Gestion des accès massifs des équipements dans les réseaux NB-IoT : une stratégie basée sur l'apprentissage par renforcement

Yassine Hadjadj-Aoul, Soraya Aït-Chellouche

Chapitre 2

Gestion des accès massifs des équipements dans les réseaux NB-IoT : une stratégie basée sur l'apprentissage par renforcement

2.1. Introduction

Les communications des objets IoT et particulièrement les communications de Machine-à-Machine (M2M) sont considérées comme l'une des plus importantes évolutions de l'Internet. La prise en charge de ces dispositifs constitue toutefois l'un des défis les plus importants auxquels les opérateurs de réseaux devraient faire face (Lin et al., 2016). En effet, le nombre considérable de dispositifs qui pourraient tenter d'accéder au réseau, en même temps, pourrait entrainer une forte congestion, voir la saturation totale, avec toutes les conséquences que cela peut occasionner. En effet, comme on peut le voir dans la Figure 2.1, un nombre très limité de dispositifs essayant simultanément d'accéder au réseau peut réduire les performances du réseau à zéro, indépendamment des possibilités d'accès disponibles (Bouzouita et al.,2016). Dans ces circonstances, il semble évident que des mécanismes de contrôle d'accès efficaces sont nécessaires pour maintenir un nombre raisonnable de tentatives d'accès.

Le groupe *Third Generation Partnership Project* (3GPP) a identifié la surcharge du réseau d'accès aléatoire, le RAN, comme une priorité à un stade précoce et a proposé plusieurs solutions. Parmi les approches suggérées, l'*Access Class Barring* (ACB), proposé dans la version 8, et son extension, l'*Extended Access Barring* (EAB), proposée dans la version 11, sont certainement les stratégies les plus efficaces (3GPP, 2011). En effet, ces approches s'attaquent au problème à sa racine en empêchant même

les tentatives d'accès au réseau. Cependant, ces approches ne font que fournir un cadre pour le contrôle de la congestion, sans donner une solution toute faite.

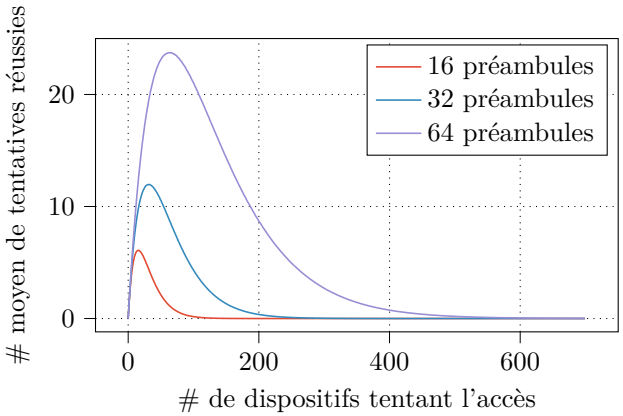

Figure 2.1. *Impacte du nombre de dispositifs IoT qui tentent l'accès simultanément sur la performance.*

L'idée introduite, dans ce papier, est assez simple, puisqu'il s'agit de calculer un facteur de blocage. Néanmoins, une bonne mise en œuvre nécessiterait une bonne connaissance du nombre de terminaux prêts à tenter l'accès pour en déduire la probabilité de blocage optimale. Cette information n'est malheureusement pas disponible dans le réseau.

Afin de résoudre ce problème, deux défis importants doivent être pris en compte : (1) concevoir une stratégie de contrôle d'accès pour la génération dynamique du facteur de blocage, et (2) estimer le nombre de dispositifs qui tentent d'accéder simultanément.

Dans ce Chapitre nous nous attaquons à ces questions en nous basant sur un estimateur qu'on a pu proposer dans un travail antérieur (Bouzouita et al., 2019). Cette estimation étant très bruitée, nous exploitons le potentiel des techniques d'apprentissage par renforcement, les plus avancées, afin de prendre en compte cette réalité complexe (c.-à-d., l'état du réseau n'est pas observable) et déduire une stratégie de contrôle sous-optimale. Nous exploitons, plus particulièrement, dans ce chapitre, l'algorithme d'apprentissage par renforcement profond *Twin Delayed Deep Deterministic policy gradient algorithm* (TD3) (Fujimoto, 2018) pour produire, à partir des estimations passées, le facteur de blocage optimal.

Le reste de ce document est organisé comme suit. La section 2.2 présente brièvement la norme NB-IoT, considérée dans ce papier. La section 2.3 donne un aperçu des principales techniques de contrôle de congestion dans les réseaux IoT, et plus particulièrement les réseaux IoT cellulaires. La section 2.4 décrit le modèle de l'accès des terminaux IoT au réseau. La section 2.5 décrit la solution de contrôle proposée, se basant sur l'algorithme TD3, adapté pour résoudre le problème du calcul du facteur de blocage. La section 2.6, décrit l'environnement de simulation de l'approche proposée et montre son efficacité par rapport à l'existant. Enfin, le document se termine par un résumé récapitulant les principaux avantages et réalisations du système proposé dans la section 2.7.

2.2. Fondamentaux de la norme NB-IoT

NB-IoT est une technologie d'accès cellulaire de type Low Power Wide Area Network (LPWAN), spécifiée dans le document Rel-13 par la 3GPP (3GPP, 2016). Des améliorations y ont ensuite été apportées dans les documents Rel-14, Rel-15 et Rel-16 (Mwakwata, 2019). Plusieurs industriels clés, tels que Ericsson, Nokia, Intel ou Huawei ont montré un vif intérêt pour cette norme et ont largement participé à sa standardisation (Agnes, 2015).

La technologie NB-IoT est également pressentie comme une candidate prometteuse pour couvrir un des trois grands piliers de la 5G, à savoir le *massive Machine Type Communication* (mMTC) (Narayanan et al., 2018) (3GPP, 2018a). En effet, l'ITU a défini différentes exigences, en termes de densité de terminaux IoT, de durée de vie de batterie, de couverture, de prix et de mécanismes et fonctionnalités supportées, auxquelles devraient répondre les technologies proposées pour mettre en œuvre les mMTC (Kafle et al.,2016). Dans ce sens, Le 3GPP a proposé plusieurs mécanismes et fonctionnalités NB-IoT pour répondre à ces exigences à travers les différentes Releases (Mwakwata, 2019). Les options d'intégration de NB-IoT avec le Réseau cœur 5G et sa coexistence avec les autres services proposés dans le cadre de cette norme sont également étudiés par le 3GPP dans (3GPP, 2019a). Une meilleure utilisation des ressources physiques via la virtualisation des fonctions réseau (NFV) ou le SDN, telle que discutée dans (Migabo et al., 2020), peut également faciliter cette intégration.

2.2.1. *Déploiement et cas d'usage*

De nombreux opérateurs, de par le monde, ont fait le choix de la norme NB-IoT, dès sa standardisation en 2016. En effet, ses propriétés inhérentes aux technologies cellulaires telles que la souplesse, l'adaptabilité, la mise à jour à distance « *over-the-*

air » des services et la maitrise des coûts ont accéléré l'adoption de cette norme. Selon une étude de GSMA *Intelligence*, le total des connexions IoT cellulaires atteindrait 3.5 milliards d'ici 2025 (GSMA, 2020). Toujours, selon la même étude, 94 réseaux commerciaux NB-IoT ont été recensés en 2020. Ces derniers sont majoritairement localisés en Chine, où la norme est promue par le constructeur Huawei. En Europe, Vodafone est le principal opérateur promouvant cette norme et en France, l'opérateur SFR a intégré NB-IoT à son offre depuis l'année dernière.

Cette adoption de NB-IoT par d'une part, les opérateurs et d'autre part, les constructeurs, donne lieu à de nombreuses applications (Huawei, 2017) (Ray, 2017), notamment dans les domaines de l'industrie 4.0, des mesures intelligentes, des villes connectées, du suivi et de la surveillance, de l'agriculture intelligente et des fermes connectées ou de la santé.

2.2.2. *Principes de transmissions*

Comme son nom l'indique, NB-IoT est basé sur la transmission de signaux sur une bande étroite. En effet, son occupation spectrale n'est que de 180 kHz, ce qui équivaut à un bloc de ressource PRB (*Physical Resource Bloc*) LTE ou encore à une porteuse GSM. De ce fait, il l'en résulte des débits très faibles (au plus 250 kbit/s), mais suffisants pour les applications pour lesquelles ce protocole est dédié.

La technologie NB-IoT s'appuie sur LTE dont elle hérite plusieurs fonctionnalités et mécanismes, notamment au niveau des couches physiques et MAC. Elle reprend ainsi les mêmes numérologies, codage canal, entrelacement, etc. Ceci a permis de réduire, d'une part, le temps nécessaire à la spécification de cette norme et d'autre part, les coûts de développement des terminaux NB-IoT. Cependant, comme la bande passante utilisée est très étroite, des modifications ont été nécessaires afin de permettre les objectifs premiers de cette norme, à savoir, le déploiement massif de connexions longue portée à complexité et coût réduits (Flore, 2016).

2.2.2.1. *Modes de déploiement*

Comme illustré dans Figure 2.3, NB-IoT peut être déployé selon trois modes différents : (i) *In-band,* au sein de la bande de fréquences LTE en se substituant à un PRB ; (ii) *Gard-band*, en utilisant les ressources spectrales inutilisées, en marge de la bande de fréquences traditionnelles LTE ; (iii) ou enfin, *Stand-alone*, sur une porteuse indépendante. Dans ce dernier cas, les fréquences du GSM sont le plus souvent indiquées pour le déploiement (Wang et al., 2017).

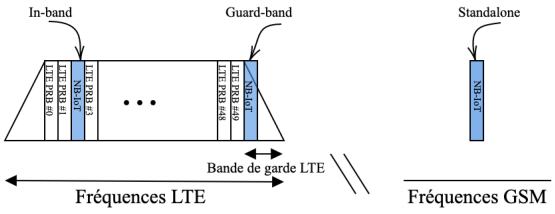

Figure 2.3. *Modes de déploiement NB-IoT.*

Les trois modes de déploiement, cités ci-dessus, sont transparents pour les terminaux non NB-IoT. Ces derniers considèrent les communications NB-IoT avoisinantes comme étant du bruit.

2.2.2.2. *Couche physique*

Sur la liaison descendante, les transmissions sont basées sur OFDM. Chaque symbole OFDM occupe 12 sous-porteuses de 15 kHz chacune, occupant ainsi l'équivalent d'un PRB LTE (180kHz). De manière également similaire à LTE, chaque trame est composée de 10 sous-trames de deux slots chacune. Les durées de la trame, de la sous-trame et du slot sont de 10 ms, 2 ms et 0.5 ms, respectivement. Chaque terminal opère sur un bloc de ressource RB (*Resource Bloc*) constitué de 7 symboles OFDM consécutifs sur la totalité du domaine fréquentiel. Concernant la modulation, seule la modulation en quadrature de phase QPSK (Quadrature Phase-Shift Keying) est supportée.

Sur la liaison montante, les transmissions sont basées sur SC-FDMA (*Single Carrier-Frequency Division Multiple Access*). Deux modes de configurations sont possibles : tonalité unique ou multitonalité. Quand la tonalité unique est utilisée, les deux numérologies, 15 kHz et 3.75 kHz, sont possibles. L'unité de ressources radio RU allouée aux terminaux correspond à l'écart d'une sous-porteuse. On dispose ainsi de 12 ou 48 RU, respectivement. En multitonalité, seul l'écart de sous-porteuse de 15 kHz peut être utilisé. L'UR peut occuper 3, 6 ou 12 sous-porteuses contiguës. Les modulations utilisées sur les liaisons montantes sont BPSK et QPSK.

Les signaux et canaux physiques utilisés dans NB-IoT sur les liaisons montantes et les liaisons descendantes sont résumés dans le Tableau 2.1 (Savaux et al., 2020).

	Canal/Signal physique	Fonction
Liaison descendante	*Narrowband Reference Signal (NRS)*	Dédié à l'estimation du canal dans le domaine fréquentiel (fournit la phase de référence). Ce signal est transmis dans toutes les sous-trames sauf celles réservées à NPSS et NSSS.
	Narrowband Primary and Secondary Synchronization Signals (NPSS and NSSS)	Hérités de LTE, ces signaux sont utilisés dans la découverte de la cellule via une synchronisation temporelle est fréquentielle. Ils permettent aussi de découvrir l'identificateur de la cellule NCellID, notamment, via le NSSS.
	Narrowband Physical Broadcast Chanel (NPBCH)	Transporte le MIB-NB (*Narrowband Master Information Block*). Ce dernier contient 34 bits de données (informations utiles au terminal, notamment au décodage du SIB1-NB, l'activation ou non de l'ACB, etc.) et 16 bits de CRC. Le MIB-NB reste inchangé pour une période de 640 ms.
	Narrowband Physical Downlink Control Channel (NPDCCH)	Contrôle les transmissions de données entre la station de base et le terminal NB-IoT. Il transporte les informations de signalisation qui régissent ces transmissions telles que l'ordonnancement des ressources sur la liaison descendante, les acquittements des données envoyées sur la liaison montante, le type de modulation utilisée, etc.
	Narrowband Physical Downlink Shared Chanel (NPDSCH)	Dédié à transmission des données en direction du terminal. Il transporte, notamment, le SIB1-NB (*Narrowband System Information Bloc*) et quelques informations de contrôle. SIB1-NB contient toutes les informations utiles pour acquérir les autres blocs SIB.
Liaison descendante	*DeModulation Reference Signal (DMRS)*	Dédié à l'estimation du canal dans le domaine fréquentiel sur la liaison montante. Contrairement au NRS, ce signal est toujours multiplexé avec les données. Il est ainsi uniquement transmis dans les RU contenant des données.
	Narrowband Physical Random Access Channel (NPRACH)	Dédié à la transmission du préambule, qui est le premier signal transmis par le terminal, afin d'établir une connexion au réseau. Le détail de ce canal est donné dans la section 2.2.3.
	Narrowband Physical Uplink Shared Chanel (NPDSCH)	Utilisé à la fois pour la transmission des données utilisateur ainsi que des données de contrôle. La distinction est faite par l'utilisation de deux formats : Format 1 et Format 2, respectivement.

Tableau 2.1. *Signaux et canaux physiques de NB-IoT.*

Afin de permettre une extension de couverture tout en limitant la puissance de transmission maximale des terminaux à 23 dBm, NB-IoT utilise des techniques améliorées de couverture telles que l'augmentation de puissance sur les liaisons descendantes ou la répétition de sous-trames à la fois sur les liaisons montantes et descendantes. Cette technique est particulièrement intéressante quand les capteurs sont déployés dans des zones distantes ou difficiles d'accès. L'objectif visé par le standard est le support d'une perte de couplage maximum MCL d'au plus 164 dB (20 dB de plus que LTE). L'étude de l'extension de couverture prévue par NB-IoT sous différentes configurations, menée dans (Adhikary et al., 2016), conclut à un support de MCL supérieurs à la limite des 164 dB, fixée par le 3GPP, à savoir, des MCL de 170.2 et 172.2 dB pour des espacements respectifs de 15 kHz et 3.75 kHz, en effectuant 128 répétitions.

Comme dans LTE, le réseau peut définir jusqu'à trois niveaux d'extension de couverture (CE0, CE1 et CE2) afin de s'adapter aux différentes conditions radio dans lesquelles opèrent les terminaux. Ces niveaux sont définis à travers deux seuils de puissance basés sur la puissance du signal de référence reçu par les terminaux NB-IoT (3GPP, 2017a). Par conséquent, deux seuils peuvent ainsi être définis par cellule pour configurer les trois niveaux de couverture. Pour chaque niveau, un nombre de répétitions du signal est défini. Plus la puissance du signal est grande (moins d'atténuation sur le trajet) plus le nombre de répétitions nécessaire est petit.

Contrairement à la norme LTE qui supporte le mode *full duplex*, le choix de FDD *half Duplex type-B* a été pris dans la Rel-13 pour NB-IoT (3GPP, 2017b). Les terminaux sont donc soit, en émission soit, en réception, mais ne peuvent effectuer les deux opérations simultanément. De plus, une sous-trame de garde est prévue pour chaque passage de la liaison montant à la liaison descendante ou vice-versa.

2.2.3. *Procédure d'accès aléatoire à la ressource radio*

Le canal NPRACH a été complètement repensé dans NB-IoT afin d'améliorer la couverture réseau et la consommation d'énergie, mais aussi, afin de s'adapter à la bande étroite de NB-IoT (Lin, 2016). En effet, dans LTE le canal PRACH occupe, à lui seul, plus de bande passante que ce qu'occupe NB-IoT dans sa totalité (1.08 MHz vs 180 kHz).

Chaque terminal NB-IoT, voulant se connecter ou se resynchroniser à la station de base sur sa liaison montante, après un long temps d'inactivité, doit exécuter une procédure d'accès aléatoire. La première étape de cette procédure consiste à transmettre une séquence de préambules sur une des bandes de fréquences allouées

périodiquement au canal NPRACH. Cette dernière est appelée opportunité d'accès aléatoire RAO (*Random Access Opportunity*).

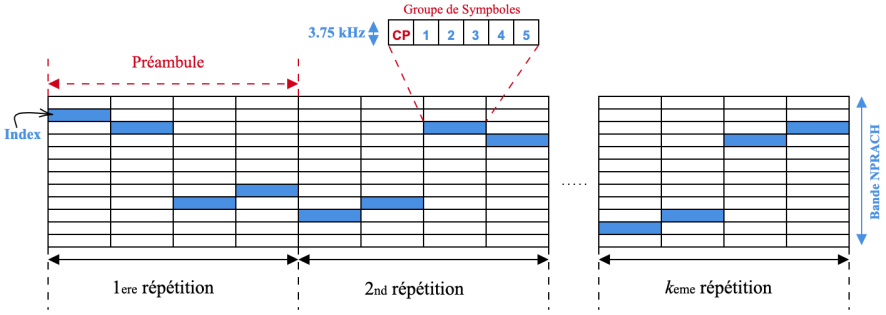

Figure 2.4. *Structure d'une séquence de préambule.*

Le préambule consiste en un ensemble de quatre groupes de symboles OFDM, comme illustré dans la Figure 2.4. C'est le MCL, visé par la norme, qui détermine ce nombre de 4. Chaque groupe de symboles est composé d'un préfixe cyclique (CP) et de K symboles de données. Afin de maintenir l'orthogonalité des transmissions d'accès aléatoire sur différentes sous-porteuses, le CP doit être assez long pour pallier des temps aller-retour longs, notamment dans des cellules aussi larges que celles visées dans NB-IoT (jusqu'à 35 km) (3GPP, 2015). Plus le nombre K est grand plus le surdébit généré par le CP est réduit. Par contre, K doit être maintenu petit afin de maitriser les interférences. Dans NB-IoT K est fixé à 5 et deux longueurs de CP sont définies pour les deux formats du canal NPRACH, à savoir 266.7 µs et 66.7 µs (3GPP, 2020) (Lin et al., 2016).

Pour des fins d'extension de couverture, le préambule peut être répété k fois ($k = 2^i, i = 0, \dots, 7$) (3GPP, 2019b). La séquence de préambules envoyée par le terminal est, par conséquent, constituée de 4×2^i groupes de symboles. Le réseau peut ainsi définir jusqu'à trois configurations différentes de la ressource NPRACH par cellule, en fonction des classes de couverture considérées. Le nombre de répétions k est donc défini pour chacune des configurations.

Chaque groupe de symboles est modulé sur une sous-porteuse différente des autres. Le canal NPRACH n'utilise que le mode tonalité unique avec un espacement de 3.75 kHz. Une bande de fréquences, allant jusqu'à 48 sous porteuses, peut ainsi être allouée à ce canal avec une bande de base de 12 sous-porteuses. Ainsi, 12, 24, 36 ou 48 sous-porteuses contiguës sont allouées à ce canal dans chaque classe de

couverture. Par conséquent, le terminal dispose de 12, 24, 36 ou 48 préambules orthogonaux et en choisit aléatoirement un à transmettre.

De plus, comme illustré dans la Figure 2.4, NB-IoT définit deux modèles de saut dans la bande de fréquences allouée au canal NPRACH : (i) un modèle fixe pour les sauts entre les différents groupes de symbole constituant le même préambule (ii) et un modèle pseudo-aléatoire pour les répétitions de ce préambule. Ainsi, au sein d'un préambule, un saut d'une fréquence est appliqué entre le premier et le second groupe de symboles et entre le troisième et le quatrième groupe de symboles. Un autre saut de six fréquences est, également, appliqué entre le second et le troisième groupe de symboles. Le modèle pseudo-aléatoire est, quant à lui, appliqué à travers le choix des index des sous-porteuses de début des différentes répétitions du préambule, en prenant comme données d'entrée l'identificateur de la cellule et le nombre de répétitions prévu (Lin et al., 2016).

La procédure d'accès aléatoire dans NB-IoT, illustrée dans la Figure 2.5, est un échange en quatre étapes entre le terminal et la station de base (3GPP, 2018b) :

- Le terminal transmet le préambule choisi à la première opportunité RAO et arme un temporisateur pour la réception de la réponse RAR ;
- Si le préambule est bien détecté par la station de base, cette dernière envoie une réponse RAR véhiculant l'avance de synchronisation et la ressource allouée ;
- Le terminal envoie alors une requête de connexion, en exploitant la ressource qui lui a été allouée et arme de nouveau un temporisateur de résolution de conflit. Cette requête, nommée msg3 véhicule notamment l'identité du terminal ;
- La station de base exécute la résolution de conflits et envoie l'identité du terminal gagnant dans le message de résolution de conflits. Si le message n'arrive pas au terminal, ce dernier reste dans son attente jusqu'à expiration du temporisateur.

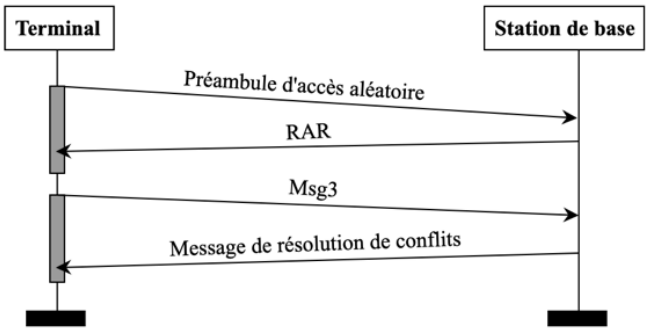

Figure 2.5. *Procédure d'accès aléatoire.*

Cette procédure échoue si le terminal ne reçoit pas l'une des deux réponses de la station de base dans les fenêtres de temps définies par les deux temporisateurs. Les collisions entre préambules, envoyés par différents terminaux, sont souvent la cause de l'échec. En effet, si deux terminaux, ou plus, choisissent le même préambule sur la même opportunité ROA, chacun d'eux voit sa tentative d'accès échouer.

Chaque terminal, pour qui la procédure d'accès a échoué, observe un temps d'attente choisi aléatoirement dans un intervalle prédéfini, puis retransmet son préambule. Le nombre de retransmissions permises dépend de la classe de couverture du terminal. Si ce nombre est atteint et que le terminal ne réussit toujours pas la procédure d'accès lancée, ce dernier se met dans la classe de couverture suivante s'il y en a ou conclut à un échec définitif de la procédure d'accès.

2.3. État de l'art

Au-delà de la généralisation du haut débit, la 5G promet d'améliorer notre qualité de vie à travers des écosystèmes connectés. Les communications massives de type machine (mMTC) joueront un rôle déterminant dans la réalisation de ces derniers. Si cette vision de l'IoT semble très attrayante, elle engendre également de nombreux défis à relever pour les constructeurs et les opérateurs de réseaux. En effet, un nombre excessivement grand de capteurs devrait être géré dans les réseaux IoT dans les prochaines années.

Comme expliqué dans la Section 2.2.3, chaque terminal voulant se connecter au réseau doit initier une procédure d'accès aléatoire. Cependant, cette dernière a été initialement conçue pour un nombre limité de terminaux et la haute densité visée par

NB-IoT peut amener, très vite, à une situation de congestion sévère. En effet, le nombre de préambule disponible à chaque RAO étant limité, plus le nombre de terminaux tentant l'accès est grand, plus le risque de collision est grand, amenant ainsi à l'échec de la procédure pour tous les terminaux ayant choisi un même préambule. Certes, les terminaux, n'ayant pas réussi l'accès, peuvent retransmettre le préambule après avoir observé un temps d'attente, mais ces retransmissions peuvent également conduire à une mauvaise utilisation des ressources spectrales, d'une part, mais aussi à augmenter la consommation d'énergie au niveau des terminaux, d'autre part (Harwahyu et al., 2019).

Vu sa criticité, la procédure d'accès aléatoire a fait l'objet de plusieurs études. Certaines études, telles que (Barakat et Brito, 2018), (Jiang et al.,2018) ou (Harwahyu et al., 2018), ont proposé des modèles analytiques pour l'optimisation de la probabilité de succès des tentatives d'accès des terminaux et du temps d'accès moyen dans différentes configurations et notamment sous des contraintes de délais (Harwahyu et al., 2018). D'autres se sont focalisés sur les retransmissions. Ainsi, dans (sun et al., 2017) un modèle, basé sur les chaines de Markov, a été proposé pour modéliser le nombre de retransmissions et dans (Harwahyu et al., 2019), les auteurs proposent un modèle pour trouver un compromis entre le nombre de répétions prévues dans la couche physique et le nombre de retransmissions prévues dans la couche MAC afin d'optimiser ces deux valeurs en exploitant la probabilité de détection du préambule. L'étude a montré que les retransmissions considérées dans NPRACH peuvent réduire le nombre de répétitions. Ces dernières ne sont nécessaires que lorsque les conditions du réseau se détériorent.

Dans (Lin et al., 2016), (Jeon et al., 2018) et (Hwang et al.,2018), le focus a été mis sur la transmission du préambule et l'estimation du temps d'arrivée. Ainsi un algorithme de détection coté récepteur, un nouveau modèle de saut dans le domaine fréquentiel NPRACH et un Framework pour la détection d'utilisateurs multiples ont été proposés, respectivement. Dans (Zhang et al., 2020), le TA des préambules qui ont subi une collision est utilisé pour améliorer les performances de la RAP.

Du point de vue de la standardisation, le contrôle de congestion au niveau de l'accès au réseau a été, très tôt, identifié comme une priorité par les organismes 3GPP et ETSI (3GPP, 2011). L'IoT cellulaire et notamment NB-IoT, bénéficie donc naturellement des solutions proposées pour les normes qui les précèdent. Parmi les solutions proposées, on retrouve L'ACB (*Access Class Barring*) et extension EAB, l'accès aléatoire à créneaux (*slotted random access*), des *backoffs* spécifiques au MTC, une allocation dynamique des ressources, etc. (Ali et al., 2017).

L'ACB et l'EAB sont celles qui s'attaquent a à la racine du problème en bloquant l'accès au réseau via la diffusion de paramètres de blocage dans les blocs SIB à chaque RAO. Les terminaux reçoivent, notamment, une probabilité de blocage p et un temps de blocage T_b pour cette opportunité. Chaque terminal voulant accéder au réseau, génère une probabilité d'accès q. Si $q < p$, le terminal a la permission d'effectuer sa tentative d'accès sinon cette dernière est différée pour un temps T_b. Ce mécanisme a été étendu. Dans l'EAB, les terminaux sont classés en fonction de leurs exigences en termes de QoS et l'algorithme EAB bloque dynamiquement les terminaux de faible priorité en fonction du taux d'arrivée en diffusant un *bitmap* dans le SIB-14.

Il parait clair que le contrôle de congestion via ces techniques repose entièrement sur la probabilité de blocage définie par le réseau. En effet, si la probabilité de blocage est trop grande alors un nombre important de terminaux passerait le contrôle d'accès entrainant ainsi des collisions et si d'autre part, cette probabilité est trop petite alors, les collisions seraient réduites, mais un grand nombre de terminaux basculeraient en mode inactif et cela conduirait à une sous-utilisation des ressources. Il est donc essentiel de calculer une probabilité de blocage optimale pour un contrôle de congestion efficace.

Une étude des performances de l'ACB et l'EAB a été menée dans (Tour et Jin, 2017). La comparaison des deux techniques via la simulation a montré que l'ACB est plus adapté aux communications avec de fortes contraintes de délai et l'EAB dans le cas de terminaux contraints en énergie. Cependant, le calcul de la probabilité de blocage optimale repose sur la connaissance de la station de base du nombre de terminaux tentant l'accès au réseau. Ce qui, en pratique, n'est pas le cas. En effet la station de base n'a pas la connaissance du nombre de terminaux dont la tentative d'accès a été bloquée.

Plusieurs mécanismes ont été proposés pour estimer le nombre de terminaux tentant l'accès au réseau (y compris les terminaux bloqués par le contrôle d'accès) afin d'en déduire la probabilité de blocage utiliser. Dans (Park et Lim, 2016), l'absence de la connaissance du nombre de terminaux bloqués, les auteurs utilisent une heuristique pour adapter la probabilité de blocage. L'algorithme proposé dans (Liu et al., 2020) fait une estimation bayésienne récursive des terminaux actifs dans chaque classe et en fonction de cette estimation, des préambules sont alloués aux différentes classes. L'algorithme a ensuite été amélioré en y assignant un facteur de blocage ACB à chacune des classes, indépendamment des autres, pour un meilleur contrôle de la congestion. Dans (Jin et al., 2017) une estimation bayésienne récursive, des terminaux actifs, basée sur le nombre de préambules non choisis, permet de calculer un facteur de blocage pour des arrivées de terminaux à caractère sporadique. Les performances de la technique EAB est étudiée dans (Cheng et al., 2015) pour les

réseaux LTE-A. Les valeurs optimales du cycle de radiomessagerie ainsi que la périodicité du SIB14 sont alors dérivées en soumettant le modèle analytique soumis à une contrainte de QoS ciblée.

Dans ce chapitre, nous nous basons sur un estimateur proposé dans un travail antérieur (Bouzouita et al., 2019) et à la différence des travaux cités, ci-dessus, nous exploitons des techniques d'apprentissage par renforcement, notamment l'algorithme TD3, pour calculer un facteur de blocage optimal à partir d'un ensemble d'estimations passées. À notre connaissance, c'est la première fois que ce type d'algorithme est utilisé dans la gestion des accès massifs des terminaux dans les réseaux NB-IoT.

2.4. Modèle pour l'accès des terminaux IoT

Le modèle proposé représente une vue d'ensemble des dispositifs IoT exécutant l'algorithme ACB. Lors de la tentative d'accès aléatoire, les dispositifs IoT se disputent les mêmes préambules disponibles. Comme l'indique la norme 3GPP, le nombre de préambules N doit être un nombre entier dans l'intervalle $[12..48]$ (ETSI,2011).

Dans chaque opportunité d'accès (c.-à-d., « Random Access CHannel » (RACH), ces préambules sont divisés en préambules réussis, choisis par un seul dispositif, en collision, choisis par deux ou plusieurs dispositifs, et libres, choisis par aucun des dispositifs.

Dans ce qui suit, nous calculons les valeurs moyennes de ces quantités que nous avons déterminées dans (Bouzouita et al., 2015). Ces dernières seront, par la suite, utilisées par nos algorithmes.

Définissons $q_N = 1 - 1/N$. Le nombre moyen de préambules réussis N_S, lors des opportunités du RACH, est donné comme suit (c'est un problème classique de lancer de balles dans des urnes) :

$$N_S = q_N^{x_2-1} x_2 \tag{1}$$

où x_2 représente le nombre de dispositifs tentant l'accès. Comme nous avons pu le démontrer, dans (Bouzouita et al., 2019), l'équation (1) est maximisée (c.-à-d., dérivée nulle) lorsque le nombre d'équipements x_2^* tentant l'accès simultanément est égale à,

$$x_2^* = -\frac{1}{\ln q_N} \tag{2}$$

Le nombre moyen de préambules libres N_I est donné par l'équation suivante :

$$N_I = N q_N^{x_2} \tag{3}$$

De (1) et (2), on obtient le nombre attendu de préambules ratés N_F:

$$N_F = N - (N_S + N_I) \tag{4}$$

Le système modélisé est une approximation de la réalité à bien des égards, notamment en ce qui concerne le nombre limité et fixe de tentatives d'accès. Toutefois, nous avons préféré simplifier le modèle pour le rendre plus tractable (voir la Figure 2.2).

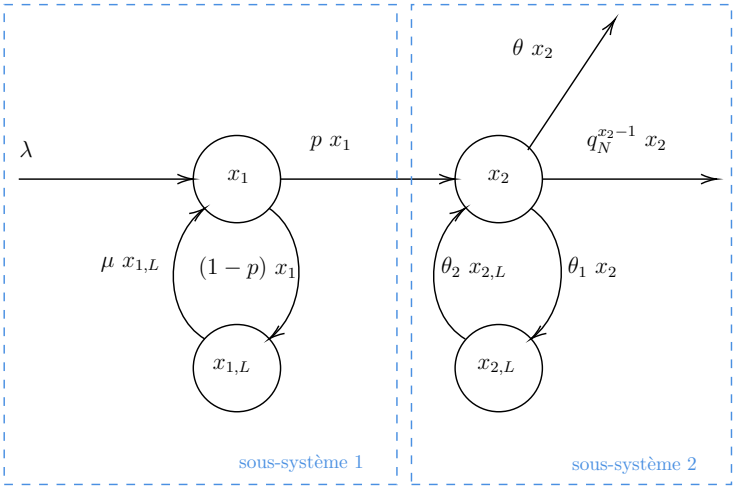

Figure 2.6. *Modèle du système.*

Le sous-système 1 représente les terminaux qui souhaiteraient se connecter ; les objets de la variable d'état x_1 représentent ceux qui peuvent essayer de se connecter avec une probabilité p, en cas d'échec, ils passent à l'état d'attente $x_{1,L}$ pour une durée déterminée. Le sous-système 2 représente les objets qui peuvent essayer de choisir un préambule. En cas de collision, ils peuvent tenter d'y accéder plusieurs fois. Ils quittent le sous-système 2 lorsqu'ils réussissent à être les seuls à avoir choisi un préambule ou lorsqu'ils atteignent le nombre maximum de tentatives (avec un taux de θ).

Le modèle proposé est fluide : les quantités concernées et les nombres entiers sont considérés comme des quantités réelles (continues). Les paramètres utilisés sont énumérés ci-dessous :

- $x_1(t)$ est le nombre de dispositifs en attente ;

- $x_{1,L}(t)$ est le nombre de dispositifs bloqués, après un échec à l'issue d'une tentative d'accès (i.e., ACB) ;
- $x_2(t)$ est le nombre total de dispositifs qui passent le contrôle ACB et attendent de démarrer la tentative d'accès aléatoire (RA) ;
- $x_{2,L}(t)$ est le nombre de dispositifs bloqués, à l'instant t, après une tentative RA ratée en attente d'une nouvelle tentative ;
- λ est le taux d'arrivée des objets IoT ;
- μ est le taux d'objets pouvant retenter l'ACB après un échec ;
- θ_1 est le taux d'échec à l'RA, qui est égal à $1 - q_N^{x_2-1}$ lorsque θ est égal à 0 (voir l'avant dernier point) ;
- θ_2 est le taux d'objet pouvant tenter l'accès après un échec ;
- θ est le taux auquel les équipements abandonnent la transmission après avoir atteint le nombre maximum de tentatives RA ; dans un système correctement dimensionné, nous devrions avoir $\theta = 0$;
- p est le facteur de blocage de l'ACB.

Nous sommes maintenant en mesure de décrire l'évolution des variables d'état $x_1(t)$, $x_{1,L}(t)$, $x_2(t)$ et $x_{2,L}(t)$ en nous basant sur le modèle représenté dans Figure 2.6. La dynamique du modèle est décrite par le système d'équations différentielles suivant :

$$\begin{cases} \frac{dx_1}{dt} &= \lambda - x_1 + \mu x_{1,L} \\ \frac{dx_{1,L}}{dt} &= (1-p)x_1 - \mu x_{1,L} \\ \frac{dx_2}{dt} &= px_1 - \left(\theta + \theta_1 + q_N^{x_2-1}\right)x_2 + \theta_2 x_{2,L} \\ \frac{dx_{2,L}}{dt} &= \theta_1 x_2 - \theta_2 x_{2,L} \end{cases} \quad (5)$$

Dans ce qui suit, nous supposons que $\theta = 0$, afin de simplifier le modèle. En effet, un système où les dispositifs atteignent souvent le nombre maximum de tentatives est un système instable, ce que nous essayons naturellement d'éviter.

Le modèle décrit, dans (5), est non linéaire et non affine, dans le contrôle. Il peut être aisément démontré que le modèle décrit est non observable, au vu de son état $[x_1 \quad x_{1,L} \quad x_2 \quad x_{2,L}]$ qui ne peut être connu précisément. Il est aussi non contrôlable, car le facteur de blocage p ne peut agir que partiellement sur l'état. Ces propriétés rendent la synthèse d'un contrôleur optimal garantissant la stabilité du système, décrit ci-dessus, très complexe.

Bien que l'état ne soit pas observable, il est possible de produire une estimation du nombre moyen de dispositifs tentants l'accès $\hat{x}_2$ en inversant les équations (1) et (3). Cela donne une mesure très bruitée, mais qui peut être, néanmoins, utile pour le blocage des IoTs, comme nous l'avons démontré dans (Bouzouita et al., 2019).

2.5. Contrôleur d'accès pour les terminaux IoT basée sur l'apprentissage par renforcement

La difficulté d'observer l'état du système, décrit dans la section précédente, nous a mené à considérer des stratégies permettant de déduire le facteur de blocage même en présence de mesures très bruitées. C'est dans ce sens, que nous nous sommes penchés sur les techniques d'apprentissage profond, qui ont démontré une très grande efficacité à extraire automatiquement les caractéristiques des systèmes « features » en présence de données entachées de bruits ou même incomplètes (Rolmnick et al., 2017).

Étant donné le manque de données, nous avons considéré la classe des techniques d'apprentissage par renforcement. Plus particulièrement, nous avons considéré la technique « Twin Delayed Deep Deterministic policy gradient algorithm » (TD3), qui peut s'attaquer à un espace d'actions continu, et qui a montré une plus grande efficacité en vitesse d'apprentissage et en performance que les approches existantes (Fujimoto et al., 2018).

Nous formulons, dans ce qui suit, le problème de l'accès dans l'IoT comme un problème d'apprentissage par renforcement, dans lequel un agent trouve itérativement un facteur de blocage sous-optimal, permettant de réduire le conflit à l'accès.

2.5.1. *Formulation du problème*

Dans l'apprentissage par renforcement (Sutton et al., 2019), nous avons deux entités principales, un environnement et un agent. Le processus d'apprentissage se fait par l'interaction entre ces entités afin que l'agent puisse optimiser un revenu total. À chaque étape t, l'agent obtient une représentation de l'état s_t de l'environnement et

choisit une action a_t, basée sur celle-ci. Ensuite, l'agent applique cette action sur l'environnement. En conséquence, l'environnement passe dans un nouvel état s_{t+1} et l'agent reçoit une récompense r_tR correspondant à cette transition ainsi que la représentation du nouvel état. Cette interaction peut être modélisée comme un processus de décision Markovien $M = (S, A, P, R)$, avec S l'espace d'état, A l'espace d'action, P la dynamique de transition et R le revenu. Le comportement de l'agent est défini par sa politique $\pi: S \longrightarrow A$ qui permet d'associer un état à une action lorsqu'il s'agit d'un système déterministe, où une distribution d'actions, lorsqu'il est probabiliste. L'objectif d'un tel système est de trouver la politique optimale $\pi^\star$ permettant de maximiser le revenu cumulé.

Dans le problème du contrôle d'accès des IoT, nous définissons un processus de décision Markovien (MDP) discret, où l'état, l'action et le revenu sont définis comme suit :

- L'État : Étant donné la non disponibilité du nombre de terminaux tentant l'accès à un instant k donné, l'état, que nous considérons, est basé sur les estimées mesurées. Une seule mesure de ce nombre étant forcément très bruitée, nous considérons une suite de plusieurs mesures, qui permettent de mieux révéler l'état présent du réseau. L'état s_k est, ainsi, défini comme le vecteur $(\hat{x}_2^k, \hat{x}_2^{k-1}, \cdots, \hat{x}_2^{k-H})$ où H représente l'horizon de mesure.

- L'action : À chaque étape, l'agent doit sélectionner le facteur de blocage p qui devra être considéré par les objets IoT. Cette valeur est continue et déterministe, dans le problème que nous considérons, c'est-à-dire qu'un même état s_k donnera toujours une même action a_k.

- Le revenu : C'est un signal que reçoit l'agent de l'environnement suite à l'exécution d'une action. Ainsi, à l'étape k, l'agent obtient un revenu r_k en conséquence de l'action a_k qu'il a réalisé dans l'état s_k. Ce revenu va permettre à l'agent de connaître la qualité de l'action exécutée. L'objectif de l'agent étant de maximiser ce revenu.

  Contrairement aux problèmes d'apprentissage par renforcement classique, l'optimum est ici connu est donné par l'équation (2). Le revenu est donné par l'équation suivante,

  $$r_k = e^{-|x_2^* - \hat{x}_2^k|} \tag{6}$$

  Le revenu est donc maximal lorsque l'action choisie permet d'obtenir un nombre d'équipements tentant l'accès $(\hat{x}_2^k)$ est égal à l'optimum x_2^*.

Cependant, comme la mesure $\hat{x}_2^k$ est entachée de bruit, cela rejaillit sur la mesure du revenu.

L'objectif d'un tel système est de trouver la probabilité de blocage permettant de maximiser la récompense moyenne. Ce qui revient à réduire la distance entre les mesures du nombre de terminaux tentant l'accès et l'optimum. Afin de réaliser cet objectif nous nous basons sur l'algorithme TD3.

L'algorithme TD3 est une approche de type Acteur-Critique, où l'acteur est un réseau de neurones qui décide dans un état particulier de l'action à prendre ; le réseau Critique permet de connaître la valeur d'être dans un état et de choisir une action particulière. TD3 permet de résoudre la problématique de la surévaluation dans l'estimation de la valeur (Thrun et Schwartz, 1993), en introduisant deux réseaux Critiques et en prenant le minimum entre les deux estimations. Cette approche est particulièrement intéressante dans notre cas vu la présence inhérente d'erreurs de mesures.

2.5.1. *Système de régulation des arrivées*

Le diagramme de la Figure 2.7 décrit le système permettant de contrôler le nombre de tentatives des objets IoT. Ce système se base sur la diffusion du facteur de blocage aux terminaux, à travers les blocs d'information du système (SIBs) qui sont propagés, et plus particulièrement, à travers le bloc SIB Type14, qui permet de diffuser les paramètres du blocage d'accès (ETSI, 2019).

Suite à la réception du facteur de blocage, les terminaux voulant effectuer une transmission exécutent l'ACB qui leurs permet le passage aux étapes suivantes avec une probabilité p, qui est calculé par notre contrôleur à base de TD3. Ces terminaux peuvent, par conséquent, tenter l'accès, en choisissant un préambule au hasard parmi les préambules disponibles. Connaissant l'état des préambules, le gNodeB peut estimer le nombre de tentatives qu'il y a eu. Ce chiffre étant très bruité, car le modèle donné ne permet d'estimer que des moyennes. On se base sur une estimée moyenne du nombre d'équipements. Nous utilisons pour cela une moyenne glissante.

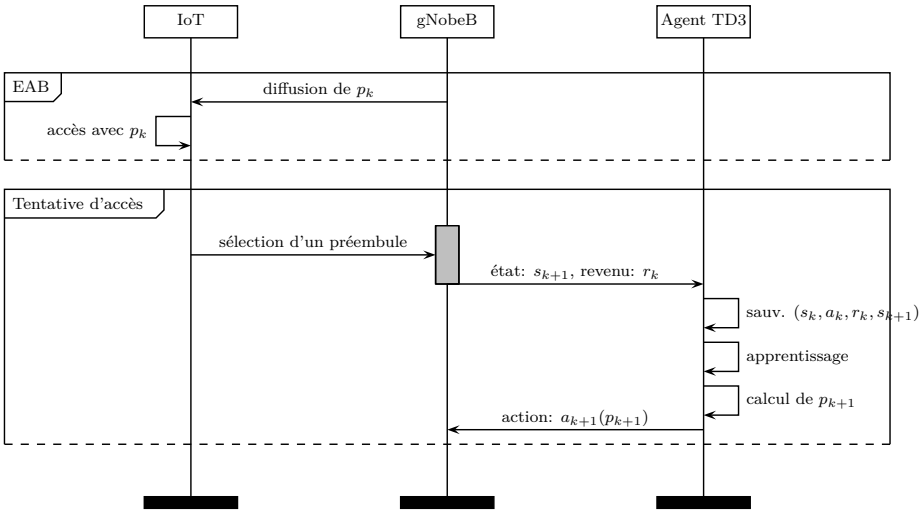

Figure 2.7. *Système de régulation des arrivées.*

Le contrôleur que nous avons proposé reçoit ces mesures, augmentées du revenu, à la fin de chaque préambule. Le revenu obtenu lui permettra de connaître la qualité des actions prises. Ces différentes données sont mises dans une mémoire des expériences passées. C'est un sous-ensemble aléatoire de cette mémoire qui lui permettra d'apprendre de façon robuste et de choisir, par la suite, une nouvelle action.

Ces différentes actions sont répétées de façon cyclique.

2.5. Évaluation des performances

Après avoir décrit notre proposition de contrôleur d'accès, nous évaluons, dans cette section, ses performances, en utilisant un environnement de simulation que nous avons construit sous (Simpy, 2020).

Nous avons considéré une antenne NB-IoT dans laquelle les demandes d'accès arrivent suivant une loi de Poisson avec un taux moyen entre deux arrivées de 0.018 s. Nous avons considéré un nombre de préambules N égal à 16, avec une fréquence d'arrivée égale à 0.1 s. Dans le système considéré, chaque équipement tentant l'accès pourra le faire pour un maximum de 16 fois. Au-delà de cette limite, le terminal abandonne la transmission.

La performance de notre contrôleur qui se base sur la technique TD3 est comparée à une approche adaptative. Nous avons considéré un horizon de mesure H égale à 10.

L'utilisation d'une fenêtre de mesure plus importante ne permet pas une amélioration significative des performances, ce qui signifie qu'une fenêtre de 10 mesures permet de suffisamment refléter l'état réel du réseau.

L'approche adaptative consiste à augmenter graduellement la probabilité de blocage lorsque le nombre de tentatives est au-delà d'un seuil prédéfini au-dessus de la valeur optimale. Lorsque la valeur est inférieure à un seuil prédéfini en dessous de la valeur optimale, la probabilité de blocage est graduellement réduite, afin de laisser plus de terminaux tenter l'accès.

Dans les Figures 2.8 et 2.9, sont exprimées la probabilité de blocage pour les deux stratégies considérées. La technique adaptative (cf. Figure 2.8) commence par une probabilité d'accès de 1 et s'adapte au fur et à mesure aux conditions de trafic, qui changent suivant une loi de Poisson. Pour la stratégie qui se base sur l'algorithme TD3, il y a une première phase, qui dure 200 s, où l'algorithme essaye d'explorer l'espace des actions suivant une loi uniforme (voir la Figure 2.5). Ce n'est qu'après cette phase que l'algorithme commence à exploiter son apprentissage, qui s'affine au fur et à mesure des expériences.

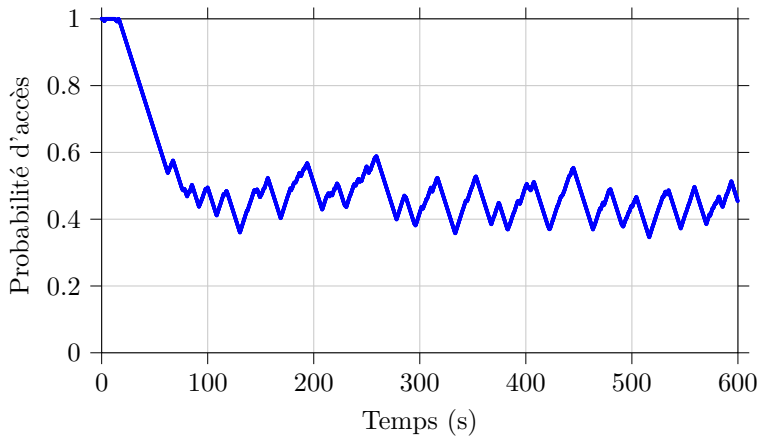

Figure 2.8. *Probabilité d'accès avec le contrôleur adaptatif.*

Nous pouvons noter que sous TD3 (cf. Figure 2.5) les actions futures n'ont pas de liens avec les actions passées, contrairement au cas adaptatif. En effet, les valeurs des actions peuvent changer du tout au tout, puisqu'elles ne dépendent que de l'état du réseau, qui peut changer très rapidement.

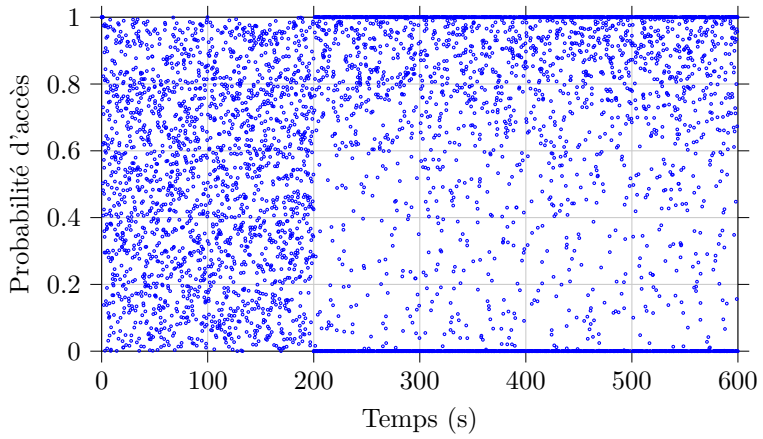

Figure 2.9. *Probabilité d'accès avec le contrôleur utilisant TD3.*

Les Figures 2.10 et 2.11 décrivent l'impact des lois de commande, décrites précédemment, sur la latence moyenne des accès. Nous ne considérons pas dans ces tracés les terminaux qui ont abandonné la transmission, des suites d'un nombre de tentatives maximales. Même si on peut remarquer, dans la Figure 2.11, quelques terminaux avec des latences légèrement supérieures à celle de la Figure 2.10, la latence est globalement du même ordre. C'est-à-dire que l'algorithme TD3 ne présente aucun avantage en termes de latence.

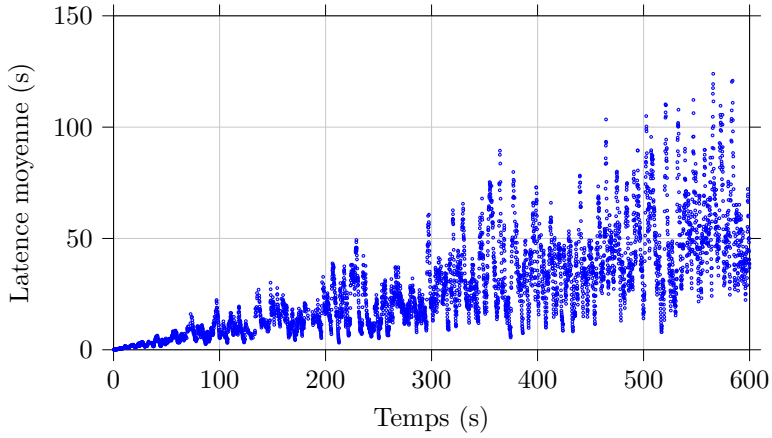

Figure 2.10. *Latence moyenne des terminaux avec le contrôleur adaptatif.*

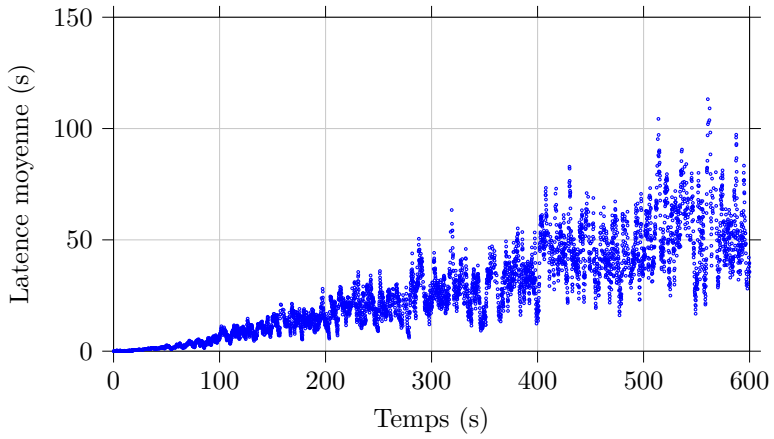

Figure 2.11. *Latence moyenne des terminaux avec le contrôleur utilisant TD3.*

Même si TD3 ne présente pas d'avantage particulier en termes de latence, on peut remarquer dans la Figure 2.13 qu'après une phase d'exploration, le revenu s'améliore de façon très significative. Cette récompense est clairement supérieure au contrôleur adaptatif, qui présente une récompense réduite et très variable (cf. la Figure 2.12). En effet, la moyenne de la récompense dans TD3 est de l'ordre de 13.91% alors que le contrôleur adaptatif présente une récompense de l'ordre de 3.6%. Cette récompense reflète le fait que sous TD3, le nombre moyen de terminaux tentant l'accès se rapproche davantage de l'optimum. Ce résultat peut-être aussi constaté dans la Figure 2.15, qui montre que le nombre de tentatives avec TD3 se rapproche davantage de l'optimum qui est égal à 15.49. En effet, le nombre moyen de tentatives en utilisant le contrôleur adaptatif est égal à 30.12 (voir la Figure 2.10), alors qu'il est égal à 19.6 pour notre approche.

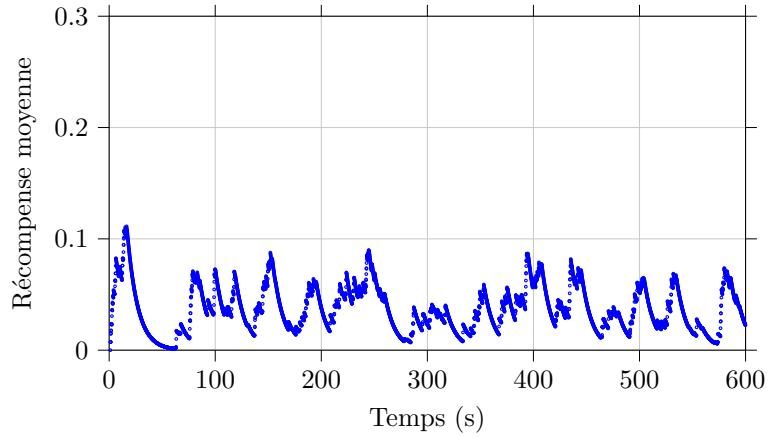

Figure 2.12. *La récompense moyenne avec le contrôleur adaptatif.*

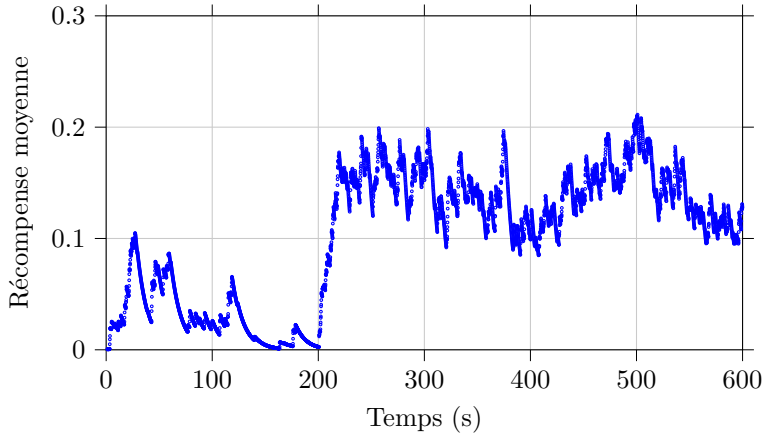

Figure 2.13. *La récompense moyenne avec le contrôleur utilisant TD3.*

Nous pouvons remarquer dans la Figure 2.14 que la technique adaptative ne permet pas de contrôler correctement le nombre de tentatives. En effet, on atteint très souvent des nombres bien plus importants que l'optimum. Cela provoque beaucoup de collisions à l'accès et des retentatives d'accès. On voit aussi que le nombre d'abandons reste relativement important, par rapport au contrôleur TD3 (cf. la Figure 2.15). Ce dernier, après la phase d'exploration, arrive à réduire significativement le nombre d'abandons. Ce qui démontre l'efficacité de l'approche proposée.

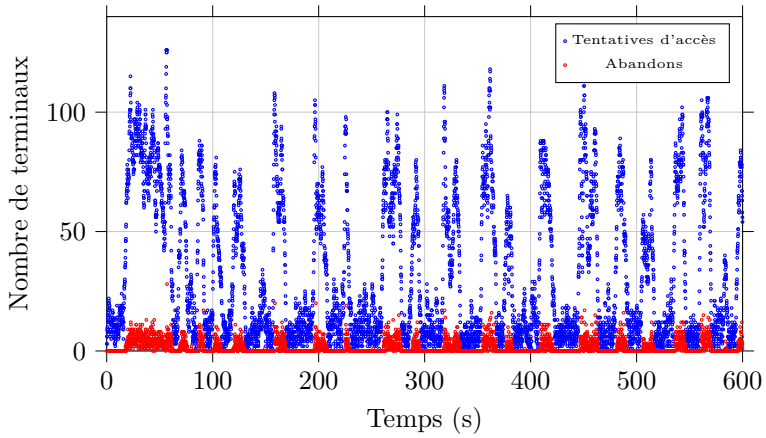

Figure 2.14. *Tentatives d'accès et abandons avec le contrôleur adaptatif.*

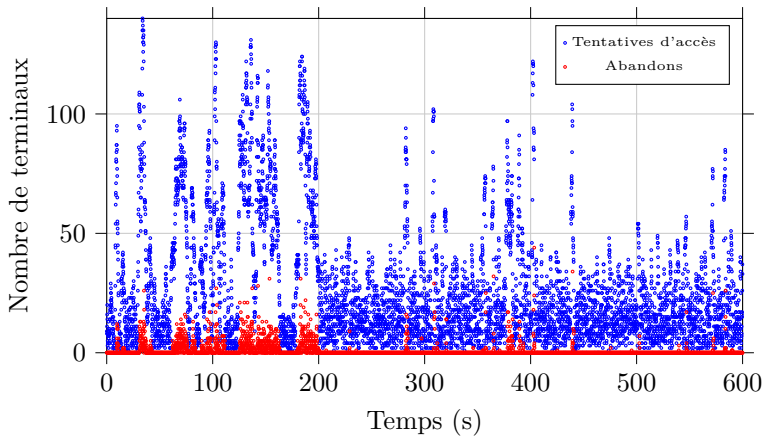

Figure 2.15. *Tentatives d'accès et abandons avec le contrôleur utilisant TD3.*

Il faut noter qu'en ayant recours à notre approche basée sur l'apprentissage par renforcement, nous avons une amélioration de performance au fur et à mesure des tentatives d'accès. La limite reste, cependant, les erreurs d'estimations qui entrainent des erreurs dans le calcul de la récompense, d'où l'importance d'avoir des estimateurs précis.

2.5. Conclusions et perspectives

Dans cet article, nous avons proposé un mécanisme pour contrôler la congestion du réseau d'accès, qui est considérée comme l'un des problèmes les plus critiques pour les objets IoT. Nous avons proposé de s'attaquer à la congestion à sa racine en gérant efficacement les accès aléatoires de ces équipements grâce à l'utilisation du mécanisme ACB.

Le mécanisme de contrôle d'accès proposé est différent des méthodes conventionnelles, qui reposent généralement sur des heuristiques simples. En effet, la technique proposée repose sur les récentes avancées en apprentissage par renforcement profond, à travers l'utilisation de l'algorithme TD3. L'approche proposée a, en plus, l'avantage d'apprendre de son environnement et pourrait donc permettre de s'adapter à la variation du schéma d'accès.

Les résultats de simulations permettent de montrer la supériorité de l'approche proposée qui arrive à maintenir un nombre de tentatives d'accès proche de l'optimum, en dépit de l'absence d'informations exactes sur le nombre de tentatives d'accès. Ce travail permet aussi de montrer le potentiel d'utiliser les techniques d'apprentissage dans les environnements où l'état ne peut être connu avec précision.

Dans le cadre de nos futurs travaux, nous prévoyons d'améliorer l'estimation du nombre de tentatives en utilisant des techniques d'apprentissage.

2.10. Bibliographie

3GPP (2011). RAN Improvements for Machine-type Communications. *Technical Report (TR) 37.868*, Version 11.0.0. [En ligne]. Disponible à l'adresse : https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2630

3GPP (2015). Cellular system support for ultra-low complexity and low throughput Internet of Things (CIoT). *TR 45.820*. [Online]. Available: https://portal.3gpp.org/ChangeRequests.aspx?q=1&versionId=46751&release=187

3GPP (2016). E-UTRA and E-UTRAN; LTE Physical Layer, General Description (Release 13). *Rapport Technique TS 36.201*, Sophia Antipolis, France. [En ligne]. Disponible à l'adresse : https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2424

3GPP (2017a). Physical layer Measurements (Release 13),). *TS 36.214, version13.5.0*. [En ligne]. Disponible à l'adresse : https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2428

3GPP (2017b). User Equipment (UE) radio transmission and reception (Release 13). *TS 36.101, version 13.9.0*. [En ligne]. Disponible à l'adresse : https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2411

3GPP (2018a). Consideration on self-evaluation of IMT-2020 for mMTC connection density. *R1-1801796*. [En ligne]. Disponible à l'adresse : http://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_92/Docs/R1-1801796.zip

3GPP (2018b). Radio Resource Control (RRC); Protocol specification. TS 36.331. [En ligne]. Disponible à l'adresse : https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2440

3GPP (2019a). Study on Cellular Internet of Things (IoT) support and evolution for the 5G System. *TR 23.724, version 16.1.0*. [En ligne]. Disponible à l'adresse : http://www.3gpp.org/DynaReport/23724.htm

3GPP (2019b). Medium Access Control (MAC) protocol specification. TS 36.321, V14.12.0, [En ligne]. Disponible à l'adresse :

https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2437

3GPP (2020). Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (Release 16). TS 36.211, Version 16.2.0, Jun. 2020. [En ligne]. Disponible à l'adresse : https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2425

Adhikary, A., Lin, X., Wang, Y.-P.-E. (2016). Performance Evaluation of NB-IoT Coverage. *the IEEE 84th Vehicular Technology Conference (VTC-Fall)*, Monréal, QC, Canada.

Agnes, S. (2015). Intel, Nokia and Ericsson collaborate on NB-LTE wireless for IoT. *Capacity Mag.*, *Tech. Rep. 915, pp. 1–5,* Helsinki, Finlande.

Ali, M. S., Hossain, E., Kim D. I. (2017), LTE/LTE-A Random Access for Massive Machine-Type Communications in Smart Cities. *IEEE Communications Magazine*, vol. 55 , no. 1, pp. 76 -83. DOI: 10.1109/MCOM.2017.1600215CM.

Baracat, G., Brito, J. (2018). NB-IoT Random Access Procedure Analysis. *the Latin-American Conference on Communications (LATINCOM)*. Guadalajara, Mexique.

Bouzouita, M., Hadjadj-Aoul, Y., Zangar, N., Rubino, G., Tabbane, S. (2015). Multiple access class barring factors algorithm for M2M communications in LTE-advanced networks. *18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, MSWiM'15, page 195–199*, New York, NY, USA.

Bouzouita, M., Hadjadj-Aoult, Y., Zangar, N., Tabbane, S. (2016). On the risk of congestion collapse in heavily congested M2M networks. *2016 International Symposium on Networks, Computers and Communications (ISNCC)*, Yasmine Hammamet, p. 1-5, doi: 10.1109/ISNCC.2016.7746063.

Bouzouita, M., Hadjadj-Aoul, Y., Zangar, N., Rubino, G. (2019). Estimating the number of contending iot devices in 5G networks: Revealing the invisible. *Transactions on Emerging Telecommunications Technologies*, 30(4):e3513. e3513 ett.3513.

Cheng, R.C., Chen, J., Chen, D. W., Wei, C. H. (2015). Modeling and Analysis of an Extended Access Barring Algorithm for Machine-Type Communications in LTE-A Networks. *IEEE Transactions on Wireless Communications,* vol. 14, no. 6, pp. 2956-2968. DOI: 10.1109/TWC.2015.2398858.

ETSI (2011). Lte; evolved universal terrestrial radio access (e-utra); physical channels and modulation. *Technical Report (TR) ETSI TS 136.211 Version 10.0.0*. [En ligne]. Disponible à l'adresse : https://www.etsi.org/deliver/etsi_TS/136200_136299/136211/15.02.00_60/ts_136211v150200p.pdf

ETSI (2019). LTE ; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification (3GPP TS 36.331 version 14.12.0 Release 14). *ETSI TS 136 331 Version 14.12.0*. [En ligne]. Disponible à l'adresse : https://www.etsi.org/deliver/etsi_ts/136300_136399/136331/13.00.00_60/ts_136331v130000p.pdf

Flore, D. (2016). 3gpp Standards for the Internet-of-Things. *gSMA MIoT*. [En ligne] Disponible à l'adresse : http://www.3gpp.org/news-events/3gpp- news/1766-iot_progress

Fujimoto, S., van Hoof, H., Meger, D (2018). Addressing Function Approximation Error in Actor-Critic Methods. *the Thirty-fifth International Conference on Machine Learning (ICML)*, Stockholmsmässan, Stockholm, Suède.

GSMA Intelligence. *Mobile IoT (LPWA)*. [En ligne]. Disponible à l'adresse : https://www.gsma.com/iot/mobile-iot/

Harwahyu, R., Cheng, R.G., Wei, C.H., Sari, R.F (2018). Optimization of Random Access Channel in NB-IoT. *IEEE Internet Things Journal*. Vol. 5, no.1, pp. 391–402. DOI: 10.1109/JIOT.2017.2786680.

Harwahyu, R., Cheng, R.G., Tsai, W.J., Hwang, J.K., Bianchi, G. (2019). Repetitions vs Retransmissions: Trade-off in Configuring NB-IoT Random Access Channels. *IEEE Internet Things Journal*. Vol. 6, no. 2, pp. 3796–3805. DOI: 10.1109/JIOT.2019.2891366.

Huawei. *NB-IoT Commercial Premier Use Case Library*. [En ligne]. Disponible à l'adresse : https://www.gsma.com/iot/wp-content/uploads/2017/12/NB-IoT-Commercial-Premier-Use-case-Library-1.0_Layout_171110.pdf

Hwang, J.K., Li, C.F., Ma, C. (2019). Efficient Detection and Synchronization of Superimposed NB-IoT NPRACH Preambles. *IEEE Internet Things*. Vol.6, pp. 1173–1182. DOI: 10.1109/JIOT.2018.2867876.

Jin*, H.,* Toor, W. T., Jung*, B. C.,* Seo, J.B. (2017). Recursive Pseudo-Bayesian Access Class Barring for M2M Communications in LTE Systems. *IEEE Transactions on Vehicular Technology*. Vol. 66, No. 9, pp 8595 - 8599. DOI: 10.1109/TVT.2017.2681206.

Jeon, W.S., Seo, S.B., Jeong, D.G. (2018). Effective frequency hopping pattern for ToA estimation in NB-IoT random access. *IEEE Transactions on Vehicular* Technology. Vol. 67, no. 10, pp. 10150 – 10154. DOI: 10.1109/TVT.2018.2857447.

Jiang, N., Deng, Y., Condoluci, M., Guo, W., Nallanathan, A., Dohler, M. (2018). RACH Preamble Repetition in NB-IoT Network. IEEE Communications Letters. Vol. 22,no. 6, pp. 1244–1247. DOI: 10.1109/LCOMM.2018.2793274.

Kafle, V.P., Fukushima, Y., Harai, H. (2016). Internet of Things standardization in ITU and prospective networking technologies. *IEEE Commun. Mag., vol. 54, no. 9, pp. 43–49*.

Lin, X., Adhikary, A., Wang, Y. -. Eric (2016). Random Access Preamble Design and Detection for 3GPP Narrowband IoT Systems. in *IEEE Wireless Communications Letters*, vol. 5, no. 6, pp. 640-643. doi: 10.1109/LWC.2016.2609914.

Liu, J., Agiwal, M., Qu, M., Jin, H. (2020). Online Control of Preamble Groups with Priority in Massive IoT Networks. *IEEE Journal on Selected Areas in Communications*. DOI : 10.1109/JSAC.2020.3018964

Migabo, E.M., Djouani, K.D., Kurien, A.M. (2020). The Narrowband Internet of Things (NB-IoT) Resources Management Performance State of Art, Challenges, and Opportunities. *IEEE Access*.

Mwakwata, CB., Malik, H., Mahtab Alam, M., Le Moullec, Y., Parand, S., Mumtaz, S (2019). Narrowband Internet of Thing (NB-IoT): from Physical (PHY) and Media Access Control (MAC) Layers Perspectives. *Sensors 19 (11), 2613*.

Narayanan, S., Tsolkas, D., Passas, N., Merakos, L. (2018). NB-IoT:Acandidate technology for massive IoT in the 5G era. *IEEE 23rd Int. Workshop Comput. Aided Modeling Design Commun. Links Netw (CAMAD). pp. 1–6*, Barcelone, Espagne.

Park, A., Lim, X. (2016). Adaptive Access Class Barring Method for Machine Generated Communications. *Mobile Information Systems.* Pp.1-6. DOI : 10.1155/2016/6923542.

Ray, B. (2017). NB-IoT case studies. *Link-labs*. [En ligne] Disponible à l'adresse : https://www.link-labs.com/blog/nb-iot-case-studies

Rolnick, D., Veit, A., Belongie, S. J., Shavit, N. (2017). Deep Learning is Robust to Massive Label Noise. *CoRR, vol. abs/1705.10694, arXiv*.

Savaux, V., Le Guen, M., Kanj, M. (2020) A Tutorial on NB-IoT Physical Layer Design. *IEEE Communications Surveys & Tutorials.*

Simpy (2020). *Discrete event simulation for Python.* [En ligne] Disponible à l'adresse : https://simpy.readthedocs.io/en/latest/

Sun, Y., Tong, F., Zhang, Z., He, S. (2017). Throughput Modeling and Analysis of Random Access in Narrow-band Internet of Things. *IEEE Internet Things Journal*. Vol.5, no. 3, pp. 1485–1493. DOI: 10.1109/JIOT.2017.2782318.

Sutton, R.S. and Barto, A. G. (2019). Reinforcement learning: An Introduction, *2nd edition, MIT press*.

Thrun, S. and Schwartz, A. (1993). Issues in using function approximation for reinforcement learning. *The 1993 Connectionist Models Summer School Hillsdale*, NJ. Lawrence Erlbaum.

Toor, W. T., Jin, H., (2017). Comparative Study of Access Class Barring and Extended Access Barring for Machine Type Communications. *International Conference on Information and Communication Technology Convergence (ICTC).* DOI: 10.1109/ICTC.2017.8191051.

Wang, Y.-P.-E., Lin, X. , Adhikary, A., Grovlen, A., Sui, Y., Blankenship, Y., Bergman, J., Razaghi, H. S. (2017). A primer on 3GPP narrowband Internet of Things. *IEEE Commun. Mag., vol. 55, no. 3, pp. 117–123*.