

Performances sur couche MAC des utilisateurs IEEE802.16e

Contents

2.1	Formulation du problème	39
2.1.1	Objectifs	39
2.1.2	Hypothèses	40
2.1.3	Environnement	40
2.2	Modèle	41
2.2.1	Analyse par Point Fixe	41
2.2.2	Performances	44
2.3	Analyses numériques	46
2.3.1	Validation du modèle	47
2.3.2	Impact des paramètres de communication	48
2.3.3	Influence du nombre de codes	50
2.3.4	Evolution du délai de réponse	52
2.3.5	Analyse globale	53
2.4	Contributions et perspectives	54

2.1 Formulation du problème

2.1.1 Objectifs

Ce chapitre vise à établir les performances sur couche MAC atteintes par les utilisateurs du standard IEEE802.16e. Cet objectif principal implique un certain nombre d'objectifs secondaires tels que de répondre au besoin de définir un modèle analytique général. Par ailleurs, ce chapitre tente d'explicitier l'influence de chacun des paramètres

de communication sur couche MAC vis à vis des performances et temps d'établissement de connexion des utilisateurs IEEE802.16e. Enfin, fort de ce nouvel outil d'analyse, le dernier enjeu sera d'apporter une critique pertinente quant au choix des valeurs préconisées et utilisées par défaut par le standard.

2.1.2 Hypothèses

Afin d'établir un modèle de performances aussi général que possible, nous utilisons un cadre hypothétique minimal. Aussi, nous nous baserons sur un faible nombre d'hypothèses de base. Premièrement, le trafic des utilisateurs est saturé. Cette hypothèse indique que chaque utilisateur de la cellule IEEE802.16e dispose constamment d'une information à transmettre. De ce fait, il n'y a aucun temps d'inactivité, ou de veille, entre la fin d'une transmission de données et le besoin d'engager une nouvelle connexion en vue d'un autre transfert. Cette hypothèse est souvent considérée comme une des hypothèses les plus fortes communément prise dans les travaux d'évaluation de performances. Néanmoins, cette hypothèse relève ici d'une importance moindre. En effet, le standard IEEE802.16 se base sur des communications orientées connexion. Aussi, seul l'intensité des connexions entrantes dans le système doivent être pris en considérations.

Deuxièmement, la section 1.3.4 précise qu'après l'envoi d'une requête de ressource, les utilisateurs engagent un compte à rebours nommé T_3 . Ce *timer* T_3 force les utilisateurs demandeurs de ressources à laisser le temps à la station de base de formaliser et de transmettre sa réponse. En temps normal, la transmission des données s'effectue dès réception de la réponse. Néanmoins, pour les besoins de l'étude, nous supposons que quels que soient les délais de réception des réponses de la station de base, les transmissions de données en résultant ne s'engagent qu'après expiration du décompte du *timer* T_3 . Cette hypothèse s'explique en considérant deux de nos objectifs. Premièrement, nous voulons définir un modèle aussi général que possible, mais en gardant une certaine maîtrise de sa complexité calculatoire. Considérez le paramètre t_r comme fixe, et non comme une variable aléatoire, répond donc à cet objectif. Deuxièmement, notre étude vise principalement à évaluer les possibilités de réduction du délai de connexion. Fixer t_r à sa valeur maximale nous assure de prendre en compte les délais maximum de connexion.

2.1.3 Environnement

Nous considérons une cellule IEEE802.16e où sont réunis un nombre fini d'utilisateur. Ces utilisateurs suivent un régime saturé. De ce fait, ils désirent continuellement engager de nouveaux appels. Aussi, chaque utilisateur engagera une nouvelle connexion du moment que sa demande précédente a été conclue.

Pour formaliser sa demande, un mobile choisit un code de modulation. Ces codes sont issus de ceux attribués au requête de bande passante (*Bandwidth Ranging Request*). Après envoi de la requête, le mobile laisse à la station de base le temps de recevoir

et traiter sa demande. Pour cela, le mobile attend l'écoulement d'un certain nombre de trames MAC. Ce procédé d'attente correspond au *timer* T_3 décrit précédemment. En cas d'absence de réponse à expiration du *timer* T_3 , le mobile engage un processus de *backoff* caractérisé par une taille minimal de fenêtre, un multiplicateur en cas d'échecs successifs et un nombre maximum de retransmissions autorisé. Par contre, comme indiqué dans les hypothèses ci-dessus, si une réponse est disponible, la transmission de données se fera à partir de l'expiration de ce *timer* T_3 . Enfin nous identifions des temps (ou *slots*) systèmes. Un temps système correspond à la durée d'une trame MAC IEEE802.16e.

2.2 Modèle

2.2.1 Analyse par Point Fixe

Principe

Notre approche théorique se base sur une analyse par Point Fixe (Kumar et al., 2006). Elle caractérise le comportement statistiques du processus de *backoff* engagé par les utilisateurs. De ces seules statistiques, l'analyse par point fixe permet d'extraire les performances globales d'accès au canal (taux de tentative, probabilité de collision, ...). Toutefois, cette analyse fut initialement développée pour le standard IEEE802.11. Aussi, nous en avons développé une nouvelle approche dont l'originalité réside dans sa parfaite adaptation au standard IEEE802.16e. L'analyse requiert que les temps d'attentes qui y sont étudiés et ceux de transmission de données soient indépendants. Ceci est vrai dans le cas du IEEE802.11, puisque durant les transmissions de données, tous les processus de *backoff* des utilisateurs en compétition sont gelés durant tout le temps de la transmission. Dans le cas du standard IEEE802.16, ce pré-requis reste vrai mais pour d'autres raisons. En effet, deux faits combinés entraînent que les processus de *backoff* étudiés soient indépendants des transmissions de données : le processus de *backoff* s'engage sur les demandes de bande passante or, celles-ci sont émises sur un canal de transmissions qui leur est propre : le *Bandwidth Request Ranging Interval*, appartenant à l'espace de temps *Ranging Subchannel* de la figure 1.2. De plus, le fait que le standard se base sur des communications orientées connexions implique que, pour un même utilisateur, le système admettra plusieurs connexions en parallèle. Une transmission de données est donc indépendante de toute autre transmission ou nouvel engagement de connexion. Ainsi, un mobile pourra engager un nouveau processus de *backoff* dès lors que sa précédente demande de connexion est terminée (qu'elle ait été satisfaite ou pas). Fort de ces spécificités propres au standard IEEE802.16e, il nous est possible d'en développer une nouvelle analyse par point fixe tout à fait adaptée au cadre de l'étude désirée.

Déroulement

La figure 2.1 décrit l'évolution du processus de *backoff* d'un mobile. Ce chronogramme a été épuré des temps de transmissions en vertu des caractéristiques énoncées précédemment. Ainsi, nous pouvons simplifier la représentation des temps de *backoff* par la concaténation des différentes fenêtres issues des transmissions et retransmissions d'un mobile. La figure montre l'état du processus de *backoff* pour trois transmissions de données. Notons que les événements d'envoi de requête sont représentés par des points. L'attente induite par les fenêtres de *backoff* ne s'engage qu'une fois le temps t_r écoulé. Les événements représentés sur cette figure sont les suivants. La première demande réussit après deux tentatives, la seconde aboutit malgré deux collisions et la troisième n'aboutit qu'après trois collisions et une quatrième tentative réussit. Remarquons que la connaissance d'un échec ou l'engagement d'un envoi de données ne commencent qu'après expiration du temps t_r , tel que défini dans nos hypothèses de départ (voir section 2.1.2).

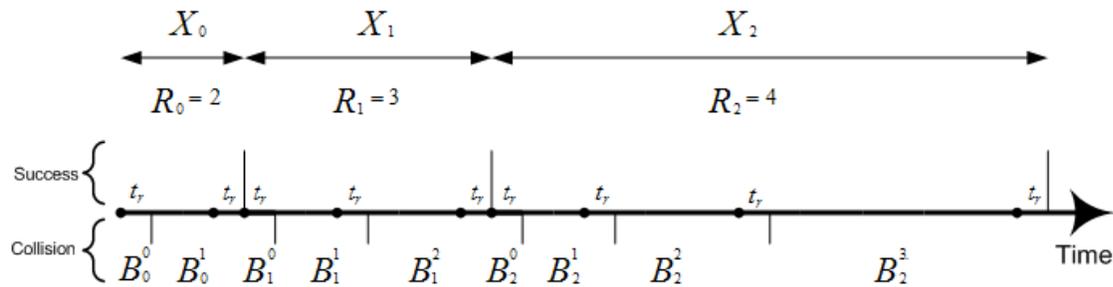


FIGURE 2.1 – Chronogramme du processus de backoff dans le IEEE802.16e

Soit X_j , la variable aléatoire du temps total requis pour transmettre la demande de ressource relative au paquet j . R_j désigne le nombre de tentatives nécessaires à la transmission des requêtes de ressources pour le paquet j . B_j^i correspond quant à lui au temps d'attente aléatoirement choisi parmi la fenêtre de *backoff* pour le paquet j et la retransmission i , enfin b_j représente le temps moyen de B_j^i pour tout paquet j . Ce temps d'attente aléatoire se caractérise entre autres par une durée minimum : CW_{min} , un nombre maximal de retransmission k , et un multiplicateur p . En conséquence des définitions de variables ci-dessus, le temps total nécessaire à l'engagement d'un transfert de données est dicté par la relation suivante :

$$X_j = R_j \cdot t_r + \sum_{i=0}^{R_j} B_j^i \quad (2.1)$$

Soit γ la probabilité de collision ressentie par un mobile, et k le nombre maximum de retransmissions autorisées. Nous pouvons maintenant définir β comme le taux moyen de tentative par temps du système. Notons ici qu'à partir du moment où les paramètres de *backoff* sont les mêmes pour toutes les stations de la cellule, leur comportement est globalement identique, et de ce fait, la collision ainsi que le taux de tentative expérimentés par chacun sont aussi les mêmes. En fait cet élément se base implicitement sur

l'approximation de découplage faites initialement par Bianchi (Bianchi, 2000). Cette hypothèse définit que toute tentative d'accès au canal est soumise à la même probabilité de collision quel que soient les stations émettrices et l'état de leur processus de *backoff*. De plus, elle décrit les opportunités de transmission (ou *slot*) comme indépendantes les unes des autres. De même, les station émettrice agissent aussi de manière indépendante. Il faut d'ailleurs noter que cette approximation est de plus en plus juste pour un nombre d'utilisateurs croissant. Or une cellule IEEE802.16e a justement pour but de servir un grand nombre d'utilisateur. Maintenant, pour les besoins du modèle, nous identifions quelques égalités triviales :

$$E\left(\sum_{i=0}^{R_j} B_j^i\right) = b_0 + \gamma b_1 + \dots + \gamma^k b_k \quad (2.2)$$

$$E(X) = E(R).t_r + E\left(\sum_{i=0}^{R_j} B_j^i\right) \quad (2.3)$$

$$E(R) = 1 + \gamma + \dots + \gamma^k \quad (2.4)$$

$$\beta = \frac{E(R)}{E(X)} \quad (2.5)$$

$$= \frac{E(R)}{E(R).t_r + E\left(\sum_{i=0}^{R_j} B_j^i\right)} \quad (2.6)$$

$$= \frac{1}{t_r + \frac{E\left(\sum_{i=0}^{R_j} B_j^i\right)}{E(R)}} \quad (2.7)$$

$$= \frac{1}{t_r + \frac{b_0 + \gamma b_1 + \gamma^2 b_2 + \dots + \gamma^k b_k}{1 + \gamma + \gamma^2 + \dots + \gamma^k}} \quad (2.8)$$

Notons que b_m , le temps moyen de *backoff* pour m retransmissions, est défini par :

$$b_m = \frac{p^m CW_{min} - 1}{2} \quad \text{et} \quad b_0 = 0 \quad (2.9)$$

Comme énoncé précédemment, tous les mobiles d'une cellule ressentent la même probabilité de collision et donc le même taux de tentative. Nous supposons alors que le nombre de tentatives effectuées par les autres mobiles est distribué suivant une loi binomiale avec pour paramètres β , $n - 1$ et N . En fait, la probabilité de collision, P_{coll} , pour un mobile tentant une demande de ressource est donnée par la relation suivante :

$$P_{coll}(\beta) = 1 - \Gamma(\beta) \quad (2.10)$$

où $\Gamma(\beta)$ représente la probabilité qu'aucun des i autres mobiles effectuant une tentative en même temps n'utilise le même code de modulation utilisé par le mobile. Cette dernière probabilité nous est donnée par la relation suivante :

$$\Gamma(\beta) = \sum_{i=0}^{n-1} \binom{n-1}{i} \beta^i (1-\beta)^{n-i-1} \left(1 - \frac{1}{N}\right)^i \quad (2.11)$$

Partant de ce constat mathématique, l'état d'équilibre du système est caractérisé par la solution de l'équation du point fixe suivante :

$$\gamma = P_{coll}(\beta) \quad (2.12)$$

Préalablement à la suite de cette étude, nous établissons que l'existence et l'unicité du point fixe peut être aisément obtenues à partir des travaux de (Kumar et al., 2006)

2.2.2 Performances

Le modèle mathématique étant maintenant défini, nous pouvons y introduire les relations de performances telles que le nombre moyen des arrivées d'appel par temps système, ainsi que la distribution du nombre des requêtes en attente au niveau de la station de base. A chaque temps système, les demandes de ressources arrivant au niveau de la station de base sont mises en file d'attente. Nous supposons ici que la file d'attente se compose d'un tampon infini. Soit H , la variable aléatoire du nombre de services effectués par la station de base durant une trame MAC.

Dès lors, nous déterminons le nombre de requêtes arrivant par trame MAC. En effet, le nombre cumulé des requêtes émises durant l'espace de contention du lien montant aboutit à un certain nombre de requêtes entrant dans la file d'attente de la station de base. Rappelons qu'une demande de ressource aboutit à la station de base uniquement si cette requête n'est pas entrée en conflit avec une autre requête modulée par le même code choisi parmi N . Soit $P(Z_t = j)$ la probabilité que la station de base reçoive j demandes de ressources sur les N codes orthogonaux, durant le temps t ; $j \in \{0, 1, \dots, N\}$. $P(X_t = i)$ représente la probabilité que i mobiles engagent une demande de ressources sur le même espace de contention t . Ces deux probabilités sont définies par les relations suivantes :

$$P(X_t = i) = \binom{n}{i} \beta^i (1 - \beta)^{n-i} \quad (2.13)$$

$$P(Z_t = j) = \sum_{i=j}^n P(Z_t = j | X_t = i) P(X_t = i) \quad (2.14)$$

La probabilité conditionnelle $P(Z_t = j | X_t = i, N)$ peut être calculée via l'expression récursive suivante :

$$\begin{aligned} P(Z_t = j | X_t = i, N) &= \sum_{k=0, k \neq 1}^i \binom{i}{k} \left(1 - \frac{1}{N}\right)^{i-k} \left(\frac{1}{N}\right)^k P(Z_t = j | X_t = i - k, N - 1) \\ &\quad + \binom{i}{1} \left(1 - \frac{1}{N}\right)^{i-1} \frac{1}{N} P(Z_t = j - 1 | X_t = i - 1, N - 1) \end{aligned} \quad (2.15)$$

La condition initiale pour $P(Z_t = j | X_t = i, N)$ étant donnée par :

$$P(Z_t = j | X_t = i, 0) = \begin{cases} 1 & \text{si } j = 0 \\ 0 & \text{autrement} \end{cases} \quad (2.16)$$

Par conséquent, le nombre moyen d'arrivées par espace de contention est donné par :

$$\lambda = \sum_{k=1}^N kP(Z = k) \quad (2.17)$$

En l'occurrence, la stabilité du système est garantie dès lors que le nombre moyen des arrivées d'appel est inférieur au nombre moyen des services effectués durant un même intervalle de temps. Aussi, la condition de stabilité est la suivante :

$$\lambda < \mu \quad (2.18)$$

Maintenant, nous désignons par M_t l'état de la chaîne à temps discret de Markov au temps t , et par Q_{ij} la probabilité de transitions entre un état $M_t = i$ et $M_{t+1} = j$. Ces probabilités sont déterminées par les relations suivantes :

$$Q_{ij} = \begin{cases} P(Z = j) & \text{si } i = 0 \\ P(H \geq i)P(Z = j) + \sum_{k=0}^{i-1} P(H = k)P(Z = j - i + k) & \text{sinon} \end{cases} \quad (2.19)$$

L'ensemble des probabilités de transition compose la matrice de transition d'état du système de la manière suivante :

$$Q = \begin{pmatrix} Q_{00} & Q_{01} & \dots \\ Q_{10} & Q_{11} & \dots \\ \dots & \dots & \dots \end{pmatrix} \quad (2.20)$$

Enfin, soit π la distribution stationnaire. Du moment que le système est ergodique, le système suivant d'équations linéaires caractérise la solution unique de distribution stationnaire du système :

$$\begin{cases} \pi = \pi Q \\ \sum_{n=0}^{\infty} \pi(n) = 1. \end{cases} \quad (2.21)$$

Le nombre moyen de requêtes dans le tampon de la station de base est défini par $S(\gamma)$ suivant :

$$S(\gamma) = \sum_{k=0}^{\infty} k\pi(k) \quad (2.22)$$

Par la formule de Little ([Gelenbe et Pujolle, 1987](#)), le temps moyen de séjour d'une requête dans la file d'attente est donné par le rapport entre le nombre moyen de requêtes $S(\gamma)$ et le taux d'arrivée μ . Néanmoins, dans le standard IEEE802.16e, le délai est rallongé d'un temps système, car les requêtes sont considérées comme intégrant le système qu'à partir de la trame MAC suivant celle ayant servi à l'envoi de la requête. Ainsi, le temps moyen de traitement d'une requête, une fois émise, est donné par la relation suivante :

$$D(\gamma) = 1 + \frac{S(\gamma)}{\mu} \quad (2.23)$$

2.3 Analyses numériques

Cette section présente plusieurs résultats de simulations obtenus à partir de plusieurs simulateurs. Le premier correspond au simulateur numérique *Matlab*. Le deuxième correspond à un programme écrit en C. Ce simulateur émule le comportement à temps discret de nombreux utilisateurs en compétition sur le slot de contention. Ainsi nous avons reproduit le comportement dynamique du système. De cela nous en extrayons les performances propres du système. En outre, le simulateur comportemental exploite les hypothèses définies dans le paragraphe 2.1.2, exceptée celle de l'approximation de Bianchi. En ce qui concerne la résolution du système d'équation décrit dans la relation (2.21), les outils logiciels ne nous permettent pas de pouvoir considérer une taille de file d'attente allant jusqu'à l'infini. Aussi, nos calculs utilisent une taille maximale de file correspondant à 500000 requêtes d'appel en attente.

Nous présentons alors différentes comparaisons entre des résultats issus du modèle théorique et d'autres obtenus par simulation comportementale. Ces comparaisons nous permettent de valider l'ensemble de nos résultats analytiques.

En se basant sur les recommandations du standard IEEE802.16e, nous avons utilisé pour ces simulations, les grandeurs désignées par défaut par le standard. Néanmoins, nous avons choisi de réduire de manière drastique la valeur initiale du compte à rebours T_3 . En effet, prenons tout d'abord en considération que la durée moyenne d'une trame MAC est de 1ms et que le *timer* T_3 est initialement défini à 50ms. Ensuite, notons que la durée de la trame MAC peut varier entre 0,5ms et 2ms. Ainsi, le *timer* T_3 peut introduire une attente allant jusqu'à cent trames MAC. Aussi, nous choisissons de donner au *timer* T_3 la valeur maximale $t_r = 10$. Ce choix vise à augmenter l'agressivité des stations appelantes. Ainsi, nous pourrions définir le comportement du système dans lequel les connexions engagées ont une plus grande exigence de réactivité. Ce choix trouvera d'ailleurs sa pleine justification avec l'analyse du délai. Par ailleurs, nous aimerions traiter ici le nombre de codes assignés aux demandes de ressources. Le standard définit un large spectre de 256 codes orthogonaux. Ces codes sont divisés en quatre familles telles que décrites dans le chapitre précédent 1.2.2. Pour les besoins de l'étude, nous devons émuler un système à lourde charge en réduisant le nombre de codes disponibles. Néanmoins, le système peut gérer efficacement un grand nombre de mobiles à partir de 32 codes disponibles pour la famille des demandes de ressources.

Dans ce chapitre, nous comparons les résultats analytiques obtenus pour le taux de tentative et la probabilité de collision expérimentés par les mobiles de la cellule IEEE802.16e. Nous traitons ici le taux de tentatives des mobiles afin de mettre en avant de quelle manière les paramètres de communications, et en particulier k , t_r et N , affectent ce taux. Par ailleurs, l'analyse du point fixe nous permet d'approfondir nos critiques des préconisations de l'IEEE. Grâce aux figures suivantes, nous pouvons aussi apprécier comment ces paramètres influent sur la probabilité de collision.

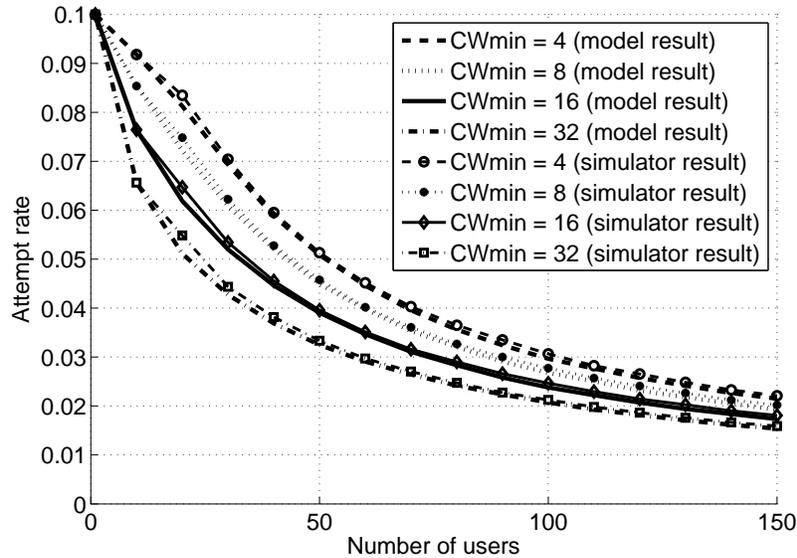


FIGURE 2.2 – Représentation de β vs. n : Taux de tentative en fonction du nombre d'utilisateurs n , pour différentes fenêtres initiales de backoff CW_{min} . $k = 16$, $t_r = 10$, $N = 4$.

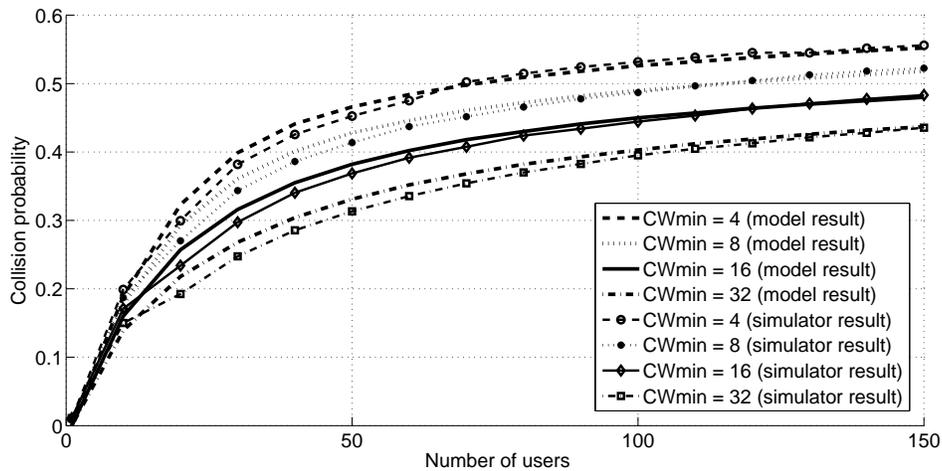


FIGURE 2.3 – Représentation de P_{coll} vs. n : Probabilité de collision en fonction du nombre d'utilisateurs n , pour différentes fenêtres initiales de backoff CW_{min} . $k = 16$, $t_r = 10$, $N = 4$.

2.3.1 Validation du modèle

Les figures 2.2 et 2.3 expriment la comparaison entre les résultats obtenus via le modèle analytique et ceux issus des simulations du standard. La comparaison des résultats montrent une forte corrélation à la fois pour le taux de tentatives et la probabilité de collision. Ces résultats nous permettent de clairement valider le modèle analytique de point fixe au travers des comportements et des valeurs qui sont similaires.

Nous faisons d'ailleurs remarquer que de légers écarts de valeurs sont observables sur les figures de validation. Ces écarts peuvent être la conséquence de l'approximation de Bianchi (Bianchi, 2000) faites dans le modèle théorique et non dans le simulateur comportemental. D'ailleurs, la forte corrélation des résultats justifie l'intégration de cette hypothèse dans le modèle de performance développé ici.

2.3.2 Impact des paramètres de communication

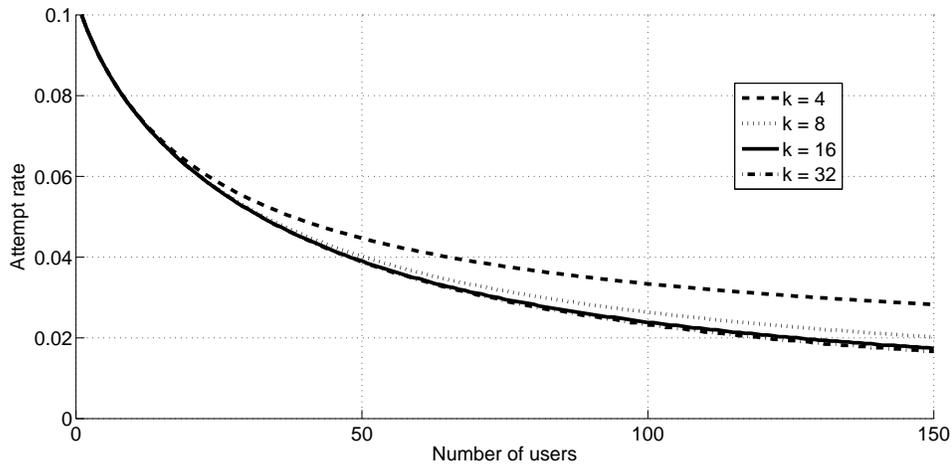


FIGURE 2.4 – Représentation de β vs. n : Taux de tentative en fonction du nombre d'utilisateurs n , pour différentes limites de retransmissions k . $CW_{min} = 16$, $t_r = 10$, $N=4$

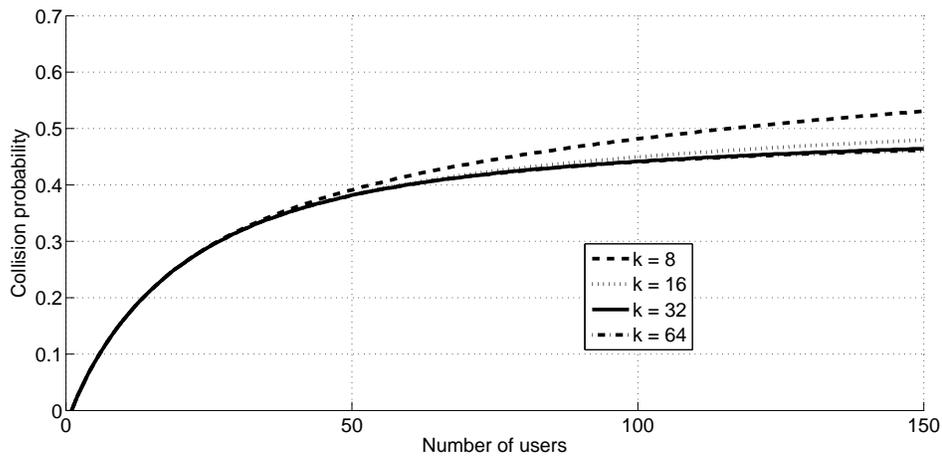


FIGURE 2.5 – Représentation de P_{coll} vs. n : Probabilité de collision en fonction du nombre d'utilisateurs n , pour différentes limites de retransmissions k . $CW_{min} = 16$, $t_r = 10$, $N=4$

Les figures 2.4 et 2.5 montrent respectivement le taux de tentative et la probabilité de collision atteints en fonction du nombre limite de retransmissions. Nous montrons que

pour une limite de retransmissions supérieure à huit essais, le taux de tentative en est très faiblement affecté. Mais pour des valeurs inférieures ($k < 8$), ce taux augmente graduellement, ainsi que la probabilité de collision. Aussi, nous jugeons ici qu'une limite de retransmission supérieure à huit essais est inutile, mais que des valeurs inférieures aboutiraient à un trop grand nombre de collisions des requêtes en entrée de la station de base.

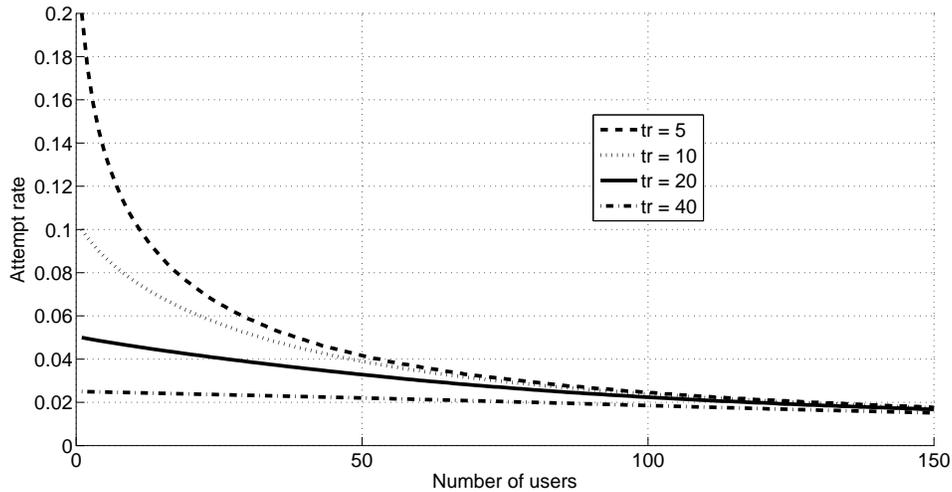


FIGURE 2.6 – Représentation de β vs. n : Taux de tentative en fonction du nombre d'utilisateurs n , pour différents délais autorisés d'attente de réponse t_r . $CW_{min} = 16$, $k = 16$, $N = 4$

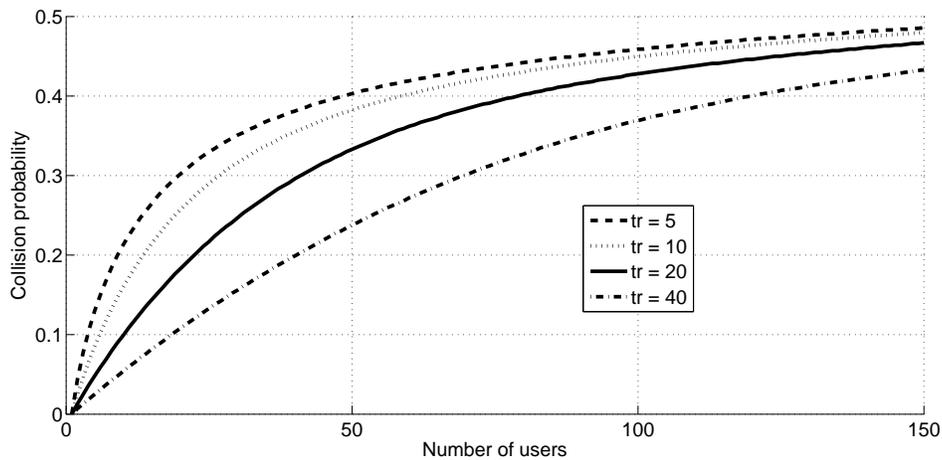


FIGURE 2.7 – Représentation de P_{coll} vs. n : Probabilité de collision en fonction du nombre d'utilisateurs n , pour différents délais autorisés d'attente de réponse t_r . $CW_{min} = 16$, $k = 16$, $N = 4$

Les figures 2.6 et 2.7 montrent l'influence du délai autorisé d'attente d'une réponse sur le taux de tentative et la probabilité de collision. La figure témoigne ainsi de l'énorme

importance de ce paramètre : puisque chaque envoi de requête doit attendre sa réponse durant un certain nombre de trames, le taux de tentative est, lui, très sensible à la durée limite d'attente (t_r). Nous observons que le taux de tentative diminue rapidement à mesure que la probabilité de collision augmente. De plus, notons que la valeur de t_r définit le taux de tentative maximum accessible par un mobile. Enfin, nous observons que la probabilité de collision suit un comportement croissant concave. Ainsi, pour un grand nombre d'utilisateurs, les probabilités de collision, pour toutes les valeurs de t_r possibles, convergent peu à peu entre elles. Ceci est dû au fait qu'avec un grand nombre d'utilisateurs, beaucoup de retransmissions auront lieu. Par conséquent, les temps d'attente t_r deviendront négligeables en comparaison des fenêtres d'attente croissante de *backoff*. Nous ferons alors remarquer que le mécanisme de *backoff* joue ainsi son rôle préventif d'un trop grand nombre de collisions. Par exemple, avec $N = 4$, la probabilité de collision de 25 utilisateurs est seulement multipliée par deux pour une valeur de t_r divisée par huit (t_r variant de manière décroissante de 40 à 5 trame MAC).

2.3.3 Influence du nombre de codes

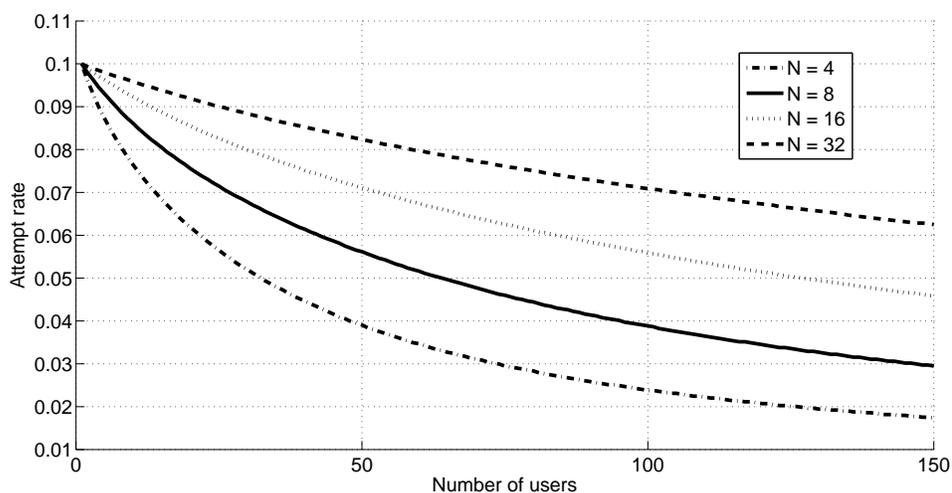


FIGURE 2.8 – Représentation de β vs. n : Taux de tentative en fonction du nombre d'utilisateurs n , pour différentes plages de codes N . $CW_{min} = 16$, $k = 16$, $t_r = 10$

Les figures 2.8 et 2.9 explicitent l'impact du nombre de codes sur le taux de tentative et la probabilité de collision. Remarquons qu'ici N n'a pas une influence directe sur le taux de tentative β en vertu de la relation (2.8), mais qu'en modifiant la probabilité de collision de par la relation (2.11), il modifie alors le point d'équilibre de la relation (2.10), et donc influence le taux de tentative. Ainsi, nous observons que l'augmentation de la plage des codes disponibles pour l'envoi des requêtes, améliore grandement les performances de ces envois. La probabilité de collision est quant à elle rapidement augmentée avec un nombre d'utilisateurs grandissant. C'est à ce stade là de l'étude que nous pensons qu'un procédé de partitionnement des codes d'envoi peut être une piste

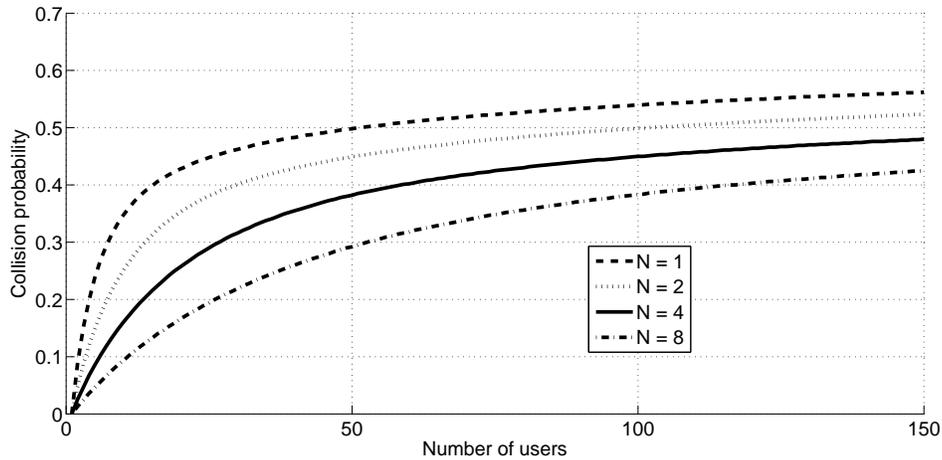


FIGURE 2.9 – Représentation de β vs. n : Probabilité de collision en fonction du nombre d'utilisateurs n , pour différentes plages de codes N . $CW_{min} = 16$, $k = 16$, $t_r = 10$

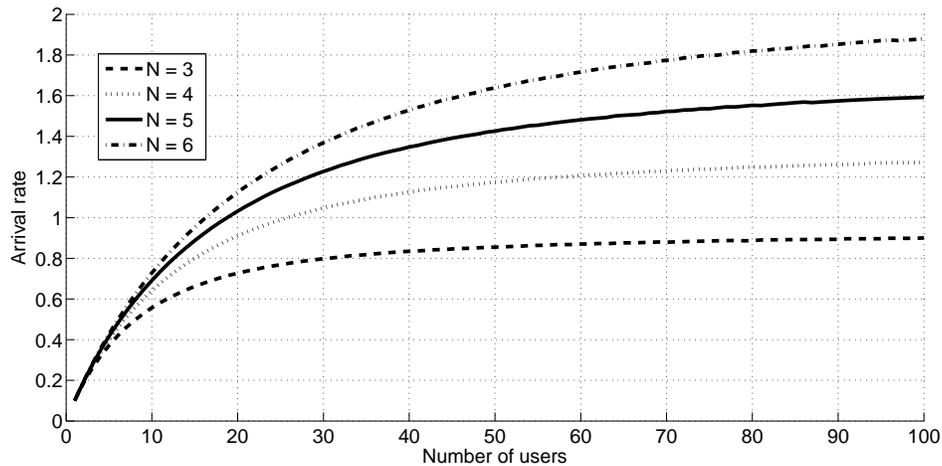


FIGURE 2.10 – Représentation de λ vs. n : Nombre moyen des arrivées en fonction du nombre d'utilisateurs n , pour différentes plages de codes N . $CW_{min} = 16$, $k = 16$, $t_r = 10$

d'étude menant à l'amélioration des performances d'établissements de connexion pour les trafics les plus sensibles à ce délai.

De là, nous analysons le nombre moyen de requêtes λ , arrivant au niveau de la station de base. La figure 2.10 montre le nombre moyen des arrivées en fonction du nombre de codes. Ce nombre moyen augmente logiquement avec le nombre de codes disponibles pour l'envoi des requêtes. Nous observons aussi sur cette figure qu'un faible nombre de codes implique une faible augmentation des arrivées pour un nombre d'utilisateurs grandissant. Par exemple, cent utilisateurs aboutissent à largement deux fois plus d'arrivées de requêtes avec six codes qu'avec seulement trois codes. En fait, il est logique qu'un plus grand nombre de codes réduise fortement la probabilité que

deux mobiles ou plus choisissent le même code pour leur envois de requête. Nous utilisons ici trois et six codes afin de limiter nos temps de simulations, mais le standard IEEE802.16e est destiné à gérer un bien plus grand nombre de codes comme défini au chapitre 1.2.2

2.3.4 Evolution du délai de réponse

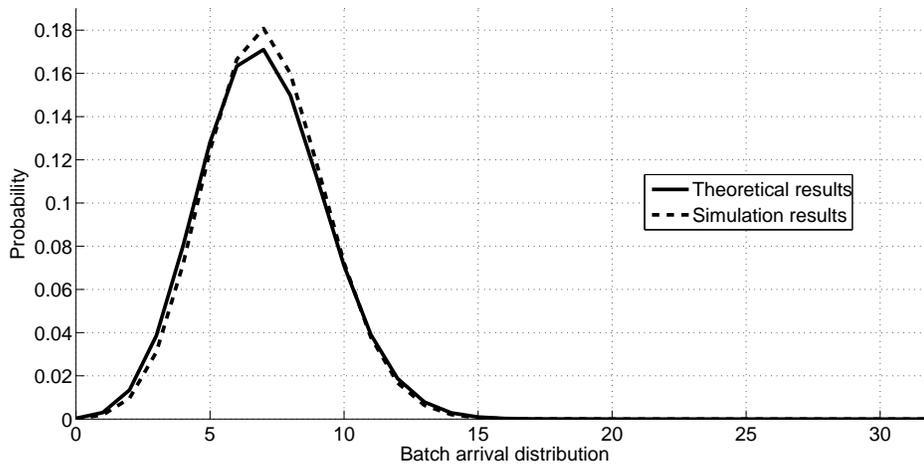


FIGURE 2.11 – Représentation de la distribution des arrivées, $P(Z = x)$, en fonction du nombre d'utilisateurs. $CW_{min} = 16, k = 16, N = 32, n = 150, t_r = 10$

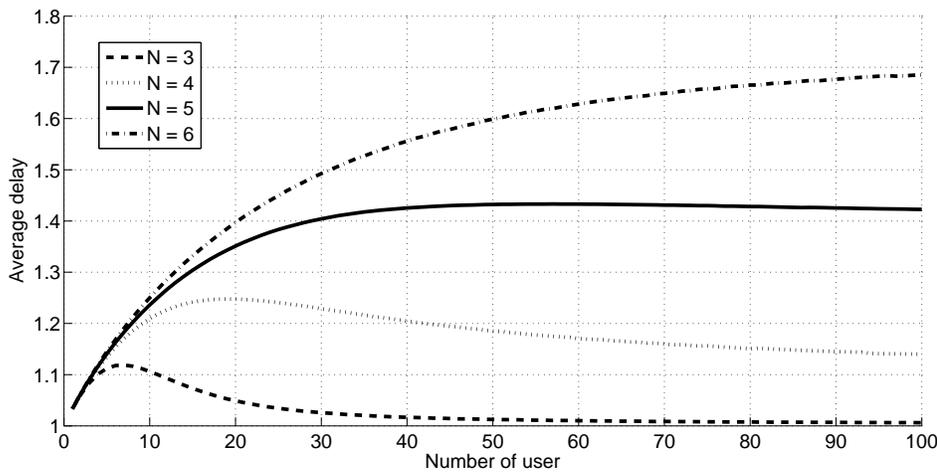


FIGURE 2.12 – Représentation $D(\gamma)$ vs. n : Délai moyen en fonction du nombre d'utilisateurs n , pour différentes plages de codes N . $CW_{min} = 16, k = 10, t_r = 10, \mu = 3$

Concernant le modèle de file d'attente développé, nous considérons que le modèle est validé dès lors que le processus d'arrivée issu du modèle et du simulateur corres-

pondent. Or, la figure 2.11 atteste de cette correspondance. Cette figure représente les résultats numériques et comportementaux de la distribution des arrivées : $P(Z = j)$. La figure montre que pour 150 utilisateurs et une plage de 32 codes d'envoi de requête, les deux sources de résultats correspondent quasi-parfaitement. Par ailleurs, nous avons observé sur nos résultats une correspondance grandissante à mesure que le nombre d'utilisateurs augmente. Nous justifions cette augmentation de corrélation entre les deux sources de résultat par l'influence de l'approximation de Bianchi comme expliqué dans la section 2.3.1.

Nous terminons notre analyse point par point des résultats par celle du délai moyen entre l'envoi d'une demande de ressources à la station de base, et la réception d'une réponse par le mobile demandeur. La figure 2.12 montre les résultats de ce délai en fonction du nombre d'utilisateurs et du nombre de codes. Premièrement, nous remarquons que le délai moyen est largement plus faible que le délai autorisé d'attente de la réponse (paramètre t_r) préconisé dans le standard. D'autres résultats, basés sur une plage de 32 codes de requêtes témoignent d'un délai maximal de 10 trames MAC. C'est d'ailleurs fort de cette première observation, que les résultats présentés ici utilisent le paramètre : $t_r = 10$. Une telle valeur implique un taux de tentative élevé. Mais même dans cet environnement à forte charge, les résultats obtenus sont loin de justifier une attente pouvant aller jusqu'aux 50 trames MAC énoncées par le standard. Donc, si une perte a lieu sur le canal lors de l'envoi d'une requête, le mobile émetteur est à même de prendre connaissance de cette perte dès les premières trames MAC suivantes, sans avoir à attendre pour cela un temps qui pourrait nuire à la QoS désirée par le service. En outre, cette figure répond à une des questions ayant motivé cette étude : comment déterminer le temps t_r approprié en fonction du nombre d'utilisateurs ? La figure 2.12 montre que pour chaque configuration, le délai moyen atteint une valeur maximale qui dépend du nombre de codes. Ainsi, nous donnons ici un moyen simple de déterminer le paramètre du *timer* T_3 adéquat à l'environnement de communication et basé sur l'observation du délai moyen.

Par ailleurs, en observant le délai moyen en fonction du nombre de codes, nous pouvons apprécier la capacité du système à supporter un grand nombre d'utilisateurs. En effet, la figure 2.12 montre un délai maximum pour un nombre d'utilisateurs n spécifique. Ce maximum augmente avec le nombre de codes. En fait, le délai diminue au delà de ce nombre n , du fait des collisions intervenant dans l'intervalle de contention. Ces collisions réduisent le nombre de requêtes entrant dans la file d'attente, et par conséquent, le délai moyen des réponses. Ainsi, le délai moyen d'envoi des réponses définit un nombre caractéristique d'utilisateurs. Ce nombre définit alors le maximum d'utilisateurs que le système peut efficacement prendre en charge.

2.3.5 Analyse globale

Les résultats obtenus dans ce chapitre montrent qu'il est possible d'obtenir de grandes améliorations de performances en réduisant uniquement le nombre limite de retransmissions autorisées k . En effet, la réduction de cette valeur conserve un taux de tentative

relativement élevé tout en conservant une probabilité de collision acceptable et peu variable. De plus, l'étude du taux de tentative révèle que le délai d'attente accepté t_r ainsi que le nombre de codes pour les envois de requêtes N , constituent les paramètres ayant les plus forts impacts sur les performances des communications IEEE802.16e. En diminuant le premier et en augmentant le second, le système assure un plus grand nombre de réponses aux requêtes émises. D'une part, le paramètre t_r permet une gestion plus efficace du délai de traitement auquel sont soumises les requêtes au niveau de la station de base. Il permet de prendre plus vite connaissance d'une perte effective d'une requête sur le canal. Donc, il est nécessaire de pouvoir adapter ce paramètre au délai réel de séjour des requêtes dans la file d'attente de la station de base. D'autre part, N caractérise la capacité globale du système à gérer l'envoi de requêtes simultanées. En fonction du nombre de codes, le système peut déterminer le nombre optimal de retransmissions possibles k . Le nombre de codes N , constitue le facteur principal de robustesse vis à vis des collisions. La figure 2.10 témoigne du fait que ce nombre de codes gagne à être évalué en fonction du nombre d'utilisateurs de la cellule. En fait, la capacité du système à offrir aux utilisateurs un certain niveau de service (GoS) dépend presque uniquement du nombre de codes N .

La figure 2.3 montre que par l'augmentation de la fenêtre initiale de *backoff* CW_{min} , nous diminuons aussi la probabilité de collision. Ce résultat est attendu après observation de la figure 2.2. Sur ce point, les constructeurs auront à trouver un compromis de performance entre le taux de tentative individuel et la probabilité globale de collision.

De l'ensemble de cette étude de performances sur couche MAC, nous considérons qu'il est possible d'introduire des mécanismes dynamiques d'adaptation à l'environnement des paramètres de communication. Ainsi, le paramètre t_r serait déterminé en fonction du nombre d'utilisateurs : en connaissance du nombre d'utilisateurs de la cellule (déclaré par un *Initial Ranging*, voir la section 1.2.2) et la plage des codes disponibles, la station de base peut calculer son délai moyen de traitement des requêtes afin de définir son propre paramètre de *timer* T_3 . De plus, les observations faites vis à vis du délai moyen et de la caractérisation de son maximum nous poussent à penser que ces éléments peuvent servir au dimensionnement des cellules IEEE802.16e. La connaissance d'une plage fixe de codes et la prévision du délai de traitement des requêtes, permettent de définir la couverture optimale d'une cellule IEEE802.16e sur une zone géographique à densité de population aussi connue.

2.4 Contributions et perspectives

Ce chapitre propose un modèle analytique complet et général des performances sur couche MAC. La réalisation de ce modèle se base sur une analyse de point fixe (FPA), initialement développée par (Kumar et al., 2006). Par sa reformulation, nous sommes parvenu à réaliser un modèle de performance simple et propre au standard IEEE802.16e.

Ce modèle fournit le comportement du taux de tentative et de la probabilité de

collision expérimentés par les mobiles d'une cellule IEEE802.16e. De plus, l'analyse numérique comprend un large ensemble de résultats obéissant aux variables d'environnement définies par le standard. Ainsi, nous pouvons critiquer l'impact de l'ensemble de ces paramètres sur différentes métriques de performances. En outre, il est important de souligner que nous fournissons à cette étude un couple de figures qui assure la validité et la cohérence de ce modèle de performance. D'une part, l'étude révèle que le délai de réponse des demandes de ressources t_r constitue le principal facteur de performance. Une adaptation de ce paramètre mènera à un gain important de performance et principalement vis à vis du délai d'établissement des connexions. Ensuite, la fenêtre initiale de congestion du *backoff* CW_{min} peut être aussi adaptée afin d'accroître le taux de tentative tout en conservant un taux de collision raisonnablement faible. De plus, l'étude prouve que le nombre limite de retransmission k a, quant à lui, un impact globalement mineur sur les performances, mais qu'une valeur voisine de 8 permettra là aussi un gain de performance appréciable.

D'autre part, nous fournissons ici les statistiques de collisions pour un large nombre d'utilisateurs, impliquant une lourde charge de trafic. Nous observons la grande influence du nombre de codes ainsi que celle du paramètre d'attente t_r . Le premier correspond au principal facteur limitant la capacité de gestion du système. D'ailleurs, ce fait montre clairement l'intérêt de définir un sous-partitionnement des codes entre les différentes classes de service, afin d'offrir à chacune une efficacité de gestion propre à leur trafics.

La seconde étape de cette étude consiste à déterminer les caractéristiques et pistes d'amélioration du nombre moyen de requêtes arrivant au niveau de la file d'attente de la station de base. Nous y observons particulièrement l'impact de la plage des codes utilisés pour la modulation des demandes de ressources. L'étude confirme les conclusions précédentes, à savoir que le nombre de codes définit effectivement la capacité du système à nourrir efficacement la file d'attente de la station de base. De plus, nous fournissons aussi la distribution des arrivées simultanées, ainsi qu'une validation de ces résultats. Ce dernier élément de contribution constitue en fait apport majeur. En effet, il apparaît que l'ensemble des travaux antérieurs considèrent exclusivement les arrivées comme des processus de Poisson. Or nous identifions dans ce chapitre le processus général résultant des influences réelles et combinées de la totalité des paramètres liés aux mécanismes opérant sur la couche MAC du IEEE802.16e.

La dernière partie de l'étude traite le délai moyen entre l'émission d'une demande de ressources et la réception de sa réponse. Le cas d'une file à tampon infini mène aux observations qui ont motivé cette étude. Le délai effectif, même dans un environnement à lourde charge de trafic, est largement inférieur à celui que le standard autorise ($D(\gamma) \ll t_r$) : les mobiles sont autorisés à attendre une réponse durant 50 trames MAC, mais nos résultats montrent que ce délai ne dépasse pas 3 trames MAC pour 50 utilisateurs, acceptant une attente de 5 trames MAC ($t_r = 5$). Notons ici que le délai réel diminue à mesure que ce temps d'attente augmente. Cette dernière remarque mène d'ailleurs à une autre observation : par l'augmentation du paramètre t_r , nous diminuons le nombre moyen d'arrivées, et donc, le délai moyen. Aussi, il est possible de mettre en place une méthode d'adaptation du paramètre d'attente des réponses en

fonction du nombre des utilisateurs et du délai moyen réel de la file de la station de base. Enfin, l'étude de ce délai au niveau de la station de base révèle une conclusion inattendue : le délai permet de déterminer le nombre de codes nécessaires à la gestion la plus efficace du nombre des utilisateurs de la cellule IEEE802.16e.