Internet [2].

La communication entre les différentes STAs formant un réseau Mobile Ad hoc NETWORK (MANET) est assurée à l'aide de liaisons sans fil et sur la base du paradigme peer-to-peer<sup>4</sup> [2, 17]. Chaque STA mobile fonctionne comme un routeur et elle est libre de se déplacer de manière aléatoire et de se connecter arbitrairement à d'autres STAs. Ainsi, la topologie du réseau peut changer rapidement et de façon imprévisible, car il peut exister un grand nombre de connexions ad Hoc indépendantes. En fait, il est possible d'avoir différentes applications fonctionnant sur le même MANET. Les MANET ont tendance à jouer un rôle important dans les nouvelles applications distribuées telles que l'informatique collaborative distribuée, les applications de détection distribuées et la réponse aux incidents sans infrastructure de communication.

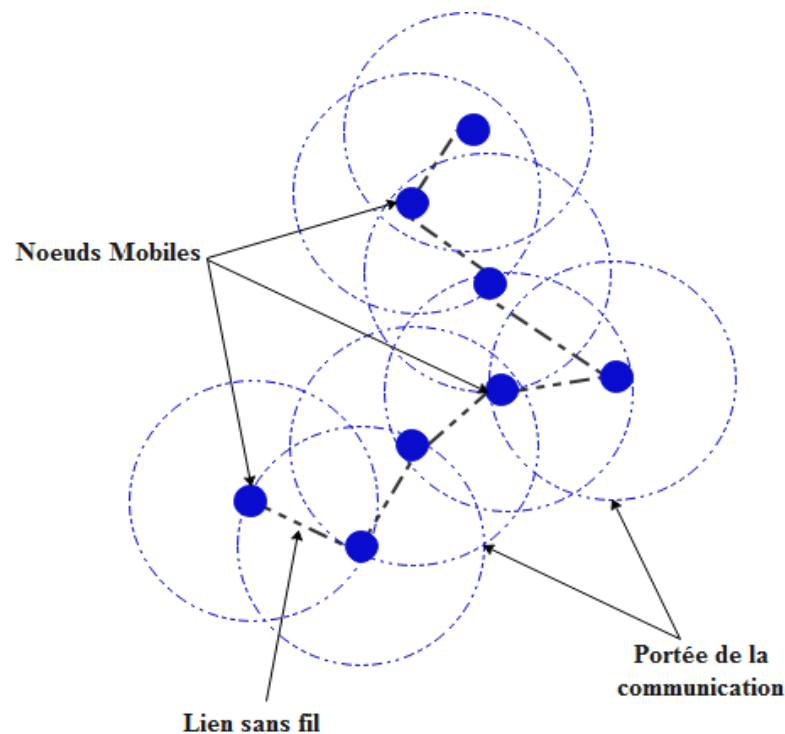
Dans un MANET, une route entre deux STAs peut consister en un saut à travers un ou plusieurs nœuds. Étant donné que la portée de transmission radio de chaque STA est limitée, il peut être nécessaire de transférer un paquet sur plusieurs sauts pour qu'il atteigne sa destination et, par conséquent, il convient d'appeler de tels réseaux « réseaux Ad Hoc sans fil multi-sauts » (Multi hop Wireless Ad Hoc networks) [3].

---

<sup>3</sup> Tout appareil contenant une interface de contrôle d'accès au support (MAC) et de couche physique (PHY) conforme IEEE 802.11 au support sans fil.

<sup>4</sup> Nommé également pair-à-pair ou P2P définit un modèle de réseau informatique utilisant une architecture distribuée, où les pairs partagent et échangent des charges de travail, sans pairs privilégiés ou périphérique administrateur central.





**Figure 2.1** : Un réseau Ad hoc

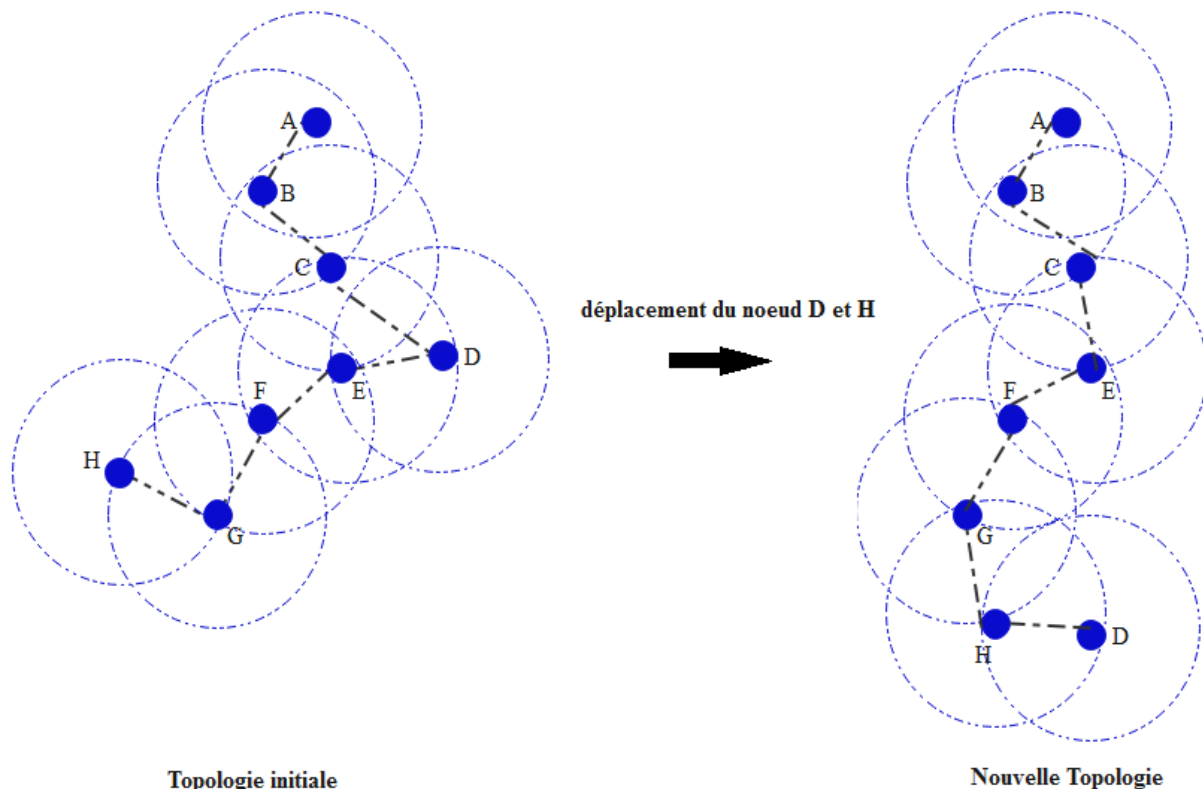
### 2.2.2. Caractéristiques

Les MANET héritent des caractéristiques communes aux réseaux sans fil en général et ajoutent des caractéristiques spécifiques aux réseaux Ad Hoc :

- *sans fil* : Les nœuds communiquent avec des interfaces sans fil et partagent le même média (radio, infrarouge<sup>5</sup>, ...etc.).
- *basé Ad Hoc* : Un réseau Ad Hoc mobile est un réseau temporaire formé dynamiquement de manière arbitraire par une collection de nœuds selon les besoins.
- *Autonome et sans infrastructure* : MANET ne dépend d'aucune infrastructure établie ou d'une administration centralisée. Chaque nœud fonctionne en mode peer-to-peer distribué, il agit comme un routeur indépendant et génère des données indépendantes. Les nœuds sont responsables d'établir et de maintenir la connectivité du réseau d'une manière continue.
- *Routing Multi saut* : Aucun routeur dédié n'est nécessaire, chaque nœud agit comme un routeur et transmet les paquets de chacun pour permettre le partage d'informations entre les nœud mobiles.

<sup>5</sup> La lumière infrarouge a une large bande passante, les débits sont relativement élevés, mais la portée est faible.

• *Mobilité* : Chaque nœud est libre de se déplacer lorsqu'il communique avec d'autres nœuds. La topologie d'un tel réseau Ad Hoc est d'une nature dynamique en raison du mouvement constant des nœuds participants, ce qui provoque le changement continu des modèles d'intercommunication entre les nœuds (figure 2.2). Les réseaux sans fil Ad Hoc éliminent les contraintes de l'infrastructure et permettent aux périphériques de créer et de rejoindre des réseaux à la volée « n'importe quand et n'importe où » pratiquement pour n'importe quelle application.



**Figure 2.2** : Changement de topologie dans un réseau Ad hoc

Les réseaux ad hoc peuvent être déployés dans un environnement quelconque, avec un coût faible d'exploitation, puisqu'ils sont sans infrastructure et sans entretien. Le déploiement d'un réseau ad hoc est simple et rapide, simple puisqu'il suffit de disposer d'un certain nombre de STAs dans un espace pour créer un réseau ad hoc, et rapide puisqu'il est immédiatement fonctionnel dès que les STAs sont présentes. Les MANETs ont une souplesse d'utilisation qui où une panne d'une STA est non pénalisante de manière globale (une station qui sert au routage peut être remplacée par une autre si elle tombe en panne).

### 2.2.3. Les Applications des MANETs

D'une façon générale, puisque les MANET sont des réseaux flexibles qui peuvent être installés n'importe où et à tout moment, sans infrastructure, ils sont utilisés dans toutes

applications où le déploiement d'une infrastructure réseau filaire est trop contraignant. Les applications dédiées aux MANET couvrent une très large gamme. Historiquement, les MANETs ont été utilisés principalement pour des applications liées au réseau tactique afin d'améliorer les communications et la capacité de survie sur le champ de bataille et pour les opérations de secours. La nature dynamique et critique des opérations militaires signifie qu'il n'est pas possible de compter sur l'accès à une infrastructure de communication préinstallée fixe sur le champ de bataille. La technologie ad hoc intéresse également la recherche des applications civiles. Le tableau 2.1 récapitule les principales applications des réseaux MANETs [19].

**Tableau 2.1 :** Applications des réseaux Ad Hoc mobiles.

<b>Application</b>	<b>Descriptions et Services</b>
Militaire	Première application des réseaux ad hoc, utilisés pour la communication militaire, opérations dans un milieu hostile, par exemple, surveiller le déplacement de l'ennemi dans une zone donnée, ou pour collecter des données dans un environnement instable où il serait dangereux d'envoyer des humains.
Réseaux de capteurs	Exploités généralement pour les applications environnementales. Leur mise en œuvre permet d'étudier et de gérer sur une longue période et sur une grande zone géographique des phénomènes complexes tels que : température, humidité, pression, bruit, etc. Comportant entre 1000-100000 nœuds, responsables de la récolte des données envoyées à un hôte centralisé pour le traitement en utilisant de faibles taux homogènes [6].
Services d'urgence	Les opérations de recherche et de sauvetage ainsi que la reprise après les catastrophes naturelles comme les sinistres, les tsunamis, etc. Par leur compacité et leur rapidité de déploiement, les MANETs permettent aux différentes équipes de secours d'établir rapidement des liaisons et d'échanger des informations.

Applications commerciales	<p>L'E-Commerce, par exemple, les différents paiements électroniques distant (Boutique, taxi).</p> <ul style="list-style-type: none"> <li>• Business : l'accès dynamique aux fichiers clients stockés dans un emplacement central à la volée fournissent des bases de données cohérentes pour tous les bureaux mobiles des agents.</li> <li>• Services de véhicules : transmission des nouvelles, des conditions routières, permettant à contribuer à des routes plus sûres et plus efficaces en fournissant aux conducteurs des informations sur la météo, l'état de la route (goulot d'étranglements, accident...).</li> </ul>
Réseau domestique et d'entreprise	<p>Réseau sans fil maison/bureau (WLAN), par exemple, application de tableau blanc partagé, utilisation de PDA <sup>6</sup> pour imprimer n'importe où, expositions commerciales. Réseau personnel (PAN<sup>7</sup>).</p>
Applications éducatives	<p>Mettre en place des salles de classe virtuelles ou des salles de conférences. Mettre en place une communication Ad Hoc lors de conférences, réunions ou cours.</p>
Divertissement	<p>Jeux multi-utilisateurs. Animaux robotiques. Accès Internet extérieur.</p>
Services prenant en charge la localisation	<p>Services de suivi, par exemple, renvoi automatique d'appel, transmission d'espace de travail réel à l'emplacement actuel. Services d'information :</p> <ul style="list-style-type: none"> <li>• push : par exemple, annoncer un service spécifique à un lieu, comme des stations-service.</li> <li>• pull : par exemple, un guide de voyage dépendant de l'emplacement ; services (imprimante, télécopieur, téléphone, serveur, stations-service) informations de disponibilité ; caches, résultats intermédiaires, informations d'état, etc.</li> </ul>

<sup>6</sup> Personal Digital Assistant : ordinateur de poche dont l'usage est prévu originalement dans un but d'organisation (gestion d'un agenda ou d'un carnet d'adresses).

<sup>7</sup> Personal Area Network, un type de réseau informatique restreint en matière d'équipements et d'espace

### 2.3. Problèmes de Conception et Contraintes

L'architecture Ad Hoc présente de nombreux avantages, tels que l'auto-reconfigurations, la facilité de déploiement, etc. Cependant, cette flexibilité et commodité ont un prix. Les réseaux sans fil Ad Hoc héritent des problèmes traditionnels des communications sans fil, tels que l'optimisation de la bande passante, le contrôle de puissance et l'amélioration de la qualité de transmission [7], tandis que leur mobilité, la nature multi-services et le manque d'infrastructure fixe créent des contraintes de conception qui sont nouvelles pour les réseaux Ad Hoc mobiles, comme indiqué ci-dessous :

1. *Ils sont sans infrastructure* : Les réseaux Ad Hoc mobiles sont des réseaux sans fil multi-sauts sans infrastructure. Ce manque d'infrastructure fixe en plus d'être sans fil, génère de nouveaux problèmes de conception par rapport aux réseaux fixes. De plus, l'absence d'une entité centralisée signifie que la gestion du réseau doit être distribuée entre les différentes entités, où la détection et la gestion des défaillances devient de plus en plus compliquée.

2. *Topologies dynamiques du réseau* : Dans les réseaux Ad Hoc mobiles, les nœuds peuvent se déplacer librement d'une manière arbitraire. La topologie du réseau, généralement multi saut, peut changer fréquemment et de façon imprévisible, entraînant des changements de routes, des déconnexions réseau fréquentes et, ainsi, des pertes de paquets [20,21].

3. *Limitation de la couche physique* : L'interface radio de chaque nœud utilise la diffusion pour la transmission du trafic et a généralement une portée de transmission sans fil limitée, ce qui génère des problèmes liés au réseau Ad Hoc mobiles comme les problèmes de terminaux cachés ou de terminaux exposés (détaillés dans les sections 2.8.1 et 2.8.2 respectivement), etc. Les collisions sont inévitables sur le support et les erreurs de transmission sont susceptibles d'entraîner une plus grande perte de paquets par rapport aux systèmes câblés.

4. *Bande passante et qualité limitées* : Une des caractéristiques primordiales des réseaux basés sur la communication sans fil est l'utilisation d'un médium de communication partagé (ondes radio). Ce partage fait que la bande passante réservée à un nœud soit modeste, variable, et nettement inférieure à celle des liaisons câblées.

5. *Variation des capacités des liens et des nœuds* : chaque nœud peut avoir une ou plusieurs interfaces radio avec des capacités d'émission / réception variables et fonctionnant sur des bandes de fréquences différentes. [22]. Cette hétérogénéité des capacités de la radio de nœud peut entraîner des liaisons asymétriques importantes. En outre, chaque nœud mobile peut avoir une configuration logicielle et matérielle différente, ce qui entraîne une variabilité de la puissance de traitement [23]. Cela conduit à la complexité de la conception de protocoles et d'algorithmes de réseau pour ce réseau hétérogène, qui exige une adaptation dynamique aux conditions changeantes de puissance et de canal, aux variations de charge/distribution du trafic, à l'équilibrage de charge, à et aux congestions.

6. *Opération sous contrainte énergétique* : chaque nœud mobile est équipé d'une batterie ayant une capacité limitée, la puissance de traitement est limitée, ce qui limite les services et les applications pouvant être pris en charge par chaque nœud. Dans les réseaux mobiles Ad hoc, cela devient un problème plus important, car chaque nœud qui sert à la fois de périphérique final et de routeur a besoin d'énergie supplémentaire pour transférer des paquets à partir d'autres nœuds. [19].

7. *Robustesse et fiabilité du réseau* : Dans MANET, la connectivité réseau est assurée par le routage et le transfert entre plusieurs nœuds. Bien que cela supprime les limites de la mise en réseau avec une infrastructure fixe, cela crée également des difficultés de conception. Un nœud ne peut pas transmettre le paquet en raison de différentes contraintes, telles qu'une surcharge, ou des connexions interrompues. Les nœuds défaillants et les liaisons médiocres peuvent avoir un effet important sur les performances globales du réseau. L'absence de points de contrôle et de gestion centralisés signifie qu'il n'est pas possible d'identifier et de distinguer ces formes de mauvais comportement rapidement et facilement, ce qui ajoute une complexité significative à la conception des protocoles.

8. *Sécurité réseau* : En général, les réseaux mobiles sans fil sont plus touchés et exposés par aux informations et les risques de paramètre sécurité physique plutôt que les réseaux câblés fixes. L'utilisation de canaux de diffusion sans fil ouverts et publics signifie que les risques de sécurité sont présentés aux nœuds avec une protection physique insuffisante [4]. De plus, puisqu'un réseau mobile Ad Hoc est un réseau sans infrastructure distribué, il repose principalement sur une solution de sécurité autonome de chaque nœud mobile, vue qu'il est difficile de mettre en œuvre, le contrôle de sécurité centralisé. Certaines exigences clés de sécurité dans les réseaux Ad Hoc comprennent :

- Confidentialité : prévenir l'écoute passive ;
- Contrôle d'accès : protection de l'accès à l'infrastructure de réseau sans fil ;
- Intégrité des données : Empêcher la manipulation du trafic (c'est-à-dire l'entrée, la modification ou l'injection de trafic) ;
- Attaques de déni de service DOS (Deny of Service) par des nœuds malveillants.

9. *Évolutivité du réseau* : Les algorithmes de gestion de réseau les plus populaires ont été principalement conçus pour fonctionner sur des réseaux sans fil fixes ou relativement limités. De larges réseaux avec des dizaines de milliers de nœuds, tels que des réseaux de capteurs et des réseaux tactiques [24], sont impliqués dans de nombreuses applications de réseautage ad hoc mobiles. Pour une mise en œuvre efficace de tels réseaux, l'évolutivité est importante. Dans des domaines tels que l'adressage, le routage, la gestion de l'emplacement, la gestion de la configuration, l'interopérabilité, la sécurité, les technologies sans fil à haute capacité, etc., l'évolution vers un grand réseau composé de nœuds aux ressources limitées n'est pas facile et pose de nombreux défis à résoudre.

10. *Qualité de service* : la qualité de service (QoS<sup>8</sup>, Quality of Service) est généralement définie comme un ensemble d'exigences de service devant être satisfaites par le réseau lors du transport d'un flux de paquets de la source à la destination [4]. Avec les besoins croissants du provisionnement de la qualité de service pour les applications temps réel évolutives telles que l'audio / la vidéo, il est préférable de prendre en charge ces services dans des environnements de réseau ad hoc. Le réseau est censé garantir à l'utilisateur un ensemble d'attributs de service spécifiques mesurables en termes de délai de bout en bout, de bande passante, de probabilité de perte de paquets, d'énergie et de retard (gigue) [25]. Les caractéristiques et contraintes de réseaux Ad Hoc sans fil et mobiles décrites ci-dessus, telles que les topologies de réseau changeante dynamiquement, la limite et la qualité de la bande passante, la variation des capacités de liaison et de nœud pose des difficultés supplémentaires pour obtenir la garantie de QoS requise.

Après avoir introduit les réseaux mobiles ad hoc, la section suivante sera consacrée au standard IEEE<sup>9</sup> 802.11, principale technologie sans fil utilisée dans les sous-couches MAC dans les réseaux ad hoc, et traite aussi les caractéristiques de transmissions sans fil ainsi que les problèmes qui leurs sont liés [26].

## 2.4. Le standard 802.11

Les réseaux locaux sans fil WLAN constituent le type de réseau le plus dominant, premièrement à cause de l'utilisation des bandes ISM<sup>10</sup> sans licence et deuxièmes par leur large adoption de la norme IEEE 802.11 définie par le comité de normalisation des réseaux locaux 802 d'IEEE, la technologie standard WLAN la plus répandue aujourd'hui. La norme originale date de 1997, elle décrit les couches MAC et physique pour un débit allant jusqu'à 2Mb/s dans la bande des 800 MHz. Plusieurs extensions visant l'amélioration du standard étaient publiées, qui se différencient principalement en la bande passante, la distance d'émission, ainsi que le débit qu'elles offrent. Les principales extensions sont résumées dans le tableau 2.2. De plus, ces technologies peuvent être réalisées à faible coût, et l'interopérabilité est garantie par la normalisation. Le protocole 802.11 définit deux contextes d'utilisation distincts pour un réseau sans fil ayant chacun des caractéristiques propres, où un groupe de STAs (équipement, utilisateur ou réseau) forme un ensemble de services de base « BSS » (Basic Service Set), les stations peuvent être connectées en mode **infrastructure** ou en mode **ad hoc** (sans infrastructure) décrit dans les paragraphes suivants. Ces deux modes de fonctionnement permettent de définir la topologie du réseau sans fil.

<sup>8</sup> Le Comité Consultatif International Téléphonique et Télégraphique (CCITT) définit la QoS comme « *l'effet général de la performance d'un service qui détermine le degré de satisfaction d'un utilisateur du service* »

<sup>9</sup> Principal organisme de normalisation des protocoles de réseaux locaux

<sup>10</sup> Fréquences radio de 2.4 GHz, avec une largeur de bande de 83.5 MHz, libres de droit pour les applications Industrielle, Médicale et Scientifique dans le monde entier ([https://fr.qaz.wiki/wiki/ISM\\_band](https://fr.qaz.wiki/wiki/ISM_band)).

### 2.4.1. Le mode avec infrastructure

Dans ce mode de fonctionnement, toutes les communications passent par une STA de base (BS : Base station) qui est appelée généralement un point d'accès (AP : Access Point)(Figure 2.2). D'une part, chaque BS se charge de contrôler un certain nombre de STAs sans fil qui sont dans sa portée de transmission, et d'autre part, elle est en mesure de se charger de certaines fonctions telles que la synchronisation. Des liens filaires peuvent être utilisés pour connecter les stations de base entre elles, ou les connectées au réseau filaire Ethernet<sup>11</sup>[18,26].

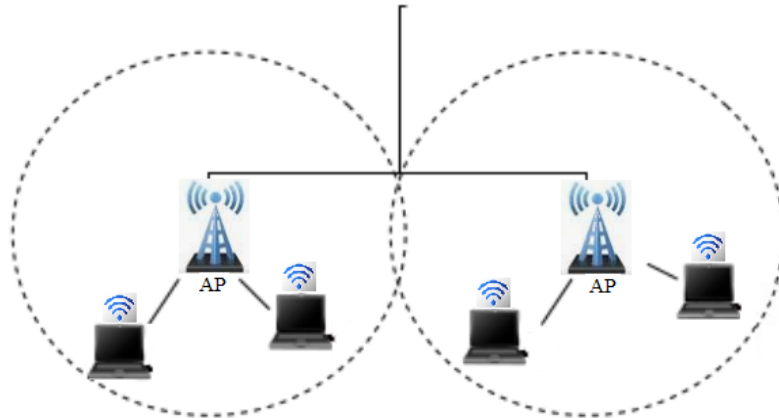


Figure 2.3: Réseau 802.11 en mode infrastructure

### 2.4.2. Le mode sans infrastructure (ad hoc)

Contrairement au mode infrastructure qui nécessite l'existence des BS entre les STAs sans fil pour qu'elles puissent communiquer. Dans le mode ad hoc, la communication entre les STAs se fait d'une manière directe, sans passer par un point d'accès [18], comme montré dans la figure 2.4.



Figure 2.4: Réseau 802.11 ad-hoc

<sup>11</sup> La technologie permettant des échanges rapides et stables, sans perturbations et sans interférences.



### 2.4.3. Présentation des protocoles MAC IEEE 802.11

Les WLANs ont été créés en 1990, avec leurs caractéristiques techniques ancrées dans la norme actuelle IEEE 802.11, la plupart des évolutions réalisées concernent l'amélioration du débit de données. Le groupes de travail IEEE 802.11 [5] a pris plusieurs années pour approuver le premier projet et ensuite le faire évoluer dans ses nombreuses normes (standard) et amendements, en particulier pour la transmission de couche PHY à plus haut débit [9]. Les normes sont IEEE 802.11b[5, 28], 802.11a[5, 27], 802.11g[5, 29], 802.11n[30] et 802.11h[34]. Pour l'amélioration de la qualité de service (QoS), il y a 802.11e[32] ; pour la sécurité, 802.11i[33] ; ...Etc. Dans le tableau ci-dessus, nous nous surlignons les informations générales de certaines des normes / amendements importants [35] qui sont étroitement liés au sujet de cette thèse.

**Tableau 2.2 :** Les différentes extensions du 802.11

La norme	Description
<b>802.11</b>	<ul style="list-style-type: none"> <li>- Définie en1997</li> <li>- Débit : 1-2 Mbps théoriques,</li> <li>- Bande de fréquence : 2,4 GHz,</li> <li>- Technique de modulation FHSS, DSSS</li> </ul>
<b>802.11a</b>	<ul style="list-style-type: none"> <li>- Définie en1999,</li> <li>- Débit : 54 Mbps théoriques, 30 Mbps réels,</li> <li>- Bande de fréquence : 5 GHz,</li> <li>- Technique de modulation : OFDM,</li> <li>- Portée :30m théorique,</li> </ul>
<b>802.11b</b>	<ul style="list-style-type: none"> <li>- Approuvée le 16 Décembre 1999,</li> <li>- Débit : 11 Mbps,</li> <li>- Bande de fréquence : 2,4 GHz,</li> <li>- Portée : 50m dans un environnement dégagé,</li> <li>- Technique de modulation DSSS étendu.</li> </ul>
<b>802.11e</b>	<ul style="list-style-type: none"> <li>- Disponible depuis 2005,</li> <li>- Bande de fréquence : 2,4 GHz, 5GHz</li> <li>- Performance en support de qualité de service (QoS) au niveau MAC,</li> <li>- Une meilleure transmission de la voix et de la vidéo.</li> </ul>
<b>802.11g</b>	<ul style="list-style-type: none"> <li>- Développée en 2003,</li> <li>- Débit : 54Mbps théoriques, 30 Mbps réels,</li> <li>- Bande de fréquence : 2,4GHz,</li> <li>- Compatible : 802.11b,</li> </ul>

	<ul style="list-style-type: none"> <li>- Portée : 55m.</li> <li>- Technique de modulation OFDM.</li> </ul>
<b>802.11h</b>	<ul style="list-style-type: none"> <li>- Datée de 2003,</li> <li>- Bande de fréquence : 5 GHz,</li> <li>- Vise à rapprocher la norme 802.11a du standard Européen (HiperLAN2),</li> <li>- Conforme avec la réglementation européenne (fréquence et économie d'énergie).</li> </ul>
<b>802.11i</b>	<ul style="list-style-type: none"> <li>- Ratifiée en juin 2004,</li> <li>- Décrit des mécanismes de sécurité des transmissions,</li> <li>- Agit en interaction avec les normes 802.11b et 802.11g,</li> <li>- Débit théorique est donc inchangé, à savoir 11 Mbps pour la 802.11b et 54 Mbps pour la 802.11g.</li> </ul>
<b>802.11n</b>	<ul style="list-style-type: none"> <li>- Publiée en 2008,</li> <li>- Débit: 600Mb/s théoriques,</li> <li>- Bande de fréquence : 2.4GHz, 5 GHz,</li> <li>- Portée : 75m, 140m,</li> <li>- Supporte la technologie MIMO<sup>12</sup> (Multiple Input Multiple Output) [13].</li> </ul>
...	- ...

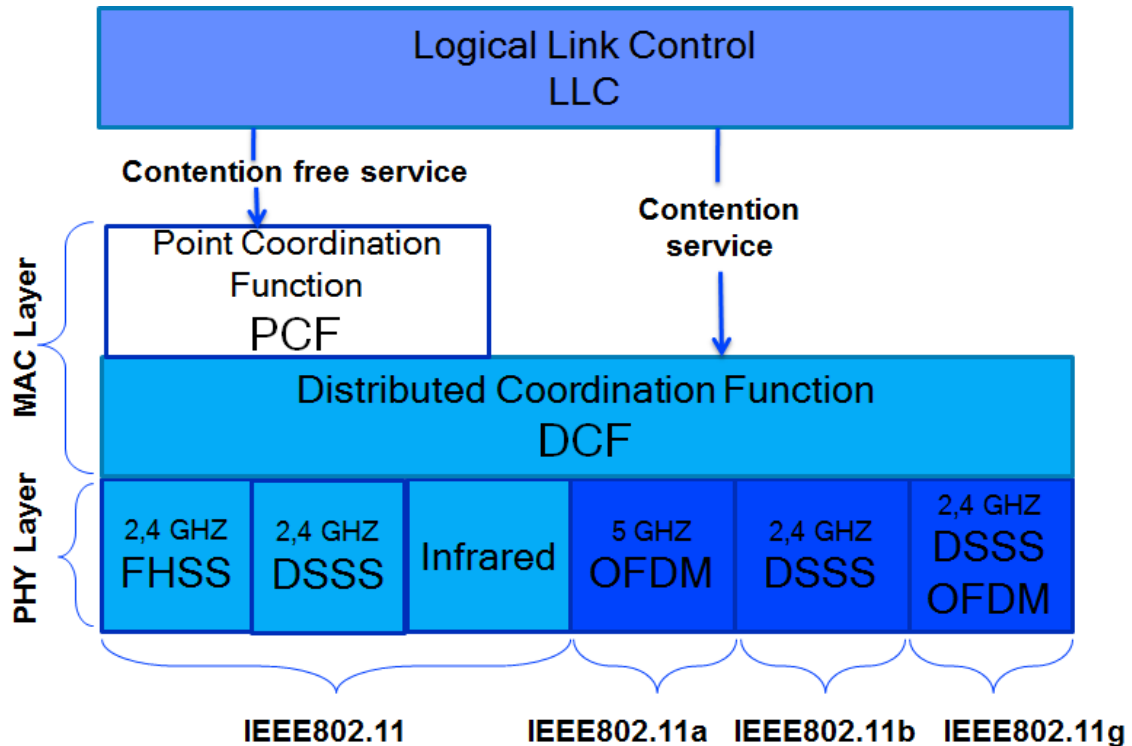
Il existe encore de nombreuses normes qui ne sont pas citées dans le tableau 2.2, comme la norme 802.11p (2010) [37] qui permet l'accès sans fil dans les environnements véhiculaires et le 802.11s (2011) [36] pour décrire le réseau maillé ; le 802.11ac [31], approuvé en 2013. Il est conçu pour fonctionner exclusivement dans la bande 5 GHz, offrant un débit agrégé allant jusqu'à 1 Gbit/s, (bandes passantes de canal de 80 et 160 MHz) et conserve l'énergie. Ainsi que la dernière norme 802.11ax (2019) [38] qui pour les réseaux denses HEW (High-Efficiency WLAN), qui modifie à la fois, la couche PHY et la couche MAC fournissant une compatibilité descendante avec les périphériques IEEE 802.11 existants. ...Etc.

## 2.5. L'architecture de 802.11

La norme 802.11 couvre les deux premières couches (basses) du modèle OSI, à savoir la couche physique (PHY), et la couche liaison de données ou plutôt la partie basse de cette couche qui s'appelle la couche MAC (Medium Accès Control). La figure (2.5) illustre l'architecture de la couche physique et la couche MAC selon la norme 802.11 comparée à celle de la norme OSI. Tel que la couche « Liaison de données » est subdivisée en deux sous-couches ; la sous couche

<sup>12</sup> Une nouvelle technologie de transfert de données utilisée dans les réseaux sans fil (wifi, LTE...) qui transmet plusieurs flux de données à la fois. MIMO double efficacement la portée d'un appareil sans fil.

LLC (Logical Link Control) et la sous couche MAC (Medium Access Control). La couche MAC du standard 802.11 utilise l'algorithme DCF pour les réseaux sans fil non centralisés, le cas des réseaux ad hoc. La DCF peut être utilisé par tous les nœuds, pour offrir un accès équitable au canal radio sans aucune centralisation de la gestion de l'accès dans un mode totalement distribué [39]. L'architecture MAC peut être décrite comme le montre la figure 2.5 comme fournissant la PCF via les services du DCF.



**Figure 2.5:** La couche PHY et la couche Liaison des données dans IEEE 802.11.

### 2.5.1. La couche physique

L'IEEE 802.11 utilise une variété de couches PHY, chacune avec des techniques de modulation différentes, dans le but d'augmenter le débit global du réseau, et de traiter un ensemble de points comme les techniques de modulation, les codes de correction des erreurs, les caractéristiques radio, la convergence de la couche physique, ainsi que d'autres problèmes liés à la signalisation. La norme IEEE 802.11 a commencé avec les trois couches PHY à savoir:

- IR (Infrarouge) :
- Le FHSS (Frequency Hopping Spread Spectrum) : désigne une technique d'étalement de bande fondée sur le saut de fréquence, la bande de fréquences est divisée en canaux et pour communiquer des données, l'émetteur et le récepteur s'accordent sur une séquence de sauts précise qui sera effectuée sur ces canaux [40].
- Le DSSS (Direct Sequence Spread Spectrum) : La modulation DSSS (Direct Sequence Spread Spectrum) est également une technique d'étalement de spectre qui divise aussi

la bande de fréquence, mais la transmission ne se fait que sur un canal donné (aucun saut n'aura lieu) contrairement au FHSS, le DSSS provoque des transitions d'état très rapides (chipping) qui tendent à étaler le spectre du signal, ceci permet avant tout de mieux résister au bruit [18, 40]. Comme la transmission ne se fait que sur un canal, les systèmes DSSS sont plus sensibles aux interférences que les systèmes FHSS, qui utilisent toute la largeur de bande [41].

Avec l'introduction de plusieurs spécifications à la famille IEEE 802.11, plus de couches physiques sont ajoutées [3], y compris

- L'OFDM (Orthogonal Frequency Division Multiplexing) : Repose sur le principe du multiplexage : permettre la transmission simultanée de plusieurs communications sur une même bande de fréquences [14].

### **2.5.2. Contrôle d'Accès au Médium**

La couche MAC (Medium Access Control) est une sous-couche de la couche liaison de données [61]. Cette couche fait l'interface avec la couche physique et comprend des protocoles définissant la manière d'allouer les canaux radios partagés entre l'ensemble des nœuds mobiles. L'objectif de la couche MAC consiste à éliminer, autant que possible, les collisions de transmissions. [39].

La couche MAC de 802.11 comprend deux fonctions de coordination pour l'accès au canal, à savoir la fonction de coordination distribuée (DCF : Distributed Coordination Function) et la fonction de coordination de point (PCF : Point Coordination Function). La technique DCF spécifie le mécanisme de contention du canal utilisée dans le cas des transmissions des données asynchrones (Service best -effort). Contrairement à DCF, la méthode PCF est basée sur un algorithme centralisé permettant de transmettre des données synchrones et asynchrones. Il s'agit d'une technique optionnelle ne pourront être utilisée qu'en mode infrastructure, par contre la technique DCF est utilisée pour les deux modes infrastructure et Ad Hoc [14, 53].

## **2.6. Fonctionnement de la couche MAC dans les réseaux ad hoc**

### **2.6.1. Fonction de Coordination de point PCF**

Un mécanisme d'accès facultatif appelé PCF peut également être intégré au MAC IEEE 802.11, qui n'est disponible que pour les configurations de réseau d'infrastructure. Cette méthode d'accès utilise un coordinateur de point (PC : Point Coordinator) travaillant au point d'accès BSS pour décider quel STA a actuellement le droit de transmettre. L'interrogation est une opération essentielle dans la fonction PCF, où les nœuds sont interrogés par le PC « maître ». Une alternance entre les modes DCF et PCF pour l'accès au canal lorsque le PCF est implémenté [11]. La période de contention (CP : Contention Period) est la durée pendant

laquelle la fonction DCF est utilisé pour l'accès au canal [20], et la durée interrogée est appelée la période sans contention (CFP : Contention Free Period), utilisé en mode PCF. La fonction PCF peut nécessiter une coordination supplémentaire, pour permettre dans un espace physique se chevauchant un fonctionnement efficace dans les cas où plusieurs BSS coordonnés par points fonctionnent sur le même canal. Un mécanisme virtuel de détection de porteuse pris en charge par un mécanisme de priorité d'entrée est utilisé par le PCF. Afin de prendre le contrôle du support en définissant le vecteur d'allocation de réseau (NAV) dans les STAs, la fonction PCF peut distribuer des informations dans des trames de gestion Beacon. L'espace inter trame (IFS), est utilisé pour toutes les transmissions de trames sous la fonction PCF, cette espace qui est plus petit que le IFS utilisé pour les trames transmises via la fonction DCF. L'utilisation d'un IFS plus petit implique que le trafic coordonné par points aura un accès prioritaire au canal sur les STAs dans les BSS se chevauchant fonctionnant selon la méthode d'accès DCF [11].

Il est à noter que la méthode d'accès PCF étant optionnelle dans le standard 802.11 et vu la complexité de sa mise en œuvre, actuellement, elle n'est plus (ou peu) implémentée.

### **2.6.2. Fonction de Coordination Distribuée DCF**

Bien que le IEEE 802.11 fournisse le mécanisme PCF, le mécanisme MAC fondamental de la norme WLAN basée sur le IEEE 802.11 est la fonction de coordination distribuée (DCF) [12]. DCF est essentiellement un mécanisme de CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) avec un algorithme de BEB (Binary Exponential Backoff) [60] (Figure 2.8). Cet algorithme est utilisé pour espacer d'une durée minimale spécifiée les retransmissions répétées du même bloc de données, souvent dans le cadre de l'évitement de congestion du réseau. Les protocoles CSMA sont bien connus dans l'industrie, le plus populaire étant l'Ethernet, qui est un protocole CSMA avec détection de collision (CSMA/CD<sup>13</sup>) [41]. Un protocole CSMA fonctionne comme suit : toute STA qui désire transmettre vérifie le médium. Si le médium est occupé (c'est-à-dire qu'une autre STA émet), la STA retarde sa transmission à un temps ultérieure. Si le support est détecté comme libre, la STA est autorisée à émettre. Ces types de protocoles sont très efficaces lorsque le support n'est pas fortement chargé, car il permet aux STAs de transmettre avec un retard minimum. Mais il y a toujours une chance que des STAs perçoivent simultanément le support comme libre et transmettant en même temps, provoquant ainsi une collision. Ces situations de collision doivent être identifiées afin que les paquets puissent être retransmis par la couche MAC, plutôt que par les couches supérieures, ce dernier cas entraînera un retard important. Afin de surmonter le problème de collision, le 802.11 utilise un mécanisme CA basé sur l'algorithme BEB [54].

---

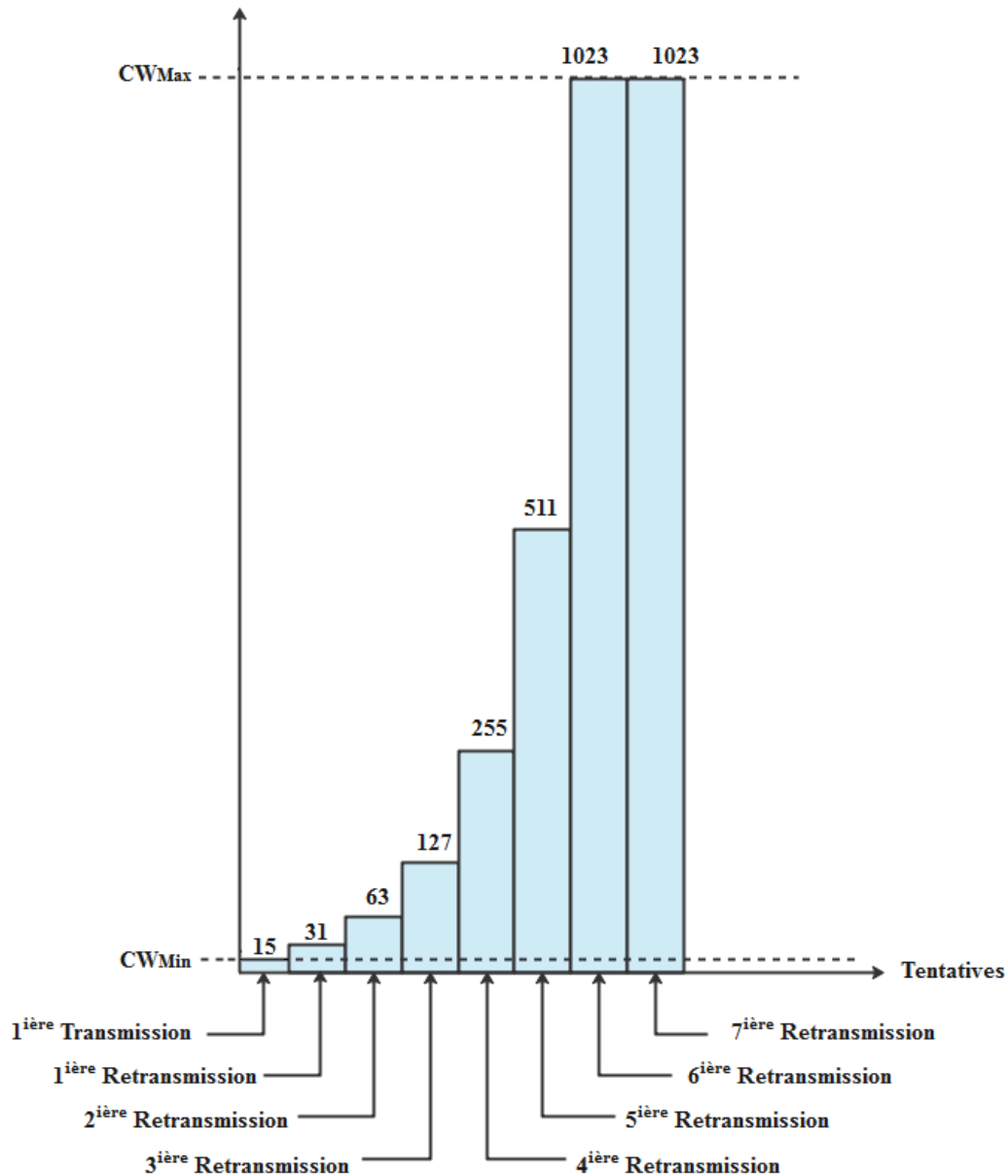
<sup>13</sup> Carrier Sense Multiple Access/Collision Detection est un protocole utilisé dans le réseau Ethernet (IEEE 802.3) pour gérer le partage de l'accès physique.

### 2.6.3. L'algorithme du backoff

Binary Exponential Backoff « BEB » (décrit dans la figure 2.8) représente un élément essentiel pour réduire la probabilité de collision sur le canal dans le mécanisme d'accès DCF [27]. Il est basé sur le calcul d'un temporisateur qui gère les transmissions et les retransmissions, de manière à réduire la chance que plusieurs STAs accèdent au canal de transmission en même temps. BEB repose sur le tirage aléatoire d'un nombre (temporisateur) appelé backoff dans un intervalle appelé fenêtre de contention (*CW*, Contention Window), par une STA qui désire transmettre sur le canal. Ce nombre est compris entre 0 et la *CW*. La durée de backoff est calculée comme suit [58] :

$$Backoff = random[0, CW - 1] \times SlotTime \quad (2.1)$$

Où *SlotTime* est la durée minimale pour déterminer l'état du canal + temps aller-retour + temps de propagation défini par la couche physique (qui vaut 20µs dans DSSS).



**Figure 2.6:** Exemple d'augmentation exponentielle de CW

La procédure de transmission suit la procédure suivante :

- Une STA qui veut transmettre des données détecte le médium. Si le support est occupé la STA diffère la transmission. Si le support est libre pendant une durée spécifiée, appelée DIFS (Distributed Interframe Space) dans la norme, la STA est autorisée à émettre.
- La STA de réception vérifie le contrôle de redondance cyclique (CRC) du paquet reçu et envoie un paquet d'accusé de réception. La réception de l'accusé de réception indique à l'émetteur qu'aucune collision n'est survenue.

- Si l'expéditeur ne reçoit pas l'accusé de réception, il retransmet la trame jusqu'à ce qu'il reçoive l'accusé de réception ou le rejette après un certain nombre de retransmissions. Selon la norme, un maximum  $R_{max}$  (égal à 7) retransmissions est autorisé avant que la trame soit abandonnée [44] (figure 2.6).
- Un intervalle d'interruption exponentielle binaire à être inséré est uniformément choisi dans  $[0, CW-1]$ .
- Le compteur de backoff est diminué tant que le canal est détecté inactive, arrêté lorsqu'une transmission est en cours, et réactivée lorsque le canal est à nouveau inactif pour plus de DIFS. Lorsque le compteur de backoff expire, la STA commence à émettre.
- Après la réception réussite de chaque trame de données, le récepteur transmet une trame d'accusé de réception (ACK) après une période SIFS (Short Interframe Space).
- La valeur de CW est fixée à sa valeur minimale,  $CW_{min}$ , dans la première tentative de transmission, et elle se fait incrémenter par des puissances entières de deux à chaque retransmission, jusqu'à une valeur prédéterminée  $CW_{max}$  (illustré dans la figure 2.6).

Le protocole décrit ci-dessus est appelé la méthode d'accès basique ou le mécanisme « *two-way handshaking* » [42] (illustré par la figure 2.7).

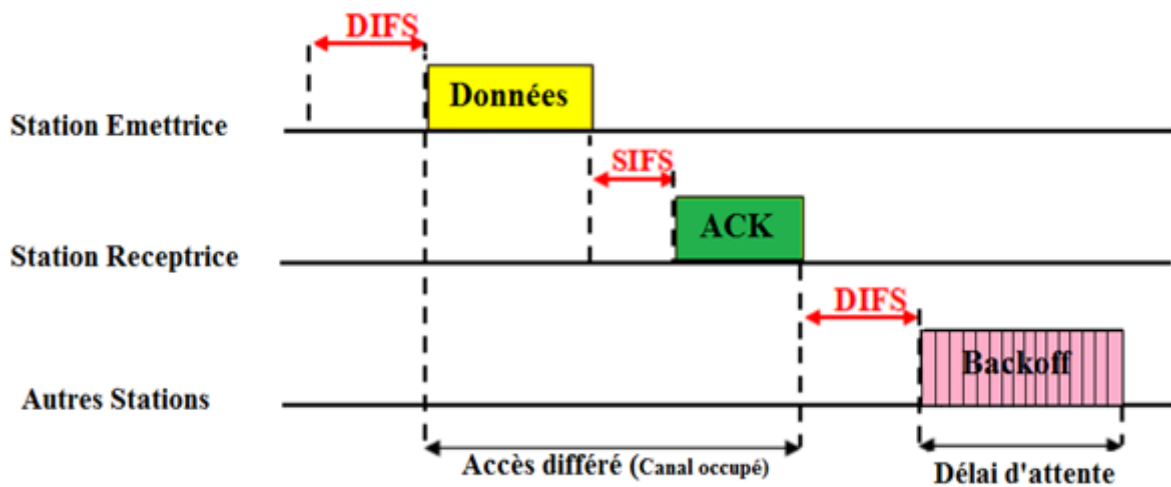


Figure 2.7: Méthode d'accès au médium DCF



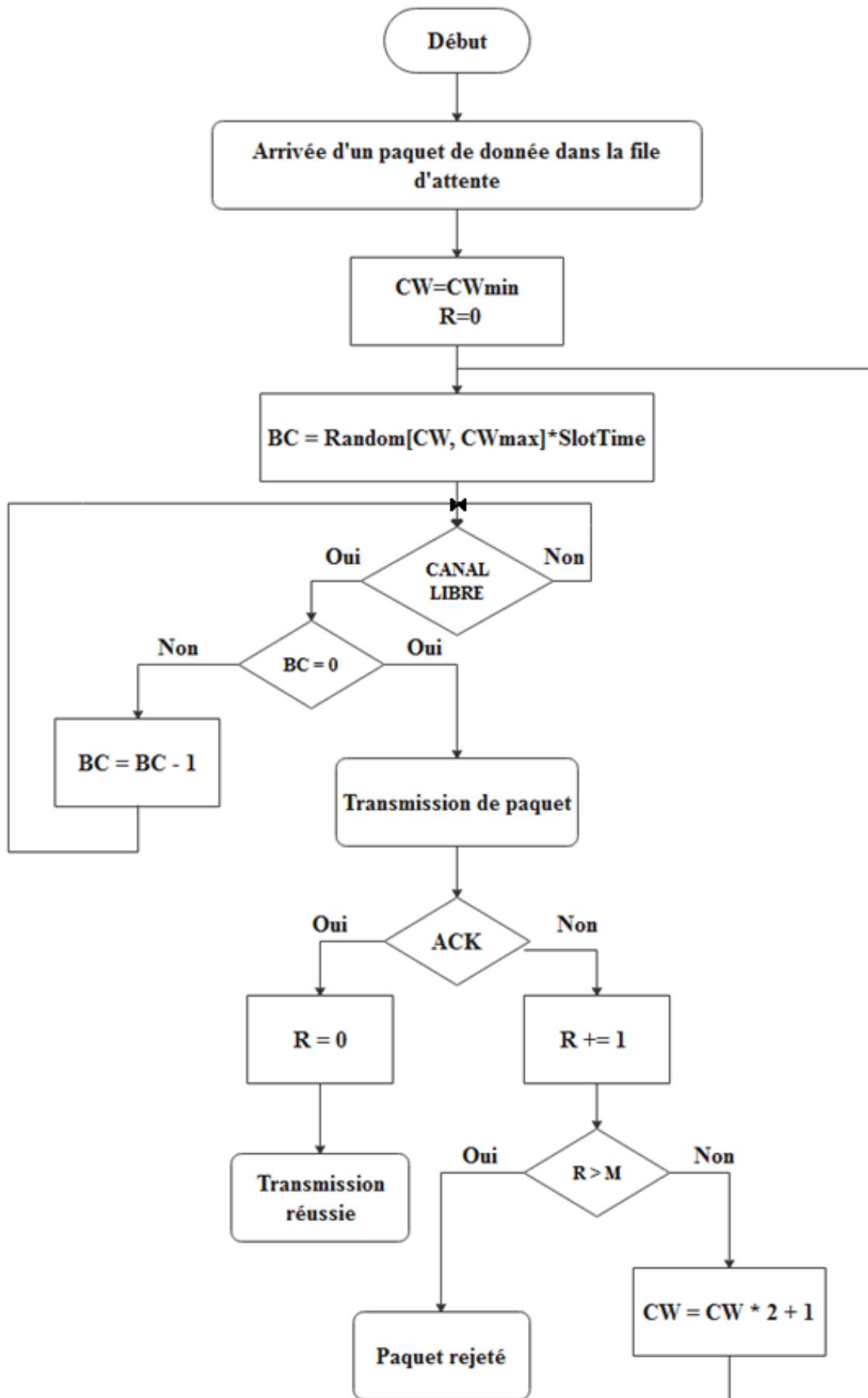


Figure 2.8: Organigramme de BEB

En outre, il existe également une autre méthode d'accès facultative dans la spécification, le « *four-way handshaking protocol* » (illustré par la figure 2.9), nommé le mécanisme RTS/CTS, ce dernier est utilisé pour réserver le canal avant la transmission de données, et fonctionnant suivant la procédure suivante :

- Une STA désirant transmettre un paquet envoie en premier lieu un paquet de contrôle court appelé Request To Send (RTS), qui comprend la source, la destination et la durée du paquet (le temps d'occupation du support de transmission).
- La STA de destination répond à son tour, par un paquet de contrôle de réponse appelé Clear To Send (CTS) après un intervalle SIFS, si le canal est libre, ce paquet inclut les mêmes informations de durées. La STA source est ensuite autorisée après une période SIFS, de commencer l'envoi de la trame de données.
- Toutes les autres STAs recevant le RTS et/ou le CTS sont informés que le support de transmission est actuellement occupé et que l'émetteur est activé pour la transmission. Elles définissent leur indicateur CS virtuel, appelé Network Allocation Vector (NAV), pour la durée donnée et utilisent cette information avec le CS physique appelée Clear Channel Assessment (CCA) [58], au moment de la détection du support pour une nouvelle tentative de transmission. CCA fournit un mécanisme basé sur la détection de toute énergie supérieure ou inférieure à un seuil donné, permettant ainsi de déterminer si le médium est inactif ou non. En examinant le NAV, une station peut éviter de transmettre même lorsque le CCA indique un support inactif, une combinaison NAV/CCA est utilisée pour indiquer l'état occupé du support. Ce mécanisme réduit la probabilité de collision car toutes les stations qui reçoivent le RTS ou le CTS mettront à jour leur NAV et les collisions potentielles peuvent être évitées.
- La station de destination confirme la réception d'une trame en envoyant un accusé de réception (ACK) à la station émettrice. Si l'expéditeur n'obtient pas cet ACK, la trame est retransmise. Il existe également des compteurs de tentatives et des délais qui limitent la durée de vie d'une trame. [42].

Il convient également de mentionner que puisque les RTS et CTS sont des trames courtes, le mécanisme réduit également le surcoût des collisions, car ces courtes transmissions permettent une détection plus rapide des collisions que ce qui serait possible pour la transmission d'un paquet entier [23]. L'objectif principal du protocole *four-way handshaking protocol* est de réduire la probabilité de collision de deux stations en raison de ne pas s'entendre, le « problème de nœud caché » détaillé dans la section 2.7.1. Puisque la trame RTS est beaucoup plus courte que celle des données, le mécanisme RTS/CTS est également utilisé pour minimiser les périodes perdues causées par les collisions.

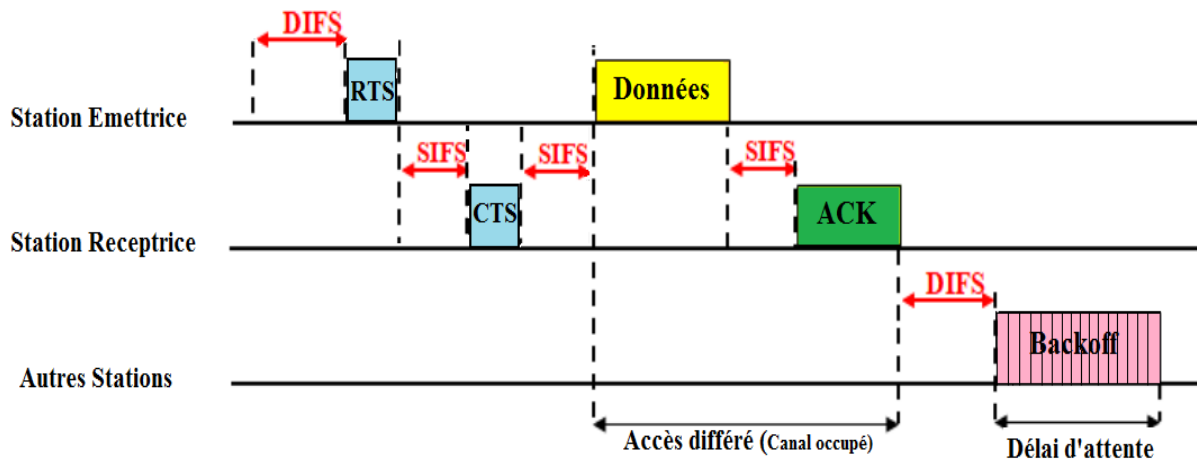


Figure 2.9: Méthode d'accès au médium avec RTS/CTS

## Espaces entre trames

La norme 802.11 définit quatre types d'espaces IFS (Inter Frame Space) entre trames, ces temporisations sont utilisées par la fonction DCF et par les autres fonctions d'accès [17], afin de varier la priorité d'accès au canal de certains paquets. Les durées IFS, de plus court au plus long (Figure 2.10), sont :

- **SIFS** (Short IFS) est le plus court IFS, il est utilisé pour la transmission des trames ACK, CTS, RTS, ..., issues d'une même STA et qu'aucune autre station ne devrait pouvoir interférer.

- **PIFS** (PCF IFS) est utilisé en mode PCF, il permet aux transmissions PCF de gagner l'accès au médium par l'utilisation d'un IFS plus petit que celui utilisé pour la transmission des trames en DCF.

- **DIFS** (DCF IFS) est le plus connu, il est utilisé en mode DCF comme un temps minimal d'attente avant la transmission.

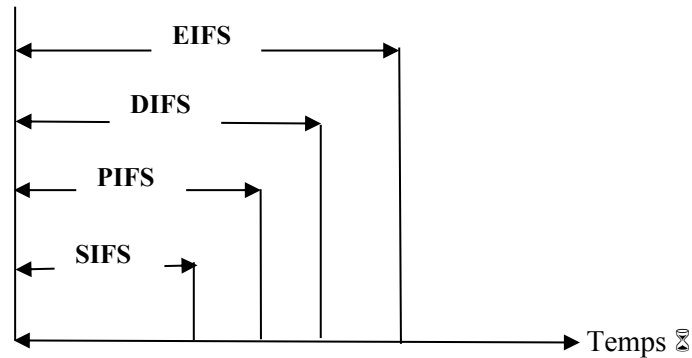
- **EIFS** (Extended IFS) est utilisé lorsqu'il y a détection de collision. Ce temps relativement long par rapport aux autres IFS, il est utilisé comme inhibiteur pour éviter des collisions en série.

Les valeurs de PIFS et DIFS sont calculées de la manière suivante :

$$PIFS = SIFS + SlotTime \quad (2.2)$$

$$DIFS = SIFS + 2 \times SlotTime \quad (2.3)$$

La valeur de SIFS est fixée par la couche physique et est calculée de telle façon que la STA émettrice sera capable de commuter en mode réception pour pouvoir décoder le paquet entrant.



**Figure 2.10:** Différence entre durées IFS

### 2.6.4. Trames MAC

Les trames de contrôle sont composées d'un en-tête et d'un champ de vérification d'erreurs.

- La trame RTS : cette trame est envoyée par le nœud source lorsque ce dernier est prêt à envoyer un paquet vers une destination connue ;
- La trame CTS : lorsque le nœud destination est prêt à recevoir une trame, en réponse à la trame RTS reçue, il envoie à la source une trame CTS. Le processus RTS/CTS permet au couple émetteur-récepteur de réserver de la bande passante afin de réduire les risques de collisions avec les autres STAs. Ce principe de réservation du canal par l'envoi de petits paquets de contrôle est appelé détection de porteuse virtuelle ou Virtual Carrier Sense. La période de réservation est stockée dans le vecteur d'allocation du réseau NAV.
- Trame ACK : C'est le paquet envoyé par la destination lorsqu'elle reçoit correctement et sans erreur une trame de données.

Les messages RTS/CTS et ACK sont prioritaires à l'accès au médium, car ils disposent d'un temps d'attente IFS (Inter Frame Space) inférieur à celui des paquets de données [23].

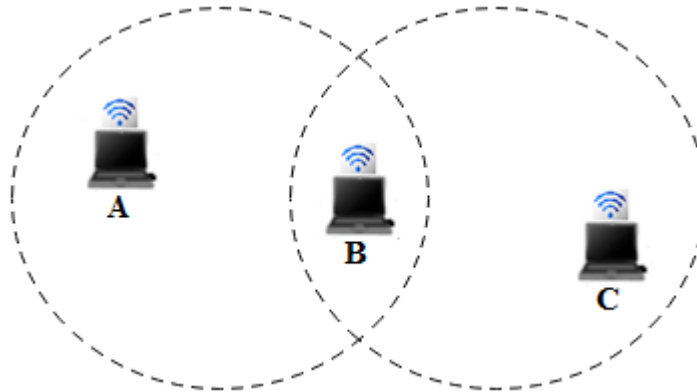
## 2.7. Problèmes dans les transmissions sans fil

Outre la limitation de la bande passante effective, le canal partagé entre les nœuds et les interférences sur le canal engendrent des problèmes de transmission dans les réseaux sans fil. Nous détaillons ci-dessous, les problèmes les plus importants connus dans les réseaux sans fil, qui sont le problème de la STA cachée et celui de la STA exposée, le problème du trou noir et le problème d'équité.

### 2.7.1. Le problème du nœud caché

Une STA (nœud) est cachée à une autre lorsqu'elle se trouve hors de portée de sa zone de transmission. Le problème de la STA cachée survient lorsque les deux STAs, essayent

d'envoyer des paquets en même temps vers la même STA, ce qui résulte en une collision au niveau de la STA destinataire des deux transmissions [27,29,30].



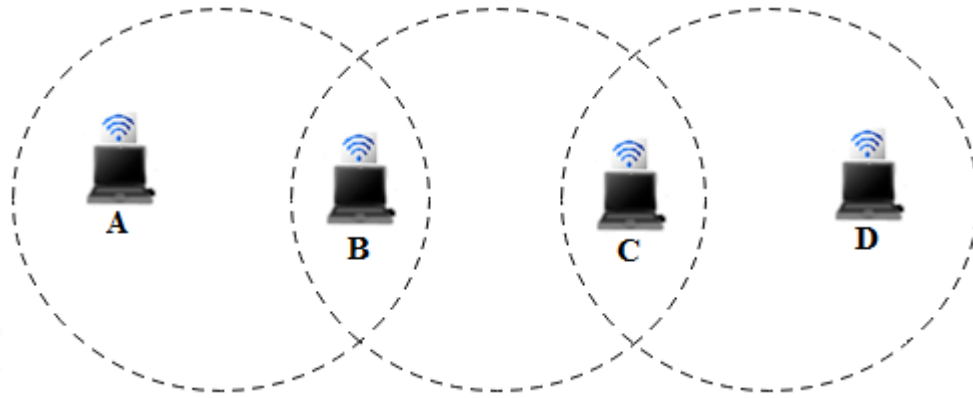
**Figure 2.11:** le problème de la station cachée.

Dans la figure (2.11), la STA *C* n'entend pas les transmissions de *A* qui émet vers *B*, *C* croit donc que le canal est libre et transmet alors vers *B*, cette transmission va entrer en interférence avec la transmission de *A* vers *B* et provoque des erreurs de données au niveau de *B*. Le fait qu'aucune des deux STAs émettrices ne sache que l'autre est sur le point d'émettre est la principale cause qui conduit à la collision. Une façon de remédier à ce problème et de mettre en place une procédure informative qui permet à la STA qui désire émettre d'informer les autres de son intention, et ne commence vraiment à émettre que lorsqu'elle reçoit une réponse de la STA destinataire, l'informant qu'elle est aussi prête à recevoir. La procédure utilisée est nommée mécanisme RTS/CTS.

### 2.7.2. Le problème du nœud exposé

Une STA (nœud) est exposée à une autre lors qu'elle se trouve dans sa portée d'émission. Ce problème survient lorsqu'une STA se trouve exposée à la STA émettrice, et en même temps, elle se trouve hors de portée de la STA destination. La position de la STA exposée la pénalise et l'empêche d'émettre quelle que soit la destination à laquelle elle veut transmettre.

Dans la figure (2.12), la STA *C* entend *B* et sait que *B* est en train de transmettre, *C* conclut que le canal est occupé et ne va donc pas transmettre à *D*, alors qu'en réalité si *C* transmet vers *D* ceci ne créera d'interférence que dans la région entre *C* et *B* mais pas d'interférence au niveau de *D*, autrement dit les données envoyées par *C* à *D* ne seront pas altérées par l'exposition de *C* à *B*. Le problème de la STA exposée peut être réduit si des antennes directionnelles sont utilisées. Dans la figure (2.12), si *B* était équipée de ce genre d'antenne, sa transmission vers *A* n'aurait pas empêché *C* d'envoyer vers *D* [23, 29].



**Figure 2.12:** Le problème de la station exposé

### 2.7.3. Problème de trou noir

L'attaque de trou noir (Blackhole) se produit lorsqu'un nœud malveillant tente de tirer autant de routes que possible pour contrôler la plupart des données circulant dans le réseau. Pour ce faire, l'attaquant doit apparaître aux autres comme étant très attractif, en présentant des routes optimales. L'attaquant se place généralement à un endroit stratégique et supprime tous les messages qu'il doit retransmettre ou bien permet la mise en œuvre d'une autre attaque, créant ainsi une sorte de puits ou "trou noir" dans le réseau [45].

### 2.7.4. Problème de l'équité

Certaines stations peuvent avoir plus de priorité que d'autres. L'équité d'accès au canal est un objectif important pour les réseaux sans fil. Le partage du support dans un réseau ad hoc repose sur les conditions de trafic et également sur la position relative des stations les unes par rapport aux autres. Précisément, certaines stations peuvent être en situation de famine au vu de la position géographique qu'elles occupent vis-à-vis de leurs voisines. L'accès au support leur est devenu quasiment impossible. Il devient alors très difficile dans ces conditions d'envisager un contrôle des ressources. Ce contrôle est particulièrement important lorsque l'on souhaite diversifier les services de communication sur des réseaux pour lesquels les capacités sont faibles et dont la qualité du canal varie significativement dans le temps. L'objectif d'équité d'accès au support est une des principales propriétés que doit offrir une méthode d'accès [46].

Bien que le protocole Ethernet câblé basé sur CSMA/CD soit connu pour être juste, sa contrepartie sans fil 802.11 basée sur CSMA/CA s'avère être injuste [43]. L'injustice des réseaux sans fil tient au fait que contrairement aux réseaux filaires, les collisions dans les réseaux sans fil sont asymétriques. Il n'est pas nécessaire dans les réseaux sans fil que toutes les stations impliquées dans la collision souffrent d'une perte de paquets. Les collisions et, par conséquent l'algorithme Binary Exponential Backoff peut se produire principalement à cause des trois raisons suivantes :

- Les transmissions provenant de deux stations interfèrent les unes avec les autres et par conséquent leurs transmissions ne sont pas reconnues. L'absence d'ACK est ensuite traitée comme une collision par les deux expéditeurs.

- Dans le mode de four-way handshake, la non réception d'un message CTS par une station pour sa requête RTS, sera traité comme une collision et le nœud émetteur double sa fenêtre de contention. Une station peut différer de renvoyer la trame CTS si une autre station dans son voisinage a réservée le canal en envoyant un RTS ou CTS à une autre station.

- Si deux stations effectuent des transmissions simultanées destinées à la même destination, l'une d'eux peut réussir à cause d'un niveau de puissance plus élevé. Ceci est appelé l'effet de capture dans les canaux sans fil [33].

La seule réelle situation de collision, est le premier scénario. Dans les deux autres scénarios l'une des deux stations réussie sa transmission et l'autre échoue. La station qui a échouée à transmettre, effectue le BEB pour la contention suivante du canal, par contre la station qui a réussie récemment à une probabilité plus élevée de gagner la contention suivante en raison de sa fenêtre inférieure. Ceci est essentiellement dû à la vue de congestion dissemblable sur le canal par différentes stations. Les stations qui réussissent généralement à accéder au médium le perçoivent qu'il est moins congestionné par rapport aux stations qui rencontrent des échecs. Cela invite les stations gagnantes à entrer dans le canal de manière excessive (en raison d'une CW inférieure) que les stations perdantes. Cela conduit à l'injustice du canal.

## 2.8. Conclusion

Ce chapitre décrit d'une manière générale les notions de base sur les réseaux mobiles ad hoc, et de manière précise la couche MAC et ses fonctions dans les réseaux ad hoc, ainsi, que les problèmes majeurs présentés au niveau de cette couche. L'une des principales préoccupations de la couche MAC est l'évitement des collisions entre les stations de réseau en raison du bon partage du support physique. L'algorithme du BEB est utilisé pour éviter les collisions, mais, les STAs souffrent de l'inégalité et l'injustice d'accès au canal due à la technique aveugle de doublement de la CW en cas de collision et de l'initialisation de la fenêtre en cas de succès. Cette technique aveugle cause d'autre dégradation des performances du réseau ad hoc tel que le débit, la consommation d'énergie, le taux de perte...etc. Dans le chapitre suivant, nous dressons un état de l'art des solutions visant à améliorer le CSMA/CA utilisé dans la couche MAC. Une attention sera donnée à l'apprentissage par renforcement, qui est une technique de l'intelligence artificielle, utilisée comme nouveau paradigme dans ce contexte, ainsi quelques solutions intelligentes basé sur le DRL seront citées.

## TRAVAUX ANTERIEURS

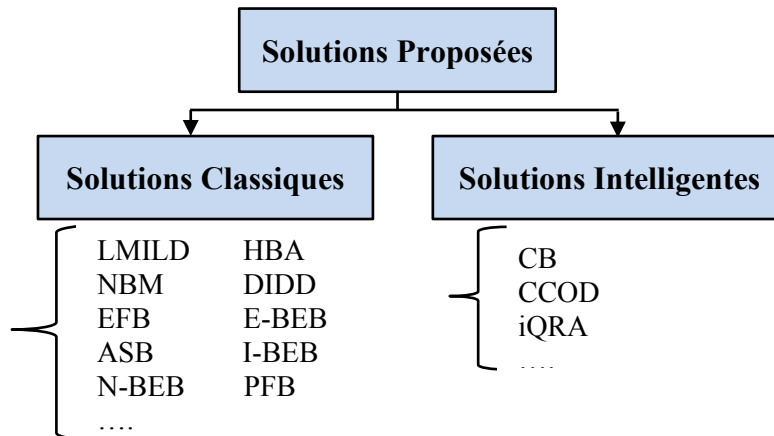
### 3.1. Introduction

Dans les MANET, le domaine Backoff est considéré comme l'un des domaines principaux des chercheurs afin d'obtenir un algorithme Backoff efficace. La méthode aveugle de calcul de la fenêtre de contention dont l'algorithme BEB du mécanisme DCF du protocole CSMA/CA utilise, a d'une part un impact considérable sur le comportement des protocoles des couches supérieures et des applications utilisant le réseau. D'une autre part, les applications temps réel (telles que l'audio/vidéo) sont sensibles aux retards et aux pertes des paquets et au débit faible. Lorsque le comportement de lien sous-jacent est injuste, certaines applications peuvent être privées de bande passante juste parce que leur part est injustement distribué ailleurs.

Puisque les problèmes d'assignation de la fenêtre de contention et l'équité sont profondément enracinés dans la couche MAC elle-même, Il est rationnel de supposer que ceux-ci peuvent être surmontés en modifiant correctement le code MAC de manière appropriée qui réalise une meilleure attribution de CW et de l'équité. A cet effet, Dans la littérature, de nombreux algorithmes de Backoff ont été proposés (BEB, LMILD, HBA, CB, ... etc.), où dans les plus récents, sont introduites des techniques de l'apprentissage automatique par renforcement, visant l'amélioration et la modification à la couche MAC [49]. Cet ensemble est illustré dans la figure 3.1. Dans ce qui suit nous présentons un ensemble de ces algorithmes qui ont été proposés afin d'éviter la collision et de résoudre les conflits entre les différentes STAs, ainsi que pour améliorer les performances réseaux tel que le débit, le taux de paquets livrés, l'énergie, l'équité...etc. Plus loin dans ce chapitre, nous mettons en évidence le contexte théorique de l'apprentissage par renforcement profond qui est une combinaison d'apprentissage



par renforcement et d'apprentissage profond, nous nous concentrerons sur les DRL des fonctions de récompense cumulée. Nous introduirons le processus de décision de Markov qui est généralement utilisé pour modéliser l'apprentissage par renforcement. Nous mettons l'accent sur le Q-Learning. Nous introduirons l'apprentissage par renforcement profond, en particulier, comme nous citons quelques travaux basés sur l'intelligence.



**Figure 3.1** : Quelques Solutions BEB dans la littérature

## 3.2. Les solutions proposées

### 3.2.1. Solutions classiques

Les algorithmes de Backoff que nous allons définir appartiennent à la catégorie des algorithmes dit dynamiques qui utilisent une variante de période d'attente de Backoff en choisissant aléatoirement la valeur du temporisateur de Backoff en fonction de l'équation 2.1. Les algorithmes exposés dans cette section n'utilisent aucune notion d'intelligence pour les modifications effectuées pour améliorer les performances des réseaux sans fil.

#### A. Binary Exponential Backoff (BEB)

Le mécanisme MAC fondamental de la norme WLAN basée sur le IEEE 802.11 est la fonction de coordination distribuée (DCF). DCF est essentiellement un mécanisme de CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) avec un algorithme de BEB (Binary Exponential Backoff). Cet algorithme est utilisé pour espacer les retransmissions répétées du même bloc de données, souvent dans le cadre de l'évitement de congestion du réseau. Dans le BEB, chaque STA réinitialise la taille de sa CW (où CW est la fenêtre de contention) à la valeur minimale ( $CW_{\min}$ ) après une transmission réussie, et double sa fenêtre après une transmission échouée, jusqu'à ce que la valeur maximale soit atteinte ( $CW_{\max}$ ). Si la STA atteint son nombre maximum de transmission  $maxR$ , le paquet sera jeté, et la transmission du prochain paquet est tentée à  $CW_{\min}$ . Par conséquent, il est fort probable qu'une station ayant

gagnée le canal et transmis avec succès va gagner le canal dans la contention suivante du canal. Un intervalle d'interruption exponentielle binaire appelé compteur de backoff est aléatoirement choisi dans  $[0, CW-1]$ , il est diminué tant que le canal est détecté inactive, arrêté lorsqu'une transmission est en cours, et réactivée lorsque le canal est à nouveau inactif. Lorsque le compteur de backoff expire, la STA commence à émettre. Le BEB fonctionne bien pour un nombre limité de STAs, mais sa performance diminue fortement lorsque le nombre de STAs devient trop important. [59, 60, 61] ; BEB calcule la fenêtre de contention selon les deux cas :

$$CW_{nouv} = \begin{cases} \min((CW_{cour} * 2) + 1, CW_{max}) & \text{En cas de collision} \\ CW_{min} & \text{En cas de succès} \end{cases} \quad (3.1)$$

### B. Linear/Multiplicative Increase and Linear Decrease (LMILD)

Proposé par Deng et al., 2004 [62], LMILD et afin de diminuer la collision, modifie dynamiquement les valeurs de CW, augmentant ainsi le débit normalisé de WLAN. Il nécessite une information supplémentaire de la couche physique pour détecter l'existence de la collision du paquet. Lorsqu'un nœud détecte le canal occupé pour une durée égale au temps de transmission du paquet, et qu'il ne reçoit ou n'entend pas un paquet, le nœud conclut la collision du paquet. La procédure suivie dans LMILD est la suivante : Si le nœud rencontre une collision, alors la CW est multipliée par le facteur  $mc$ , sinon si le nœud surprend une collision, alors la CW est incrémentée linéairement par des unités  $lc$  sinon si le paquet a été transmis avec succès, la CW est décrémentée linéairement par des unités  $ls$ . Les paramètres  $mc$  et  $lc$  augmentent leurs CW en cas de collision pour réduire la probabilité de collision. La probabilité d'accès au canal est réduite pour les plus grandes valeurs de  $lc$  et les plus petites valeurs de  $ls$ . En raison de l'augmentation de la taille des CW, la probabilité de collisions est augmentée dans le cas inverse. Sur la base de l'estimation de l'indice d'équité, les valeurs des paramètres  $mc$ ,  $ls$ , et  $lc$  sont choisis comme 2, 2, et 6 respectivement.

### C. Double Increment Double Decrement (DIDD)

L'approche DIDD est proposée par Chatzimisios et al., 2007 [65], pour améliorer le débit de saturation à la fois dans la méthode d'accès simple et dans la méthode d'accès RTS / CTS en utilisant l'approche de probabilité conditionnelle élémentaire. DIDD double la CW exactement comme BEB lorsque la collision est observée pour diminuer la probabilité de collision. Si le paquet a été transmis avec succès, la CW est divisée par deux pour réduire la collision de paquet. Alors qu'en BEB, CW se réinitialise à  $CW_{min}$ , ce qui, avec un trafic accru, peut provoquer une dégradation significative des performances. Pour éviter cela, après une transmission réussie, DIDD diminue la probabilité de collision avec une CW plus élevée. Dans cette approche, à condition que les paquets ne soient pas supprimés et atteignent leur nombre total de tentatives de retransmission, un délai plus élevé peut être supérieur à BEB. À partir de la simulation, il a été montré que DIDD augmente le taux de saturation., étant donné que les paquets ne sont pas

supprimés et atteignent leur nombre maximal de tentatives de retransmission, un délai plus élevé peut être supérieur à BEB. De la simulation, il a été démontré que DIDD augmente le débit de saturation. DIDD calcule la CW comme suit :

$$CW_{nouv} = \begin{cases} CW_{cour} \times 2 & \text{En cas de collision} \\ \frac{(CW_{cour})}{2} & \text{En cas de succès} \end{cases} \quad (3.3)$$

#### D. Hybrid Backoff Algorithm (HBA)

Dans le HBA proposé par Peng et al., 2007 [63], la taille de la CW est incrémentée linéairement après plusieurs collisions. Basé sur le modèle de chaîne de Markov, la performance du HBA, en termes de débit saturé, de probabilité et de délai de chute de paquets, est évaluée et comparée avec BEB. Basé sur les phases de backoff la CW est exponentiellement augmentée ( $CW_i = 2^i CW$  ;  $0 \leq i \leq m_1$ ) au début de la phase de backoff ou elle est augmentée linéairement au milieu de la phase de backoff ( $CW_i = 2^{m_1} CW + a(i - m_1)$  ;  $m_1 \leq i \leq m_2$ ) et aussi à la fin de la phase de backoff la CW est augmentée linéairement. ( $CW_i = 2^{m_1} CW + a(m_2 - m_1)$  ;  $i > m_2$  et  $0 \leq i \leq 7$ ) alors que  $m_2$  indique la valeur maximale du seuil pour la phase de backoff et  $m_1$  donne la valeur minimale du seuil pour la phase de backoff. La valeur de  $m_1$  et  $m_2$  doit être inférieure ou égale à 7 et  $m_2$  doit être supérieure à  $m_1$ . Après plusieurs collisions, c'est-à-dire au milieu de la phase de backoff, HBA sélectionne une taille de CW plus grande afin de réduire la probabilité de collision en choisissant de manière appropriée le paramètre  $m_1$ . Si  $m_1$  est grand, alors le nombre de slots de backoff est augmenté pour réduire la probabilité de collision et la probabilité de perte de paquets. Mais cela augmentera le retard dû à un temps d'attente plus long.

En récapitulatif, HBA calcule la CW comme suit :

$$CW_{nouv} = \begin{cases} 2^i \times CW_{cour}; 0 \leq i \leq m_1 & \text{Au début du backoff} \\ 2^{m_1} CW_{cour} + a(i - m_1); m_1 \leq i \leq m_2 & \text{Au milieu du backoff} \\ 2^{m_1} CW_{cour} + a(m_2 - m_1); i > m_2 \text{ \& } 0 \leq i \leq 7 & \text{A la fin du backoff} \end{cases} \quad (3.2)$$

#### E. New Backoff Method (NBM)

Cette approche de backoff proposée par Minooei et Nojumi, 2007 [64], est modélisée pour augmenter le débit plus que celui donné par BEB avec un modèle de chaîne de Markov en temps discret. NBM choisit le temps de backoff entre les tailles de la CW précédente et actuelle, au lieu entre zéro et la CW actuelle. Afin de réduire la contention entre les nœuds lorsque le réseau n'est pas saturé, l'intervalle  $[CW_{i-1}, CW_i]$  est choisi plutôt que  $[0, CW_i]$ . Une fois le paquet transmis avec succès, le compteur de backoff est décrémenté d'une unité au lieu d'assigner  $CW_{min}$ . Il a été montré que le débit global de saturation est presque égal à BEB lorsque le

nombre de STAs augmente. L'algorithme NBM est donc idéal pour un faible trafic avec un petit nombre de STAs.

## F. Improved Backoff Algorithm (I-BEB)

Dans l'Improved Backoff Algorithm proposé par Kun Wang et Changming Sun, 2009 [69], afin d'améliorer la capacité de débit, L'algorithme I-BEB a le même concept que BEB classique dans les cas de collision, donc reporter les transmissions. Les modifications ont été apportées en cas de transmission réussie en utilisant des variables constantes pour contrôler la CW (la figure 3.2 illustre l'organigramme d'I-BEB en cas de succès de transmission). La valeur de la fenêtre de contention doit être rapidement réduite sous forme de multiple. Un seuil de fenêtre de contention  $G$  est conçu, lorsque la valeur de la fenêtre de contention diminue vers une certaine mesure  $G$ , alors la valeur de la fenêtre de contention est juste égale à  $G$ . Cependant, après que cette procédure s'exécute pour une période de temps, l'algorithme I-BEB produira un résultat analogue à celui de BEB, donc un dispositif de comptage est ajouté pour enregistrer le nombre de fois d'envoi avec succès. Le principe de l'ajout d'un dispositif de comptage est quand ce compteur atteint une certaine valeur, la fenêtre de contention augmentera sous forme linéaire, d'autres nœuds auront une plus grande chance d'accéder au canal. Puis, le but est atteint. Les auteurs d'I-BEB ont fixé les valeurs pour les paramètres  $C$ ,  $D$ ,  $E$  et  $G$  comme suit :  $C = 12$ ,  $D = 4$ ,  $E = 8$ ,  $G = 0,125 * CW_{min}$ .

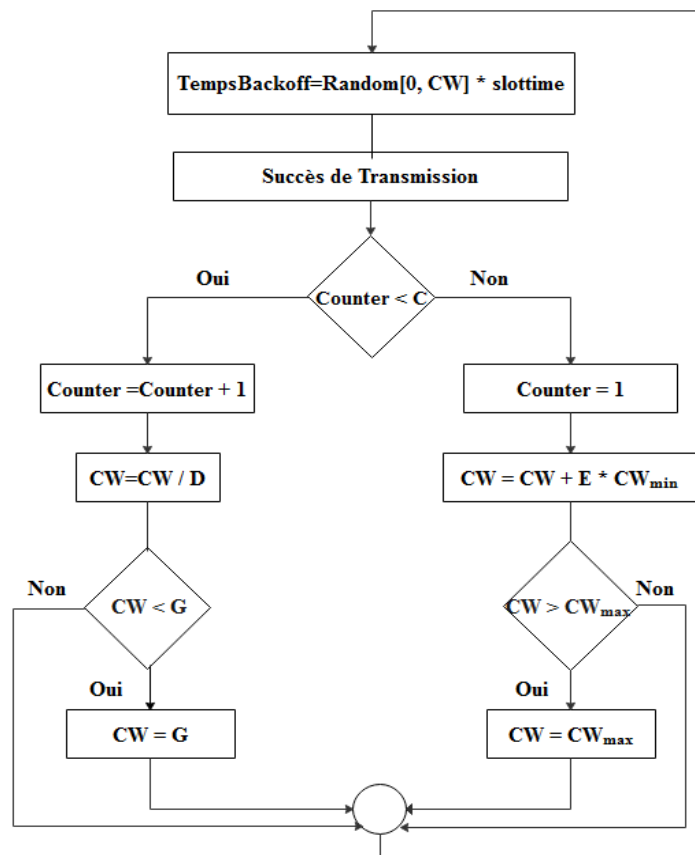


Figure 3.2 : Organigramme de l'I-BEB « cas de succès de transmission »

## G. Enhanced Fibonacci Backoff (EFB)

L'EFB, proposé par Al Oqaily et al., 2010 [66], sélectionne la valeur CW optimale basée sur la méthode de Fibonacci. Chaque fois que la collision se produit, CW est mis au prochain numéro de Fibonacci depuis la valeur CW actuelle (nombre de Fibonacci  $\leq CW_{max}$ ). Sinon, il est défini sur le nombre Fibonacci précédent qui doit être supérieur ou égal à  $CW_{min}$ . L'EFB donne des incréments plus petits dans la plus grande fenêtre de contention, suite à la nature de la suite de Fibonacci. A partir des résultats de la simulation, il a été conclu que l'EFB donne un meilleur rapport de livraison de paquets et réduit le délai d'accès et le surcharge de routage pour un nombre différent de nœuds à des vitesses de mobilité variables. L'équation 4.4 montre comment l'EFB calcule la CW :

Suite de Fibonacci  $\rightarrow F(n) = F(n - 1) + F(n - 2) \quad // F(0) = 1 \quad F(1) = 1 \quad n \geq 0$

$$CW_{cour} = F(n)$$

$$CW_{nouv} = \begin{cases} F(n + i) & | F(n + i) < CW_{max} & \text{En cas de collision} \\ F(n - i) & | F(n - i) > CW_{min} & \text{En cas de succès} \end{cases} \quad (3.4)$$

## H. Enhanced BEB (E-BEB)

Proposé par Mohammed Al-Hubaishi et al., 2012 [67], sur la base du nombre de transmissions réussies dans un environnement stable, E-BEB qui augmente ou diminue la CW pour fournir une plus grande équité d'accès au canal et ainsi un débit supérieur à BEB et I-BEB (présenté dans la section 4.2 9). Si une collision se produit, CW est doublé exactement comme BEB. La valeur du compteur est incrémentée lorsque la transmission est réussie, sur la base de cette valeur, la nouvelle valeur CW est calculée comme suit :

$$CW_{nouv} = \begin{cases} CW \times 2 & \text{En cas de collision} \\ \begin{cases} \frac{1}{\sqrt{CW_{min}}} \times CW_{min} \\ CW + \left(\frac{CW_{max}}{CW}\right) \times CW_{min} \end{cases} & \text{En cas de succès} \end{cases} \quad (3.5)$$

Où  $CW_{min} \leq CW_{cur} \leq CW_{max}$ , basé sur les résultats des simulations, il a été montré que l'algorithme E-BEB donne à la fois un indice d'équité élevé et un débit plus élevé que ceux donnés par les autres algorithmes.

## I. Optimized Pessimistic Fibonacci Backoff (PFB)

Proposé par Bani Yassein et al., 2012 [71], PFB augmente la taille de la fenêtre de contention de manière exponentielle et un échec de transmission survient pour prolonger le temps d'attente avant la retransmission, puis PFB augmente le temporisateur de manière cubique afin d'éviter d'augmenter extrêmement le délai backoff. Après un nombre fixe d'incrémentations cubiques, PFB commence à augmenter le temporisateur, en utilisant une suite de Fibonacci pour obtenir une croissance moins dramatique de la taille de la fenêtre de contention permettant ainsi aux nœuds d'effectuer davantage des essais pour accéder au canal. De la simulation ; il a été démontré que PFB donne un taux de paquets délivrés PDF, un délai de bout en bout et une charge de routage meilleurs.

## J. New Binary Exponential Backoff (N-BEB)

Proposé par Mohammad Shurman et al, 2014 [70], Dans N-BEB, si la transmission réussie, la CW est décrétementée pour les neuf premières transmissions d'une valeur égale à  $\sqrt{2}$ , en suivant cette valeur est incrémentée d'une valeur égale à 1,5 et le compteur est remis à 1. En cas d'échec de transmission, l'algorithme incrémente la CW d'une valeur égale à  $\sqrt{2}$  pour les deux premiers échecs, puis il décrémente cette valeur de 1,5 et le compteur est remis à 1. La figure 3.3 illustre l'organigramme de N-BEB dans les deux cas succès et collision respectivement.

## K. Adaptively Scaled Backoff (ASB)

Proposé par Nurullah Shahin et al., 2018 [68], l'ASB est similaire au BEB en cas de transmissions réussies, mais la différence consiste dans le calcul de la taille de la CW, qui multiplie la CW par le  $S_{factor}$  en cas de collision :

$$CW_{nouv} = \begin{cases} \min((CW_{min} * S_{factor}), CW_{max}) & \text{En cas de collision} \\ CW_{min} & \text{En cas de succès} \end{cases} \quad (3.6)$$

Où :  $S_{factor} = S_{factor} + \text{Round} \left( \frac{n * P_c}{S_{factor}} \right)$  et  $P_c$  est la probabilité de collision

D'après les simulations, SBA offre de meilleures performances que BEB en termes de débit, délai et équité.

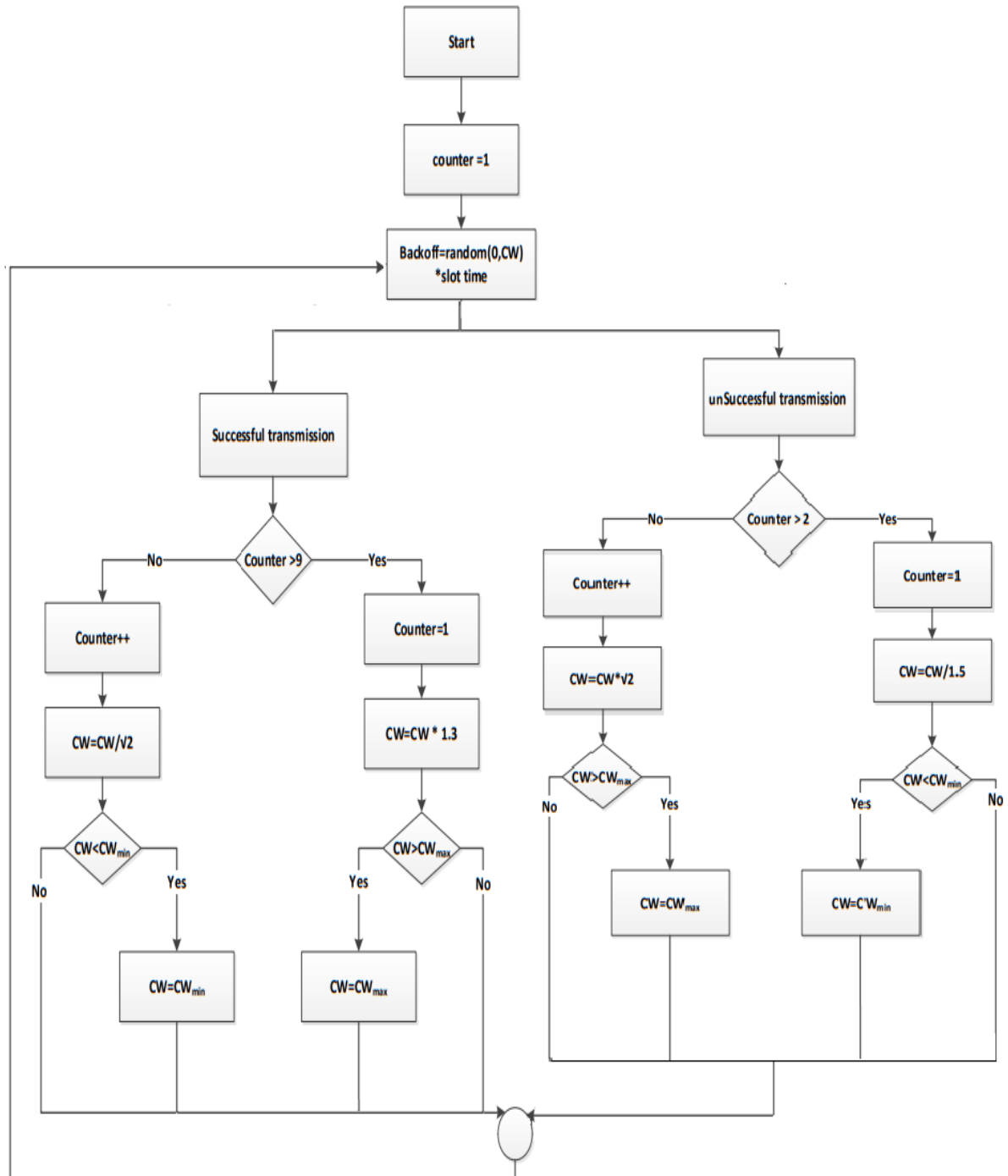


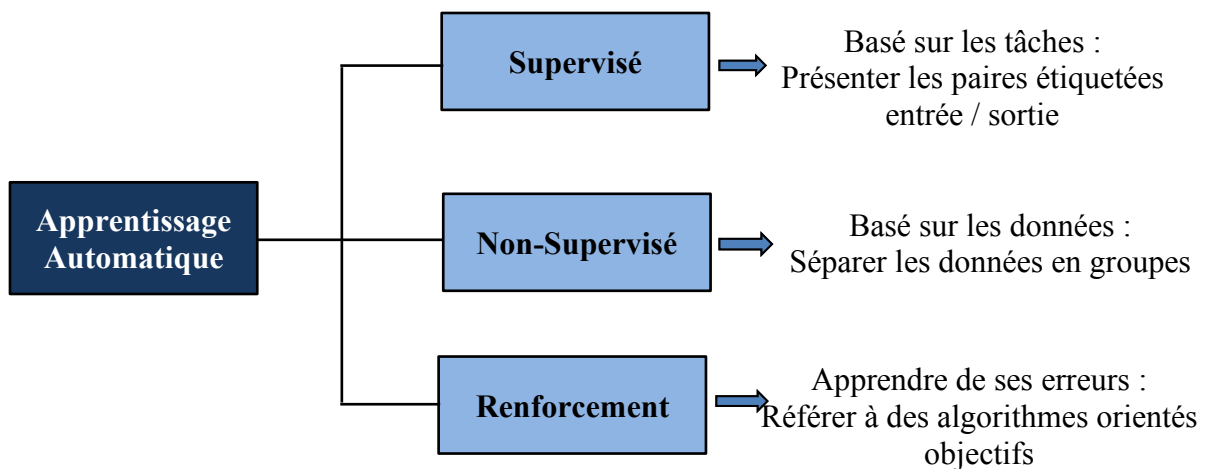
Figure 3.3 : organigramme de N-BEB

### 3.2.2. Solutions Basées sur l'intelligence

Avant de citer les solutions intelligentes proposées dans la littérature, en premier lieu nous allons mettre l'accent sur l'apprentissage par renforcement (RL : Reinforcement Learning), qui fait référence à un ensemble de problèmes d'apprentissage automatique (ML : Machine Learning), utilisé dans ces solutions.

L'objectif principal de RL est d'élaborer des politiques de contrôle de processus dont le modèle d'évolution est inconnu. L'apprentissage est dirigé par une grandeur scalaire (appelée récompense ou renforcement), cette valeur permet de mesurer la qualité de la dernière transition (action) entreprise par le processus (i.e. état précédent – action – nouvel état). Avec certaines conditions, l'apprentissage converge vers une politique de décision maximisant la somme des renforcements futurs (politique optimale).

Dans RL, un agent peut périodiquement prendre des décisions, observer les résultats, puis ajuster automatiquement sa stratégie pour atteindre la politique optimale. Le but de l'agent est de résoudre le défi en recherchant une séquence idéale d'actions.



**Figure 3.4 :** Types d'apprentissage automatique

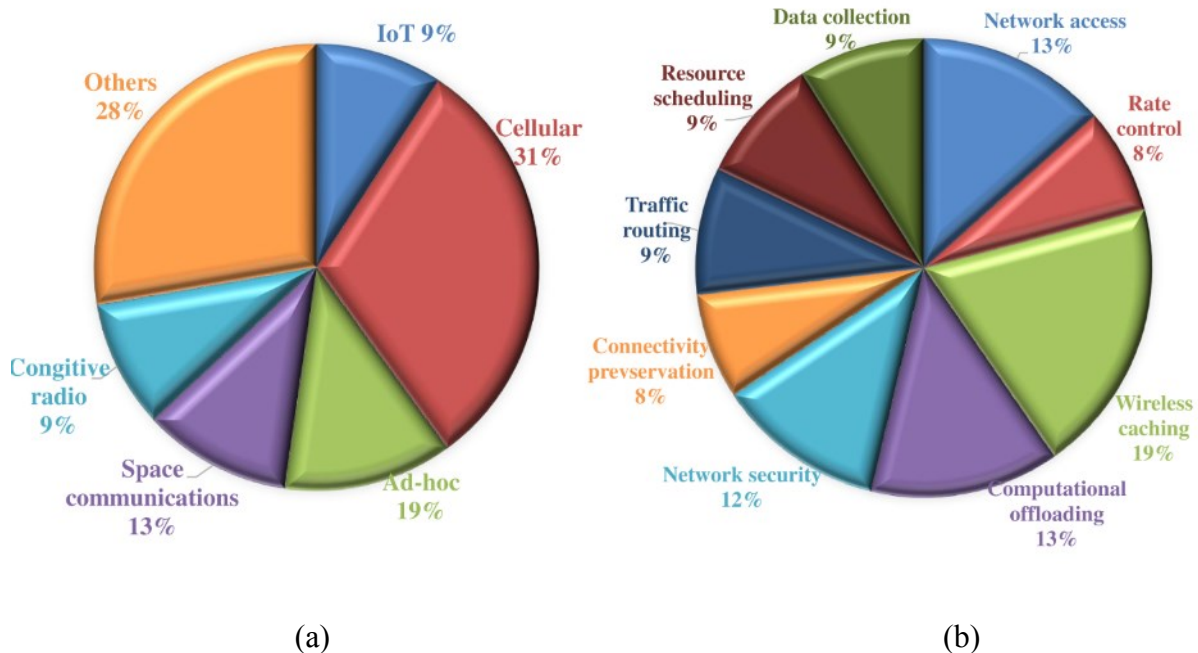
Bien que RL soit un des trois domaines de l'apprentissage automatique (ML : Machine Learning), il est fondamentalement différent des deux approches d'apprentissage automatique standard (supervisé et non-supervisé) dans différents aspects (figure 3.4). Premièrement, l'apprentissage par renforcement n'est pas basé sur l'acquisition de données. Au lieu de cela, dans RL, l'agent apprend à fonctionner par sa propre expérience créée lors de son interaction avec l'environnement. L'agent entreprend d'abord des actions et, par conséquent, l'état changera en fonction des récompenses gagnées ou perdues et ne dépend pas d'un superviseur. Deuxièmement, l'apprentissage par renforcement se concentre sur la recherche d'une politique optimale plutôt que sur l'analyse des données.

Le processus d'apprentissage prend beaucoup de temps pour parvenir à la bonne politique, alors qu'il s'est avéré converger, ce qui rend RL inapproprié et inapplicable aux grands réseaux. De plus, dans la pratique, les implémentations d'apprentissage par renforcement sont très limitées. L'apprentissage en profondeur DL (Deep Learning) a récemment été implémenté en tant que nouvel outil révolutionnaire qui peut surmonter les limites de RL et ainsi donner naissance à l'apprentissage par renforcement, à savoir l'apprentissage par renforcement profond DRL (Deep Reinforcement Learning) [48], cette nouvelle notion fera l'objectif de la section suivante.



### 3.2.2.1. Apprentissage par renforcement profond DRL

DRL, est une catégorie d'apprentissage automatique et d'intelligence artificielle qui a récemment été utilisée comme un outil émergent pour résoudre efficacement divers problèmes et défis dans le domaine des réseaux et des communications (Internet de objets IoT, réseaux hétérogène HetNet, Ad hoc ...) (voir figure 3.5). Dans DRL, les machines intelligentes peuvent apprendre de leurs actions de la même manière que les humains apprennent de l'expérience. Inhérent à ce type d'apprentissage automatique est qu'un agent est récompensé ou pénalisé en fonction de ses actions, c-à-d, les actions qui les amènent au résultat visé sont récompensées (renforcées). En DRL, l'agent doit observer l'environnement pour augmenter la récompense globale dans un Processus de Décision de Markov (MDP) (Markov Decision Process), par exemple, un agent doit prendre des décisions locales et autonomes, par exemple l'accès au spectre, la sélection du débit de données, le contrôle de la puissance de transmission et l'association des stations de base, pour atteindre les objectifs de différents réseaux y compris, par exemple, la maximisation du débit et la conservation d'énergie dans des environnements incertains et stochastiques. Le DRL est une approche DL qui apprend l'environnement, ce qu'il faut faire et comment décrire les circonstances du comportement actuel, afin d'optimiser une récompense numérique. En général, l'agent n'a pas d'information des actions à entreprendre, mais grâce à une série d'essais il doit apprendre quelle action donne une récompense maximale.



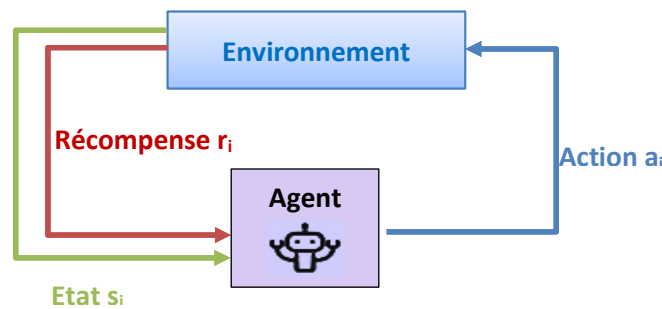
**Figure 3.5 :** Taux de travaux connexes pour (a) différents réseaux et (b) différents problèmes dans les réseaux [50]

Les MDP fournissent dans des situation spécifiques un contexte précis pour modéliser la prise de décision, où les conséquences sont comparativement aléatoires, et le décideur contrôle en partie les conséquences.

Le Q-Learning et les MDP partiellement observable POMDP (Partially Observable MDP) sont des exemples de DRL. Le QL pourrait être appelé pour découvrir une stratégie optimale pour agir à partir de n'importe quel MDP fini, en particulier lorsque l'environnement est inconnu. POMDP est une technique RL qui ne suit pas de modèle et peut également être combinée avec des modèles MDP, ces paradigmes créent des outils pour soutenir la prise de décision dans l'IoT (où les appareils IoT sont les agents et le réseau sans fil est l'environnement) [50].

Les problèmes de communication et de réseaux que les DRL tentent de résoudre sont : l'accès au réseau, le contrôle du débit de données, la mise en cache sans fil, le déchargement des données, la sécurité du réseau, la préservation de la connectivité, le routage du trafic et la collecte de données. En outre, les pourcentages d'œuvres liées aux DRL pour différents réseaux et différents problèmes dans les réseaux sont indiqués sur la Figure 3.5(a) et (b), respectivement [50].

L'apprentissage par renforcement peut être décrit par le diagramme de la Figure 3.6.



**Figure 3.6** : diagramme de l'apprentissage par renforcement

Les définitions essentielles de l'apprentissage par renforcement profond sont [52] :

**Agent** : Un mécanisme logiciel / matériel qui prend certaines mesures en fonction de son interaction avec l'environnement. L'algorithme est l'agent.

**Action** : une action est l'un des mouvements (transitions) possibles que l'agent peut effectuer. Une action est presque explicite, mais il convient de noter que les agents choisissent généralement parmi une liste d'actions discrètes et possibles.

**État** : Un état est une situation concrète et immédiate dans laquelle se trouve l'agent, c'est-à-dire un lieu et un moment spécifiques, une configuration instantanée qui met l'agent en relation avec d'autres choses importantes.

**Environnement** : Le monde à travers lequel l'agent se déplace et qui répond à l'agent. L'environnement prend l'état actuel et l'action de l'agent comme entrée, et renvoie comme sortie la récompense de l'agent et son état suivant.

**Récompense** : une récompense est le retour d'information par lequel nous mesurons le succès ou l'échec des actions d'un agent dans un état donné.

### 3.2.2.2. Processus de décision de Markov

La majorité des problèmes de l'apprentissage par renforcement peuvent être présentés sous forme d'un MDP. Tous les états de MDP satisfont la propriété de « Markov », se référant au fait que l'avenir ne dépend que de l'état actuel, pas de l'histoire des états [56]. Les MDP sont des modèles mathématiques utilisés pour modéliser une variété de problèmes d'optimisation et sont résolus via la programmation dynamique (DP) et l'apprentissage par renforcement (RL) [51].

Un MDP peut être défini comme un quadruplet  $MPD = (S, A, R, T)$ [52], où

**S** est l'espace d'états dans lequel évolue le processus, composé de tous les états possibles du système.

**A** est l'ensemble de toutes les actions possibles qui contrôlent la dynamique de l'état.

**R** est la fonction de récompense  $S \times A \rightarrow R$ , obtenue par l'exécution de l'action  $a$  à l'état  $s$ .

**T** est la fonction de transition de probabilité  $S \times T \times S \rightarrow [0, 1]$  où :

$T[s'|s, a]$  Est la probabilité de transition de l'état courant  $s \in S$  à  $s' \in S$  après la prise de l'action  $a \in A$ .

Le problème dans les MDP est de trouver une "*politique*" notée  $\pi$  pour le décideur, qui est une stratégie employée par l'agent pour déterminer l'action suivante en fonction de l'état actuel. Une politique mappe l'ensemble des états  $S$  sur l'ensemble des actions  $A$ , les actions qui offrent la plus haute récompense.

$\pi : S \rightarrow A$ . La politique peut être déterministe ou stochastique. Une fois qu'un processus de décision de Markov est combiné avec une politique de cette manière, il fixe l'action pour chaque état, et la combinaison résultante se comporte comme un processus de récompense de Markov.

Spécifiquement, supposons que l'environnement est un système stochastique à temps discret et à états finis. Soit  $S = (s_1, s_2, \dots, s_n)$  l'espace d'états  $S$  et  $A = (a_1, a_2, \dots, a_m)$  l'espace d'actions  $A$ . Si à l'épisode  $k$ , l'agent RL se trouve sur un état  $s_k \in S$ , il choisit l'action  $a_k \in A$  en fonction de la politique  $\pi$  afin d'interagir avec son environnement. Ensuite il passe à l'état  $s_{k+1} \in S$  avec une probabilité  $P(a|ss')$  et fournit à l'agent une récompense de retour notée  $r_k(s, a)$  [47]. Le processus est ensuite répété. L'objectif de l'agent RL est de maximiser la récompense ou la valeur d'état actualisée attendue à plus ou moins long terme, en tenant de moins en moins compte du futur, connue sous le nom de retour  $R_t$  :

$$R_t = \sum_{i=t}^{\infty} R(s_i, a_i, s_{i+1}) \quad (3.7)$$

Où  $a_i = \pi(s_i)$  et  $\gamma$  est le facteur d'actualisation et satisfait  $0 \leq \gamma \leq 1$ .

Maintenant que nous avons défini l'objectif et ce qui définit un comportement optimal, nous pouvons commencer à introduire des méthodes pour trouver ce comportement :

Pour les cas dans lesquels la dynamique du MDP et les fonctions de récompense sont connues (basées sur un modèle), ce problème peut être résolu en utilisant la programmation dynamique.

Pour les cas dans lesquels la dynamique du MDP ou la fonction de récompense n'est pas disponible (sans modèle), il existe un besoin de méthodes d'apprentissage par renforcement. Les méthodes d'apprentissage par renforcement peuvent être divisées en deux catégories principales : Algorithmes basés sur la fonction de valeur et algorithmes basés sur des politiques [51].

### 3.2.2.3. Algorithmes basés sur la fonction de valeur

La fonction de valeur  $V^\pi(s)$  attribue des valeurs aux états. La valeur de l'état  $s$  sous la politique  $\pi$  est par définition le rendement à long terme attendu  $R_t$ , avec la remise attendue à partir de l'état  $s$  suivant la politique  $\pi$ , par opposition à la récompense à court terme et définie par l'équation de Bellman comme :

$$V^\pi(s) = R(s, \pi(s)) + \gamma \sum_{s' \in S} P(\pi(s) | ss') V^\pi(s') \quad (3.8)$$

Où  $R(s, \pi(s)) = E\{r(s, \pi(s))\}$  est la valeur moyenne de la récompense  $r(s, \pi(s))$ .

Cette équation est appelée équation de Bellman de  $V^\pi(s)$  [47]. Les équations de Bellman font référence à un ensemble d'équations qui décomposent la fonction de valeur en récompense immédiate plus les valeurs futures actualisées. Il est évident que nous pouvons aussi calculer l'équation de Bellman de  $Q^\pi$  [52], la valeur  $Q$  est similaire à la valeur  $V$ , sauf qu'elle prend l'action actuelle comme un paramètre supplémentaire. Il se réfère au retour à long terme  $R_t$  d'une action, prenant une action  $a_t$  spécifique, dans le cadre d'une politique spécifique à partir de l'état actuel  $s_t$ .  $Q(s_t, a_t)$  est défini par l'équation suivante :

$$Q^\pi(s, a) = R(s, a) + \gamma \sum_{s' \in S} P(\pi(s) | ss') V^\pi(s') \quad (3.9)$$

## A) Recherche de la valeur optimale

Grâce à l'équation de Bellman, nous avons donc réussi à établir une formule qui nous permet de calculer objectivement la valeur  $V$  associée à un état  $s$  en suivant la Policy  $\pi$ , et cela en calculant récursivement la valeur de tous les états enfants possibles. Trouver le meilleur état

successeur, celui qui nous guidera le plus efficacement à la solution. Nous définissons donc  $V^*(s)$ , la fonction valeur optimale [51].

Pour une politique optimale :  $V^*(s) = \text{supp}_{\pi} V^{\pi}(s)$ , l'utilité d'une politique optimale peut s'écrire de manière récursive sous la forme de l'équation de Bellman :

$$V^*(s) = R(s, \pi(s)) + \gamma \max_{a \in A(s)} \sum_{s' \in S} P(\pi(s)|ss') V^*(s') \quad (3.10)$$

Et ce  $\forall a \in A$ , l'ensemble des actions disponibles. Et pour la politique, il suffit, simplement, une fois  $V^*$  calculé, de choisir à chaque état l'action qui mène au  $V^*(s_0)$  le plus grand (Politique gourmande "greedy Policy"). De la même manière, nous pouvons définir  $Q^*(s, a)$ , la fonction qui nous donnera le retour attendu en effectuant l'action  $a$  à l'état  $s$  et en suivant par la suite la politique optimale (détaillée dans la section suivante).

Où  $\gamma \in [0, 1]$  est le facteur d'actualisation (Discount factor) qui est un multiplicateur. Les récompenses futures, telles que découvertes par l'agent, sont multipliées par ce facteur afin d'atténuer l'effet cumulatif de ces récompenses sur le choix d'action actuel de l'agent. C'est au cœur de RL, c'est-à-dire réduire progressivement la valeur des futures récompenses pour donner plus de poids aux actions récentes.

## B) Algorithme RL

Les techniques basées sur un modèle et sans modèle sont deux types différents de techniques RL. L'agent basé sur un modèle construit un modèle de transition de l'environnement et des plans (par exemple par anticipation) à l'aide du modèle. Dans les algorithmes basés sur un modèle, s'il y a suffisamment d'échantillons de chaque paramètre d'état, les estimations de la récompense et de la probabilité de transition convergent vers le MDP, la fonction de valeur et la politique corrects. Cependant, l'obtention d'un nombre suffisant d'échantillons reste un défi à résoudre. Un inconvénient de la méthode basée sur un modèle est que le modèle MDP réel doit être créé lorsque la taille de l'état est trop grande. En outre, une approche RL basée sur des politiques recherche directement la politique optimale  $\pi^*$  qui est la politique qui obtient une récompense future maximale. Alors que, l'approche RL basée sur la valeur estime la fonction de valeur optimale  $Q^*(s, a)$ , qui est la valeur maximale pouvant être atteinte sous n'importe quelle politique.

Les agents des algorithmes sans modèle tels que le Q-Learning [53] peuvent apprendre directement l'action et la politique. En outre, une approche d'apprentissage par renforcement basée sur les politiques recherche directement la politique optimale  $\pi^*$ , qui est la politique qui obtient une récompense future maximale. De plus, l'approche RL basée sur la valeur estime la fonction de valeur optimale  $Q^*(s, a)$  qui est la valeur maximale réalisable sous n'importe quelle

politique. Différence temporelle (TD)<sup>14</sup>, State-Action-Reward-State-Action (SARSA)<sup>15</sup> [51] et Q-Learning sont quelques exemples d'algorithmes RL sans modèle fonctionnant sur la base de la différence temporelle.

### C) Q-Learning

Dans cette thèse, nous nous intéresserons plus particulièrement à la méthode Q-Learning. Le Q-Learning [53, 57] fait partie des méthodes d'apprentissage par renforcement RL sans modèle. C'est aussi un algorithme hors politique en raison de l'apprentissage avec une stratégie gloutonne  $a = \max_a(Q(s, a))$  et il garantit une exploration suffisante des états en raison du suivi d'une distribution de comportement. Cette distribution de comportement est choisie en utilisant un algorithme  $\epsilon$ -gourmand ( $\epsilon$ -greedy), qui sera expliqué dans la section suivante. Il s'agit d'apprendre par l'expérience les actions à effectuer en fonction de l'état actuel. Q-Learning tente de déterminer la politique, en l'absence de la fonction de transition de probabilités et de la fonction de la récompense. Dans Q-Learning, les politiques et la fonction valeur sont représentées par une matrice indexée par des paires état-action. Formellement, pour chaque état  $s$  et action  $a$  [55], Une autre grandeur  $Q(s, a)$ , peut être utilisée, obtenue à partir de  $V(s)$ . Elle va fournir la "qualité" de l'action  $a$  effectuée dans l'état  $s$ . A long terme,  $Q(s, a)$  sous la politique  $\pi$  converge vers la valeur  $Q$  optimale définie comme étant  $Q^*$  et s'écrit donc :

Soit  $Q^*(s, a) = \max_{\pi} Q^{\pi}(s, a)$  étant la fonction d'action optimale sous la politique  $\pi$ , la fonction de valeur optimale peut être réécrite grâce à  $Q^*$  (équation de Bellman) comme :

$$V^*(s) = \max_{a \in A(s)} (Q^*(s, a)) \quad (3.11)$$

La politique optimale  $\pi$  peut être exprimé :

$$\pi^* = \operatorname{argmax}(Q^*(s, a)) \quad a \in A(s) \quad (3.12)$$

La récompense future totale est  $R_t$  écrite comme suit :

$$R_t = \sum_{i=t}^n r_i \quad (3.13)$$

Où  $r_i$  est la récompense de chaque état.

Etant donné que l'environnement est stochastique, il existe une incertitude quant aux augmentations futures au cours des étapes de temps d'exécution. En conséquence, le calcul de

<sup>14</sup> Proposé par Richard Sutton en 1988 (pour plus de détails : Sutton, R.S. (1988). Learning to predict by the method of temporal differences. Machine Learning, 3, p 9-44)

<sup>15</sup> SARSA suit exactement une politique  $\epsilon$ -gourmande (on-policy)

$R_t$  n'est pas possible et, par conséquent, la récompense future actualisée est calculée au lieu de  $R_t$  comme suit :

$$R_t = r_t + \gamma r_{t+1} + \dots + \gamma^{n-t} r_n \quad (3.14)$$

La fonction  $Q$  est la récompense future maximale actualisée dans l'état  $s$  et l'action  $a$ . Par conséquent, la fonction  $Q$  peut être exprimée comme la somme de la récompense  $r$  et de la récompense future maximale pour l'état suivant  $s'$  et l'action  $a'$ , et réécrire d'une façon récurrente :

$$Q^*(s, a) = R(s, a) + \gamma \sum_{s' \in S} P(a|ss') \max_{a' \in A(s')} Q^*(s', a') \quad (3.15)$$

Comme la récompense peut facilement être illimitée, un facteur d'actualisation  $\gamma$  ( $0 < \gamma < 1$ ), est utilisé. Dans l'algorithme QL,  $Q(s, a)$  estime la récompense comme la récompense cumulative (approximation universelle de  $Q^*$  comme point fixe de l'équation précédente) et est mise à jour comme suit:

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha \Delta Q_t(s_t, a_t) \quad (3.16)$$

où  $\alpha \in [0, 1]$  est le taux d'apprentissage. L'apprentissage se produit rapidement sur la base d'une estimation d'apprentissage améliorée  $\Delta Q_t(s_t, a_t)$  appelée la différence temporelle [57], exprimée comme :

$$\Delta Q_t(s_t, a_t) = \left[ r(s_t, a_t) + \gamma \max_{a' \in A(s_t)} Q_t(s_{t+1}, a'_t) - Q_t(s_t, a_t) \right] \quad (3.17)$$

#### **D) Exploration vs exploitation (Algorithme $\epsilon$ -greedy)**

L'exploration permet à un agent d'améliorer ses connaissances actuelles sur chaque action, ce qui, entraînera des avantages à long terme. L'amélioration de la précision des valeurs d'action estimées permet à un agent de prendre des décisions plus éclairées à l'avenir.

L'exploitation, quant à elle, choisit l'action gourmande pour obtenir le plus de récompense en exploitant les estimations actuelles de la valeur d'action de l'agent.

Lorsqu'un agent explore, il obtient des estimations plus précises des valeurs d'action. Et quand il exploite, il pourrait obtenir plus de récompense. Il ne peut cependant pas choisir de faire les deux simultanément, cette situation est appelée le dilemme exploration-exploitation.

Pour obtenir la récompense future maximale, les agents doivent trouver un équilibre entre exploitation (gourmande) et exploration (algorithmes  $\epsilon$ -gourmands) [56].

La sélection d'action  $\epsilon$ -Greedy se fait comme suit

$$a_t = \begin{cases} \text{action aléatoire } a & \text{avec probabilité } \varepsilon \\ \max_t Q(a) & \text{avec probabilité } 1 - \varepsilon \end{cases} \quad (3.18)$$

Le pseudo code de l'algorithme Q-Learning est donné par l'algorithme 3.1 suivant :

**Algorithme 3.1: Q-learning**

Préconditions :  $\alpha$  : taux d'apprentissage,  $\gamma$  : facteur d'actualisation,  $\varepsilon$  : probabilité

Initialisation : matrice  $Q(s, a) = 0 \quad \forall s, a$

**Pour** chaque épisode **faire**

Initialiser  $s$

**Répéter**

Choisir  $a$  de  $s$  selon une politique dérivée de  $Q$  (e.g.  $\varepsilon$ -greedy)

$a \leftarrow \pi(s)$

Exécuter l'action  $a$  et observer la récompense  $r$  et l'état suivant  $s'$

$\Delta \leftarrow r(s, a) + \gamma \max_{a' \in A(s')} Q(s', a') - Q(s, a) \quad //$  Erreur de différence temporelle

$Q(s, a) = (1 - \alpha)Q(s, a) + \alpha\Delta \quad //$  Mise à jour de  $Q$

$s \leftarrow s' \quad //$  Mise à jour de l'état

**Jusqu'à**  $s$  est terminal

**Fin pour**

### 3.2.2.4. Algorithmes basés sur des politiques

Les méthodes de recherche de politiques fonctionnent selon une approche plus directe. Au lieu de trouver la valeur de chaque état possible puis de dériver la politique optimale, des méthodes basées sur des politiques (par exemple la programmation dynamique, ou l'algorithme SARSA)) cherchent à trouver directement une politique  $\pi$  qui maximise le rendement attendu. Le cœur des algorithmes basés sur des politiques (on policy) fait référence à des algorithmes qui tentent d'évaluer et d'améliorer la politique de comportement, c'est-à-dire la politique utilisée pour prendre des décisions.

### 3.2.2.5. Quelques Solution intelligentes

Après avoir introduit les DRL et leurs principes, ci-dessous, quelques solutions basées sur l'apprentissage par renforcement.

#### A. Cognitive Backoff Mechanism (CB)

Proposé par Nurullah Shahin et al., 2019 [72], dans CB, un mécanisme basé sur les chaînes de Markov à temps discret (DTMC : Discrete-Time Markov Chain Model), qui détermine de manière adaptative la CW pour éviter efficacement les collisions avec un débit élevé et un faible délai de transmission. La probabilité de collision conditionnelle  $p_{ck}$  mesurée et le nombre de stage backoff  $i$  déterminent la CW.



$$CW_{nouv} = \begin{cases} \min \left( (2^i (CW_{min} + 1)^{(pck+1)} - 1), CW_{max} \right) & \text{En cas de collision} \\ CW_{min} & \text{En cas de succes} \end{cases} \quad (3.19)$$

D'après les simulations, CB offre une meilleure performance que BEB en termes de débit, délai et équité.

### B. Centralized Contention Window Optimization with DRL (CCOD)

Proposé par Witold Wydmanski et al., 2020[73], une méthode d'application de DRL dans les réseaux multicanaux qui prend en charge deux algorithmes entraînés, ayant la tâche d'optimiser le débit de saturation des réseaux 802.11ax en prédisant correctement les valeurs CW, tout en maintenant un coût de calcul bas. L'agent est localisé dans le AP, qui a une vue globale du réseau, pour calculer d'une manière centralisée les informations de DRL exigées.

### C. intelligent QI-based Resource Allocation (iQRA)

Proposé par Rashid Ali et al., 2018 [74], Un mécanisme intelligent d'allocation des ressources basé sur QL était proposé pour l'accès aux canaux de la couche MAC dans les WLAN denses. Les simulations ont montré que par rapport aux protocoles MAC non intelligents conventionnels, Les performances du iQRA donnent un meilleur débit, délai d'accès au canal et d'équité.

## 3.3. Conclusion

Dans ce chapitre nous avons réalisé une synthèse sur quelques travaux d'évaluation de la norme IEEE 802.11. Cette synthèse nous a permis de constater que la majorité des travaux se concentrent sur l'amélioration du mécanisme CSMA/CA utilisé pour éviter les collisions et exactement le calcul de la fenêtre de contention dans l'algorithme BEB. Nous avons vu qu'il existe des solutions basées sur la modification des calculs de la CW, d'une manière linéaire, exponentielle, logarithmique..., tandis que les travaux récents introduisent des techniques intelligentes pour la sélection adéquate de la CW. Cette synthèse qui nous a permis d'orienter et d'extraire les problèmes du standard, comme elle nous a motivé pour introduire une amélioration au mécanisme BEB, dans le but de remédier à ces problèmes et d'aboutir à des résultats satisfaisants afin d'améliorer les performances du protocole MAC 802.11 dans différents environnements. Une précieuse attention est donnée à l'algorithme de Q-Learning, qui est un algorithme DRL sans modèle basé valeur, où, sa structure fonctionnelle principale était décrite, et dans une sous-section ultérieure (dans le chapitre 6), l'utilisation de Q-Learning comme futur paradigme pour le mécanisme de backoff dans une fonction de coordination distribuée (DCF) est suggéré pour les réseaux MANETs.

La prochaine partie sera consacré aux différentes contributions de cette thèse.

## O-BEB UN NOUVEL ALGORITHME ADAPTATIF ET EQUITABLE

### 4.1. Introduction

Le développement technologique qu'a connu le monde d'aujourd'hui a touché tous les domaines, particulièrement le secteur de la communication qui connaît une évolution remarquable par l'apparition de la technologie sans fil. Cette dernière permet l'établissement des communications sans fil dans des environnements mobiles qui offrent une grande flexibilité d'emploi. En particulier, Les réseaux mobiles ad-hoc (MANET) sont un nouveau paradigme des réseaux mobiles. Ils se construisent par l'interconnexion de différentes entités mobiles inconnues et ne reposent sur aucune infrastructure fixe ou un contrôle centralisé. Aucune STA de base n'entre en jeu, ils sont des réseaux locaux utilisant le médium radio. Cependant le canal radio est limité, la bande passante disponible pour la communication dans un tel réseau est également limitée, les nœuds sont mobiles et se déplacent librement, ce qui rend la topologie du réseau fréquemment changeante et de manière imprévisible. L'accès à ce médium partagé devrait être contrôlé de telle manière que tous les nœuds reçoivent une part équitable de la quantité de bande passante disponible, et que la bande passante soit utilisée efficacement sans épuiser les ressources du réseau, ce contrôle est effectué par le protocole MAC (Medium Access Control), et en raison de tous les obstacles posés par les MANETs, il est difficile d'obtenir un contrôle d'accès au médium simple, efficace, équitable et économe en énergie.

Le protocole MAC de l'IEEE 802.11 est la norme pour les réseaux locaux sans fil. Cependant, ce protocole n'a pas été conçu pour les réseaux multi-sauts. Bien qu'il puisse prendre en charge certaines architectures de réseau Ad Hoc, il n'est pas destiné à prendre en charge le réseau Ad Hoc mobile, dans lequel la connectivité multi-sauts est l'une des fonctionnalités les plus importantes.

L'efficacité des protocoles MAC peut être mesurée à l'aide de deux paramètres : la probabilité de collision et l'équité dans l'attribution du canal aux nœuds concurrents. L'IEEE 802.11 tente de résoudre le problème de collision en suivant l'algorithme de Binary Exponential Backoff (BEB). La procédure de backoff utilisée dans presque tous les protocoles d'accès au médium sans fil (basés sur la contention) est essentiellement empruntée à l'Ethernet câblé où la nature non uniforme du média n'existe pas. Ainsi, la procédure de Binary Exponential Backoff qui fournit un accès équitable au médium dans les réseaux câblés devient la cause d'injustice dans les réseaux sans fil, elle souffre à la fois de l'équité et de l'efficacité. L'injustice du MAC a un impact considérable sur le comportement des protocoles de couche supérieure et des applications utilisant le réseau. Puisque le problème de l'équité est profondément enraciné dans la couche MAC, il est raisonnable de conclure qu'elle peut être résolue en modifiant le code MAC de manière appropriée.

A cet effet, Dans ce chapitre, nous présentons un nouvel algorithme de backoff, dénommé Backoff Exponentiel Binaire Oppose (**O-BEB : Opposite Binary Exponential Backoff**) qui peut surmonter les limites des solutions MAC existantes. Dans l'algorithme O-BEB, une augmentation linéaire ou une réduction selon l'indice de la suite de Fibonacci sont effectués dans le cas de succès ou de collision de transmission respectivement. Les résultats des différentes simulations montrent que l'algorithme proposé atteint, lorsqu'il est utilisé dans un environnement mobile ad hoc, un meilleur débit, une énergie résiduelle et un débit de paquets délivrés plus élevé que le BEB utilisé dans la norme IEEE802.11 DCF, l'I-BEB, et le N-BEB.

Dans la suite du chapitre, nous présentons en détail dans la section suivante, le nouvel algorithme O-BEB. La section qui suit présente les différentes simulations et résultats obtenus. Enfin, la section 5 conclut ce chapitre.

## 4.2. Backoff Exponentiel Binaire Opposé (O-BEB)

Ce travail se concentre sur l'environnement réseau non-AP (sans point d'accès) similaire au mode ad hoc décrit dans la norme IEEE 802.11. Dans notre proposition Opposite Binary Exponential Backoff (O-BEB), contrairement au BEB qui réinitialise la CW à  $CW_{min}$  après chaque transmission réussie, et la double après une collision, l'O-BEB fonctionne de la manière suivante :

### A. En cas de succès

En cas de transmission réussie, l'algorithme O-BEB augmente linéairement la CW à une certaine valeur de seuil (threshold) égal à  $round((\log_2(CW_{min}) * 2))$ . Réinitialisée la CW à  $CW_{min}$  après un succès de transmission donne la possibilité au nœud de cette transmission réussie sur des nœuds concurrents d'accéder au médium dans la contention suivante d'une part, et d'autre part, elle lui donne aussi la même chance que de nouveaux nœuds joints. Bien au

contraire, l'augmentation linéaire dans la valeur de CW dans l'O-BEB donne de meilleures chances aux nouveaux nœuds et empêche le nœud de transmission réussite de monopoliser le canal. Mais même un ajout d'un seuil, sous la forme d'une valeur log et malgré sa petite taille, s'accumulera, conduisant à une grande valeur pour la CW et ainsi le nœud perdra toute contention pour accéder au médium, ce qui conduira à de mauvaises performances. Pour cela, un compteur *succ\_counter* pour les transmissions réussies doit être comparé au seuil *succ\_Threshold* relatif, si le compteur dépasse le seuil, la CW sera affecté à sa valeur minimum  $CW_{min}$ .

**Fonction** Reset\_log\_CW()

**Debut**

succ\_counter = succ\_counter + 1

**Si** (succ\_counter > succ\_Threshold ) **alors**

succ\_counter = 1

cw\_nouvelle =  $CW_{Min}$

**Sinon**

cw\_nouvelle = cw\_actuelle +  $(\log_2(CW_{Min}) * 2)$

**Si** (cw\_nouvelle >  $CW_{Max}$ ) **alors**

cw\_nouvelle =  $CW_{Max}$

**Fin si**

**Fin si**

Retourne cw\_nouvelle

**Fin**

## B. En cas de collision

En cas de collision, O-BEB incrémente la CW comme dans BEB, pour un faild\_counter nombre d'échec de transmission. Après que ce compteur faild\_counter dépasse un seuil de collision prédéfini Faild\_Threshold, la CW diminuera linéairement pour chaque collision en soustrayant l'indice de Fibonacci correspondant au terme de Fibonacci le plus proche à la CW courante.

L'incrémement de la CW des nœuds pour chaque collision jusqu'à ce qu'elle atteigne le  $CW_{max}$  conduira à une injustice dans l'accès au médium et la famine de ces nœuds. Au lieu d'aller jusqu'à  $CW_{max}$ , l'approche O-BEB arrête d'incrémenter la CW avant d'atteindre sa

dernière phase  $CW_{max}$ , et commence à diminuer la CW par des petites valeurs. Cette approche maintiendra la valeur de la CW dans la phase intermédiaire pendant une période relativement longue et, par conséquent, en conjonction avec le schéma d'augmentation de la CW plus haut, en théorie, l'O-BEB donnera à tous les nœuds une chance équitable d'accéder au médium sans sacrifier le débit autant que le BEB. L'algorithme d'O-BEB est bien illustré dans La Figure 4.1.

**Fonction** Incr\_fibo\_CW()

**Debut**

fail\_counter = fail\_counter + 1;

**Si** (fail\_counter < fail\_Threshold) **alors**

    cw\_nouvelle = 2 \* CW\_actuelle + 1

**Si** (cw\_nouvelle > CWMax) **alors**

        cw\_nouvelle = CW<sub>Max</sub>

**Finsi**

**Sinon**

    fibo\_index\_counter = fibo\_index\_counter + 1

    cw\_nouvelle = cw\_actuelle - fibo\_index\_counter

**Si** (cw\_nouvelle < CW<sub>Min</sub>)

        fibo\_index\_counter = 0

        fail\_counter = 1

        cw\_nouvelle = CW<sub>Min</sub>

**Finsi**

**Finsi**

**Finsi**

Retourne cw\_nouvelle

**Fin**

L'équation suivante correspond au calcul de la fenêtre de contention CW par O-BEB :

$$CW_{nouv} = \begin{cases} \begin{cases} CW_{cour} + \text{round}((\log_2(CW_{min}) * 2)) & \text{Si } succ\_counter < succ\_Threshold \\ CW_{min} & \text{sinon} \end{cases} & \text{En cas de succès} \\ \begin{cases} CW_{cour} - \text{Fibonacci}_{index}(CW_{cour}) & \text{Si } fail\_counter > fail\_threshold \\ CW_{cour} \times 2 & \text{sinon} \end{cases} & \text{En cas de collision} \end{cases} \quad (4.1)$$

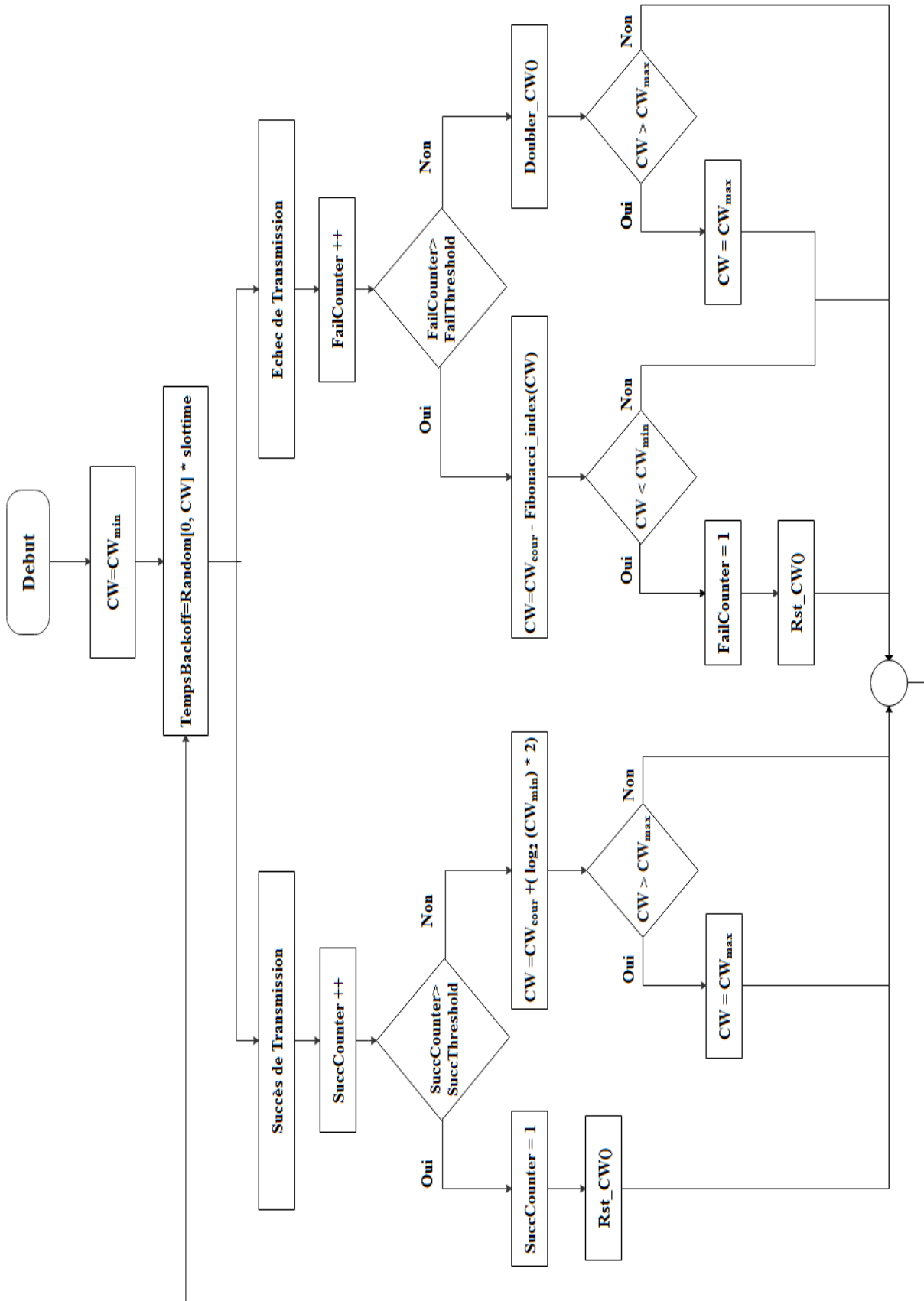


Figure 4.1 : Organigramme de l'O-BEB

### 4.3. L'implémentation de l'I-BEB et de N-BEB

Pour pouvoir bien comparer notre proposition, nous avons implémenter les deux approches mentionnées en chapitre 3 section F (l'I-BEB [69]) et section J (N-BEB [70]) respectivement. Nous avons voulu analyser leurs comportements dans un environnement Ad Hoc simulé, et comparer ses performances avec celles de BEB de l'IEEE 802.11. Les différents résultats et analyses sont présentés dans la section suivante.

### 4.4. Simulation et analyse des résultats

NS2.35 [80] est utilisé pour évaluer l'algorithme proposé. Le protocole MAC standard a été modifié pour mettre en œuvre notre proposition. Les fichiers concernés par la modification sont principalement mac802\_11.h et mac802\_11.cc. Les protocoles I-BEB, et N-BEB sont également implémentés sous NS2, pour pouvoir les comparer avec le nôtre et le standard BEB. Plusieurs scénarios dans un environnement hautement mobile, sans RTS / CTS, ont été réalisés pour tester soigneusement l'algorithme, utilisant les paramètres résumés dans le tableau 4.1.

**Tableau 4.1 : Paramètres de simulation**

Paramètre	Valeur affectée
Surface	1000 x 800 m <sup>2</sup>
Temps de simulation	100 s
Nombre de nœuds mobiles	25, 50, 75, 100
Modèle de mobilité	Random walk point <sup>16</sup>
Vitesse moyenne	8,5 m/s
Temps de pause	0
Protocole de routage	OLSR <sup>17</sup>
Protocole de transport	UDP <sup>18</sup>
Type de trafic	CBR <sup>19</sup>
Taille de paquet	512 B
Temps de slot	20 μs
DIFS	40 μs
SIFS	10 μs

<sup>16</sup> C'est le modèle le plus simple pour décrire un mouvement aléatoire. Dans le développement et l'évaluation par simulation des algorithmes de communication il est nécessaire de reproduire les mouvements des utilisateurs de façon réaliste, Il est la représentation en termes de mobilité sur deux dimensions du mouvement d'une particule suspendue dans un liquide (mouvement Brownien, formalisé par Albert Einstein en 1905).

<sup>17</sup> Optimized Link State Routing : est un protocole de routage destiné aux réseaux maillés, sans fil ou mobiles.

<sup>18</sup> User Datagram Protocol, protocole de la couche de transport du modèle OSI, qui permet la transmission de données sous forme de datagrammes

<sup>19</sup> Constant Bit Rate (Débit binaire constant), décrit le mode d'encodage d'une piste vidéo ou audio

$CW_{min}$	31
$CW_{max}$	1023
Energie initiale	20j
Succ_Threshold	9
Faild_Threshold	5

La notion de qualité de service (QoS) dans les réseaux se traduit en termes de paramètres (appelés aussi métriques ou critères) mesurant les performances de transmissions des données de la source vers la destination. Les paramètres qui s'avèrent fondamentaux dans les réseaux ad hoc, et qu'ils doivent être prise en compte lors du déploiement des mécanismes pour l'amélioration de qualité de service dans ce type de réseaux et les plus importants sont utilisés pour mesurer les performances de l'O-BEB parmi lesquels :

### a) Débit moyen

Le débit (Throughput) est le rapport entre la quantité totale de données qui atteint le récepteur depuis la source et le temps pris par le récepteur pour recevoir le dernier paquet [76]. Il est représenté en paquets par seconde ou en bits par seconde.

### b) Délai moyen de bout en bout

Le délai moyen de bout en bout est le temps moyen nécessaire aux paquets pour traverser le réseau de la source vers la destination, il peut être affecté par le retard de transmission, la mise dans les files d'attente de mémoire tampon, les échanges de contrôle MAC et les délais produits par les activités de routage [77].

$$MoyE2Edelai = \frac{\sum TempsR(i) - TempsS(i)}{\sum paquet\_reçu} \quad (4.2)$$

Ou TempsR(i) : est le moment ou le paquet de données i est reçu par la destination.

TempsS(i) : est le moment ou le paquet de donnée i est émis par la source

### c) Energie résiduelle moyenne

L'énergie résiduelle moyenne, représente la quantité d'énergie moyenne restante à chaque nœud du réseau, elle est exprimée comme la différence entre l'énergie initiale et l'énergie consommée par un nœud :

$$MoyResEnergie = \frac{\sum_{i=1}^N Energie_{initiale} - Energie_{consommée}}{N} \quad (4.3)$$



### d) Taux de paquets délivrés

Le taux de paquets délivrés (PDF : Packet Delivery Fraction) est le nombre de paquets de données reçus par les nœuds de destination par le nombre total de paquets de données transmis par le nœud source [78].

$$PDF = \frac{\text{Nombre\_Total\_de\_paquets\_reçus}}{\text{Nombre\_Total\_de\_paquets\_emis}} \quad (4.4)$$

### e) Équité

L'indice d'équité (FI : Fairness Index) entre les nœuds concurrents dans les réseaux Ad Hoc est une mesure importante du système. Plusieurs mesures d'indices d'équité ont été proposées dans la littérature technique, l'équation bien connue est celle de l'indice de Jain [79]. Un indice d'équité est un nombre réel qui mesure le degré de partage équitable ou injuste de la ressource entre les nœuds concurrents.

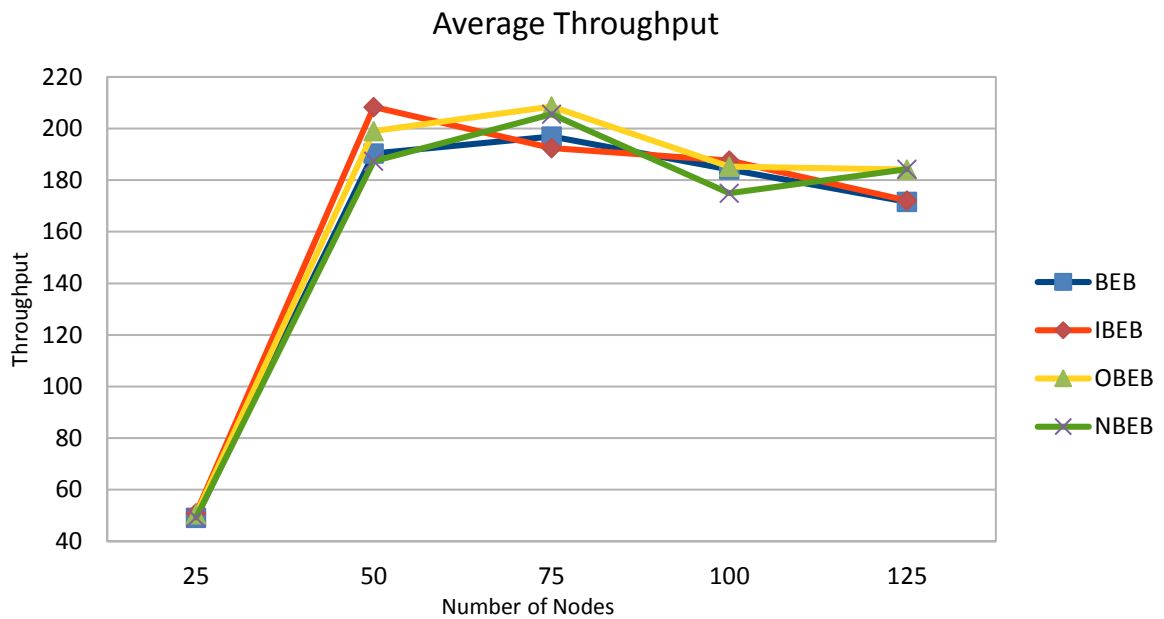
$$FI = \frac{(\sum_{i=1}^n x_i)^2}{N \times \sum_{i=1}^n x_i^2} \quad (4.6)$$

Où  $x_i$  est la portion d'accès du nœud  $i$  parmi  $N$  nœuds concurrents

**Tableau 4.2 :** Index d'équité

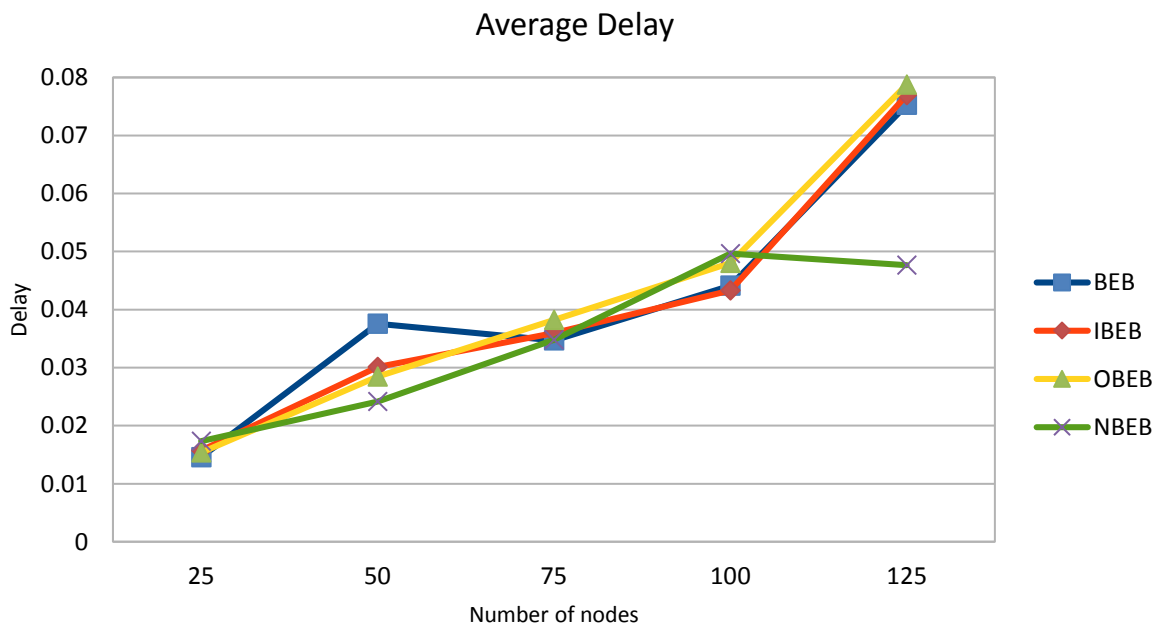
Protocoles \ Nœuds	Nœuds				
	25	50	75	100	125
BEB	0.996	0.994	0.996	0.995	0,996
I-BEB	0.997	0.995	0.997	0.996	0,996
N-BEB	0.995	0.996	0.997	0.996	0.996
O-BEB	0.997	0.996	0.997	0.996	0,996

L'indice d'équité des quatre protocoles est presque identique (100%)(comme le montre le tableau 4.2), avec un léger avantage pour O-BEB et N-BEB, comme le montre la figure 4.5.



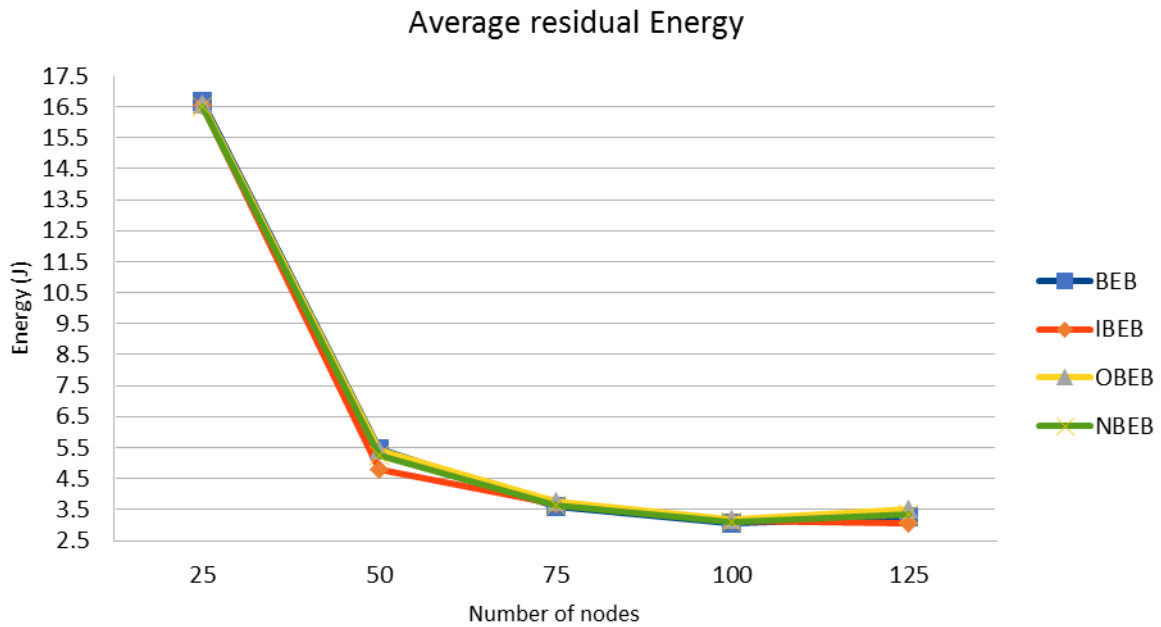
**Figure 4.2** : Débit moyen

La figure 4.2 montre que l'algorithme O-BEB atteint presque le meilleur débit à partir de 75 nœuds, où il est meilleur de 11,63 kbps, 16,06 kbps et 2,99 kbps par rapport à BEB, I-BEB et N-BEB respectivement. Et dans le cas d'un nombre maximum de nœuds, O-BEB est meilleur de 12,59 kbps et 12,03 kbps par rapport respectivement à BEB et I-BEB, et presque identique au débit donné par N-BEB.

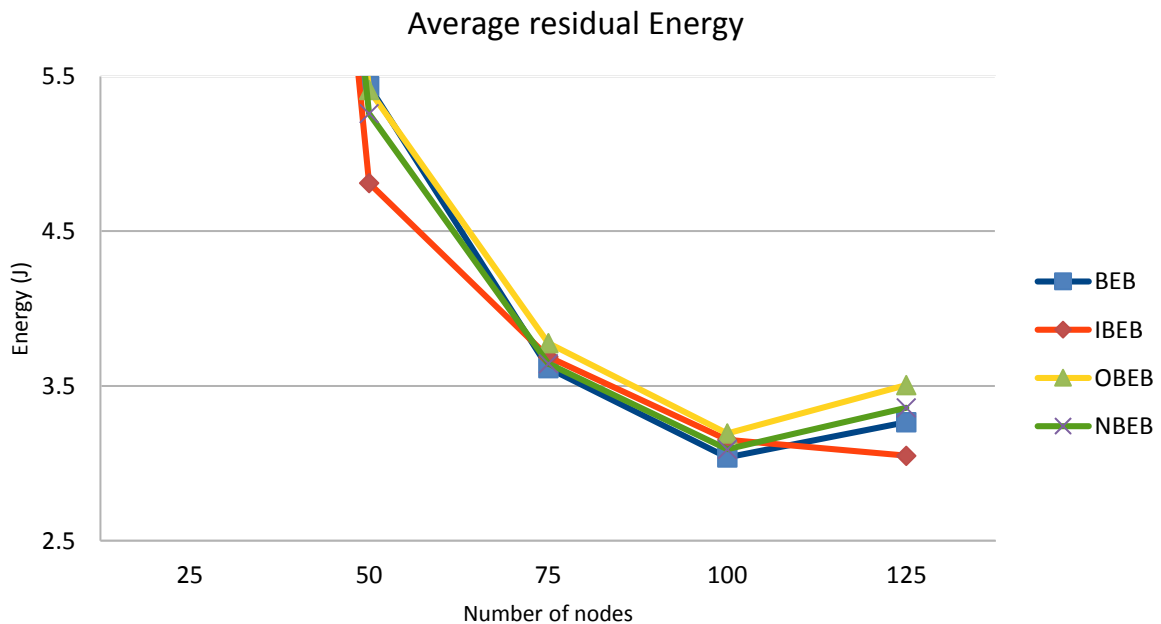


**Figure 4.3** : Délai moyen

La figure 4.3 illustre le délai moyen de bout en bout des quatre protocoles BEB, I-BEB, N-BEB et O-BEB sur un nombre croissant de nœuds, de 25 jusqu'à 125 nœuds, chaque nœud ayant une vitesse de déplacement pouvant atteindre 8,5 mètres par seconde. La figure montre que l'O-BEB donne un délai plus important par rapport aux autres pour 75 et 125 nœuds, par contre il donne le meilleur délai lorsque le nombre de nœuds est de 25.



**Figure 4.4 :** Energie résiduelle moyenne



**Figure 4.4 bis:** Energie résiduelle moyenne (zoom à partir de 50 nœuds)

La figure 4.4 montre l'énergie résiduelle sur un nombre croissant de nœuds jusqu'à 125 nœuds, la figure montre qu'O-BEB fournit toujours la meilleure énergie résiduelle à partir de 50 nœuds (comme le montre la figure 4.4 bis). O-BEB fournit une énergie plus élevée de 4,23%, 2,30% et 3,48% par rapport à BEB, I-BEB et N-BEB respectivement pour 75 nœuds. Et dans le cas de 125 nœuds, il dépasse BEB, I-BEB et N-BEB de 6,83%, 13,01% et 4,19% respectivement.

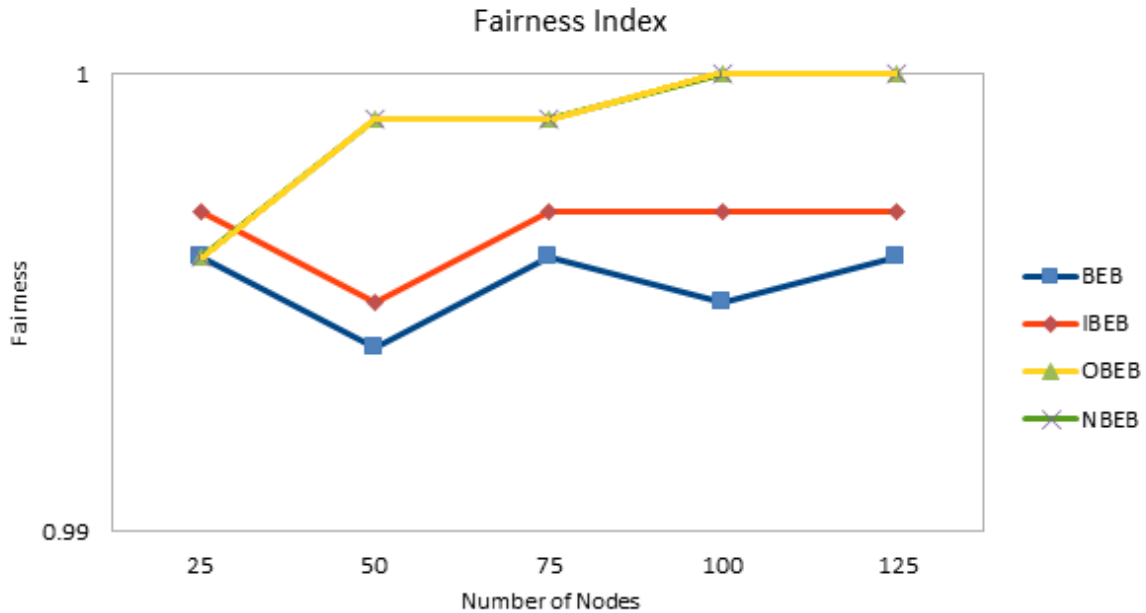


Figure 4.5 : Indice d'équité

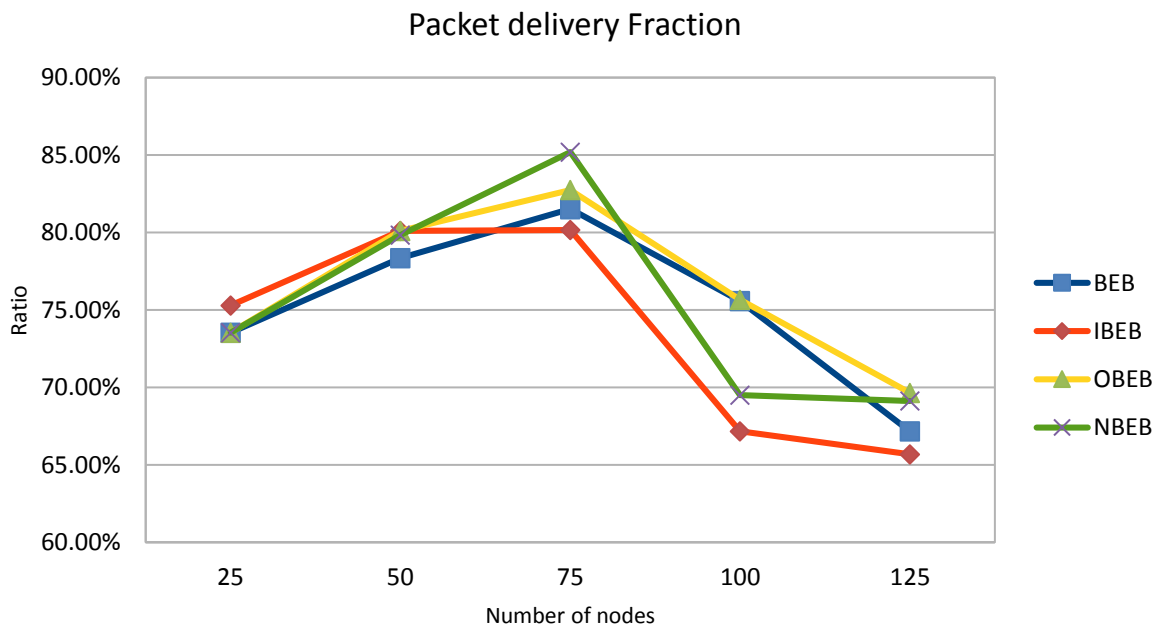
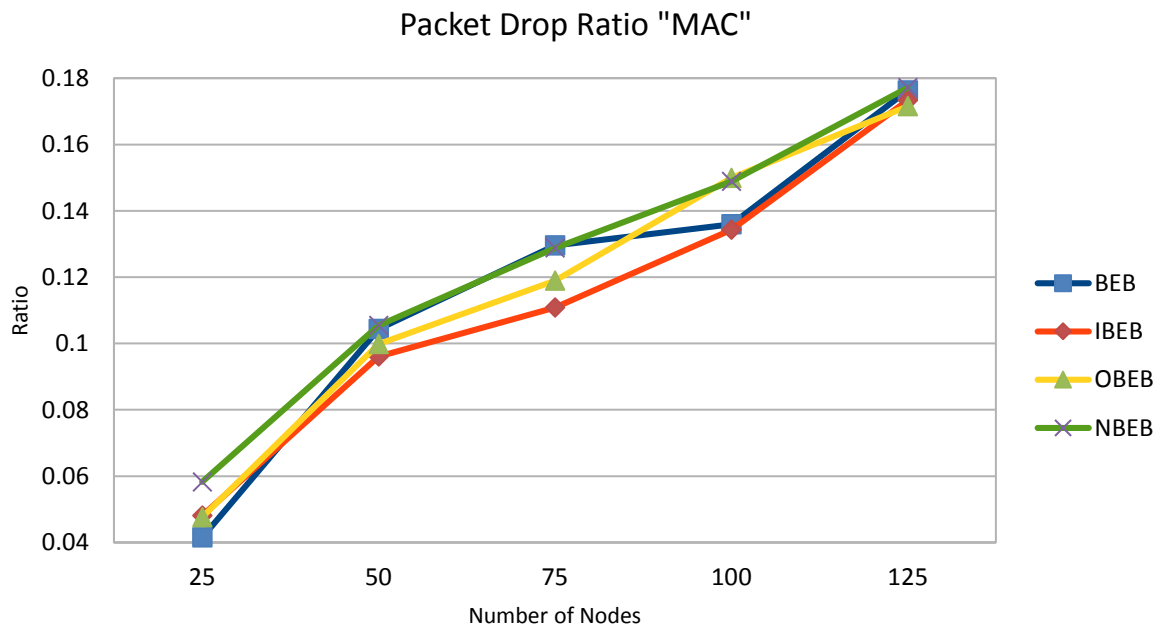


Figure 4.6 : Taux de paquets délivrés

La figure 4.6 illustre le taux de livraison des paquets des protocoles BEB, I-BEB, N-BEB et O-BEB. Le graphique montre qu'O-BEB fournit plus de paquets que les autres pour 50, 100 et 125 nœuds. Par exemple, dans le cas de réseau de 100 nœuds, l'O-BEB est presque identique au BEB et dépasse respectivement l'I-BEB et le N-BEB de 7,11% et 4,76%. Et dans un nombre maximum de nœuds (125 nœuds), l'O-BEB dépasse BEB de 2,49%, I-BEB de 3,97% et N-BEB de 0,53%.

La figure 4.7 montre le taux de perte au niveau de la couche MAC, où ce taux est presque identique (<1%) pour les quatre protocoles quel que soit le nombre de nœuds



**Figure 4.7 :** Taux de perte

## 4.5. Conclusion

Dans ce chapitre, nous avons introduit notre première contribution, qui est un nouvel algorithme nommé O-BEB qui adopte un schéma inversé pour calculer la valeur CW après une transmission réussie ou une collision, que celui utilisée dans le BEB de la norme IEEE 802.11 DCF. L'incrémentation linéaire de la valeur CW pour un nœud après une transmission réussie donne aux nouveaux nœuds plus de chances de transmettre leurs données, obtenant ainsi plus de débit. La diminution de la valeur CW après une collision empêche le nœud de transmission réussi de monopoliser le canal. Le protocole O-BEB proposé offre des performances améliorées en termes de débit, de taux de livraison de paquets, de consommation d'énergie et donne un meilleur indice d'équité, mais avec un coût supplémentaire pour les délais de bout en bout. Les différentes simulations réalisées sous NS2, avec un nombre variable de nœuds mobiles allant

de 25 à 125 nœuds, ont montré qu'O-BEB est plus performant que le BEB utilisé dans le standard IEEE 802.11 DCF, l'Improved BEB et le New BEB.

Pour améliorer l'efficacité de cette approche en termes de délai de bout-en-bout, nous allons essayer d'en faire une deuxième solution (M-BEB), qui sera l'objectif du chapitre suivant.

## **M-BEB : REDUCTION EXPONENTIELLE BINAIRE AMELIOREE ET EQUITABLE**

### **5.5. Introduction**

Les MANET sont principalement utilisés pour des applications d'urgence telles que les opérations de secours en cas de catastrophe, la surveillance de l'habitat animal et les opérations militaires où un réseau de communication est nécessaire immédiatement après un incident majeur ou une nécessité temporaire, comme une conférence ou un séminaire dans un nouvel endroit où il n'y a pas une infrastructure réseau et une solution alternative sont nécessaires. L'émergence des applications en temps réel et l'utilisation généralisée des appareils sans fil et mobiles ont généré le besoin de fournir une prise en charge de la qualité de service (QoS) dans les environnements de réseaux sans fil et mobiles. Il est très important de déterminer la qualité de service du réseau qui dépend principalement du réseau. Dans MANET, il existe plusieurs paramètres qui influencent l'amélioration de la qualité de service du réseau tels que le débit, le délai de bout en bout, le taux de livraison de paquets...etc.

Un paramètre de qualité de service important est le délai moyen de bout en bout, reflétant le temps moyen entre l'envoi d'un paquet et sa réception. Le délai de bout-en-bout comprend le retard de routage et d'autres retards divers, tels que, le retard de transmission dû aux échanges de contrôle au niveau de la couche MAC, le retard de propagation et le retard de la file d'attente.

La solution O-BEB présentée dans le chapitre précédent, comparée avec le standard BEB, L'I-BEB et le N-BEB offre des performances améliorées en termes de débit, PDF, consommation d'énergie, et d'équité, mais elle donne en délai de bout en bout supérieur de celui donné par les trois autres algorithmes à partir de 75 nœuds (figure 4.3).

Nous présentons dans ce chapitre, un nouvel algorithme de backoff modifié (M-BEB : Modified Binary Exponential Backoff), qui améliore le délai de bout en bout par rapport à l'algorithme O-BEB, tout en préservant les autres paramètres de QoS.

Le reste du chapitre est organisé comme suit : Section 2 présente le nouvel algorithme M-BEB. La section 3 présente les différentes simulations et résultats obtenus. Enfin, la section 3 conclut ce chapitre.

## 5.2. Backoff Exponentiel Binaire modifié (M-BEB)

Dans la nouvelle proposition **Modified Binary Exponential Backoff Algorithm (M-BEB)**, nous focalisons notre travail sur les environnements fortement mobiles ad hoc. Alors que l'algorithme BEB réinitialise la fenêtre de contention  $CW$  à  $CW_{\min}$  après chaque transmission réussite et double la fenêtre ( $CW \times 2$ ) en cas de collision, M-BEB adapte un schéma linéaire basé sur deux valeurs de seuil  $S1$  et  $S2$ , et un compteur de transmissions réussies  $C$ , pour calculer la valeur de la fenêtre de contention dans les deux cas succès ou collision. L'organigramme de l'algorithme M-BEB est bien illustré dans La Figure 5.1.

### A. En cas de succès

La proposition M-BEB calcule la  $CW$  en cas de transmission réussie comme suit :

Un compteur  $C$  compte le nombre de transmissions réussies d'une STA, tant que le compteur  $C$  est inférieur à seuil  $S1$ , et exactement comme BEB, M-BEB réinitialise la  $CW$  à  $CW_{\min}$ . Si le compteur  $C$  est entre seuil  $S1$  et seuil  $S2$ , la  $CW$  augmente linéairement et reçoit la valeur courante de  $CW$  additionner avec la valeur  $(CW_{\min}+1)/4$ , si la nouvelle valeur de  $CW$  dépasse  $CW_{\max}$ , alors  $CW_{\max}$  est affecté à  $CW$ . Si le compteur  $C$  atteint le seuil  $S2$ , il est réinitialisé à 1.

Réinitialiser  $CW$  à  $CW_{\min}$  après un succès de transmission donne au nœud de cette transmission réussie la même possibilité sur les autres nœuds concurrents d'accéder au canal dans la contention suivante, et c'est la même chance donnée aux nouveaux nœuds joints. Bien au contraire, l'augmentation linéaire dans la valeur de  $CW$  dans le M-BEB donne de meilleures chances aux nouveaux nœuds et empêche le nœud de transmission réussite de monopoliser le canal.

La fonction `Reset_CW()` ci-dessous, représente le calcul de la  $CW$  dans le cas de transmission réussie de la proposition M-BEB



**Fonction Reset\_CW()**

**Début**

**Si** (C < 7) **alors**

C = C+1

CW<sub>nouvelle</sub> = CW<sub>min</sub>

**Sinon**

CW<sub>nouvelle</sub> = CW<sub>courante</sub> + (CW<sub>min</sub>+1) /4

C = C + 1

**Si** (CW<sub>nouvelle</sub> > CW<sub>max</sub>) **alors**

CW<sub>nouvelle</sub> = CW<sub>max</sub>

**Fin si**

**Si** (C = 12) **alors**

C = 1

**Fin si**

**Fin si**

**Retourner** CW<sub>nouvelle</sub>

**Fin**

**B. En cas de collision**

En cas de collision **M-BEB** incrémente la CW comme dans **BEB**, avec une petite réduction d'une constante égale à seuil S1. Le calcul de la CW dans le cas de collision dans la proposition M-BEB, est donnée dans la fonction incre\_CW() ci-après.

**Fonction increm\_CW()**

**Début**

CW<sub>nouvelle</sub> = 2 \* CW<sub>courante</sub> - S1

**Si** (CW<sub>nouvelle</sub> > CW<sub>max</sub>) **alors**

CW<sub>nouvelle</sub> = CW<sub>max</sub>

**Fin si**

**Retourner** CW<sub>nouvelle</sub>

**Fin**

L'équation suivante correspond au calcul de la fenêtre de contention par le M-BEB :

$$CW_{\text{nouv}} = \begin{cases} CW_{\text{min}} & \text{si } C < \text{seuil1} \\ CW_{\text{cour}} + \frac{CW_{\text{min}} + 1}{4} & \text{si } \text{seuil1} \leq C < \text{seuil2} \\ 2 * CW_{\text{cour}} - S1 & \text{En cas de collision} \end{cases} \quad \begin{matrix} \text{En cas de succès} \\ \text{En cas de collision} \end{matrix} \quad (5.1)$$

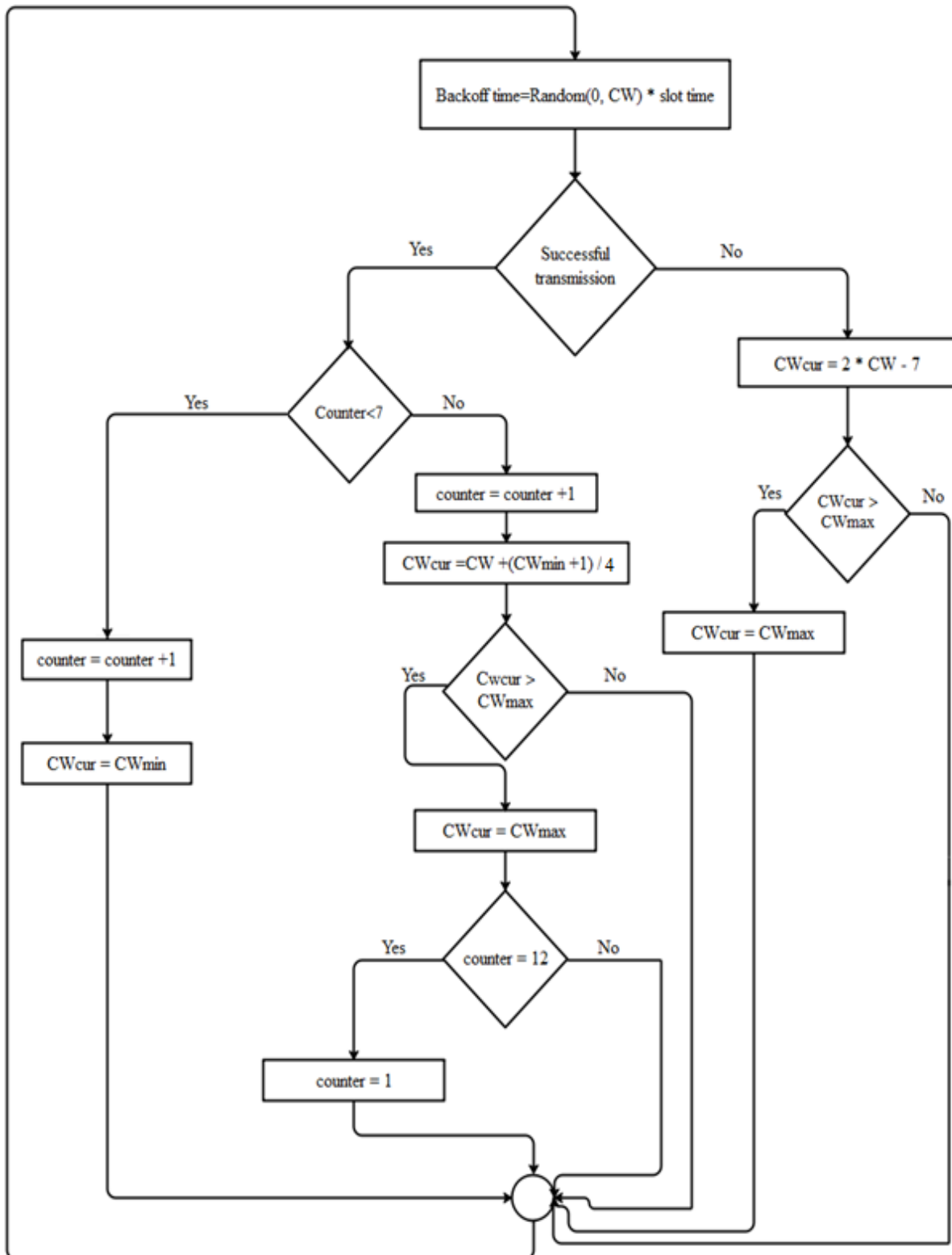


Figure 5.1 : Organigramme de M-BEB

### 5.3. Simulation et analyse des résultats

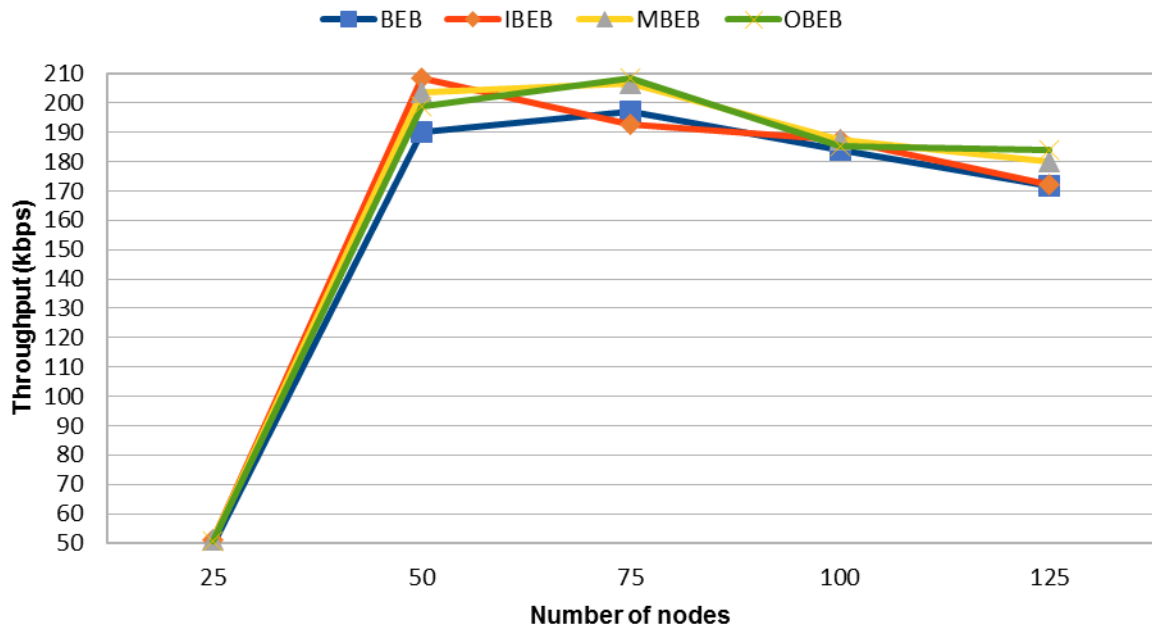
Pour évaluer l’algorithme proposé, nous avons utilisé NS2.35 [80]. Les modifications étaient apportées au protocole standard MAC pour implémenter notre proposition. Les fichiers concernés par la modification sont principalement mac802\_11.h et mac802\_11.cc, le protocole I-BEB [69] et O-BEB [75] étaient aussi implémenté pour pouvoir les comparer avec le M-BEB et le BEB standard. Plusieurs scénarii dans un environnement où tous les nœuds sont mobile avec les paramètres de simulation donnés dans le tableau 5.1, ont été réalisés pour tester attentivement l’algorithme d’une part, et pour fixer les deux seuils seuil1 et seuil2, d’une autre part.

**Tableau 5.1** : Paramètres de simulation

Paramètre	Valeur affectée
Surface	1000 x 800 m <sup>2</sup>
Temps de simulation	100 s
Nombre de nœuds mobiles	25/50/75/100/125
Modèle de mobilité	Random walk point
Vitesse moyenne	8,5 m/s
Protocole de routage	OLSR
Protocole de transport	UDP
Type de trafic	CBR
Taille de paquet	512 B
$CW_{min}/ CW_{max}$	31/1023
Energie initiale	20j
Seuil1/ Seuil2	7/12

Les paramètres suivants sont utilisés pour mesurer la performance de M-BEB.

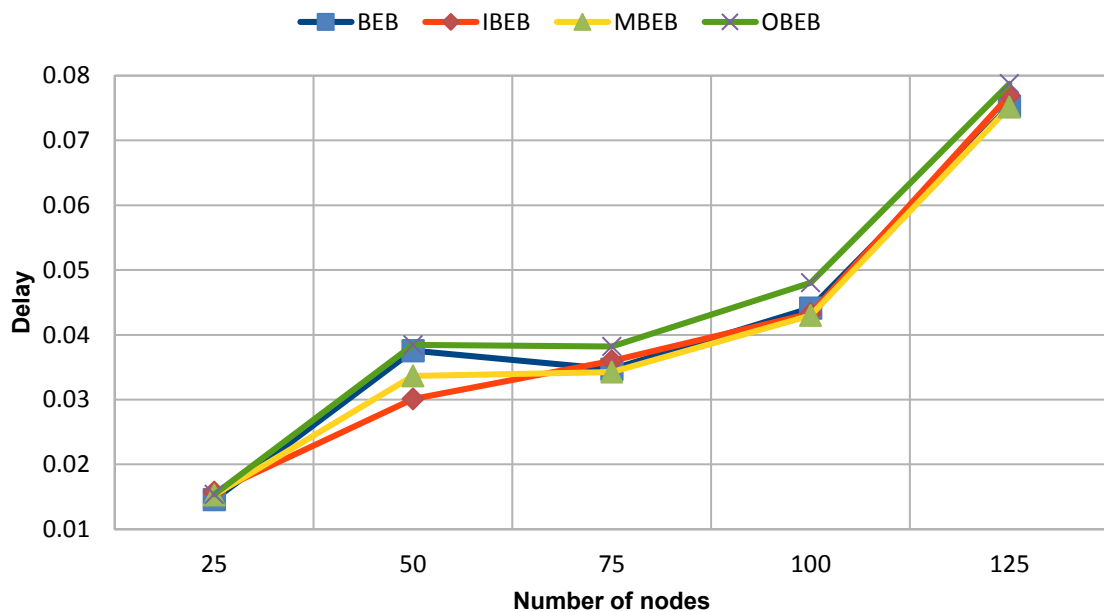
**a) Débit moyen**



**Figure 5.2 : Débit moyen**

La figure 5.2, présente le débit moyen en fonction de nombre de nœuds. Notre proposition M-BEB donne un débit quasi stable et égal aux débit donnés par les trois autres algorithmes à partir de 75 nœuds, le débit donné par M-BEB est meilleur de 10% et 14% que celui donné par BEB et I-BEB respectivement dans le cas de 75 nœuds.

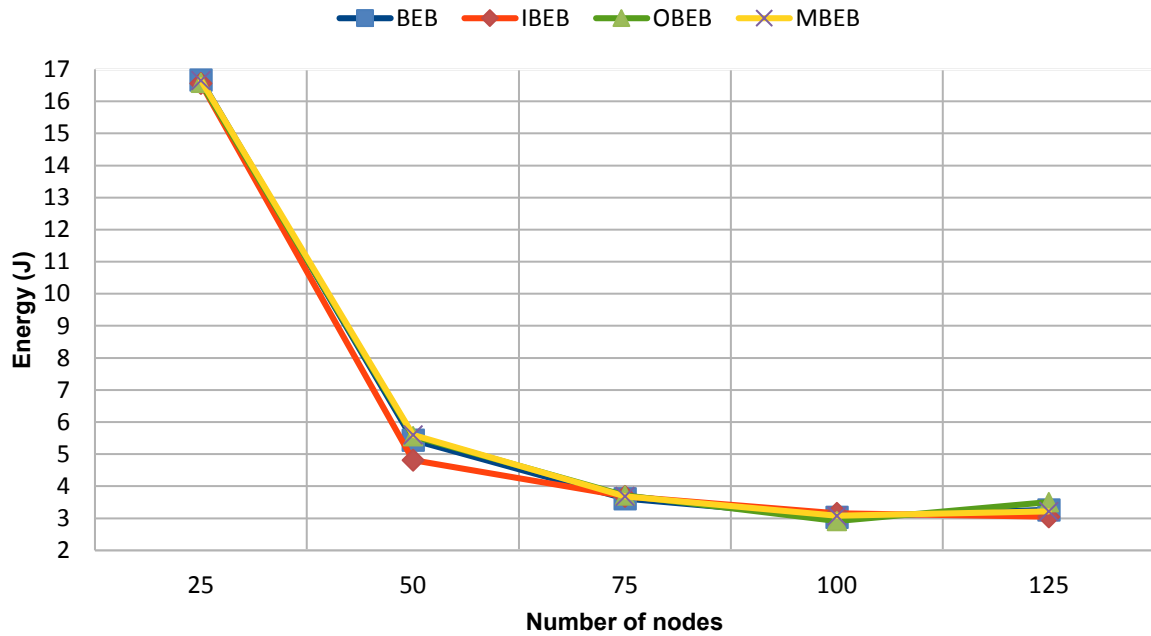
**b) Délai moyen de bout en bout**



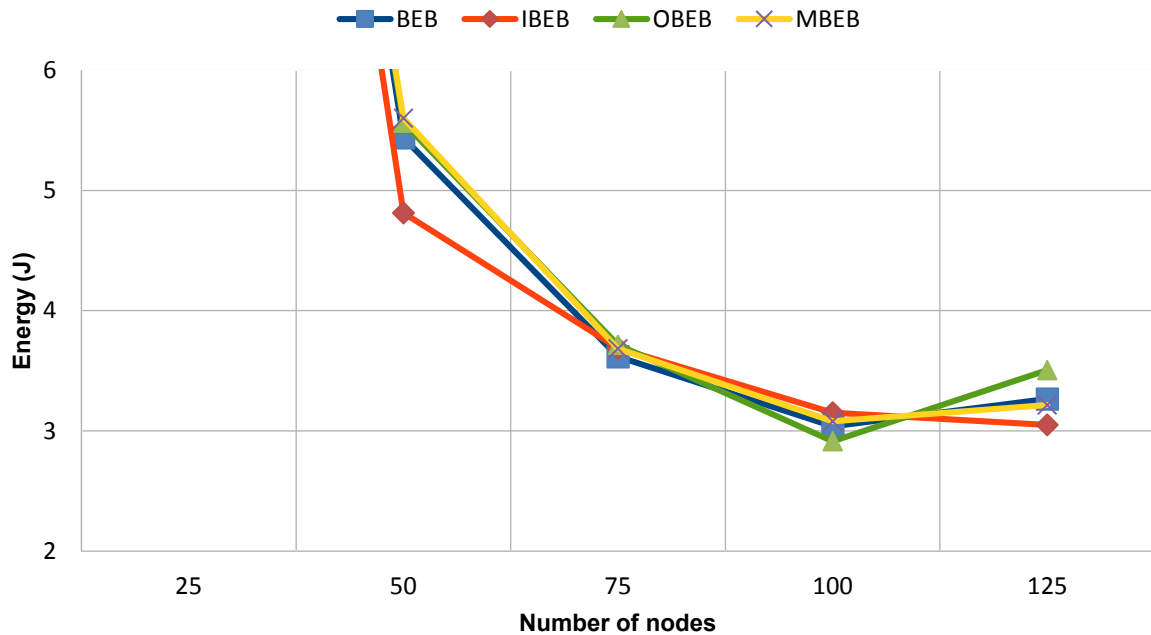
**Figure 5.3 : Délai moyen**

L'O-BEB [16], souffrait de délai de bout en bout supérieur à celui donné par les trois autres protocoles. De la figure 5.3, M-BEB donne le moindre délai moyen de bout en bout à partir de 75 nœuds.

**c) Energie résiduelle moyenne**



**Figure 5.4 : Energie résiduelle moyenne**



**Figure 5.4 bis : Energie résiduelle moyenne (zoom à partir de 50 nœuds)**

La figure 5.4 donne l'énergie résiduelle restante sur un nombre croissant de nœuds de 25 jusqu'à 125. Comme le montre la figure les énergies résiduelles sont presque identiques avec un avantage de 0.5% à 1.65% du M-BEB dans le cas où le nombre de nœuds est égal à 25, 50 et 75 (comme le montre la figure 5.4 bis).

#### d) Taux de paquets délivrés

La figure 5.5, montre le taux de paquets délivrés avec succès, Le M-BEB donne le meilleur taux dans le cas de 50, 75 et 125 nœuds, dépassant ainsi BEB, O-BEB et I-BEB de 2.45%, 0.68% et de 3.81% respectivement dans le cas de 50 nœuds, de 3.29%, 3.35% et de 4.62% respectivement dans le cas de 75 nœuds, et de 2.16% 3.39% et 3.64% respectivement dans le cas de 125 nœuds.

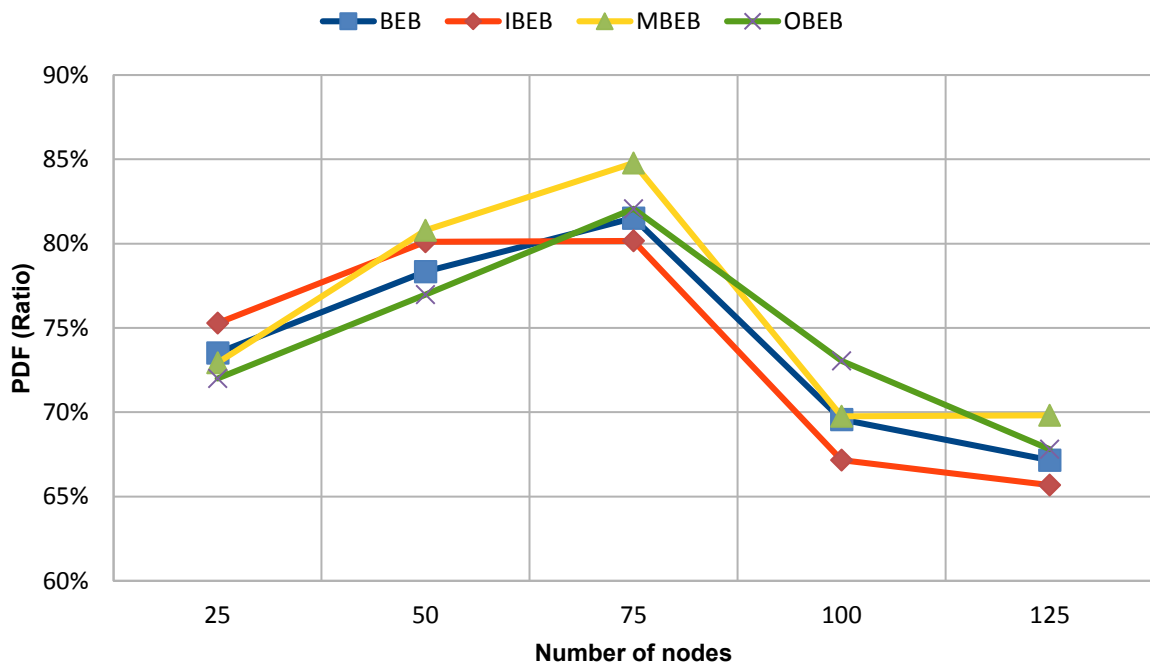


Figure 5.5 : Taux de paquets délivrés

#### e) Indice d'équité

L'indice d'équité donné par les quatre protocoles est presque identique, avec un avantage de M-BEB qui donne un indice égal à 100% à partir de 50 nœuds, comme le montre la figure 5.6.

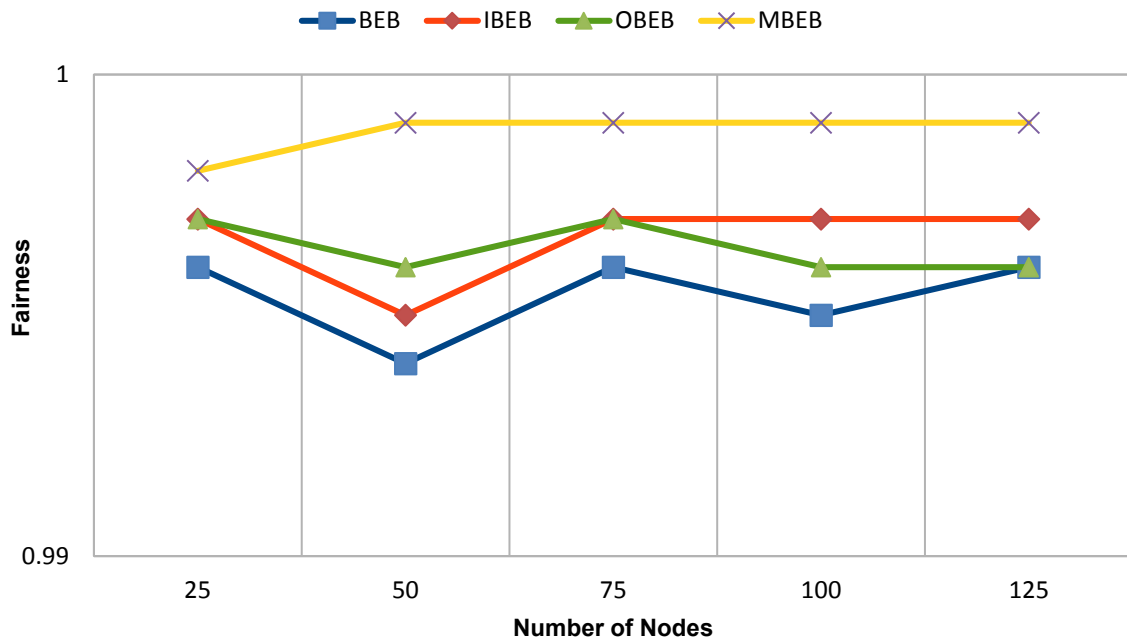


Figure 5.6 : indice d'équité

#### 5.4. Conclusion

Nous avons proposé, dans ce chapitre, un nouvel algorithme M-BEB qui adopte un schéma différent pour calculer la valeur de la CW après une transmission réussie ou une collision que celui utilisé dans le BEB dans la norme IEEE 802.11 DCF. L'incrémentaire linéaire de la valeur de CW, pour un nœud après un nombre de transmissions réussies donne aux nouveaux nœuds plus de chance de transmettre des données. Diminuer la valeur de CW après une collision par le seuil, diminue la monopolisation du canal par le nœud de la transmission réussie. M-BEB, améliore les performances de BEB en termes de débit, de taux de livraison de paquet, de consommation d'énergie, d'équité, et le délai de bout en bout. Les simulations réalisées sous NS2, avec un nombre croissant (de 25 à 125) de nœuds mobiles ont montré que M-BEB est plus performant que le BEB, l'I-BEB et O-BEB, principalement en termes d'équité, de délai de bout en bout, et d'énergie résiduelle.

Dans le prochain chapitre une autre solution basée sur l'intelligence sera exposée.

## MISQ : UN MECANISME INTELLIGENT DE SELECTION DE CW BASE SUR Q-LEARNING

### 6.1. Introduction

Actuellement, l'apprentissage profond (DL) est devenu populaire, grâce aux applications réussies dans différents domaines de recherche, tels que la reconnaissance vocale, le traitement du langage naturel et la vision par ordinateur, cette popularité a conduit au développement de l'utilisation de l'apprentissage profond (DL), dans le domaine des technologies de communications sans fil, pour bénéficier des avantages de l'intelligence artificielle dans ce domaine. L'utilisation des principes DL dans les réseaux sans fil ont une longue histoire et ont atteint de nombreuses réalisations, en particulier dans les couches supérieures de communication, comme dans les cognitive radio networks (CRNs) et pour la gestion des ressources de la couche MAC [27]. Un autre champ de défis pour le DL, c'est la couche physique, par exemple la reconnaissance de modulation, le codage et le décodage...etc. Beaucoup de chercheurs pensent que l'introduction de DL dans les réseaux sans fil peut optimiser les performances de ces réseaux. Une technique de DL qui est l'apprentissage par renforcement profond (DRL) et qui est motivée par la sensibilité comportementaliste et la philosophie de contrôle, dans laquelle un apprenant peut atteindre un objectif en interagissant avec son environnement [57]. Le DRL utilise des modèles d'apprentissage spécifiques, parmi lesquels le processus de décision de Markov (MDP), le MDP partiellement observé (POMDP) et le Q-Learning (QL) [74]. QL est inspiré par psychologie comportementaliste, qui permet de découvrir une stratégie optimale d'action pour tout MDP fini, principalement lorsque l'environnement est inconnu [47].



Le mécanisme DCF de la couche MAC basé sur QL peut être utilisé pour guider les futurs MANETs et pour allouer le canal sans fil plus efficacement. Lorsqu'un périphérique sans fil est déployé dans un nouvel environnement, aucune donnée n'est généralement disponible sur les scénarios historiques. Par conséquent, les algorithmes QL sont le meilleur choix pour observer et apprendre l'environnement pour une sélection optimale des politiques. Dans un MANETs, les collisions au niveau du canal est le problème le plus vital causant une dégradation des performances. Puisque QL trouve des solutions grâce à l'expérience d'interaction et d'apprentissage avec un environnement, il est proposé de l'utiliser pour modéliser la fenêtre de contention optimale dans le CSMA/CA basé DCF. En d'autres termes, un apprenant (STA) contrôle intelligemment la sélection CW à l'aide de l'algorithme basé sur QL. Dans un mécanisme de backoff basé sur DCF, les STA peuvent être équipés avec l'algorithme intelligent QL. La politique est la décision d'un STA de changer la taille CW (c'est-à-dire de prendre une action pour rester à la même étape de backoff sans changement de la CW, ou de augmenter la taille CW et passer à la CW suivante, ou de passer à la CW précédente ou éventuellement de réinitialiser la CW à sa valeur initiale. La fonction de récompense est une fonction clé dans l'apprentissage automatique, elle capture le gain de chaque action effectuée dans n'importe quel état. Par exemple, une récompense peut être la probabilité de collision de canal, le délai d'accès au canal ou le taux de perte de paquets (Packet Loss Ratio) subi par le STA à un état spécifique (étape de backoff).

Dans ce chapitre, Nous appliquons l'apprentissage par renforcement basé sur le Q-Learning pour apprendre le schéma du backoff optimal dans le protocole MAC basé sur la contention (sans point d'accès) afin d'améliorer l'utilisation du canal réseau en tenant compte en plus des collisions, l'état du trafic produit par les STAs (par exemple, trafic élevé si le taux d'occupation de la file d'attente des paquet à transmettre par cette STA dépasse 50%, ou moins élevé sinon), et pour améliorer les performances du CSMA / CA dans les MANETs.

Le reste du chapitre est organisé comme suit : Section 2 discute l'apprentissage par renforcement et le processus de Markov. La section 3 présente le cœur de ce chapitre, l'introduction du Q-Learning pour l'optimisation de la fenêtre de contention. Les différentes simulations et résultats obtenus sont exposés dans la section 4. Enfin, la section 5 conclut ce travail.

## **6.2. Processus de décision de Markov et apprentissage par renforcement**

Un MDP est un outil mathématique qui permet de modéliser des agents RL. Il est défini comme un quadruple  $MDP = (S, A, R, T)$  (détaillé en section 3.2.2.2), où S représente tous les états du système dans lequel le processus fonctionne et A est l'ensemble de toutes les actions possibles qui contrôlent la dynamique de l'état. R est la fonction de récompense  $S \times A \rightarrow R$ ,

obtenue en exécutant l'action  $a$  dans l'état  $s$  et  $T$  est la fonction de probabilité de transition  $S \times T \times S \rightarrow [0,1]$ .

Q-Learning (détaillé en section 3.2.2.3/C) fait partie des méthodes RL sans modèle. Il s'agit d'apprendre par l'expérience quelles actions entreprendre en fonction de l'état actuel. Le Q-Learning tente de déterminer la politique  $\pi$ , en l'absence de la fonction de transition de probabilité et de la fonction de récompense. L'un des principaux intérêts de Q-Learning est qu'il peut être exprimé comme une reformulation de l'algorithme d'itération de la valeur. Or il a été prouvé que ce dernier, pour un MDP fini (avec  $S$  et  $A$  finis), converge vers une politique optimale  $Q^*$ .

Dans la section suivante, nous discutons l'algorithme d'optimisation de la CW concernant le délai d'accès au canal, le débit des paquets livrés et le débit en tant que MDP. Notre objectif est de parvenir à une politique d'accès optimale au canal tout en définissant correctement la fonction de récompense et l'algorithme RL.

### 6.3. Estimation de la fenêtre de contention CW

Les collisions de paquets se produisent lorsque plusieurs STAs accèdent au canal simultanément. Ainsi, l'un des objectifs du protocole MAC basé sur la contention est d'éviter les collisions. La STA doit donc vérifier l'état du canal pour éviter les collisions. Un mécanisme de backoff est nécessaire pour la transmission après un certain délai lorsque le canal est occupé. La détermination de la durée de backoff nécessite une CW, et l'efficacité de l'accès au canal est déterminée par la sélection correcte de la taille de la CW. Bien qu'il soit difficile de concevoir un mécanisme d'accès au canal optimal, il est essentiel de choisir une méthode CW appropriée.

Comme décrit précédemment, Q-Learning a déjà été largement appliqué dans les réseaux sans fil. Dans les sous-sections suivantes, la structure fonctionnelle principale de l'algorithme Q-Learning est décrite, et l'utilisation de Q-Learning comme futur paradigme pour le mécanisme de backoff dans une fonction de coordination distribuée (DCF) est suggérée pour les réseaux MANETs.

Ainsi, nous avons adopté un schéma de sélection CW intelligent, tout en définissant une fonction de récompense à maximiser lorsque le nombre de collisions ou le nombre de paquets générés par la STA est important pour un accès rapide au canal. L'espace d'états contient les tailles CW et les actions déterminent la taille CW dans l'étape  $k$  à partir de l'étape  $k-1$  (détaillé ultérieurement dans les sections B.1 et B.2 respectivement).

Nous discutons la fonction de récompense définie et de l'algorithme utilisé pour permettre à l'agent d'apprendre la politique de sélection de fenêtre de contention optimale. Le délai d'accès

au canal et le débit sont deux problèmes critiques pour les réseaux sans fil. Par conséquent, il ne serait pas souhaitable de minimiser le temps d'accès au prix d'un débit inacceptable.

### A. La fonction récompense

La récompense exprime l'objectif d'un algorithme QL. Dans notre proposition, en fonction des deux situations, l'agent Q-Learning sélectionne une CW qui maximise la récompense dans l'état  $s_i$ . La fonction de récompense comporte deux composants, le premier considère les collisions  $C_k$  effectuées par la STA  $s_i$  dans un épisode  $k$  (On désigne par épisode le temps depuis que le paquet est mis en file d'attente jusqu'à sa transmission réussie ou son rejet après que le nombre de retransmissions  $Rmax$  autorisées soit atteint). Le deuxième composant analyse les informations  $T_k$  du nombre de paquets à transmettre générées par la STA  $s_i$ . Le taux d'occupation de la file d'attente de transmission de la STA  $s_i$  est défini comme.

$$T_k = \left( \frac{NBP_k}{TailleQueue} \right) * 100 \quad (6.1)$$

Où :  $NBP_k$ : représente le nombre de paquets dans la file d'attente de la STA  $s_i$ , en épisode  $k$ ,

$TailleQueue$ : Est la taille de la file d'attente.

Le taux de collisions commises par la STA  $s_i$  dans un épisode  $k$  est le suivant :

$$C_k = \frac{\sum_{i=0}^m col_i}{Rmax} * 100 \quad (6.2)$$

Où :  $col_i$  : est une collision de transmission d'un paquet par la STA  $S_i$ .

$Rmax$  : est le nombre de retransmissions autorisées pour un paquet.

Une distinction entre les deux composants de la récompense est effectuée par une fonction fitness qui calcule le rapport collision/trafic comme suit :

$$Fitness = \alpha T_k + (1 - \alpha) C_k \quad (6.3)$$

Ainsi, le calcul de la fonction de récompense s'effectue comme :

$$R_T = \begin{cases} \frac{C_k}{Rmax} & \text{si } fitness > \text{seuil et } T_k \text{ élevé} & \text{en cas de collision} \\ \frac{NBP_k}{SizeQueue} & \text{si } fitness < \text{seuil et } T_k \text{ élevé} & \text{en cas de succès} \\ 0 & & \text{sinon} \end{cases} \quad (6.4)$$

Où  $Seuil$  est la moyenne de la valeur maximum et de la valeur minimum de la fonction  $Fitness()$  :

$$Seuil = moy(\min(Fitness), \max(Fitness)) \quad (6.5)$$

Ainsi le temps backoff est choisi aléatoirement comme suit :

$$TempsBackoff = random(0, CW_{t-1} - 1) \quad (6.6)$$

## B. L'algorithme MISQ

L'objectif du Q-Learning est de trouver une politique optimale, c'est-à-dire la sélection de tailles optimales de CW pour un état  $s$  qui optimise la récompense  $r$ . Le mécanisme proposé comprend un algorithme distribué d'apprentissage en ligne permettant aux STAs d'apprendre leur politique en temps réel. Plus précisément, toutes les STAs déterminent leur temps d'attente pour chaque épisode. Tous les paramètres utilisés dans des procédures telles que  $s_i$ ,  $a_i$  sont locaux, c'est-à-dire utilisés uniquement au niveau de la STA et non partagés avec d'autres STAs.

### B.1. Les Etats

Dans l'algorithme MISQ proposé, l'ensemble fini des tous les états possibles d'environnement  $S = \{s_0, s_1, s_2, s_3, s_4, s_5, s_6\}$ , où :

$$s_0 = 15, s_1 = 31, s_2 = 63, s_3 = 127, s_4 = 255, s_5 = 511, s_6 = 1023.$$

Chaque état correspond à une taille de la fenêtre de contention correspondant à la norme 802.11 allant de  $CW_{min} = 15$  à  $CW_{max} = 1023$  pour chacune des deux situations  $St$  ( $St=succès$  ou  $St=collision$ ). L'ensemble  $S$  d'états est représenté sur le schéma de la figure 6.1 par les sommets.

### B.2. Les Actions

Les actions sont les différentes décisions basées sur une méthode de sélection d'action dans laquelle, après chaque tentative de transmission indépendamment de la collision ou du succès, une STA peut soit explorer l'environnement en choisissant une action au hasard, soit en suivant une stratégie gourmande. L'algorithme peut transiter vers une paire différente  $(s, a)$  et en tirer de l'expérience (récompense) dans l'approche aléatoire. Sinon, la stratégie gourmande exploite son expérience acquise jusqu'à présent et choisit l'action qui donne la valeur  $Q$  max pour son état actuel, présentée par l'équation (3.18).

L'ensemble des actions possibles dans MISQ est  $A = \{rester, augmenter, diminuer, initialiser\}$ , présenté dans la figure 6.1 par les arêtes  $a_0, a_1, a_2, a_3$  respectivement, où *rester*

détermine l'action de rester sur le même état en cas une collision et un taux élevé d'occupation de la file d'attente. La transition *augmenter* détermine l'action de passer à l'état suivant en cas de collision et de taux faible d'occupation de file d'attente. Cependant, la transition *diminuer* définit l'action de retour à l'état précédent en cas de transmission réussie et de taux faible d'occupation de file d'attente, et *initialiser* signifie l'action de réinitialisation à l'état  $s_0$  en cas de transmission réussie et un taux élevé d'occupation de file d'attente.

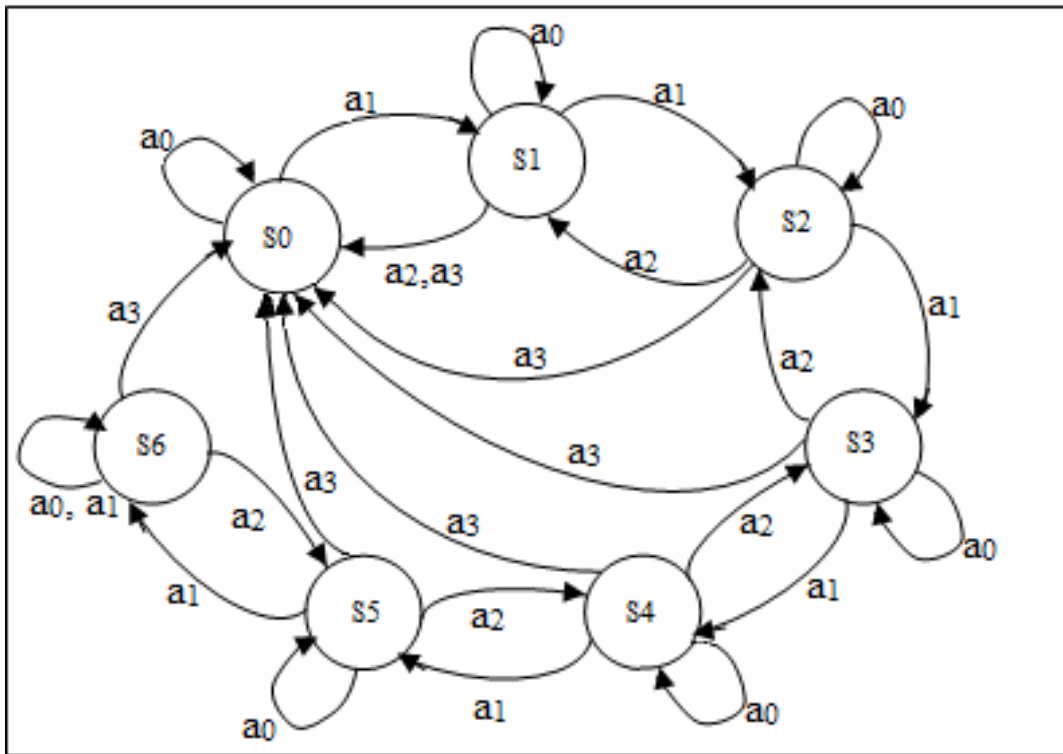


Figure 6.1 : Diagramme états transitions de la proposition

**Remarques :**

- Le trafic d'une STA est considéré élevé si  $T_k$  le pourcentage d'occupation de la file d'attente de la STA  $s_i$  est supérieur à 50% et moins élevé sinon.
- Et que chaque transition dans Le diagramme de la figure 6.1 est récompensée par une récompense  $r_i$  (équation 6.4)

**Représentation Mathématique de l'espace d'états/transition**

Pour représenter mathématiquement le graphe de la figure 6.1, nous utilisons une table de transition  $T(S_i, A_j)$ , ou  $S_i, i \in [0,6]$  et  $A_j, j \in [0,3]$  représentent respectivement les états et les actions. Le tableau se présente comme suit :

$$T(S_i, A_j) = \begin{array}{c|cccc} & \underline{a0} & \underline{a1} & \underline{a2} & \underline{a3} \\ \hline s0 & s0 & s1 & & \\ s1 & s1 & s2 & s0 & s0 \\ s2 & s2 & s3 & s1 & s0 \\ s3 & s3 & s4 & s2 & s0 \\ s4 & s4 & s5 & s3 & s0 \\ s5 & s5 & s6 & s4 & s0 \\ s6 & s6 & s6 & s5 & s0 \end{array}$$

Par exemple :  $T(s3, a3) = T(127, 'initialiser') = s0 = 15$ .

### B.3. Algorithme proposé

L'algorithme MISQ fonctionne essentiellement de la manière suivante : une STA émet un paquet puis reçoit le résultat S (succès ou collision) de cette transmission, déterminé par la réception ou non d'un paquet ACK dans un délai raisonnable. L'agent Q-Learning adapte alors la valeur CW de la STA en tenant compte des deux paramètres  $T_k$  et  $C_k$  (représentant le taux d'occupation de la file d'attente et le nombre de collision de la STA respectivement en épisode  $k$ ) avant d'envoyer le paquet suivant, et le processus est répété. Le fonctionnement de l'algorithme MISQ est illustré à la figure 6.2.

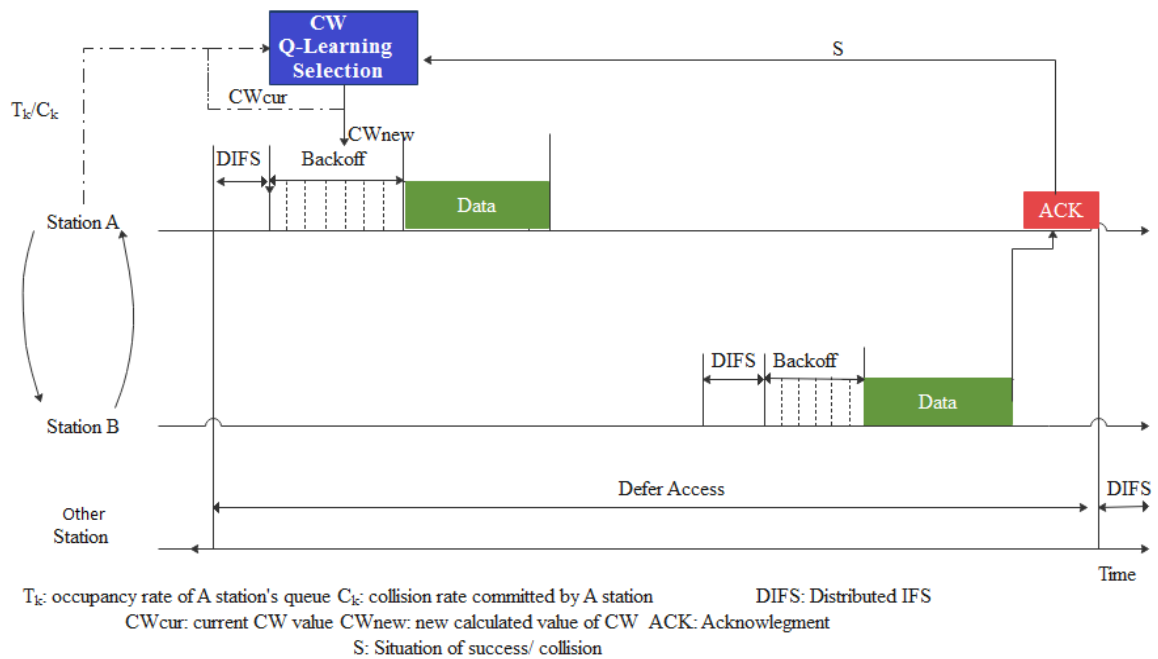


Figure 6.2: Protocole MAC basé sur Q-Learning

**Algorithme : CW Optimisation utilisant MISQ**

**Variable Globale :** Q\_matrix //pour retrouver des valeurs déjà apprises.

**Fonction Q-Learning actions set**

**Début**

**Si** action = stay **alors**

$CW_{t+1} = CW_t$  //pas de changement de CW

**Sinon**

**Si** action = increase and  $CW_t \neq S6$  **alors** //  $S6 = CW_{max}$

$CW_{t+1} = CW_t + 1$

**Sinon**

**Si** action = decrease and  $CW_t \neq S0$  **alors** //  $S0 = CW_{min}$

$CW_{t+1} = CW_t - 1$

**Sinon**

$CW_{t+1} = CW_{min}$

**Fin si**

**Fin si**

**Fin si**

**Retourner**  $CW_{t+1}$

**Fin**

**Fonction  $\epsilon$  –greedy Get-action**

**Début**

Générer  $p = \text{valeur\_aleatoire} \in [0, 1]$

**Si**  $p < \epsilon$  **alors**

Sélectionner aléatoirement une action  $a_i \in A$

**Sinon**

Sélectionner  $a_i^*$  selon l'équation (3.12)

**Fin si**

**Retourner**  $a_i$

**Fin**

**Fonction CW Q-Learning selection**

**Entrées :** nbpk, nbcolk, Situation // (succès ou collision)

**Sortie :** CW // Fenêtre de contention CW Optimisée

**Initialiser** R\_matrix,  $\alpha$ ,  $\gamma$ ,  $\epsilon$ .

**Début**

- Evaluer  $T_k$ ,  $C_k$ , la fonction Fitness et le seuil, équations (6.1), (6.2), (6.3) et (6.5) respectivement.
- Calculer la fonction de récompense selon l'équation (6.4) selon la situation  $S$  succès ou collision.
- Mettre à jour la matrice  $R\_matrix$  pour le couple  $(s, a)$  avec la récompense calculée dans le point précédent.
- Choisir une action  $a$  aléatoire pour explorer/exploiter (fonction  $\epsilon$ -greedy Get-action)
- Calculer  $\Delta Q(s, a)$  selon l'équation (3.17).
- Mettre à jour la matrice  $Q\_matrix$  selon l'équation (3.16) pour le couple  $Q(s, a)$
- Trouver l'action optimale  $a^*$  selon l'équation (3.12).
- Mise à l'échelle de CW selon l'action optimale  $a^*$  (Fonction Q-Learning actions set).
- Retourner CW.
- Choisir une valeur aléatoire de temps backoff selon l'équation (6.6).

**Fin**

L'algorithme ci-dessus montre le mécanisme de sélection de fenêtre de contention basé sur Q-Learning (MISQ) qui sélectionne la valeur CW adéquate en fonction de l'expérience acquise à partir de son interaction avec l'environnement.

La STA  $s$  vérifie périodiquement le canal si son taux d'occupation de file d'attente  $T_k$  est supérieur à 0. Une fois que la STA a des paquets à transmettre, le temps d'attente est effectué en sélectionnant la  $CW_t$ , compte tenu du succès de la situation ou de la collision. La fonction de fitness et le seuil sont évalués, et ainsi, la fonction de récompense est calculée. Une action  $a$  est choisie par la fonction «  $\epsilon$ -greedy Get-action » et exécutée pour obtenir le nouvel état  $s'$  et la récompense  $r$ . Après le calcul de l'estimation d'apprentissage améliorée  $\Delta Q(s, a)$ , la matrice  $Q\_Matrix$  est mise à jour pour le couple  $Q(s, a)$ . Sur la base de la valeur  $a^*$ , la valeur  $Q$  maximale pour toutes les actions de la matrice  $Q\_Matrix$  apprise correspondante, la CW est mise à l'échelle en appelant la Fonction « Q-Learning actions set ». Enfin, une valeur aléatoire est choisie pour le temps d'attente à l'aide de l'équation (6.6).

## 6.4. Evaluation des performances

### A. Configuration de la simulation

Dans nos simulations, nous avons utilisé les paramètres du réseau donnée dans le tableau 6.1, principalement tirée de la norme 802.11 [14]. Les simulations sont effectuées à l'aide de Python [81]. La topologie de la simulation est aléatoire et le nombre de nœuds varie de 5 à 50, et les nœuds émettent sans mécanisme RTS / CTS.



**Tableau 6.1** : Paramètres de simulation

Paramètre	Valeur
Packet payload	8184 bits
MAC Header	272 bits
PHY Header	128 bits
Taille file d'attente	10,20,40
Itération	1500
Slot time ( $\sigma$ )	50 $\mu$ s
DIFS/SIFS	128 $\mu$ s/28 $\mu$ s
Max backoff stage (m)	7
Retry limit	4
$CW_{\min}$	15
$CW_{\max}$	1023

Pour évaluer les performances de notre proposition, nous avons :

1. Programmer le principe de l'algorithme proposé dans [74], que nous avons nommé « QL\_BEB » dans nos figures et qui consiste à utiliser deux actions dans l'espace d'états  $A = \{0, 1\}$  ;  
Où l'action 0 : indique une décrémentation de la CW en cas de succès de transmission et, l'action 1 : indique une incrémentation de la CW en cas d'une collision de transmission.
2. Comparer les résultats des simulations avec le BEB classique, ainsi que l'algorithme simulé en 1.
3. Varier la taille de la file d'attente de chaque STA, taille égale à 10, 20 ou 40 paquets, en supposant que chaque nœud possède au moins un paquet et au maximum taille de la file d'attente paquets à transmettre.
4. Effectuer beaucoup de simulations, afin de déterminer les paramètres de l'algorithme Q-Learning pour fournir de bonnes performances à notre proposition.

## B. Paramètres Q-Learning

Pour bien sélectionner, et évaluer les paramètres Q-Learning (taux d'apprentissage  $\alpha$ , facteur de réduction  $\gamma$  et la probabilité d'exploration/exploitation epsilon  $\epsilon$ ) nous avons simulé un réseau de 25 nœuds concurrents, tout en variant les trois paramètres : petit, moyen et large, afin de déterminer les bonnes valeurs utilisées pour le reste des simulations.

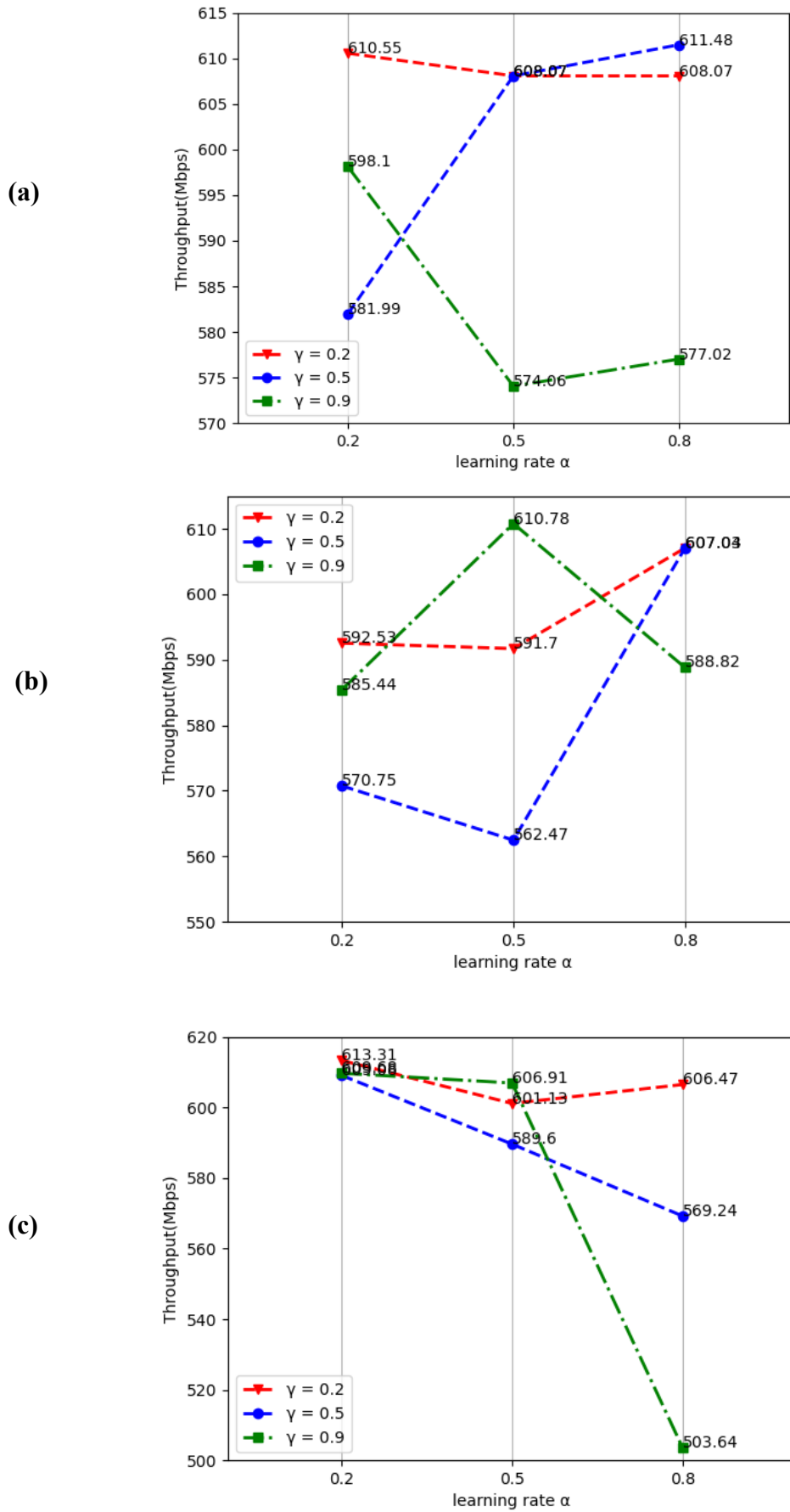
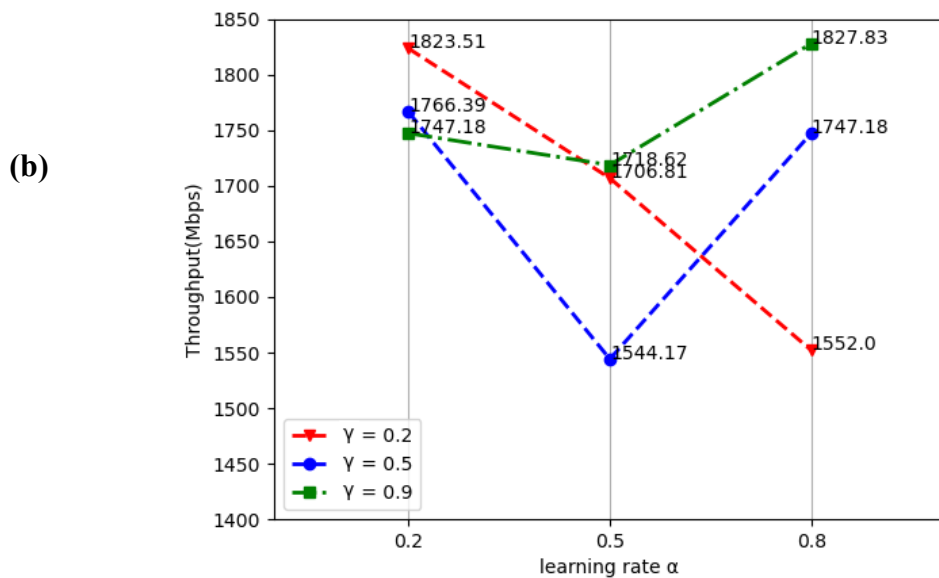
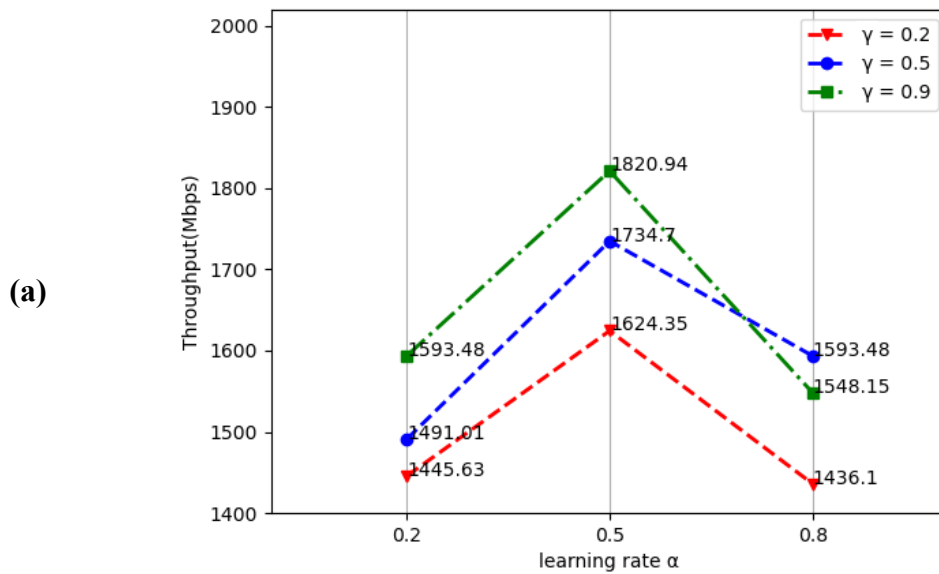
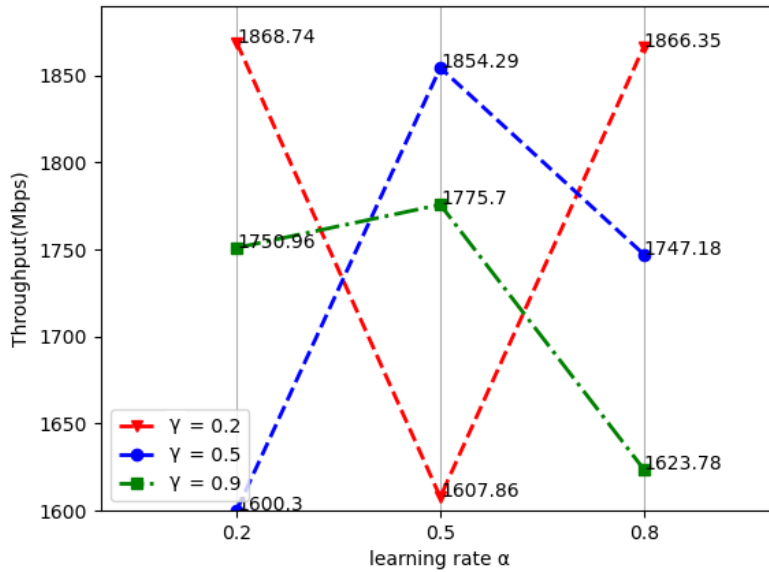


Figure 6.3 : comparaison de débit dans un réseau de 25 nœuds de taille file d'attente =10 avec (a)  $\epsilon = 0.3$  (b)  $\epsilon = 0.6$  et (c)  $\epsilon = 0.8$

La figure 6.3 montre l'effet des paramètres sur le débit pour un réseau de 25 nœuds de taille de file d'attente égale à 10. La figure 6.3 (a) où  $\varepsilon$  est mis à 0.3,  $\alpha=0.8$ ,  $\gamma=0.5$  donne le meilleur débit, alors qu'une valeur  $\alpha=0.2$  dégrade le débit et une valeur  $\alpha=0.5$  montre un petit changement de débit. Nous voyons aussi qu'une valeur de  $\gamma=0.2$ , donne un débit quasi stable quel que soit la valeur de  $\alpha$ . La figure 6.3 (b) montre qu'avec une valeur moyenne de  $\alpha$  et  $\varepsilon$  (0.5), il suffit de mettre  $\gamma$  à sa valeur maximale ( $\gamma=0.9$ ) pour avoir le meilleur débit. Cependant, pour  $\alpha=0.2$  et  $\varepsilon=0.9$ , il suffit de mettre  $\gamma$  à sa petite valeur ( $\gamma=0.2$ ) pour améliorer le débit comme montrer dans la figure 6.3(c).



(c)



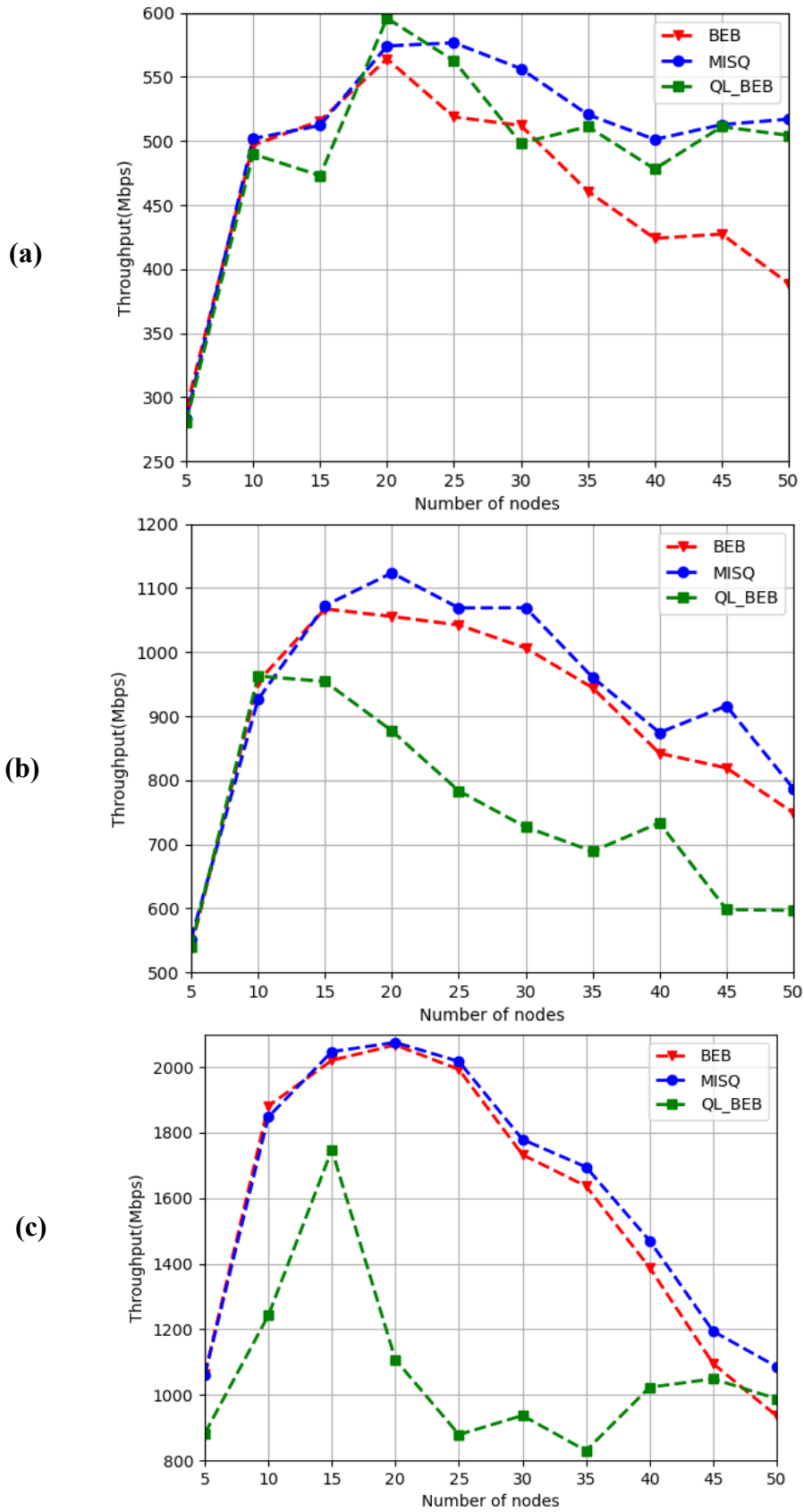
**Figure 6.4** : comparaison de débit dans un réseau de 25 nœuds de taille file d'attente =40 avec (a)  $\epsilon = 0.3$  (b)  $\epsilon = 0.6$  et (c)  $\epsilon = 0.8$

La figure 6.4(a) montre qu'une valeur moyenne de  $\alpha = 0.5$ , une valeur large de  $\gamma = 0.9$  et une valeur petite de  $\epsilon = 0.3$  sont efficace pour donner le meilleur débit dans le cas où la taille de la file d'attente est de 40. La figure 6.4 (c) montre qu'avec une valeur large de  $\epsilon = 0.8$ , petite de  $\gamma = 0.2$ , des valeurs petites ou larges de  $\alpha$  ( $\alpha = 0.2$  ou  $\alpha = 0.8$ ) sont efficace pour avoir un meilleur débit. D'après les différentes et multiples simulations effectuées, nous avons conclu qu'une combinaison d'une valeur large de  $\gamma$  et moyenne de  $\alpha$  et petite de  $\epsilon$  peuvent être efficace pour les différents réseaux avec différentes tailles de file d'attente.

## 6.5. Paramètres de simulation

Les paramètres suivants sont utilisés pour mesurer les performances de notre proposition MISQ.

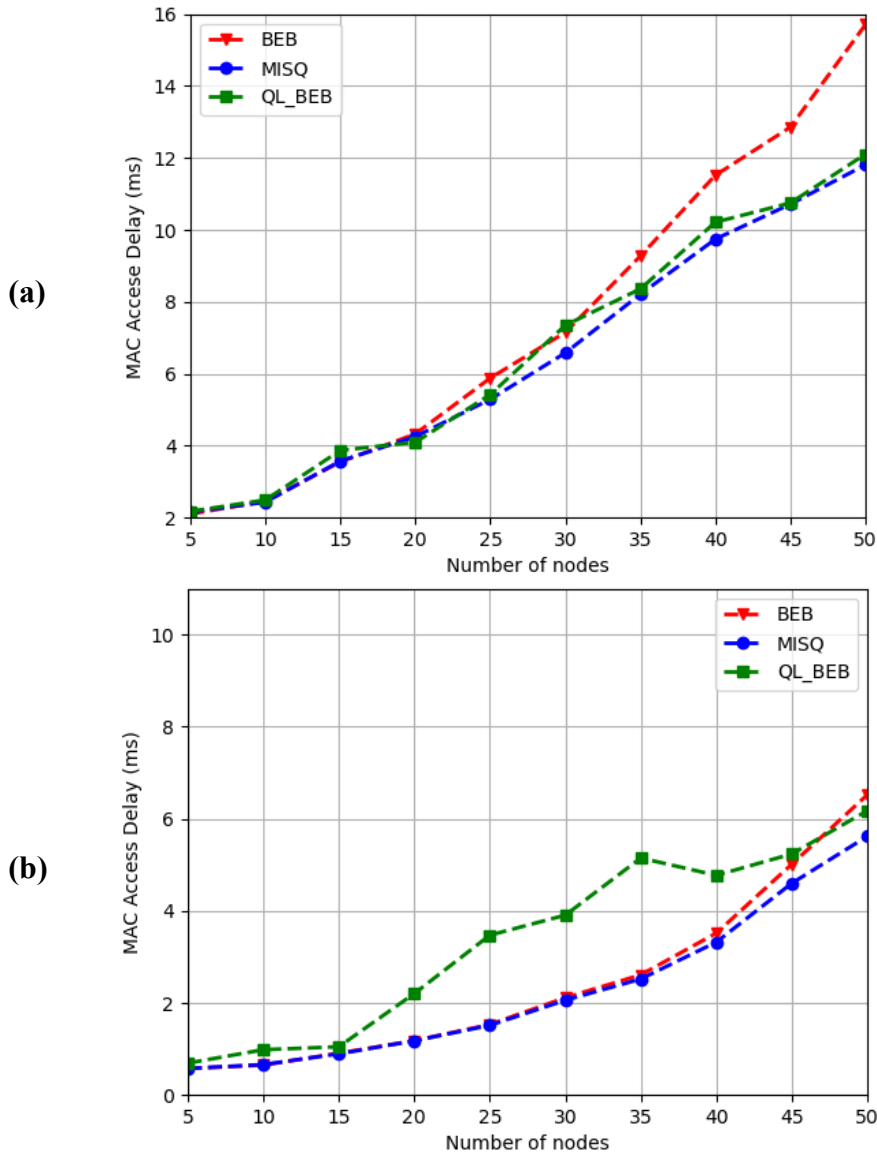
### 6.5.1. Débit moyen

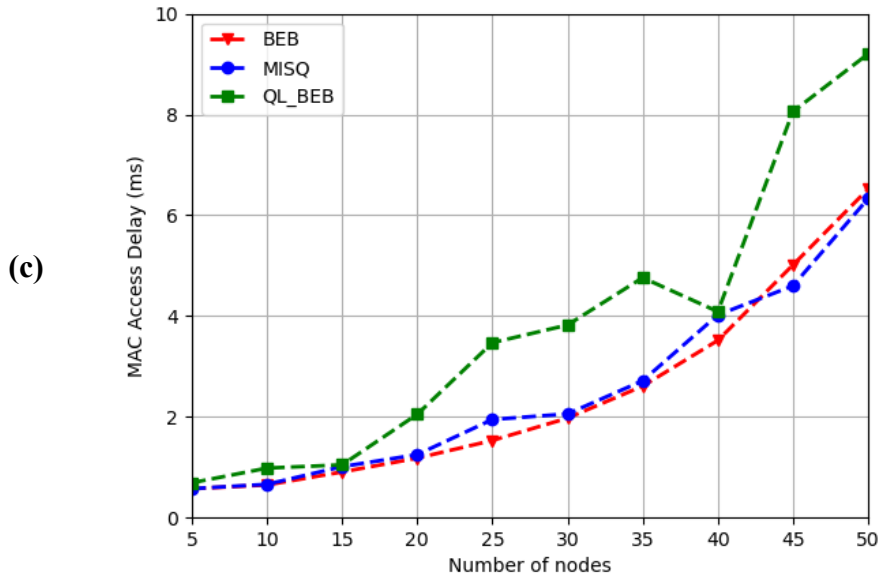


**Figure 6.5:** Débit moyen dans (a) taille file d'attente =10, (b) taille file d'attente =20, (c) taille file d'attente =40

La figure 6.5, donne le débit de l’algorithme MISQ dans (a) le cas où la file d’attente est égale à 10, (b) le cas où la taille de la file est égale à 20 et (c) le cas où la taille de la file est égale à 40. Comparé avec les deux autres algorithmes, notre proposition donne le meilleur débit en particulier à partir de 25 nœuds quel que soit la taille de la file d’attente, cette optimisation est due aux récompenses cumulatives apprises par le réseau ad hoc, ce qui n’est pas le cas où le nombre de nœuds est 5 ou 10.

### 6.5.2. Délai d’accès MAC moyen

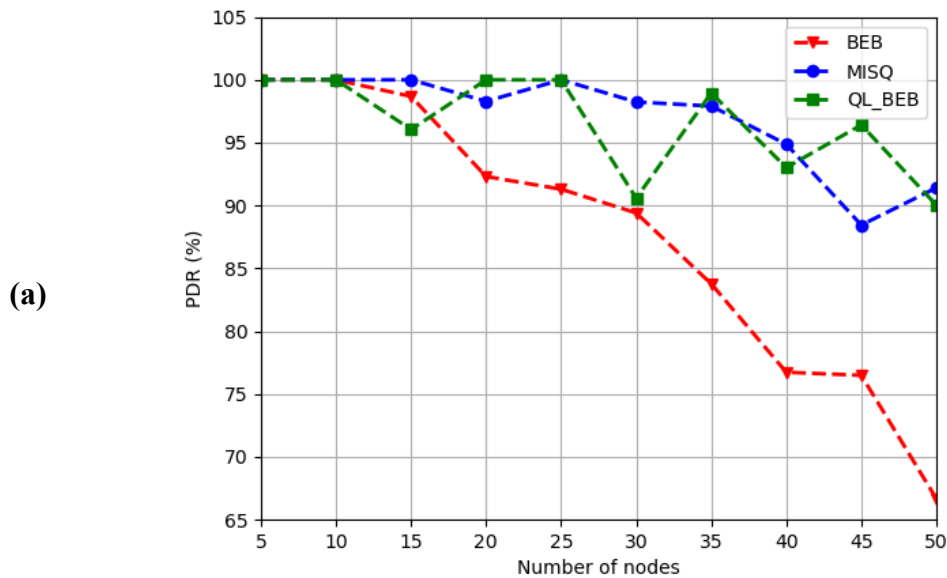


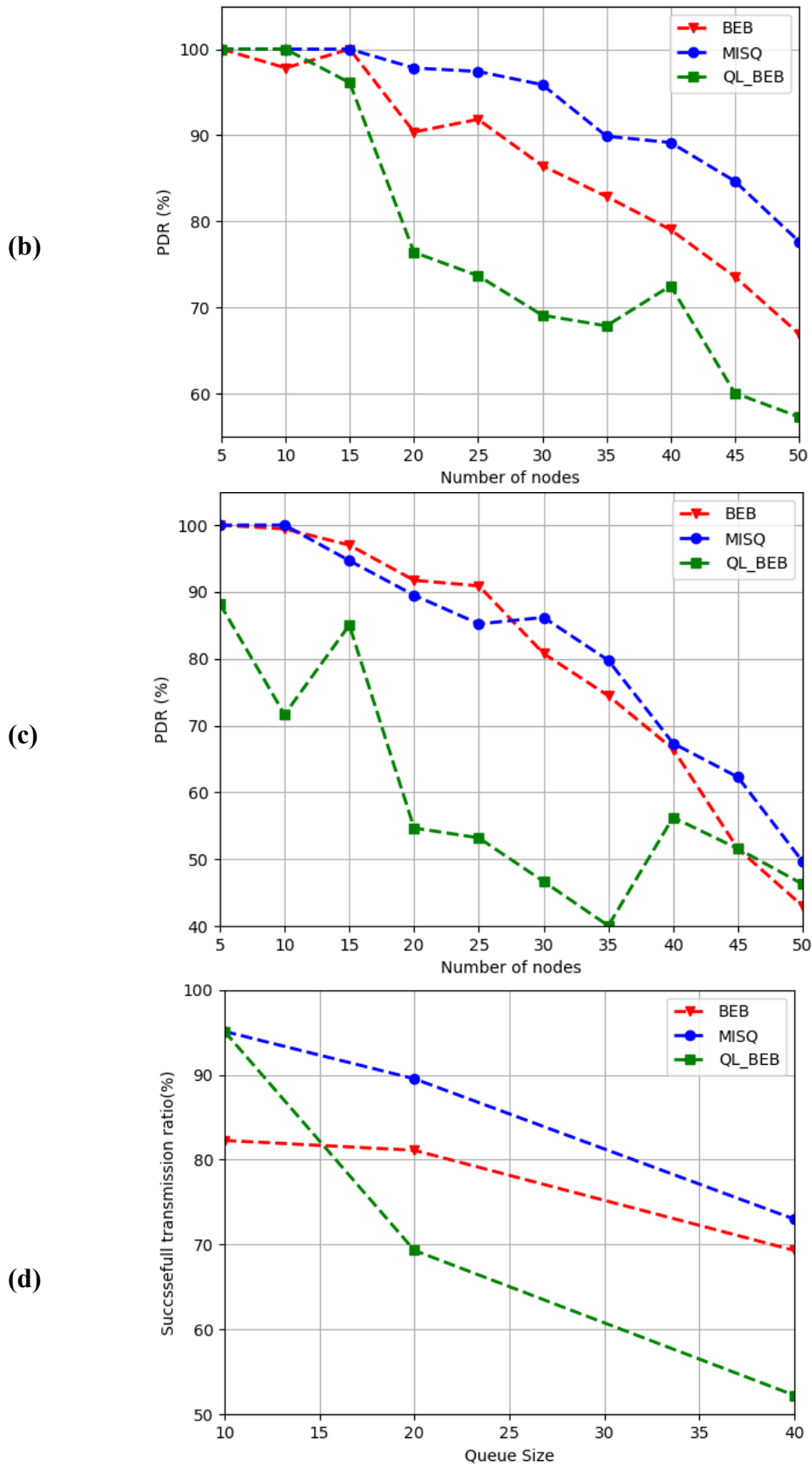


**Figure 6.6:** Délai d'accès MAC pour (a) : taille file d'attente =10, (b) taille file d'attente =20, (c) taille file d'attente =40

La figure 6.6, montre que l'algorithme MISQ proposé donne le moindre délai ou égal au délai donné par BEB, dans le cas où la file d'attente égale à 10 (figure 6.6 (a)) ou égale à 20 (figure 6.6 (b)), par contre dans la figure 6.6 (c) notre algorithme donne le moindre délai à partir de 45 nœuds, il est essentiel de noter que notre algorithme et l'autre simulé ont des retards d'accès au canal dû aux opérations supplémentaires d'apprentissage de l'environnement.

### 6.5.3. Taux de paquets délivrés





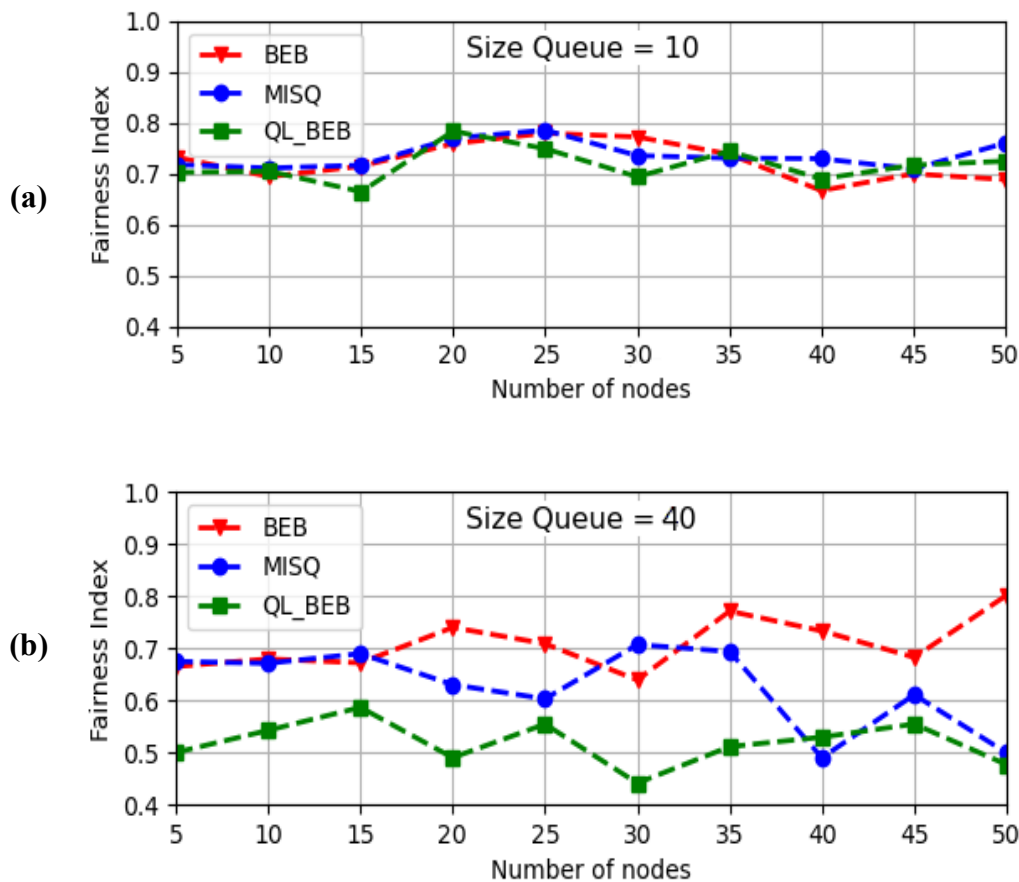
**Figure 6.7:** Packet Delivery Ratio pour (a) : taille file d'attente =10, (b) taille file d'attente =20, (c) taille file d'attente =40, (d) : moyenne de transmission réussie



La figure 6.7 expose le taux de paquets délivrés avec succès, le cas de la taille de la file d'attente égale à 20 c'est-à-dire la figure 6.7 (b) montre une optimisation visuelle de l'algorithme MISQ proposé par rapport aux deux autres à partir de 15 nœuds. Dans le cas où la file égale 40, la proposition donne le meilleur taux à partir de 30 nœuds.

En moyenne la proposition donne le meilleur taux moyen de transmissions réussites quel que soit la taille de la file d'attente, comme mentionné dans la figure figure6.7 (d).

#### 6.5.4. Indice d'équité



**Figure 6.8:** Indice d'équité de Jain, (a) taille file d'attente =20, (b) taille file d'attente = 40

L'indice d'équité des trois protocoles est presque identique dans le cas où la taille de la file d'attente est égale à 10 (figure 6.8(a)), avec un indice quasi stable pour MISQ. Par contre, dans le cas où la taille de la file est égale à 40 (figure 6.8 (b)), BEB donne le meilleur indice, cette baisse d'indice dans MISQ est due à la stratégie suivie par la proposition, préférant des STAs qui ont réussies leurs transmissions et qui ont un taux élevé de paquets à transmettre, d'une part, et d'autre part favoriser les STAs qui font des collisions ayant aussi un taux élevé de paquets à transmettre.

## 6.6. Conclusion

Dans ce chapitre, Nous avons présenté un algorithme, qui exploite les principes d'apprentissage par renforcement profond pour apprendre les paramètres CW corrects pour le standard 802.11, pour optimiser la fenêtre de contention. Motiver par les caractéristiques de l'apprentissage par renforcement profond (DRL) dans les réseaux sans fil, nous avons utilisé une de ses techniques, qui est le Q-Learning, comme paradigme d'allocation de la ressource canal dans la couche MAC. La proposition que nous avons nommée « **Mécanisme Intelligent de Sélection de CW basé sur Q-Learning** » MISQ, se focalise sur l'optimisation de la fenêtre de contention après chaque transmission réussie ou collision, tout en préservant les autres paramètres réseaux tel que le débit, le taux de paquet délivrés et le temps d'accès au canal. Les résultats des différentes simulations montrent que l'algorithme proposé MISQ atteint un débit meilleur, un délai d'accès à la couche MAC acceptable et un taux de paquets délivrés supérieur à celui du BEB utilisé dans la norme IEEE802.11 DCF, en particulier, lorsque la taille de la file d'attente est inférieure ou égale à 20 quel que soit le nombre de stations (10, ..., 50).

## CONCLUSION

Les réseaux sans fil ad hoc sont largement utilisés ces dernières années. L'énorme tâche d'allouer efficacement les ressources de la couche MAC dans les MANET est l'un des obstacles à cette performance en raison de leur nature distribuée basée sur les contentions. Actuellement, les MANETs basés sur CSMA / CA utilise un mécanisme d'interruption exponentielle binaire (BEB), qui augmente et diminue aveuglément la fenêtre de contention après des collisions et des transmissions réussies, respectivement. Pour gérer le défi de la dégradation des performances causée par la densité croissante des MANETs, un mécanisme inversé (O-BEB) basé sur une augmentation et une réduction de la fenêtre de contention dans les deux cas succès de transmission et collision respectivement est proposé dans cette thèse. O-BEB surmonte la limitation du BEB pour atteindre un rendement et une robustesse élevés dans les MANET, et améliore les performances du CSMA / CA dans les réseaux denses. Cependant, O-BEB, ne satisfait pas le paramètre important « Délai de bout en bout » est donne un délai proche ou supérieur à celui donné par BEB.

Le délai étant considéré comme paramètre important de QoS, une autre approche (M-BEB) permettant la satisfaction de ce dernier ainsi que d'autres paramètres comme le débit, le PDF, l'énergie et l'équité, est proposée aussi dans cette thèse. Cependant, avec l'aide de l'inférence sophistiquée des conditions de canal sans fil, il est prédit que les futurs WLAN peuvent accéder de manière autonome aux meilleurs outils de canal pour répondre aux diverses exigences des WLAN denses.

Motivé par les applications et les fonctionnalités potentielles de l'apprentissage par renforcement profond (DRL) dans les réseaux sans fil, comme le déploiement de la radio cognitive. Dans cette thèse, l'une des techniques DRL, Q-Learning, est proposée comme un paradigme intelligent pour l'allocation de la ressource canal de la couche MAC dans les MANET. Le paradigme DRL proposé utilise l'inférence intelligente basée sur QL appelée

Mécanisme Intelligent de Sélection de CW basée sur QL (MISQ), qui prend en compte en plus de l'information de l'état de la transmission (succès ou collision) deux informations supplémentaires le taux d'occupation de la file d'attente et le nombre de collision pour la sélection de la CW approprié. Les résultats de la simulation montrent que le mécanisme MISQ proposé optimise les performances de BEB dans les environnements de réseau ad hoc. Les résultats de la simulation montrent également que le mécanisme MISQ proposé améliore également les performances des MANET compatibles QoS, en particulier pour les environnements de réseau denses.

Les considérations de recherche futures incluent la formulation d'un modèle mathématique pour le mécanisme proposé MISQ. Les travaux futurs comprennent également des évaluations des performances de MISQ avec d'autres propositions, dans divers modèles de mobilité.

## REFERENCES

- [1] Basagni, S., Conti, M., Giordano, S., & Stojmenovic, I. (Eds.). (2004). Mobile ad hoc networking. John Wiley & Sons. DOI:10.1002/9781118511305.
- [2] Boukerche, A. (Ed.). (2005). Handbook of algorithms for wireless networking and mobile computing. CRC Press.
- [3] Braun, T., Kassler, A., Kihl, M., Rakocevic, V., Siris, V., & Heijenk, G. (2009). Multihop wireless networks. In Traffic and QoS management in wireless multimedia networks (pp. 201-265). Springer, Boston, MA.
- [4] Frikha, M. (2010). Réseaux ad hoc : routage, qualité de service et optimisation. Hermès Science.
- [5] Perahia, E., & Stacey, R. (2013). Next generation wireless LANs: 802.11 n and 802.11 ac. Cambridge university press.
- [6] Chlamtac, I. (1996, October). Issues in Mobile Computing. In Plenary Address, Seventh IEEE International Conference on Personal, Indoor, and Mobile Radio Communications (PIMRC'96), Taiwan.
- [7] Stojmenovic, I. (2002). Handbook of wireless networks and mobile computing. New York : Wiley.
- [8] Meraihi, R. (2005). Quality of service support and topology control in mobile ad hoc networks (Doctoral dissertation, Télécom ParisTech).
- [9] Govil, J., & Kumar, B. (2006). Wireless LAN and IEEE standards. IETE Technical Review, 23(1), 47-60.
- [10] Ali, R., Zikria, Y. B., Kim, B. S., & Kim, S. W. (2020). Deep reinforcement learning paradigm for dense wireless networks in smart cities. In Smart cities performability, cognition, & security (pp. 43-70). Springer, Cham. [https://doi.org/10.1007/978-3-030-14718-1\\_3](https://doi.org/10.1007/978-3-030-14718-1_3).
- [11] IEEE Computer Society LAN MAN Standards Committee. (1999). Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. ANSI/IEEE Std. 802.11-1999. <http://www.di-srv.unisa.it/~vitsca/RC-0809I/IEEE-802-11.pdf>
- [12] Wu, H., & Pan, Y. (2008). Medium access control in wireless networks (Vol. 8). Nova Publishers.
- [13] Cho, Y. S., Kim, J., Yang, W. Y., & Kang, C. G. (2010). MIMO-OFDM wireless communications with MATLAB. John Wiley & Sons. <https://doi.org/10.1002/9780470825631>
- [14] Du, K. L., & Swamy, M. N. (2010). Wireless communication systems: from RF subsystems to 4G enabling technologies. Cambridge University Press. doi:10.1017/CBO9780511841453.001.

- 
- [15] Moon, J., & Lim, Y. (2017). A reinforcement learning approach to access management in wireless cellular networks. *Wireless Communications and Mobile Computing*, 2017. <https://doi.org/10.1155/2017/6474768>.
- [16] Williams, B., & Camp, T. (2002, June). Comparison of broadcasting techniques for mobile ad hoc networks. In *Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing* (pp. 194-205). <https://doi.org/10.1145/513800.513825>.
- [17] Zerguine, N., Mostefai, M. (2006, May). Un état de l'art sur : Les réseaux ad hoc et la qualité de service, *Conférence Internationale sur l'Informatique et ses Applications (CIIA06)*, Saida, Algérie.
- [18] Information technology - Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications (2003), ANSI/IEEE Std 802.11, 1999 Edition (Reaffirmed 2003).
- [19] Basagni, S., Conti, M., Giordano, S., & Stojmenovic, I. (Eds.). (2013). *Mobile ad hoc networking: cutting edge directions* (Vol. 35). John Wiley & Sons.
- [20] Chlamtac, Imrich & Redi, Jason. (2000). *Mobile Computing: Challenges and Potential*. [https://www.researchgate.net/publication/2240484\\_Mobile\\_Computing\\_Challenges\\_and\\_Potential](https://www.researchgate.net/publication/2240484_Mobile_Computing_Challenges_and_Potential).
- [21] Vaidya, N. H. (2004, July). Mobile ad hoc networks: routing, MAC and transport issues. In *Proceedings of the IEEE International Conference on Computer Communication INFOCOM*.
- [22] Chlamtac, I., & Lerner, A. (1987). Fair algorithms for maximal link activation in multihop radio networks. *IEEE Transactions on Communications*, 35(7), 739-746. DOI:10.1109/TCOM.1987.1096847.
- [23] Banerji, S., & Chowdhury, R. S. (2013). On IEEE 802.11: wireless LAN technology. *arXiv preprint arXiv:1307.2661*. DOI: 10.5121/ijmnet.2013.3405.
- [24] Freebersyser, J. A., & Leiner, B. (2001). A DoD perspective on mobile ad hoc networks. In *Ad hoc networking* (pp. 29-51).
- [25] Zhu, H., Zeng, G., & Chlamtac, I. (2003, May). Control scheme analysis for multimedia inter-and intra-stream synchronization. In *IEEE International Conference on Communications, 2003. ICC'03.* (Vol. 1, pp. 7-11). IEEE.
- [26] Gast, M. (2006). *802.11 wireless networks: The definitive guide* (p. 159). Nanjing, JiangSu, China: Southeast University Press.
- [27] IEEE Computer Society LAN MAN Standards Committee. (1999). Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: High-speed physical layer in the 5GHz band. *IEEE std 802.11 a-1999*. <https://pdos.csail.mit.edu/archive/decouto/papers/802.11a.pdf> dernier accès : 21/01/20121
- [28] IEEE 802.11 Working Group. (1999). Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications: higher-speed physical layer extension in the 2.4 GHz band. *ANSI/IEEE Std 802.11*. <https://pdos.csail.mit.edu/archive/decouto/papers/802.11b.pdf> dernier accès : 21/01/20121

- 
- [29] IEEE 801.11g, (2003), Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 4: Further Higher Data Rate Extension in the 2.4 GHz Band. IEEE Computer Society LAN MAN Standards Committee. <http://www.diegm.uniud.it/tonello/MATERIAL/STANDARDS/802.11/802.11g-2003.pdf> dernier accès: 21/01/2021.
- [30] IEEE 801.11n, (2008), Draft supplement to Part 11: Enhancement for higher throughput. IEEE Computer Society LAN MAN Standards Committee.
- [31] IEEE 801.11ac, (2013), Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications--Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz. IEEE Computer Society LAN MAN Standards Committee.
- [32] IEEE 801.11e, (2005), Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements, IEEE 802.11e-2005. (Amendment to IEEE std 802.1, 1999 Edition, (Reaff 2003)). IEEE Computer Society LAN MAN Standards Committee.
- [33] Singh, R., & Sharma, T. P. (2015). On the IEEE 802.11 i security: a denial-of-service perspective. *Security and Communication Networks*, 8(7), 1378-1407. <https://doi.org/10.1002/sec.1079>.
- [34] IEEE 801.11h, (2003), IEEE Std 802.11h-2003 Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 5: Spectrum and Transmit Power Management Extensions in the 5 GHz band in Europe, <https://www.sis.se/api/document/preview/907801/>
- [35] Hanzo, L., Akhtman, Y., Akhtman, J., Wang, L., & Jiang, M. (2011). MIMO-OFDM for LTE, WiFi and WiMAX: Coherent versus non-coherent and cooperative turbo transceivers. John Wiley & Sons.
- [36] IEEE P802.11, "REVmd/D2.0 ", (December 2018) - IEEE Draft Standard for Information Technology -- Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks -- Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, in P802.11-REVmd/D2.0, pp.1-4601.
- [37] IEEE 802.11p. (2010), "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments". IEEE 802.11p published standard. IEEE. Retrieved August 10, 2011. <https://standards.ieee.org/getieee802/download/802.11p-2010.pdf>
- [38] IEEE 802.11ax : (2018), la sixième génération du Wi-Fi, Livre blanc technique, Document public de Cisco. [https://www.cisco.com/c/dam/global/fr\\_fr/products/wireless/white\\_paper\\_c11\\_740788.pdf](https://www.cisco.com/c/dam/global/fr_fr/products/wireless/white_paper_c11_740788.pdf)
- [39] Bianchi, G. (1998). IEEE 802.11-saturation throughput analysis. *IEEE communications letters*, 2(12), 318-320.
- [40] Géron, A. (2009). WiFi Professionnel : La norme 802.11, le déploiement, la sécurité. Dunod.
- [41] Michel Terré. (2007), WiFi Le Standard 802.11 Couche physique et couche MAC (cours en ligne), <http://easytp.cnam.fr/terre/images/WiFi.pdf>.

- 
- [42] Lee, Y., Chung, M. Y., & Lee, T. J. (2008). Performance analysis of IEEE 802.11 DCF under nonsaturation condition. *Mathematical Problems in Engineering*, 2008. <https://doi.org/10.1155/2008/574197>.
- [43] Koksai, C. E., Kassab, H., & Balakrishnan, H. (2000, June). An analysis of short-term fairness in wireless media access protocols. In *Proceedings of the 2000 ACM SIGMETRICS international conference on Measurement and modeling of computer systems* (pp. 118-119).
- [44] Kang, J., & Eom, D. S. (2019). Offloading and transmission strategies for IoT edge devices and networks. *Sensors*, 19(4), 835. <https://doi.org/10.3390/s19040835>,
- [45] Li, Y., & Thai, M. T. (Eds.). (2008). *Wireless sensor networks and applications*. Springer Science & Business Media.
- [46] Harivelo, F., & Anelli, P. (2007, January). Équité pour réseau ad hoc Wi-fi. In *Huitièmes Journées Doctorales en Informatique et Réseaux (JDIR'07)* (pp. 77-86).
- [47] Bellman, R. (1957). A Markovian decision process. *Journal of mathematics and mechanics*, 6(5), 679-684.
- [48] Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep Learning* (Vol. 1, No. 2). Cambridge : MIT press.
- [49] Xie, N., Zhang, L., Wu, J., & Wang, H. (2015). An adaptive backoff algorithm based on channel utilization for IEEE 802.11. *Wireless Personal Communications*, 83(4), 2965-2973. DOI 10.1007/s11277-015-2576-y
- [50] Luong, N. C., Hoang, D. T., Gong, S., Niyato, D., Wang, P., Liang, Y. C., & Kim, D. I. (2019). Applications of deep reinforcement learning in communications and networking: A survey. *IEEE Communications Surveys & Tutorials*, 21(4), 3133-3174. doi : 10.1109/COMST.2019.2916583.
- [51] Sigaud, O., & Buffet, O. (2008). *Processus décisionnels de Markov en intelligence artificielle*.
- [52] Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press. [Online]. Available: <http://www.incompleteideas.net/book/RLbook2018trimmed.pdf>
- [53] Watkins, C. J. C. H. (1989). *Learning from delayed rewards*.
- [54] Syed, I., Shin, S. H., Roh, B. H., & Adnan, M. (2017). Performance improvement of QoS-enabled WLANs using adaptive contention window backoff algorithm. *IEEE Systems Journal*, 12(4), 3260-3270. DOI : 10.1109/JSYST.2017.2694859.
- [55] Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.
- [56] Pierre Gérard. (2008), *Apprentissage par Renforcement 'Apprentissage Numérique'*, support de cours, Université de Paris 13 LIPN, <https://lipn.univ-paris13.fr/~gerard/docs/cours/ar-cours-support.pdf>.
- [57] Watkins, C. J., & Dayan, P. (1992). Q-learning. *Machine learning*, 8(3-4), 279-292. DOI 10.1007/BF00992698.
- [58] Wu, H., & Pan, Y. (2008). *Medium access control in wireless networks* (Vol. 8). Nova Publishers.



- 
- [59] Bianchi, G. (2000). Performance analysis of the IEEE 802. 11 distributed coordination function. *IEEE Journal on selected areas in communications*, 18(3), 535-547.
- [60] Bianchi, G., Fratta, L., & Oliveri, M. (1996, October). Performance evaluation and enhancement of the CSMA/CA MAC protocol for 802.11 wireless LANs. In *Proceedings of PIMRC'96-7th International Symposium on Personal, Indoor, and Mobile Communications (Vol. 2, pp. 392-396)*. IEEE.
- [61] Bernardos, C. J., Soto, I., & Banchs, A. (2008). IEEE 802.11 standards. *Medium Access Control in Wireless Networks*, 235.
- [62] Deng, J., Varshney, P. K., & Haas, Z. J. (2004). BO algorithm for the IEEE 802.11 distributed coordination function. *Electrical Engineering and Computer Science*.
- [63] Peng, X., Jiang, L., & Xu, G. (2007, September). Performance analysis of hybrid backoff algorithm of wireless LAN. In *2007 International Conference on Wireless Communications, Networking and Mobile Computing (pp. 1853-1856)*. IEEE.
- [64] Minooei, H., & Nojumi, H. (2007). Performance evaluation of a new backoff method for IEEE 802.11. *Computer Communications*, 30(18), 3698-3704.
- [65] Chatzimisios, P., Vitsas, V., C. Boucouvalas, A., & Tsoulfa, M. (2007). Achieving performance enhancement in IEEE 802.11 WLANs by using the DIDD backoff mechanism. *International Journal of communication systems*, 20(1), 23-41.
- [66] Yassein, M. B., Al Oqaily, O., Min, G., Mardini, W., Khamayseh, Y., & Manaseer, S. S. (2010, June). Enhanced Fibonacci backoff algorithm for mobile Ad-hoc network. In *2010 10th IEEE International Conference on Computer and Information Technology (pp. 749-754)*. IEEE.
- [67] Al-Hubaishi, M., Abdullah, T., Alsaqour, R., & Berqia, A. (2012). E-BEB algorithm to improve quality of service on wireless ad-hoc networks. *Research Journal of Applied Sciences, Engineering and Technology*, 4(7), 807-812.
- [68] Shahin, N., Ali, R., Kim, S. W., & Kim, Y. T. (2018, April). Adaptively scaled back-off (ASB) mechanism for enhanced performance of CSMA/CA in IEEE 802.11 ax high efficiency WLAN. In *NOMS 2018-2018 IEEE/IFIP Network Operations and Management Symposium (pp. 1-5)*. IEEE.
- [69] Wang, K., & Sun, C. (2009, December). An improved backoff algorithm of Ad hoc networks. In *2009 International Conference on Computational Intelligence and Software Engineering (pp. 1-4)*. IEEE.
- [70] Shurman, M., Al-Shua'b, B., Alsaadeen, M., Al-Mistarihi, M. F., & Darabkh, K. A. (2014, May). N-BEB: New backoff algorithm for IEEE 802.11 MAC protocol. In *2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (pp. 540-544)*. IEEE.
- [71] Yassein, M. B., Alomari, M. A., & Mavromoustakis, C. X. (2012). Optimized Pessimistic Fibonacci Back-off Algorithm (PFB). *International Journal of Advanced Computer Science and Applications*, 3(9).
- [72] Shahin, N., Ali, R., Kim, S. W., & Kim, Y. T. (2019). Cognitive backoff mechanism for IEEE802. 11ax high-efficiency WLANs. *Journal of Communications and Networks*, 21(2), 158-167.

- 
- [73] Wydmański, W., & Szott, S. (2020). Contention window optimization in IEEE 802.11 ax networks with deep reinforcement learning. arXiv preprint arXiv:2003.01492. arXiv preprint arXiv:2003.01492.
- [74] Ali, R., Shahin, N., Zikria, Y. B., Kim, B. S., & Kim, S. W. (2018). Deep reinforcement learning paradigm for performance optimization of channel observation-based MAC protocols in dense WLANs. *IEEE Access*, 7, 3500-3511.
- [75] Zerguine, N., Aliouat, Z., Mostefai, M. (2019, November). Enhanced Collisions rate by a new Adaptive and Fair Binary Exponential Backoff. In 1st International Conference on Innovative Trends in Computer Science CITCS. Guelma, Algeria.
- [76] Peng, X., Jiang, L., & Xu, G. (2007, September). Performance analysis of hybrid backoff algorithm of wireless LAN. In 2007 International Conference on Wireless Communications, Networking and Mobile Computing (pp. 1853-1856). IEEE.
- [77] Williams, B., & Camp, T. (2002, June). Comparison of broadcasting techniques for mobile ad hoc networks. In Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing (pp. 194-205).
- [78] Tseng, Y. C., Ni, S. Y., Chen, Y. S., & Sheu, J. P. (2002). The broadcast storm problem in a mobile ad hoc network. *Wireless networks*, 8(2), 153-167.
- [79] Jain, R. K., Chiu, D. M. W., & Hawe, W. R. (1984). A quantitative measure of fairness and discrimination. Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA.
- [80] NS2. <https://www.isi.edu/nsnam/ns/>
- [81] « python version3.7 »<https://www.python.org/downloads/release/python-370/>