Partitionnement des codes modulant de connexion

Contents

3.1 Formulation du problème

3.1.1 Motivation et Objectifs

Nos motivations à entreprendre la présente étude s'articulent autour de trois enjeux d'amélioration. Le premier vise à mettre en place un premier élément de contrôle d'admission, directement sur la couche MAC des utilisateurs : le partitionnement des codes sur les classes de service, si il est opéré de façon dynamique, permet de favoriser l'accès aux ressource du système. Cette prioritisation des classes de services se ferai en fonction des ressources encore disponibles au niveau de la station de base. Le deuxième enjeux consiste à offrir une réduction du délai de connexion aux appels temps réel. Les travaux portant sur le délai des communications dans le standard IEEE802.16 atteste de l'importance primordiale de ce délai d'établissement vis à vis du délai global des communications temps réel. Aussi, un partitionnement des codes entre classes permettrait aux classes de services sensibles au délai d'accéder plus rapidement aux ressources du système. Enfin, le dernier enjeux consiste à fournir aux communications un élément de contrôle de congestion des requêtes en entrée du système. En effet, le partitionnement offre la possibilité de limiter le nombre d'arrivées de requêtes. De ce fait, la station de base profiterait d'un allégement notable de sa charge. En outre, ce mécanisme permettrait d'éviter tout débordement de capacité des requêtes en attentes de traitement. De plus, un tel allègement de charge réduirait là encore les délais de traitement des requêtes, et par conséquent, le délai d'établissement des appels temps réel.

Ce chapitre vient en complément des résultats obtenus dans le chapitre précédent. Les conclusions de ce dernier chapitre énoncent l'importance majeure du nombre de codes dans les performances générales de connectivité des utilisateurs IEEE802.16e. Aussi, il nous semble nécessaire d'étudier les performances accessibles par un système associant les classes de service du standard avec un mécanisme de partitionnement des codes utilisés pour l'envoi des demandes de ressources. Ce chapitre vise donc à établir les gains de performances apportés par un mécanisme orignal de partionnement des codes. Partant du modèle initial, nous réalisons ici l'intégration de ce nouveau mécanisme aboutissant à la représentation des gains de performances au travers de plusieurs métriques. Enfin, fort de ce nouveau modèle élargi, le dernier enjeu sera d'apporter une critique pertinente quant au format optimal de ce partitionnement de code.

3.1.2 Environnements

L'environnement conserve l'intégralité des éléments définis dans le chapitre précédent (voir section 2.1.3).

En sus, nous mettons en place ici un mécanisme de partitionnement des codes servant à la modulation des requêtes de bande passante. Ainsi, la plage initiale N des codes associée aux envois de ces requêtes se subdivise en deux sous-plages : N_1 et N_2 ($N_1 + N_2 = N$). La première est entièrement dédiée aux trafics sensibles au délai : UGS et rtPS (voir section 1.3.5). Les codes composants cette sous-plage ne pourront être utilisés que par les mobiles désirant ouvrir une connexion répondant à une de ces deux classes de services prioritaires. Nous identifierons cette catégorie de trafic par la classe « 1 » ou « RT » en tant que classe la plus prioritaire. La deuxième plage de codes sera également disponible aux trafics de type 1, mais elle sera surtout la seule disponible aux trafics dits « 2 » ou « NRT » réunissant les connexions insensibles aux variations du délai : nrtPS et BE (voir section 1.3.5). Ainsi, nous établissons le principe de partionnement en interdisant aux trafics les moins prioritaires d'accéder à une partie des codes disponibles pour les demandes de ressources. Afin de rendre ce mécanisme et son modèle associé aussi général que possible, nous définissons ici α comme la probabilité qu'un trafic de classe 1 choisisse un code de la sous-plage N_1. Naturellement, $1 - \alpha$ correspond à la probabilité que ce même trafic choisisse un des codes de la sous-plage N_2. Notons que cette caractéristique nous permet de faire varier la nature du principe de partionnement : depuis un partionnement rigide où chaque trafic dispose de sa propre sous-plage de code (cas où $\alpha = 1$) ; vers une séparation plus flexible où les trafics les plus prioritaires disposent de ressources supplémentaires, mais impliquant une compétition avec les trafics moins prioritaires. Remarquons que le cas particulier où $\alpha = \frac{N_1}{N}$ correspond à une distribution uniforme du choix de code pour la classe 1. Par ailleurs, signalons qu'à l'intérieur de chaque sous-plage N_1 ou N_2, les codes sont choisis d'une manière uniforme, quelle que soit la probabilité α.

Enfin, nous définissons respectivement r et $1 - r$ comme les probabilités qu'un mobile engageant une nouvelle connexion le fasse respectivement dans la classe de trafic 1 et 2.

3.2 Modèle

3.2.1 Analyse par Point Fixe

Afin de développer une nouvelle analyse par Point Fixe propre au principe de partitionnement des codes, certaines données doivent être définies :

Soit $R_{j,k}$, le nombre de tentatives nécessaires à la transmission des requêtes de ressources pour le paquet j d'une classe k, $B^i_{j,k}$ correspond quant à lui au temps d'attente aléatoirement choisi parmi la fenêtre de *backoff* pour la i-ième retransmission du paquet j de la classe k. Enfin, soit $b_{i,k}$, le temps moyen de $B^i_{j,k}$ pour tout paquet j de la classe k.

Le standard IEEE802.16e définit $b_{i,k}$ à travers la relation (3.1). p_k est le facteur multiplicateur pour la classe k qui augmente la taille de la fenêtre de *backoff* à chaque tentative d'envoi de requête. i correspond au numéro actuel de la tentative de connexion ($0 \leq i \leq m_k$) où m_k est le maximum de retransmissions autorisées pour la classe k. Enfin, $CWmin_k$ constitue la fenêtre de *backoff* initiale pour la classe k.

$$b_{i,k} = \frac{p_k^i.CWmin_k - 1}{2} \qquad b_{0,k} = 0 \tag{3.1}$$

Nous rappelons au travers de la figure 3.1 le schéma du principe sur lequel se base l'étude temporelle du Point Fixe.

De là, nous déterminons X^k_j, le nombre total de trames requises pour l'envoi, jusqu'à réception ou abandon, de la demande de ressources :

$$X^k_j = R_{j,k}t_r + \sum_{i=0}^{R_{j,k}} B^i_{j,k} \qquad k = \{1,2\} \tag{3.2}$$

Soit β_k le taux de tentative pour les trafics de classe k, $k \in \{1,2\}$ et γ_k la probabilité de collision ressentie par un mobile pour la classe k.

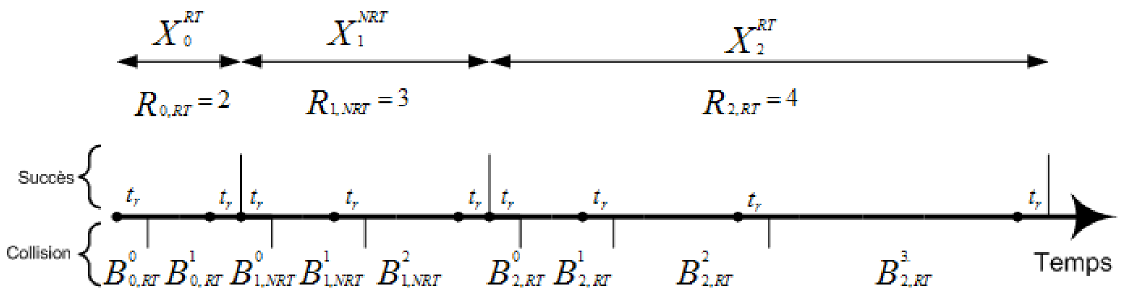

FIGURE 3.1 – *Chronogramme du processus de* backoff *dans le IEEE802.16e avec différenciation de classe de trafic*

Premièrement, nous calculons comme suit le taux de tentative des connexions de classe k :

$$\beta_k = \frac{E(R_k)}{E(X)}, \quad k = \{1,2\} \tag{3.3}$$

$$\begin{aligned}
E(R_k) &= 1 + \gamma_k + \gamma_k^2 + ... + \gamma_k^{m_k} && (3.4)\\
E(X_k) &= E(R_k)t_r + E\Big(\sum_{i=0}^{R_{j,k}} B_{j,k}^i\Big) && (3.5)\\
E\Big(\sum_{i=0}^{R_{j,k}} B_{j,k}^i\Big) &= b_{0,k} + \gamma_k b_{1,k} + \gamma_k^2 b_{2,k} + ... + \gamma_k^{m_k} b_{m_k,k} && (3.6)\\
E(X) &= rE(X_1) + (1-r)E(X_2) && (3.7)
\end{aligned}$$

$E(X_1)$ et $E(X_2)$ correspondent aux temps moyens nécessaires à la finalisation d'un envoi de requête respectivement pour les classes 1 et 2.

De plus, le taux de tentative général β d'un mobile, quelque soit la classe de trafic est donné simplement par la somme des taux de tentatives respectifs à chaque classe :

$$\beta = \beta_1 + \beta_2 \tag{3.8}$$

Soit Γ_k, la probabilité qu'une tentative de classe k ne collisionne avec aucune autre tentative émise au même moment sur le même code. Cette probabilité correspond à la relation suivante :

$$\Gamma_1 = \sum_{i=0}^{n-1} \binom{n-1}{i} (1-\beta)^{n-i-1} \sum_{j=0}^{i} \binom{i}{j} \beta_1^j \beta_2^{i-j} \times \tag{3.9}$$
$$\Big(\sum_{l=0}^{j} \binom{j}{l} \alpha^{l+1}(1-\alpha)^{j-l}(1-\frac{1}{N_1})^l + \sum_{l=0}^{j} \binom{j}{l} (1-\alpha)^{l+1}\alpha^{j-l}(1-\frac{1}{N_2})^{i-j+l} \Big)$$

$$\Gamma_2 = \sum_{i=0}^{n-1} \binom{n-1}{i} (1-\beta)^{n-i-1} \sum_{j=0}^{i} \binom{i}{j} \beta_1^j \beta_2^{i-j} \times \tag{3.10}$$
$$\Big((1-\frac{1}{N_2})^{i-j} + \sum_{l=0}^{j} \binom{j}{l} \alpha^{j-l}(1-\alpha)^l(1-\frac{1}{N_2})^l \Big)$$

$$\gamma_1 = 1 - \Gamma_1(\beta_1, \beta_2) \quad \gamma_2 = 1 - \Gamma_2(\beta_1, \beta_2) \tag{3.11}$$

Au vu de ces relations nous mettons en valeur par la fonction G_k les dépendances liant les grandeurs β_k et γ_k entre elles. Ces fonctions G_k sont en fait les fonctions résultantes des relations étroites existant entre les taux de tentative et les probabilités de collision caractérisant les deux classes de service. Cette relation est explicitée au travers des égalités définissant la relation (3.3).

$$\beta_1 = G_1(\gamma_1, \gamma_2) \quad \beta_2 = G_2(\gamma_1, \gamma_2) \tag{3.12}$$

Suivant l'analyse de Point Fixe, le point d'équilibre du système correspond à la solution du système d'équations défini par les relations (3.11) et (3.12).

3.2.2 Performances

Nous déterminons le nombre de requêtes arrivant par trame IEEE802.16e. 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 la plage des codes accessibles à sa classe de trafic. Afin de fournir le plus grand champ d'étude de performance possible à notre modèle, nous formulons aux travers des relations suivantes la distribution des arrivées pour chaque plage de codes, ainsi que pour chaque classe de trafics.

Pour cela, soit Z_k le nombre de requêtes de classe k reçues avec succès. Notons d'ailleurs que $Z_k \in \{0, 1, .., N\}$ et $Z_1 + Z_2 \leq N$. Soit Z^s le nombre d'arrivées sur la sous-plage de code N_s. Naturellement, Z_k^s correspond au nombre de requêtes réussies pour la classe k sur la sous-plage N_s. Suivant la même nomenclature, nous définissons $X_k^s \in \{0, 1, .., n\}$ le nombre de mobiles transmettant simultanément une requête de classe k sur la sous-plage de codes s. $X_1^1 + X_2^1 + X_1^2 + X_2^2 \leq n$. Enfin, notons que $X_2^1 = 0$ et $X_2^2 = X_2$.

Arrivées par sous-plage « s »

Cette partie définit la distribution des arrivées sur chacune des sous-plages de codes : N_1 et N_2. Notez que ces distributions dépendent de l'étendue de chacune de ces sous-plages. Toutefois, nous excluons la représentation de ces valeurs dans le but de facilité la lecture des relations.

Soit Z^1, $Z^1 \leq N_1$, la variable aléatoire du nombre de tentatives engagées sur la sous-plage N_1.

$$P(Z^1 = i) = \sum_{j=i}^{n} P(Z^1 = i | X_1 = j) P(X_1 = j) \tag{3.13}$$

$$P(X_1 = i) = \begin{pmatrix} n \\ i \end{pmatrix} \beta_1^i (1 - \beta_1)^{n-i} \tag{3.14}$$

$$P(Z^1 = i | X_1 = j) = \sum_{k=i}^{j} P(Z^1 = i | X_1^1 = k) P(X_1^1 = k | X_1 = j) \tag{3.15}$$

$$P(X_1^1 = k | X_1 = j) = \begin{pmatrix} j \\ k \end{pmatrix} \alpha^k (1 - \alpha)^{j-k} \tag{3.16}$$

La probabilité $P(Z^1 = i | X_1^1 = k, N_1)$ peut être calculée en se basant sur la relation (2.15) définie dans la section du chapitre précédent : section 2.2.2. Ce calcul aboutit à la relation suivante :

$$\begin{aligned} P(Z^1 = j | X_1^1 = i, N_1) = & \sum_{k=0, k \neq 1}^{i} \begin{pmatrix} i \\ k \end{pmatrix} (1 - \frac{1}{N_1})^{i-k} (\frac{1}{N_1})^k P(Z^1 = j | X_1^1 = i - k, N_1 - 1) \\ & + \begin{pmatrix} i \\ 1 \end{pmatrix} (1 - \frac{1}{N_1})^{i-1} \frac{1}{N_1} P(Z^1 = j - 1 | X_1^1 = i - 1, N_1 - 1) \end{aligned} \tag{3.17}$$

La condition initiale au calcul récursif est donnée par :

$$P(Z^1 = j | X_1^1 = i, 0) = \begin{cases} 1 & \text{si } j = 0 \\ 0 & \text{sinon} \end{cases} \tag{3.18}$$

Enfin, le nombre moyen d'arrivées sur la sous-plage de codes 1 est donné par la relation suivante :

$$\lambda^1 = \sum_{x=0}^{N_1} x P(Z^1 = x) \tag{3.19}$$

Soit Z^2, $Z^2 \leq N_2$, la variable aléatoire du nombre de tentatives engagées sur la sous-plage N_2.

$$P(Z^2 = i) = \sum_{j=i}^{n} P(Z^2 = i|X = j)P(X = j) \tag{3.20}$$

$$\begin{aligned} P(Z^2 = i|X = j) \quad = \quad & \sum_{k=i}^{j} P(Z^2 = i|X_1^2 + X_2 = k, X = j) \times \\ & P(X_1^2 + X_2 = k|X = j) \end{aligned} \tag{3.21}$$

La probabilité $P(Z^2 = i|X_1^2 + X_2 = k, X = j)$ peut être elle aussi déduite de la relation (2.15), où le dernier terme de la dernière équation est donné par :

$$P(X_1^2 + X_2 = k|X = j) = \sum_{l=0}^{k} \binom{k}{l} \left(\frac{\beta_2^l \beta_1^{j-l}}{\beta^j}\right) \binom{j-l}{k-l} \alpha^{j-k}(1-\alpha)^{k-l} \tag{3.22}$$

Enfin, le nombre moyen d'arrivées sur la sous-plage de codes 2 est donné par la relation suivante :

$$\lambda^2 = \sum_{x=0}^{N_2} xP(Z^2 = x) \tag{3.23}$$

Arrivée par classe « k »

Soit Z_1, $Z_1 \leq N$, la variable aléatoire du nombre de tentatives de classe 1 (temps réel) entrant dans le système.

$$P(Z_1 = i) = \sum_{j=0}^{i} P(Z_1^2 = j)P(Z^1 = i - j) \tag{3.24}$$

La probabilité $P(Z^1 = i - j)$ nous est donnée par la relation (3.13).

$$P(Z_1^2 = i) = \sum_{j=i}^{n} P(Z_1^2 = i|X = j)P(X = j) \tag{3.25}$$

La probabilité $P(X = j)$ est donnée par :

$$P(X = j) = \binom{n}{j} \beta^j (1-\beta)^{n-j} \tag{3.26}$$

$$P(Z_1^2 = i|X = j) = \sum_{k=0}^{j-i} P(Z_1^2 = i|X_2 = k, X = l)P(X_2 = k|X = j) \tag{3.27}$$

$$P(X_2 = k|X = j) = \begin{pmatrix} j \\ k \end{pmatrix} \frac{\beta_2^k \beta_1^{j-k}}{\beta^j} \tag{3.28}$$

$$\begin{aligned} P(Z_1^2 = i|X_2 = k, X = l) &= \sum_{l=i}^{j-k} P(Z_1^2 = i|X_1^2 = l, X_2 = k, X = j) \times \\ & P(X_1^2 = l|X_2 = k, X = j) \end{aligned} \tag{3.29}$$

$$P(X_1^2 = l|X_2 = k, X = j) = \begin{pmatrix} j-k \\ l \end{pmatrix} \alpha^{j-k-l}(1-\alpha)^l \tag{3.30}$$

La probabilité $P(Z_1^2 = i|X_1^2 = l, X_2 = k, X = j)$ est elle aussi obtenue via la relation récursive (2.15).

Enfin, le nombre moyen d'arrivées des appels temps réel est donné par la relation suivante :

$$\lambda_1 = \sum_{x=0}^{n} xP(Z_1 = x) \tag{3.31}$$

Soit Z_2, $Z_2 \leq N_2$, la variable aléatoire du nombre de tentatives de classe 2 (non temps réel) entrant dans le système.

$$P(Z_2 = i) = \sum_{j=i}^{n} P(Z_2 = i|X = j)P(X = j) \tag{3.32}$$

La probabilité $P(X = j)$ est donnée par la relation (3.26)

$$P(Z_2 = i|X = j) = \sum_{k=i}^{j} P(Z_2 = i|X_2 = k, X = l)P(X_2 = k|X = j) \tag{3.33}$$

La probabilité $P(X_2 = k|X = j)$ est donnée par la relation (3.28)

$$\begin{aligned} P(Z_2 = i|X_2 = k, X = l) &= \sum_{l=0}^{j-k} P(Z_2 = i|X_1^2 = l, X_2 = k, X = j) \times \\ & P(X_1^2 = l|X_2 = k, X = j) \end{aligned} \tag{3.34}$$

La probabilité $P(X_1^2 = l|X_2 = k, X = j)$ est donnée par la relation (3.30)

La probabilité $P(Z_2 = i|X_1^2 = l, X_2 = k, X = j)$ est elle aussi obtenue via la relation récursive (2.15).

Enfin, le nombre moyen d'arrivées des appels non temps réel est donné par la relation suivante :

$$\lambda_2 = \sum_{x=0}^{n} xP(Z_2 = x) \tag{3.35}$$

Soit Z la variable aléatoire qui représente le nombre total des arrivées. Dès lors que Z^1 et Z^2 sont indépendantes, la distribution de Z est obtenue par le produit de toutes les configurations possibles d'arrivées entre les requêtes de type RT et NRT.

$$P(Z = j) = \sum_{k=0}^{j} P(Z^1 = k)P(Z^2 = j - k) \tag{3.36}$$

Par la suite, nous définissons les caractéristiques du système quant au traitement des requêtes arrivant en entrée 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 supposerons ici que la file d'attente se compose d'un tampon infini. Soit H, la variable aléatoire du nombre de service effectués par la station de base durant une trame MAC.

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 :

$$\lambda_1 + \lambda_2 < \mu \tag{3.37}$$

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 probabilité sont déterminées par les relations suivantes :

$$Q_{ij} = \begin{cases} P(Z = j|N_1, N_2) & \text{si } i = 0 \\ P(H \geq i)P(Z = j) + \sum_{k=0}^{i-1} P(H = k)P(Z = j - i + k) & \text{sinon} \end{cases} \tag{3.38}$$

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

$$Q = \begin{pmatrix} Q_{00} & Q_{01} & \dots \\ Q_{10} & Q_{11} & \dots \\ \dots & \dots & \dots \end{pmatrix} \tag{3.39}$$

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 :

$$\begin{cases} \pi = \pi Q \\ \sum_{n=0}^{\infty} \pi(n) = 1. \end{cases} \tag{3.40}$$

Le nombre moyenne de requêtes en attentes, ainsi que le temps moyen de séjour dans la file suivent respectivement les relations (2.22) et (2.23) du chapitre précédent.

3.3 Analyse numérique

Le protocole et les outils d'analyses numériques correspondent à ceux utilisés et décrits dans le chapitre précédent à la section 2.3.

3.3.1 Validation du modèle

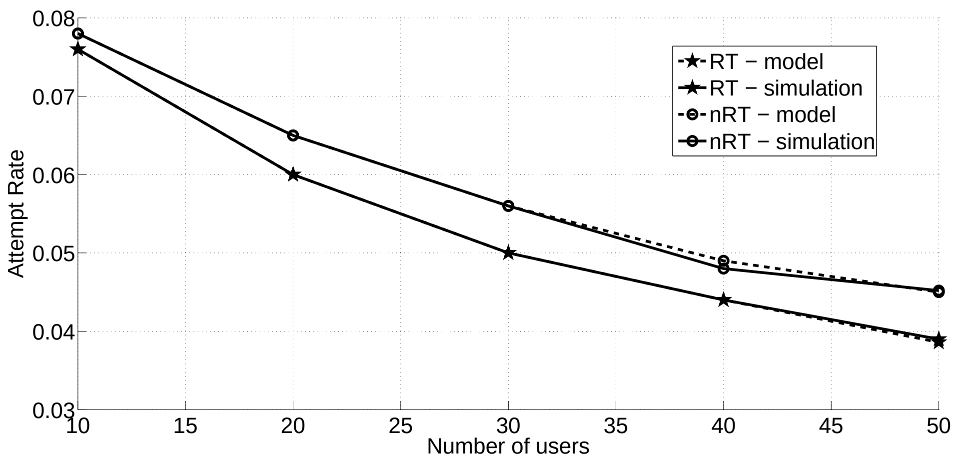

FIGURE 3.2 – *Taux de tentative RT et NRT en fonction du nombre d'utilisateurs n.* $CWmin_k = 16$, $m_k = 10$, $tr = 5$, $N_1 = 8$, $N_2 = 8$, $r = 0.5$

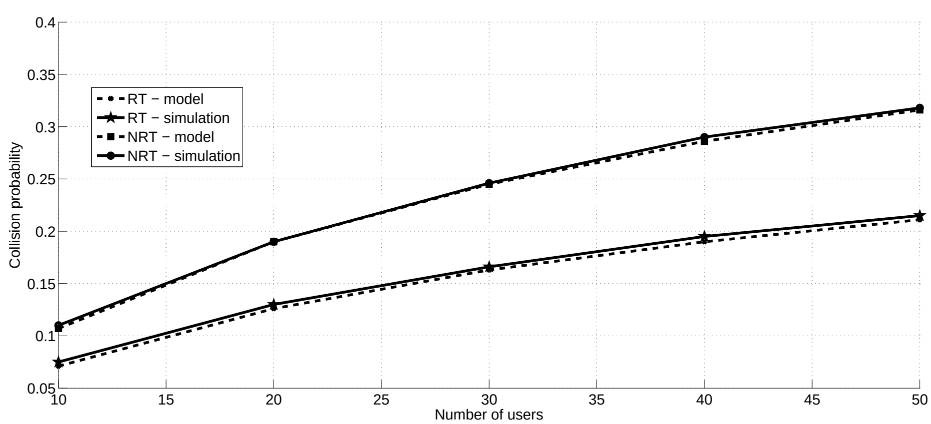

FIGURE 3.3 – *Probabilité de collision RT et NRT en fonction du nombre d'utilisateurs n.* $CWmin_k = 16$, $m_k = 10$, $tr = 5$, $N_1 = 8$, $N_2 = 8$ *et* $r = 0.5$

Nous comparons en premier lieu les résultats numériques avec ceux de notre simulateur comportemental. La figure 3.2 représente la comparaison des taux de tentative pour chaque classe en fonction du nombre d'utilisateurs. Nous observons sur cette figure que les résultats concordent parfaitement, nous permettant de valider notre mo-

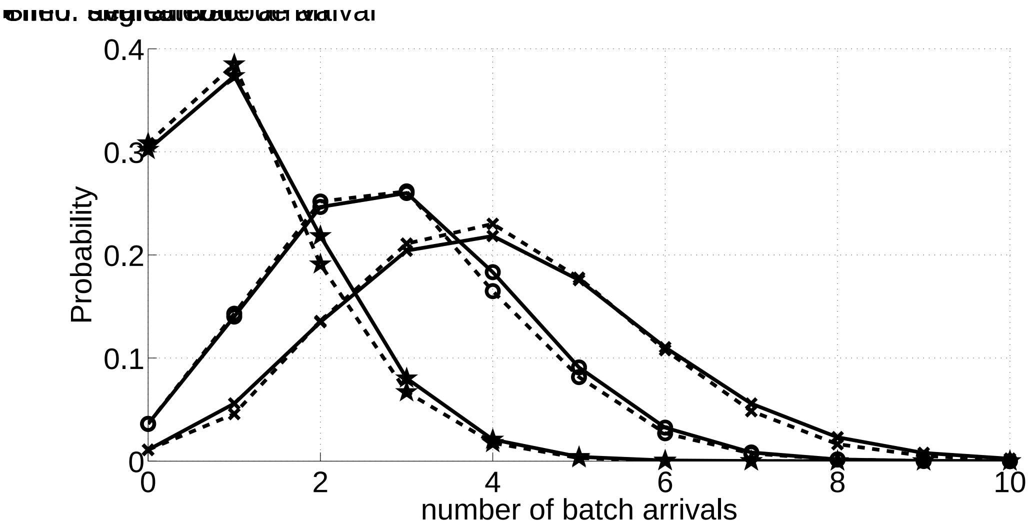

FIGURE 3.4 – *Distributions des arrivées globales et par sous-plage de codes en fonction du nombre d'utilisateurs n. $CWmin_k = 16$, $m_k = 10$, $tr = 5$, $N = 24$, $N_2 = 12$, $n = 75$ et $r = 0.5$*

dèle théorique. Cette observation se confirme sur la figure 3.3 présentant cette même comparaison pour la probabilité de collision. De plus, la figure 3.4 présente là aussi une figure comparant les distributions des arrivées par plage Z^1 et Z^2 issues du modèle théorique et des simulations comportementales des stations IEEE802.16e. La très forte corrélation liant ces deux types de résultats nous confirme la pertinence de notre modèle des arrivées.

3.3.2 Apport du partitionnement

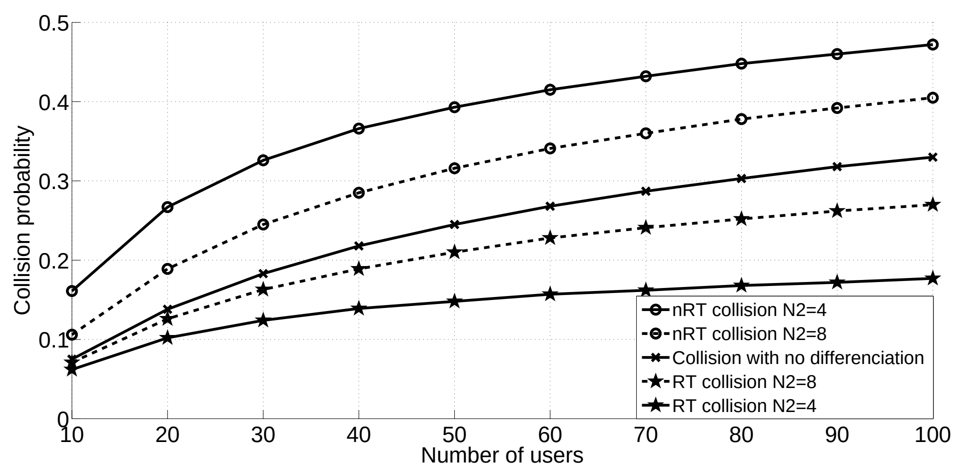

FIGURE 3.5 – *Probabilité de collision RT et NRT en fonction du nombre d'utilisateurs n pour différents profils de partitionnement N_2. $CWmin_k = 16$, $m_k = 10$, $tr = 5$, $N = 16$ et $r = 0.5$*

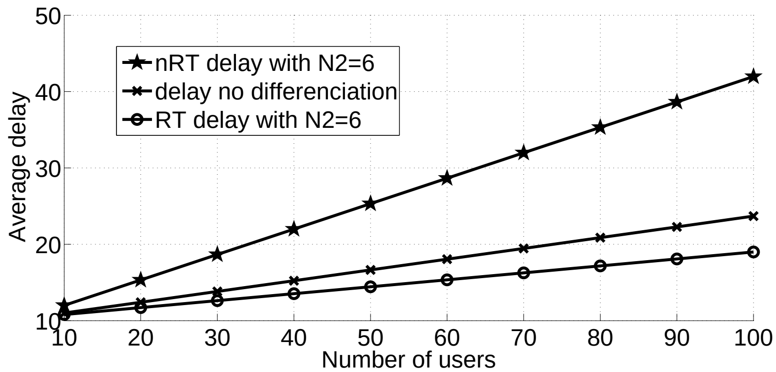

FIGURE 3.6 – *Délai RT en fonction du nombre d'utilisateurs n, pour différents profils de partitionnement N_2. $CWmin_k = 16$, $m_k = 10$, $tr = 5$, $N = 16$ et $r = 0.5$*

De là, nous comparons sur la figure 3.5 les probabilités de collision obtenues avec et sans l'utilisation du partitionnement des codes. Ainsi, nous observons que le principe de partitionnement réduit fortement les collisions des trafics temps réel en offrant un espace de codes élargi et dédié à ces trafics. Par conséquent, une telle mesure facilite l'accès au canal pour les requêtes de ressources. La figure 3.6 montre explicitement la réduction du délai d'établissement de connexion pour ce type de trafic.

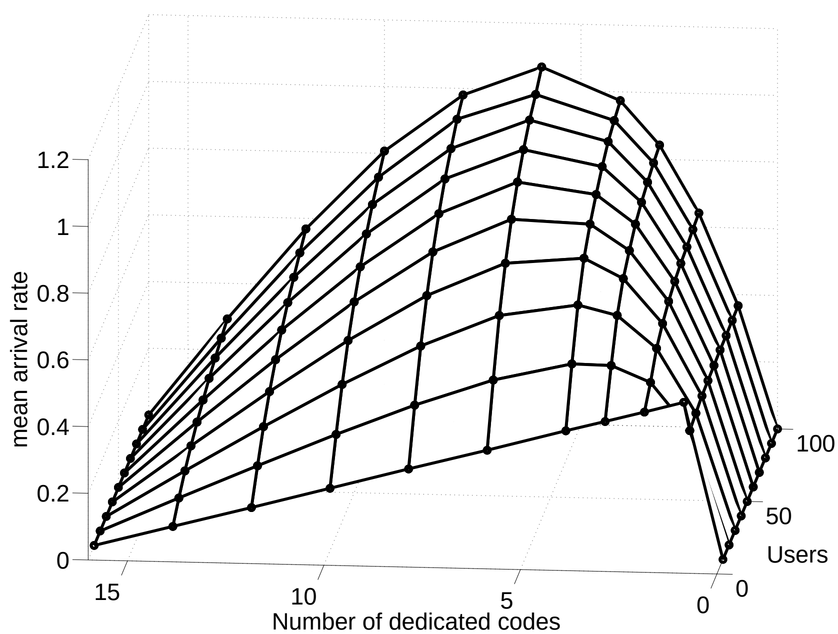

FIGURE 3.7 – *Arrivée moyenne sur la sous-plage de code dédiée au trafic RT (N_1) en fonction du nombre d'utilisateurs n et pour différents profils de partitionnement N_1. $CWmin_k = 16$, $m_k = 10$, $tr = 5$, $N = 16$ et $r = 0.50$*

Au travers de la figure 3.7, nous observons le comportement moyen du nombre des arrivées suivant l'importance du partitionnement opéré sur les codes. La figure témoigne que le partitionnement offre de meilleurs performances à condition que la configuration (choix de N_2) se fasse en considérant le nombre d'utilisateurs. En effet, il est primordial de remarquer qu'une trop large plage de codes N_1 dédiée aux trafics temps réel engendre un effondrement des performances des trafics non temps réel. En conséquence, les mobiles désirant engager une connexion non temps réel, demeure longuement en attente de fin de *backoff*. Aussi, les envois de requêtes des trafics temps réels se font de plus en plus rare. Dans l'état, ceci n'affecte pas réellement les performances de délai une fois que le mobile engage une première tentative de trafic temps réel. Par contre, ce ralentissement devient catastrophique si l'on considère que les trafics temps réel demeurent plus longuement en attente au niveau de l'ordonnanceur de trafic de chaque mobile. Ainsi, le principe de partitionnement des codes doit se faire en accord avec le nombre d'utilisateurs : pour un nombre d'utilisateurs donné, le nombre moyen des arrivées atteint un maximum pour une sous-plage de codes dédiés N_1 spécifique. Notons que cette plage de codes optimale augmente légèrement à mesure que le nombre d'utilisateurs grandit aussi.

3.3.3 Gains de performance

Sur les figures 3.8 à 3.11, nous explicitons le gain obtenu par le partitionnement sur les probabilités de collision et délai des trafics temps réel et non temps réel. Ce gain se base sur l'expression (3.41) où v_{ori} et v_{new} sont respectivement les valeurs originales obtenues dans le chapitre précédents et celles exploitant le partitionnement de codes. Les résultats en ordonnée correspondent au pourcentage de gain par rapport aux résultats obtenus dans le chapitre précédent (voir la section 2.3). Cette figure témoigne aussi de l'impact du paramètre α définissant le type de distribution de probabilité suivi dans le choix des sous-plages de codes pour les trafics temps réel.

$$\text{Gain} = \frac{v_{ori} - v_{new}}{v_{ori}} \tag{3.41}$$

La figure 3.8 montre le gain obtenu sur la probabilité de collision pour les trafics temps réel. Elle montre que notre mécanisme permet de réduire les collisions des requêtes jusqu'à 70 %. De plus, nous identifions le couple de valeurs $\{N_2, \alpha\}$ qui détériore particulièrement les performances des requêtes temps réel. Afin, de pouvoir mieux juger du couple de valeurs optimale $\{N_2, \alpha\}$, nous proposons dans la figure 3.9 l'observation de ce même gain pour les requêtes de type non temps réel. Pour ce type de trafic, nous observons naturellement que les performances des trafics non temps réel augmentent à mesure que l'étendue de codes disponibles, N_2, s'élargit et la probabilité α grandit. Néanmoins, la comparaison des résultats contenus par ces deux figures montre qu'un compromis entre les gains atteignables par chacun des trafics est aussi possible. Par exemple, pour $N_2 \geq \dfrac{N}{2}$ et $\alpha \leq 0,5$ on observe une diminution significative des collisions pour les deux types de trafic. Ainsi, en fonction de ses objectifs commerciaux

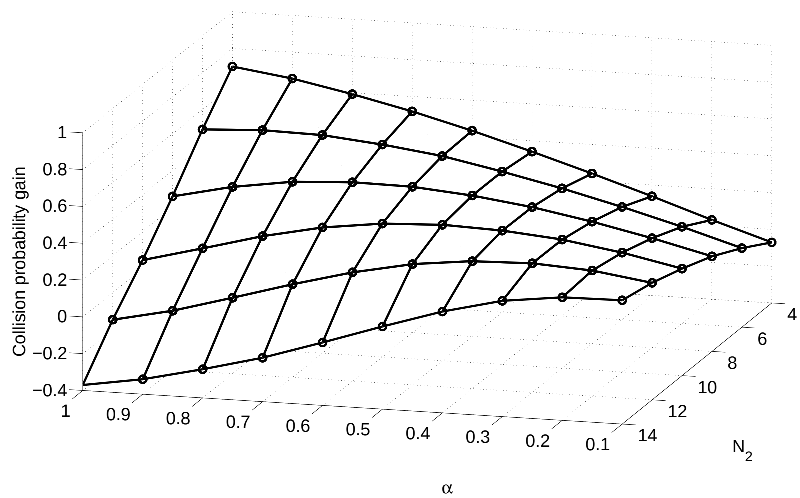

FIGURE 3.8 – *Gain de la probabilité de collision des trafics RT en fonction du profil de partitionnement N_2 et de la probabilité α. $CWmin_k = 16$, $m_k = 10$, $tr = 5$, $N = 16$, $n = 50$ et $r = 0.5$*

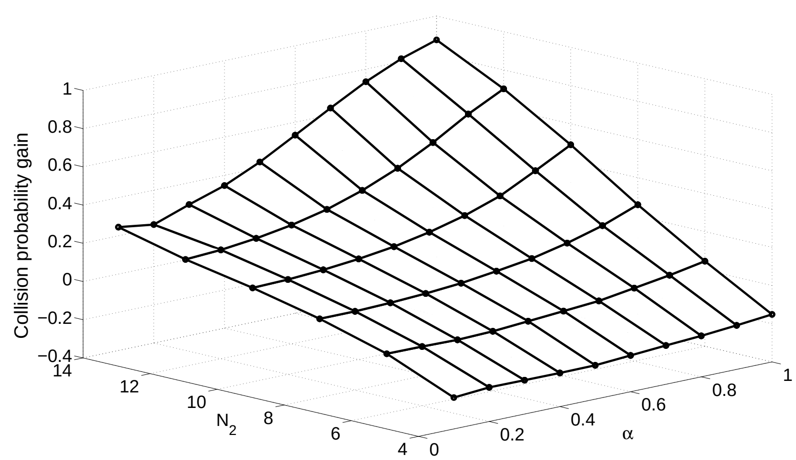

FIGURE 3.9 – *Gain de la probabilité de collision des trafics NRT en fonction du profil de partitionnement N_2 et de la probabilité α. $CWmin_k = 16$, $m_k = 10$, $tr = 5$, $N = 16$, $n = 50$ et $r = 0.5$*

et tarifaires, un fournisseur d'accès IEEE802.16e pourra définir ses propres règles de prioritisation entre ces deux types de flux de données.

En complément aux résultats déjà présentés, nous fournissons ici les figures représentatives des gains accessibles pour le délai d'établissement de connexions pour les deux types de trafic. Premièrement, la figure 3.10 montre le gain, toujours exprimé en pourcentage, pour le délai des trafics temps réel. A l'inverse des comportements observés précédemment, les gains les plus forts s'obtiennent ici par l'utilisation de valeurs élevées pour α et faibles pour N_2. Concernant les gains possibles pour le délai des trafics

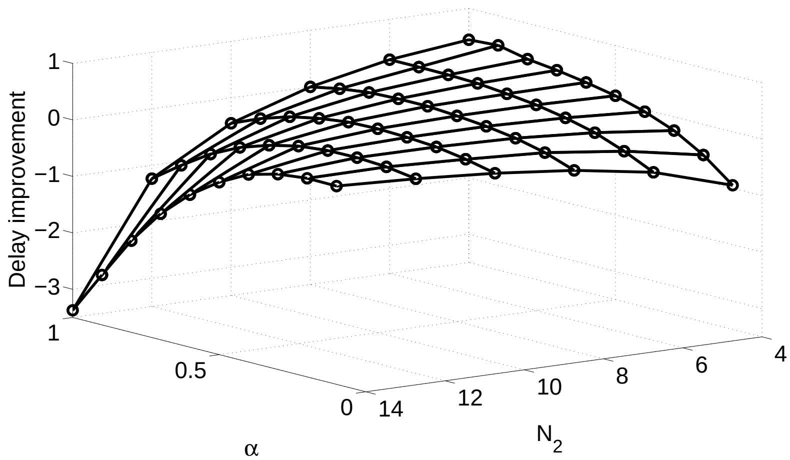

FIGURE 3.10 – *Gain de délai des trafics RT en fonction du profil de partitionnement N_2 et de la probabilité α. $CWmin_k = 16$, $m_k = 10$, $tr = 5$, $N = 16$, $n = 50$ et $r = 0.5$*

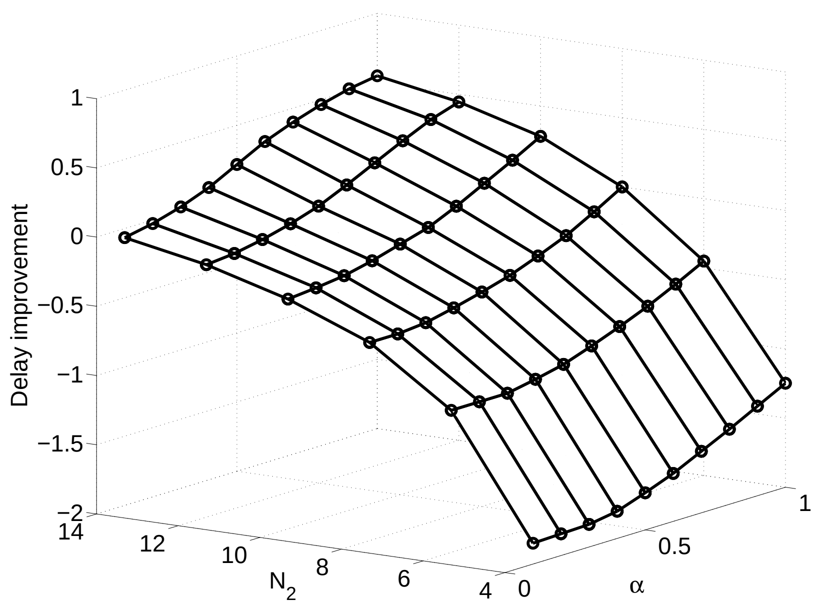

FIGURE 3.11 – *Gain de délai des trafics NRT en fonction du profil de partitionnement N_2 et de la probabilité α. $CWmin_k = 16$, $m_k = 10$, $tr = 5$, $N = 16$, $n = 50$ et $r = 0.5$*

non temps réel, le comportement est similaire à la figure 3.9.

De manière générale, l'étude des gains révèle que de fortes améliorations sont possibles à condition qu'on établisse une préférence claire entre la probabilité de collision et le délai d'établissement de connexion. Néanmoins, l'observation précise des résultats montre qu'un gain mutuel est possible pour un petit éventail de valeurs du couple

$\{N_2, \alpha\} : 8 \leq N_2 \leq 10 \quad \text{et} \quad 0,4 \leq \alpha \leq 0,6.$

3.3.4 Dépendance et approximation des distributions d'arrivée

Nous terminons cette analyse numérique par la discution de la dépendance implicite existante entre les arrivées de chaque trafic. En effet, nous avons remarqué précédemment que quelque soit le nombre des arrivées RT et NRT, il existe des relations de dépendance : $Z_1 \leq N$, $Z_2 \leq N_2$ et $Z_1 + Z_2 \leq N$.

Par exemple, la figure 3.12 montre la distribution jointe des arrivées des requêtes relatives aux appels temps réel et non temps réel. Ces résultats ont été obtenus pour un groupe de 50 utilisateurs exploitant 8 codes $N = 8$. Ces codes sont partitionnés à raison de 4 codes dédiés aux trafics temps réel $N_1 = 4$ et 4 autres codes partagés entre les trafics temps réel et non temps réel $N_2 = 4$. Les autres paramètres suivent le standard IEEE802.16e (Forum, 2005).

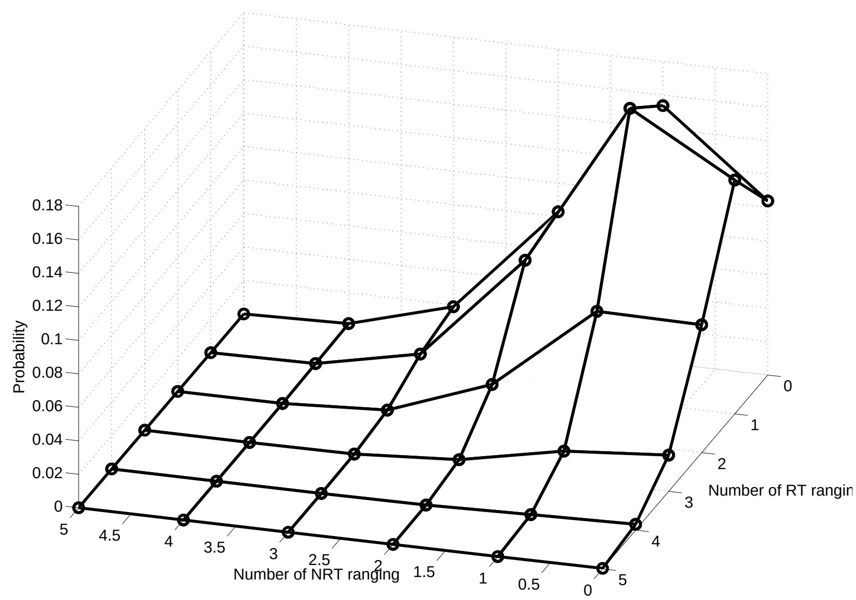

FIGURE 3.12 – *Distributions des arrivées temps réel et non temps réel pour 50 utilisateurs, 4 codes RT dédiés et 4 codes NRT partagés*

Au vu de ces résultats nous jugeons que cette distribution peut pertinemment être approximée par le produit de gaussiennes. Afin d'appuyer ce fait, nous comparons la distribution de ces arrivées avec d'autres, obtenues par combinaison de deux distributions gaussiennes. Ainsi, nous définissons deux variables aléatoires ψ_1 et ψ_2 indépendantes de distributions gaussiennes discrétisées tronquées et normalisées. Ces distributions se définissent à travers les moyennes respectives μ_1, μ_2 et les variances σ_1, σ_2

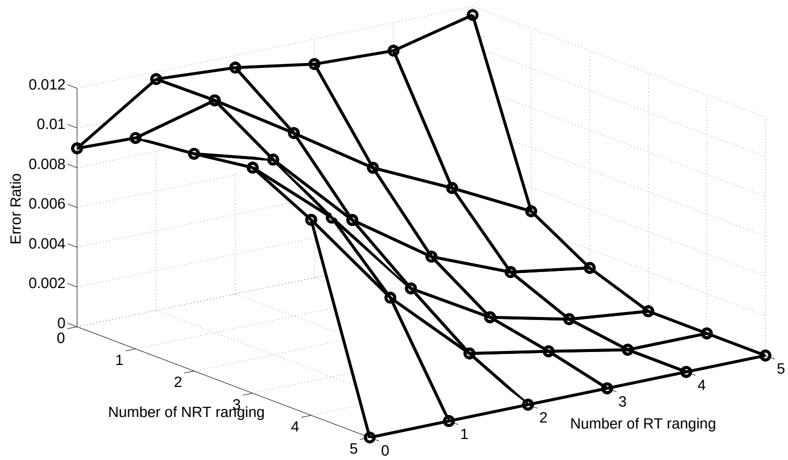

FIGURE 3.13 – *Erreur relative entre l'approximation par deux distributions gaussiennes et celle des arrivées par classe de service (figure 3.12)*

respectivement. Nous considérons leur produit suivant la relation suivante :

$$P(Z_1^{\psi} = i, Z_2^{\psi} = j) = P(\psi_1 = i)P(\psi_2 = j) \tag{3.42}$$

En considérant un large spectre de valeurs pour les moyennes et variances de ces deux gaussiennes, nous comparons la distribution de la relation (3.42) avec celle obtenue par notre simulateur comportemental. Or nous obtenons une approximation fine de celle d'origine. Ainsi, la figure 3.13 présente l'erreur relative calculée à partir de la relation (3.41) entre les différentes probabilités présentées par la figure 3.12 et celles issues de notre approximation. L'expérience a porté sur deux distributions gaussiennes de moyenne $\mu_1 = 0.5575$ et $\mu_2 = 0.7564$ et de variance $\sigma_1 = 1.2739$ et $\sigma_2 = 1.0749$ respectivement. Ces distributions suivent la loi normale qui à tout réel x, y associe une probabilité y suivant la relation (3.43). Dans cette configuration, l'erreur relative d'approximation ne dépasse jamais les 0.012. Autrement dit, en tout point de notre distribution des arrivées temps réel et non temps réel, l'approximation par un produit de deux gaussiennes n'excède pas les 1.2 %. Cette expérience empirique nous porte à croire que le comportement des arrivées par classe de service dans un système IEEE802.16e peut être approximé par le somme de deux variables aléatoires indépendantes de distributions gaussiennes. En outre, cet observation nous amène aussi à considérer la faisabilité d'exploitation d'une nouvelle hypothèse. L'expérience menée ici se base sur un grand nombre de codes ($N = 8$) vis à vis du nombre d'utilisateur ($n = 50$). Aussi, les probabilités d'avoir un grand nombre d'arrivées simultanées par classe est extrêmement faible (voir la figure 3.4 par exemple). Dès lors, nous pouvons dresser l'hypothèse, que dans un tel environnement, les processus d'arrivée des deux classes de trafics sont indépendants.

$$y = f(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{Z\sigma^2}} \quad (3.43)$$

3.4 Conclusion et perspectives

Cette partie de l'étude développe un élargissement original du modèle mis en place dans le chapitre précédent. Nous y développons un nouveau modèle théorique introduisant un mécanisme nouveau de partitionnement adaptatif des codes alloués exclusivement aux trafics temps réel. Ce mécanisme est caractérisé par le couple de valeurs $\{N_2, \alpha\}$. Il a pour but d'apporter une granularité supplémentaire favorisant les trafics sensibles au délai d'établissement de connexion avec la station de base. En outre, nous proposant dans ce chapitre un outil simple ouvrant un large champs d'application sur les mécanismes d'établissement de connexion de la couche MAC.

Ainsi, le principe de partitionnement établit un élément de contrôle d'admission en amont de celui opéré par la station de base. Ainsi certains types de service se verraient refuser l'envoi de requêtes de ressources par une absence de codes disponibles à leur attention. Cette mesure a pour but d'empêcher un service de consommer les ressources radio utiles à l'envoi d'un « ranging request », alors que la station de base est dans l'incapacité de lui attribuer la moindre ressource. Par exemple, nous pouvons définir un algorithme adaptatif définissant les plages de codes accessibles pour chaque type de trafic en fonction des ressources encore disponibles pour ces flux. Ce mécanismes diminuerait le nombre de codes alloués aux trafics temps réel, dans la mesure où ces derniers occupent déjà la quasi-totalité des ressources du système. En outre, les trafics permettant le partage dynamique des ressources auront de plus grandes chances d'accès au système IEEE802.16. Par ailleurs, la station de base peut de la même manière limiter l'engorgement des requêtes à son entrée. En conséquence, les trafics temps réel ont maintenant la possibilité de réduire de manière forte le délai d'engagement de leur transmissions. Enfin, ce principe peut à terme mener à une réduction globale des temps de traitement des requêtes au niveau de la station de base.

Autre contribution majeure de cette étude : la caractérisation des arrivées par rafale. En effet, à ce jour, il nous semble que cette étude est la première à proposer un calcul abouti de la distribution des arrivées discrètes par classe de service. En l'occurrence, ce calcul exploite la totalité des paramètres relatifs à l'engagement des communications dans le IEEE802.16. Aussi, ces distributions constituent une contribution majeure. Elles définissent un processus d'arrivée réaliste, générale et entièrement paramètrable pour tout travaux futur. Par ailleurs, l'étude montre que la distribution des arrivées par classe peut être approximée par la combinaison de deux distributions gaussiennes. Ce fait introduit la possibilité et la justification d'exploiter des arrivées de type gaussien pour tout travaux futurs portant ou utilisant un processus d'arrivée des requêtes dans le standard IEEE802.16.

Les résultats obtenus dans cette étude expriment de manière qualitative et quantitative la nature des gains donnés par ce nouveau mécanisme de partitionnement des codes. Ce mécanisme offre une maîtrise complète des améliorations de performances accessibles pour chaque type de trafic. Ainsi, tout fournisseur d'accès pourra se définir sa propre politique de différenciation de flux. Ce choix pourra aller de pair avec une tarification adéquate. En effet, au travers du choix du couple de paramètres $\{N_2, \alpha\}$, un fournisseur de service IEEE802.16e pourra choisir le couple en fonction de la métrique de performance devant faire l'objet d'une amélioration particulière.

Dans une vision plus égalitaire d'amélioration conjointe, les mobiles IEEE802.16e peuvent améliorer leur performances globales de connexions (collision et délai) par l'utilisation d'un principe de partitionnement où les codes sont répartis suivant le nombre de requêtes respectif pour chaque trafic et une distribution uniforme du choix des codes pour les trafics temps réels. Soit en d'autre terme :

$$r = \frac{N_2}{N} \quad \text{et} \quad \alpha = \frac{N_1}{N} \tag{3.44}$$

Partant de l'idée fondatrice de ce mécanisme, différentes pistes de recherche sont identifiables. Premièrement, les classes de trafic prises indépendamment : nrtPS, rtPS et UGS pourraient se voir allouer une partie dédiée de la plage des codes tout en conservant l'accès au seuls codes accessibles au trafic *Best Effort*. La formalisation de ce principe généralisé de partitionnement permettrait une différenciation extrêmement fine et modulable de différenciation des services entre eux. Par ailleurs, nous avons montré que le processus d'arrivée des requêtes temps réel et non temps réel peut être approximé par le produit de deux variables aléatoires indépendantes de distributions gaussiennes . Aussi, il nous importe de pouvoir caractériser les paramètres de ces deux distributions en fonction de ceux de communications. Ainsi, nous serions à même d'identifiant les relations liant les moyennes et variances de ces distributions gaussienne avec l'ensemble des paramètres des mécanismes de connexion du standard IEEE-802.16e.