

Évaluation des performances par la Perturbation Singulière

Sommaire

3.1 Introduction	79
3.2 Formulation du problème	81
3.2.1 Hypothèses	81
3.3 Série de Taylor	83
3.3.1 Définitions et Hypothèses	83
3.3.2 Représentation analytique de la distribution stationnaire	86
3.3.4 Temps de séjour d'appels <i>best effort</i>	88
3.4 Résultats numériques	89
3.5 Conclusion	93

3.1 Introduction

Dans le chapitre précédent, nous avons proposé une analyse très puissante et exacte de la capacité du réseau de troisième génération UMTS en s'appuyant sur la technique d'accès multiple WCDMA. Certes, cette analyse intitulée Analyse spectrale, nous a permis d'évaluer les performances du système qui jouent un rôle très important dans la garantie de la qualité de service demandée par les utilisateurs du réseau. Néanmoins, l'analyse spectrale demande un temps de calcul considérable vu la complexité de ses équations mathématiques quand le nombre d'appels de la classe temps réel (RT) devient grand. Cette complexité devient compliquée avec la diversité de service qu'offre le système aux utilisateurs mobiles.

Nous observons dans la plupart des cas, que les transitions du nombre d'appels RT se produisent beaucoup plus rapidement que les transitions du nombre d'appels BE. Dans ce cas particulier des comportements de transitions, nous définissons deux échelles de

transitions différentes, une rapide (nombre d'appels RT qui se produisent rapidement) et une lente (nombre d'appels BE qui se produisent lentement). Ce qui nous a permis de proposer une approximation simple et efficace en temps de calcul dans les conditions de la distribution stationnaire du système.

En effet, cette approximation appelée perturbation singulière est traitée dans ce chapitre. Nous considérons le système WCDMA en présence de deux classes de service : temps réel (RT) et *best effort* (BE). Les services de la classe RT ont des ressources dédiées et donc un contrôle d'admission, et les services de la classe BE n'ont pas de contrôle d'admission. Le premier objectif de ce chapitre est de calculer la distribution stationnaire de manière approximative des états d'équilibre du système en nombre d'appels RT et BE. Après cela, nous fournissons explicitement les performances du système. Néanmoins, le fournisseur de service ou l'opérateur peut utiliser cette approche basée sur l'approximation pour évaluer en temps réel avec un calcul rapide des métriques de performances d'appels RT et BE.

La perturbation singulière a été définie dans la littérature comme une faible interaction entre les groupes sachant que la forte interaction entre les états de chaque groupe est rapide. La perturbation singulière joue un rôle important dans les chaînes de Markov et surtout pour pouvoir déterminer une approximation simple des probabilités d'états d'équilibre dans lesquels nous pouvons identifier deux échelles de temps : une rapide et une lente. Les transitions entre les groupes sont beaucoup moins fréquentes que celles entre les états dans un même groupe ou bien l'inverse. La perturbation singulière a fait l'objet de plusieurs travaux de recherche et nous citons par exemple les travaux (Avrachenkov, 2004, 1999; Philips et al., 1981; Altman et al., 2004b; Avrachenkov et Haviv, 2004). Les auteurs dans (Philips et al., 1981), montrent que les chaînes de Markov finies admettant des interactions faibles et fortes, peuvent être décomposées en deux échelles de temps, rapide et lent, en utilisant l'approche de la perturbation singulière. Ils considèrent que les faibles interactions perturbent le système. Ainsi, Les auteurs dans (Altman et al., 2004a), étudient la perturbation singulière et régulière des chaînes de Markov sur un espace d'états dénombrable. Ils considèrent que les probabilités de transition entre les classes ergodiques perturbent le système (transitions rares) pour appliquer la perturbation singulière. Ils trouvent une représentation de la distribution stationnaire sous forme d'une série de Taylor en fonction du paramètre de la perturbation. Puis, ils étendent leurs résultats vers le processus quasi-naissance et mort dans les files d'attente.

Notre approche est basée sur la modélisation du système par une chaîne de Markov à deux dimensions, dont le premier correspond au nombre d'appels temps réel et le deuxième au nombre d'appels *best effort*. Afin d'obtenir la distribution stationnaire d'états d'équilibre approximativement, nous utilisons la perturbation singulière avec un processus quasi de naissance et de mort QBD (Latouche et Ramaswami, 1999). Nous considérons deux échelles de temps : une rapide et l'autre lente. Le premier représentant le nombre d'appels RT change rapidement et le deuxième représentant le changement d'appels BE lentement. La perturbation singulière nous a permis de représenter les probabilités d'états d'équilibre comme une série de Taylor. Les termes de cette série donnent une approximation de la distribution stationnaire que nous avons trouvé

par l'analyse spectrale dans le premier chapitre. Le premier terme de la série donne déjà une très bonne approximation quand le système identifie bien les deux échelles de temps citées auparavant.

3.2 Formulation du problème

3.2.1 Hypothèses

Les métriques de performances (comme le débit, le temps de séjour et la probabilité de blocage) sont très difficiles à calculer explicitement. Pour cela, dans le chapitre 2, nous avons utilisé l'analyse spectrale pour obtenir la forme explicite de la distribution stationnaire. Certes, la solution exacte trouvée dans le chapitre précédent, nous a permis d'évaluer les performances du système en fonction de certains paramètres du système. Mais il reste difficile à calculer la distribution stationnaire et surtout quand le nombre d'appels RT augmente, car les performances trouvées par l'analyse spectrale dépendent directement de ceux en RT. Donc, quand le nombre d'appels RT augmente, le système va prendre un temps important pour le calcul des performances et s'il y a des erreurs dû aux interférences des évanouissements, d'effet de masque, etc., le système doit refaire tout le calcul mathématique. Ce qui va rendre le système plus lourd au niveau temps de calcul.

Dans ce chapitre, nous introduisons une approche simple et rapide permettant d'évaluer en temps réel les performances du système. L'approche est basée sur l'hypothèse suivante : Dans le système WCDMA supportant les services RT et BE, nous supposons que selon les statistiques du canal, les appels BE restent longtemps durant leurs services (comme le cas du service Chat) et ceux en RT prennent généralement moins de temps de service (comme le cas du service voix). Donc, nous construisons deux groupes d'appels dans le système, un représente les transitions rapides causées par les appels RT et l'autre représente les faibles transitions causées par les appels BE. Ensuite, nous formulons mathématiquement une approximation de ces transitions. Ces approximations touchent la matrice des taux de transitions (\mathbf{Q}) de la chaîne de Markov décrivant le système, le taux d'arrivée d'appels BE (λ_{BE}) et le taux de service μ_{BE} . Nous définissons ces approximations comme suit

$$\mathbf{Q}(\epsilon) = \mathbf{Q}_0 + \epsilon\mathbf{Q}_1, \quad (3.1)$$

$$\lambda_{BE} = \epsilon\bar{\lambda}_{BE}, \quad (3.2)$$

$$\mu_{BE} = \epsilon\bar{\mu}_{BE}, \quad (3.3)$$

où ϵ est un paramètre strictement positif très petit représentant les faibles transitions, $\bar{\lambda}_{BE}$ est un paramètre strictement positif lié aux taux des arrivées, $\bar{\mu}_{BE}$ est un paramètre strictement positif dépendant des taux de services, $\mathbf{Q}(\epsilon)$ est un générateur correspondant à la chaîne de Markov perturbée; \mathbf{Q}_0 est un générateur non perturbé qui correspond à des interactions fortes et $\epsilon\mathbf{Q}_1$ est le terme perturbé correspond aux interactions

faibles. Nous avons les structures matricielles suivantes

$$\mathbf{Q}_0 = \begin{bmatrix} \mathbf{Q}_{RT} & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{Q}_{RT} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_{RT} & \dots \\ \mathbf{0} & \mathbf{0} & \ddots & \ddots \end{bmatrix},$$

$$\mathbf{Q}_1 = \begin{bmatrix} -\mathbf{A}_0 & \mathbf{A}_0 & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{A}_2 & -\mathbf{A}_2 - \mathbf{A}_0 & \mathbf{A}_0 & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{A}_2 & -\mathbf{A}_2 - \mathbf{A}_0 & \mathbf{A}_0 & \dots \\ \mathbf{0} & \mathbf{0} & \ddots & \ddots & \ddots \end{bmatrix}.$$

où \mathbf{Q}_0 a une infinité de blocs fini \mathbf{Q}_{RT} et \mathbf{A}_0 et \mathbf{A}_2 sont des matrices carrées d'ordre $M_{RT} + 1$, tel que

$$\begin{aligned} \mathbf{A}_0 &= \text{diag}(\bar{\lambda}_{BE}), \\ \mathbf{A}_2 &= \text{diag}(\nu(i); \quad i = 0, 1, 2, \dots, M_{RT}). \end{aligned}$$

Ainsi, la matrice \mathbf{Q}_{RT} est carrée d'ordre $M_{RT} + 1$ et représente les arrivées et les départs d'appels RT

$$\begin{aligned} \mathbf{Q}_{RT}[i, i-1] &= i\mu_{RT}, \\ \mathbf{Q}_{RT}[i, i] &= -(\lambda_{RT} + i\mu_{RT}), \quad i = 0, \dots, M_{RT} - 1, \\ \mathbf{Q}_{RT}[M_{RT}, M_{RT}] &= -M_{RT}\mu_{RT}, \\ \mathbf{Q}_{RT}[i, i+1] &= \lambda_{RT}. \end{aligned}$$

Soit $\pi^{(\epsilon)}(i, j) \geq 0$ ($i = 0, \dots, M_{RT}$ et $j \geq 0$) la distribution de probabilité des états du nombre d'appels RT et BE à l'équilibre. Puisque $\mathbf{Q}(\epsilon)$ est irréductible ($\epsilon > 0$), alors cette probabilité invariante de mesure existe, unique et satisfait le système suivant :

$$\pi^{(\epsilon)}\mathbf{Q}(\epsilon) = \mathbf{0}, \quad (3.4)$$

$$\pi^{(\epsilon)}\mathbf{1} = \mathbf{1}, \quad (3.5)$$

où $\mathbf{1}$ est un vecteur unité de même dimension que le vecteur $\pi^{(\epsilon)}$. La chaîne de Markov a un espace d'états

$$E = \{(i, j), 0 \leq i \leq M_{RT}, j \geq 0\},$$

Qu'on peut partitionner en sous ensembles $l(j)$ tels que :

$$F = \{l(0), l(1), l(2), \dots\},$$

où $l(j) = \{(0, j), \dots, (M_{RT}, j)\}$ pour $j \geq 0$.

La matrice \mathbf{Q}_0 contient plusieurs classes ergodiques $l(j)$, chacune associée à un générateur \mathbf{Q}_{RT} qui est un bloc de la matrice diagonale \mathbf{Q}_0 .

Remarque 1 Nous signalons que, notre analyse est autant appliquée, dans le cas où les appels RT se produisent plus fréquemment que dans le cas de ceux en BE. Nous identifions dans ce cas, deux échelles d'interactions, une représente l'interaction rapide d'appels RT à intra-groupes et l'autre représente l'interaction moins fréquente inter-groupes d'appels BE. D'où l'approximation de ces interactions formulée en fonction des taux de transitions

$$\mathbf{Q}(\epsilon) = \frac{1}{\epsilon} \mathbf{Q}_0 + \mathbf{Q}_1, \quad (3.6)$$

$$\lambda_{RT} = \frac{1}{\epsilon} \bar{\lambda}_{RT}, \quad (3.7)$$

$$\mu_{RT} = \frac{1}{\epsilon} \bar{\mu}_{RT}, \quad (3.8)$$

où $\frac{1}{\epsilon}$ exprime la forte interaction d'appels RT qui perturbent le système, $\bar{\lambda}_{RT}$ est un paramètre strictement positif lié aux taux des arrivées, $\bar{\mu}_{RT}$ est un paramètre strictement positif dépendant des taux de services.

L'objectif de la section ci-après, est de trouver la représentation analytique de la distribution stationnaire $\pi(\epsilon)$, de la chaîne de Markov perturbée $\mathbf{Q}(\epsilon)$, en fonction du paramètre de la perturbation ϵ .

3.3 Série de Taylor

3.3.1 Définitions et Hypothèses

Premièrement, nous avons besoin d'introduire quelques définitions de base. Soit un vecteur fixe dénombrable δ de la forme $\delta = (\delta_0, \delta_1, \dots)$ tel que $\delta_i \geq 1$. Le vecteur δ et l'opérateur associé ont la même dimension. Notons $I_{M_{RT}} = \{0, 1, \dots, M_{RT}\}$ et $I_\infty = \{0, 1, \dots\}$, deux sous ensembles, alors la norme δ -norme est définie selon l'opérateur associé comme suit :

1. La norme δ -norme du vecteur v est définie comme suit :

$$\|v\|_\delta = \max_{i \in I} \frac{|v_i|}{\delta_i}, \text{ où } v_i = (v_0, v_1, \dots).$$

2. La norme δ -norme de l'opérateur fini $B = (b_{ij})$ est donnée par

$$\|B\|_\delta = \max_{i \in I_{M_{RT}}} \frac{\sum_{j=0}^{M_{RT}} |b_{ij}| \delta_j}{\delta_i}.$$

3. L'opérateur $D = (d_{jk})$ d'infinité de blocs (d_{jk}) est donnée par la formule suivante :

$$\|D\|_\delta = \max_{j \in I_\infty} \frac{\sum_{k \in I_\infty} \|d_{jk}\|_\delta \delta_k}{\delta_j}.$$

La chaîne de Markov doit vérifier les hypothèses du lemme suivant :

- Lemme 6**
1. La chaîne de Markov imperturbée a plusieurs classes ergodiques. Ces classes sont $l(j)$ pour tout $j \in I_\infty$.
 2. Chaque sous chaîne de Markov qui correspond à une classe ergodique de \mathbf{Q}_0 est uniformément stable au sens de Lyapunov.
 3. La chaîne de Markov agrégée est irréductible et stable au sens de Lyapunov.
 4. Le générateur \mathbf{Q}_1 est δ -borné ($\|\mathbf{Q}_1\|_\delta < \infty$).

Démonstration du Lemme 6 1. A partir des analyses précédentes, la chaîne de Markov non perturbée est constituée de plusieurs classes $l(j)$ pour $j \in I_\infty$ et il n'y a pas d'états de transit. Or, puisque les valeurs λ_{RT} et μ_{RT} sont positives, alors la classe $l(j)$ est une classe finie ergodique.

2. L'opérateur de transition de la chaîne de Markov imperturbée est $\mathbf{P}^{(0)} = r_0^{-1}\mathbf{Q}_0 + \mathbf{I}$ dont la forme est

$$\mathbf{P}^{(0)} = \mathbf{I} + \begin{pmatrix} r_0^{-1}\mathbf{Q}_{RT} + \mathbf{I} & \mathbf{0} & \dots \\ \mathbf{0} & r_0^{-1}\mathbf{Q}_{RT} + \mathbf{I} & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix},$$

où \mathbf{I} est une matrice unité d'ordre $M_{RT} + 1$, \mathbf{I} matrice unité par bloc et $r_0^{-1} = \frac{1}{\lambda_{RT} + M_{RT}\mu_{RT}}$.

Posons $\mathbf{P}_{RT} = \mathbf{I} + r_0^{-1}\mathbf{Q}_{RT}$ matrice de probabilité de transition. Puisque nous avons $P_{RT}(0,0) = 1 - \frac{\lambda_{RT}}{r_0} > 0$, alors la chaîne correspond aux classes ergodiques de la chaîne de Markov non perturbée, est fortement apériodique, ce qui implique qu'il existe des constantes positives c et β telles que $c > 0$, $0 < \beta < 1$ (voir (Altman et al., 2004a) pour plus de détails).

$$\|\mathbf{P}_{RT}^n - \mathbf{1}\tilde{q}_{l(i)}\|_\delta \leq c\beta^n, \forall n \geq 0 \text{ et } \beta < 1.$$

3. Soit $\tilde{\mathbf{Q}}_1$ un générateur de la chaîne de Markov agrégée sur l'espace F où chaque groupe $l(j)$ est remplacé par un état agrégé j (Par simplicité, nous remplaçons $l(j)$ par j).

$$(\tilde{\mathbf{Q}}_1)_{i,j} = \tilde{q}_i \mathbf{Q}_1[i,j] \underline{\xi}, \quad \forall (i,j) \in I_\infty \times I_\infty,$$

où $\underline{\xi} = (1, \dots, 1)^t \in \mathfrak{R}^{(M_{RT}+1) \times 1}$. Cette matrice agrégée représente les transions entre les groupes. Chaque colonne dans $\tilde{\mathbf{Q}}_1$ correspond à un groupe dans F . La matrice $\tilde{\mathbf{Q}}_1$ est tridiagonale :

$$\begin{aligned} \tilde{\mathbf{Q}}_1[j, j+1] &= \tilde{q}_j \mathbf{A}_0 \underline{\xi} \\ \tilde{\mathbf{Q}}_1[j, j-1] &= \tilde{q}_j \mathbf{A}_2 \underline{\xi} \\ \tilde{\mathbf{Q}}_1[j, j] &= -\tilde{q}_j (\mathbf{A}_0 + \mathbf{A}_2) \underline{\xi}, (\forall j > 0) \\ \tilde{\mathbf{Q}}_1[0, 0] &= -\tilde{q}_0 (\mathbf{A}_0) \underline{\xi} \end{aligned}$$

où $\tilde{q}_j \mathbf{A}_n \underline{\xi} = \sum_{i=0}^{M_{RT}} \tilde{q}_j(i) (\mathbf{A}_n)_{(i,i)}$ pour $n = 0, 2$. La distribution invariante \bar{v} correspondant à ce générateur, est obtenue en résolvant le système d'équations suivant

$$\bar{v}\tilde{\mathbf{Q}}_1 = \mathbf{0}, \sum_{i=0}^{\infty} \bar{v}_i = 1 \quad (3.9)$$

où $\bar{v} = [\bar{v}_0, \bar{v}_1, \dots]$ est le vecteur de probabilité $\tilde{q}_j = [\tilde{q}_j(0), \dots, \tilde{q}_j(M_{RT})]$ est une solution unique de la forme

$$\tilde{q}_j \mathbf{Q}_{RT} = \mathbf{0}, \tilde{q}_j \underline{\xi} = 1, \quad (3.10)$$

Toute fois, le vecteur \tilde{q}_j est donné par

$$\tilde{q}_j(n) = \frac{(\rho_{RT})^n}{n!} \left(\sum_{i=0}^{M_{RT}} \frac{(\rho_{RT})^i}{i!} \right)^{-1}, 0 \leq n \leq M_{RT} \quad (3.11)$$

Le générateur $\tilde{\mathbf{Q}}_1$ devient

$$\tilde{\mathbf{Q}}_1 = \begin{bmatrix} -\bar{\lambda}_{BE} & \bar{\lambda}_{BE} & 0 & \dots \\ c & -\bar{\lambda}_{BE} - c & \bar{\lambda}_{BE} & \ddots \\ 0 & c & -\bar{\lambda}_{BE} - c & \bar{\lambda}_{BE} \\ 0 & \ddots & \ddots & \ddots \end{bmatrix},$$

où

$$c = \sum_{i=0}^{N_{RT}} v(i)\tilde{q}(i) + \sum_{i=N_{RT}+1}^{M_{RT}} v(i)\tilde{q}(i)$$

Posons

$$\alpha(M_{RT}) = \frac{\bar{\lambda}_{BE}}{\sum_{i=0}^{N_{RT}} v(i)\tilde{q}(i) + \sum_{i=N_{RT}+1}^{M_{RT}} v(i)\tilde{q}(i)}.$$

Le générateur de la chaîne de Markov agrégée correspond au générateur d'un processus QBD homogène du modèle M/M/1 dont le taux d'arrivée est $\bar{\lambda}_{BE}$ et le taux de départ est c . Nous pouvons vérifier que la chaîne de Markov agrégée est irréductible ergodique et fortement apériodique si $\alpha(M_{RT}) < 1$.

Soient $\tilde{\mathbf{P}} = r_{BE}^{-1}\tilde{\mathbf{Q}}_1 + \mathbf{I}$, la matrice des probabilités de transition associée à la chaîne de Markov agrégée et $r_{BE}^{-1} = \frac{1}{\bar{\lambda}_{BE} + 2v(0)}$. Nous remarquons que l'état $(0,0)$ est fortement apériodique, car la probabilité suivante à cet état est strictement positive $\tilde{\mathbf{P}}(0,0) = 1 - \frac{\bar{\lambda}_{BE}}{r_0} > \frac{v(0)}{r_0} > 0$, alors la fonction de Lyapunov δ est vérifiée $\tilde{\mathbf{P}}\delta \leq \gamma\delta + b\mathbf{1}_{(0,0)}$, où $b = \sqrt{c}(1 - \sqrt{\alpha(M_{RT}))}\delta_0, \delta_{l(i)} = (\sqrt{\alpha(M_{RT}))}^i$ et $\gamma = 1 - (\bar{\lambda}_{BE} - c)^2$ (pour plus de détail voir (Altman et al., 2004a; Meyn et Tweedie, 1994)).

Nous rappelons que les états de la chaîne agrégée sont les classes ergodiques de la chaîne originale, en présence de la perturbation, elle représente exactement les faibles transitions entre les groupes.

4. La norme de l'opérateur \mathbf{Q}_1 est majorée par une constante positive, car nous avons

$$\|\mathbf{Q}_1\|_\delta \triangleq \max_{j \in I_\infty} \frac{\sum_{k \in I_\infty} \|\mathbf{Q}_1(i, k)\| \delta_k}{\delta_j} = \max_{i \in I_\infty} \left\{ \frac{\|\mathbf{A}_2(i, \cdot)\|_\delta (\delta_0 + \delta_1)}{\delta_0}, \right. \\ \left. \frac{\|\mathbf{A}_0\|_\delta \delta_{i-1} + \|(\mathbf{A}_2 + \mathbf{A}_0)\|_\delta \delta_i + \|\mathbf{A}_0\|_\delta \delta_{i+1}}{\delta_{i-1}} \right\},$$

où

$$\|\mathbf{A}_0\|_\delta = \bar{\lambda}_{BE} \|I\|_\delta = \bar{\lambda}_{BE}, \text{ et } \|\mathbf{A}_2\|_\delta = v(0).$$

Puisque nous avons

$$\max(v(i)) = v(0), \text{ et } \|(\mathbf{A}_2 + \mathbf{A}_0)\|_\delta = \bar{\lambda}_{BE} + v(0)$$

Alors nous obtenons

$$\|\mathbf{Q}_1\|_\delta = \max_{i \in I_{MRT}} \left\{ \bar{\lambda}_{BE} \left(1 + \frac{\delta_1}{\delta_0}\right), \bar{\lambda}_{BE} \left(1 + \frac{\delta_i}{\delta_{i-1}}\right) + v(0) \left(1 + \frac{\delta_{i+1}}{\delta_i}\right) \right\}.$$

Par conséquent, le générateur \mathbf{Q}_1 est δ -borné.

3.3.2 Représentation analytique de la distribution stationnaire

Nous présentons maintenant la distribution stationnaire des états d'équilibre de la chaîne de Markov perturbée, par une fonction analytique de paramètre ϵ . A partir du lemme 6, les hypothèses du Théorème 4.1 dans (Altman et al., 2004a) sont satisfaites, nous obtenons les résultats suivants

Théorème 4 La forme analytique de la probabilité de mesure invariante $\pi^{(\epsilon)}$ en fonction du paramètre ϵ est

$$\pi^{(\epsilon)} = \pi^{(0)} + \pi^{(1)}\epsilon + \pi^{(2)}\epsilon^2 + \pi^{(3)}\epsilon^3 + \dots \quad (3.12)$$

avec $0 < \epsilon \leq \min(\epsilon_{max}, \frac{1}{\|\mathbf{U}\|_\delta})$ et la norme $\|\mathbf{U}\|_\delta$ est donnée par la proposition 3.3.3.

Les coefficients probabilistiques de $\pi^{(\epsilon)}$ peuvent être calculés d'une manière recursive par les formules suivantes

$$\pi^{(0)} \mathbf{Q}_0 = \mathbf{0}, \quad (3.13)$$

$$\sum_{(i=0, j=0)}^{(M_{RT}, \infty)} \pi^{(0)}(i, j) = 1. \quad (3.14)$$

$$\pi^{(n)} \mathbf{Q}_0 + \pi^{(n-1)} \mathbf{Q}_1 = \mathbf{0}, \quad (3.15)$$

$$\sum_{(i=0, j=0)}^{(M_{RT}, \infty)} \pi^{(n)}(i, j) = \mathbf{0}, \quad \forall n \geq 1. \quad (3.16)$$

Tous les états de la chaîne de Markov non perturbée ne communiquent pas entre eux. Pour tout $\epsilon > 0$ suffisamment petit, les états de la chaîne de Markov perturbée communiquent entre eux, c'est à dire qu'il y a une discontinuité de $\pi(\epsilon)$ à $\epsilon = 0$. Donc la distribution stationnaire de la chaîne non perturbée n'est pas unique, en particulier, elle s'écrit en fonction des solutions stationnaires de toutes les classes ergodiques. Par exemple nous avons :

$$\pi_j^{(0)} = \bar{v}_j \tilde{q}_j,$$

où \bar{v}_j est la solution de l'équation (3.9) et sa forme est :

$$\bar{v}_i = \alpha(M_{RT})^i \bar{v}_0, \text{ avec } \bar{v}_0 = 1 - \alpha(M_{RT}), \forall i \in I_\infty \quad (3.17)$$

Pour trouver la forme explicite de la distribution stationnaire du système, nous introduisons quelques variables. Soit $\xi_{l(i)}$ le vecteur propre colonne, qui correspond à la valeur propre 1 de la matrice de transition d'une chaîne de Markov imperturbée, qui fournit la probabilité que le processus soit absorbé par la classe $l(i)$, sachant qu'il commence à partir de l'état i ($\mathbf{Q}_0 \xi_{l(i)} = 0$). Dans notre modèle, l'expression de ce vecteur est donnée pour $\forall j \in I_{M_{RT}}$ et $\forall k \in I_\infty$ par :

$$\xi_{l(i)}(k, j) = \begin{cases} 1 & \text{si } k = i; \\ 0 & \text{sinon.} \end{cases}$$

Soient \mathbf{N} la matrice des probabilités d'absorption par des classes ergodiques, c'est à dire nous avons $\mathbf{N} = [\xi_{l(0)} \quad \xi_{l(1)} \quad \dots]$ et \mathbf{M} matrice des probabilités stationnaires de la chaîne de Markov \mathbf{Q}_0 dont la forme est $\mathbf{M} = \text{diag}(\tilde{q}_{l(0)}, \tilde{q}_{l(1)}, \dots)$.

Proposition 3.3.3 *La solution d'équations (3.13)-(3.15) en utilisant les conditions de normalisation est donnée de manière récursive pour :*

1. Le premier coefficient, par

$$\pi_{(i,j)}^{(0)} = (1 - \alpha(M_{RT})) \alpha(M_{RT})^j \frac{(\rho_{RT})^i}{i!} \frac{1}{\sum_{k=0}^{M_{RT}} \frac{(\rho_{RT})^k}{k!}}. \quad (3.18)$$

2. Le coefficient d'ordre n ($n > 0$), par

$$\pi^{(n)} = \pi^{(0)} \mathbf{U}^n, \mathbf{U} = \mathbf{Q}_1 \mathbf{H} [\mathbf{I} + \mathbf{Q}_1 \Lambda \tilde{\mathbf{H}} \mathbf{M}], \quad (3.19)$$

avec

$$\|\mathbf{U}\|_\delta = g_1 \|\mathbf{H}\|_\delta (1 + g_1 \|\Lambda \tilde{\mathbf{H}} \mathbf{M}\|_\delta).$$

Les matrices $\mathbf{H}^0, \tilde{\mathbf{H}}$ sont des matrices de déviation de la chaîne imperturbée et de la chaîne agrégée, respectivement. Aussi nous avons

$$\Lambda = \begin{pmatrix} \underline{\xi} & \underline{0} & \underline{0} & \underline{0} & \dots \\ \underline{0} & \underline{\xi} & \underline{0} & \underline{0} & \dots \\ \underline{0} & \underline{0} & \underline{\xi} & \underline{0} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}, \text{ avec } \underline{\xi} = \begin{bmatrix} 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} \text{ et } \underline{0} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}.$$

Ainsi la matrice \mathbf{H} est donnée par

$$\mathbf{H} = [-\mathbf{Q}_0 + \mathbf{NM}]^{-1} - \mathbf{NM} = \text{diag}(\mathbf{H}_0, \dots)$$

où

$$\mathbf{H}_0 = [q^* - \mathbf{Q}_{RT}]^{-1} - q^*, \quad q^* = \underline{\zeta} \tilde{q},$$

Car nous avons $\tilde{q} = \tilde{q}_j$, d'où le δ -norme de la matrice de déviation \mathbf{H}^0 est donnée par :

$$\|\mathbf{H}\|_\delta = \|H_0\|_\delta, \quad \tilde{\mathbf{H}} = [-\tilde{\mathbf{Q}}_1 + \tilde{\mathbf{Q}}_1^*]^{-1} - \tilde{\mathbf{Q}}_1^*, \quad \tilde{\mathbf{Q}}_1^* = \zeta_{\mathbf{B}} \tilde{q}, \quad \zeta_{\mathbf{B}} = (1, \dots, 1) \in \mathfrak{R}^{\infty \times 1}$$

Les blocs matricielles $U_{i,j}$ de la matrice $\mathbf{U} = \{U_{i,j}\}_{i,j \in F}$ sont obtenus par la relation :

$$U_{i,j} = \mathbf{Q}_{i,j}^1 \mathbf{H}_j + \sum_{l \in F} \mathbf{Q}_{i,l}^1 \mathbf{H}_l \sum_{k \in B} \mathbf{Q}_{l,k}^1 \tilde{\mathbf{H}}_{k,j} \mathbf{1} q_j.$$

Nous constatons ici que notre approche de la perturbation singulière est plus simple que l'approche d'analyse spectrale que nous avons utilisé dans (Ibrahimi et al., 2007). Nous avons juste besoin de calculer la distribution stationnaire pour chaque classe récurrente de la chaîne de Markov non perturbée \mathbf{Q}_0 dans (3.10). Aussi on calcule la distribution invariante de la chaîne agrégée $\tilde{\mathbf{Q}}_1$ par (3.17).

Lemme 7 *La distribution stationnaire du nombre d'appels en BE selon l'ordre des coefficients de la série de Taylor est donnée en distinguant les deux cas suivants :*

1. Pour le premier coefficient $\pi^{(0)}$, nous trouvons

$$P_\epsilon^{(0)}(Y = j) = \pi_j^{(0)} \mathbf{1} = (1 - \alpha(M_{RT})) \alpha(M_{RT})^j.$$

2. Pour les autres coefficients $\pi^{(n)}$ ($n \geq 1$), nous trouvons

$$P_\epsilon^{(n)}(Y = j) = \pi_j^{(0)} \mathbf{1} + \epsilon \pi_j^{(1)} \mathbf{1} + \dots + \epsilon^n \pi_j^{(n)} \mathbf{1},$$

$$\text{où } \pi_j^{(n)} = \sum_{i \in B} \pi_i^{(0)} \mathbf{u}_{i,j}^{(n)} \text{ et } \mathbf{u}^n = \{\mathbf{u}_{i,j}^{(n)}\}.$$

Après avoir trouvé les formes explicites des probabilités marginales, nous déduisons certains métriques de performance par la suite.

3.3.4 Temps de séjour d'appels *best effort*

Maintenant, nous pouvons déterminer les métriques de performance du système de manière approximative. En effet, le nombre moyen marginal $E_\epsilon(Y)$ d'appels BE dans le système est donné par

$$E_\epsilon(Y) = \sum_{j=0}^{\infty} j P_\epsilon^{(n)}(Y = j), \quad n \geq 0. \quad (3.20)$$

Le temps de séjour moyen d'appels BE dans le système est donné par la formule de Little

$$T_\epsilon^{BE} = \frac{1}{\lambda_{BE}} \sum_{j=0}^{\infty} j P_\epsilon^{(n)}(Y = j), \quad n \geq 0. \quad (3.21)$$

En particulier, pour le premier coefficient de la série de Taylor, c-à-d pour $n = 0$, nous trouvons les performances suivantes

$$\begin{aligned} E_\epsilon(Y) &= \frac{\alpha(M_{RT})}{1 - \alpha(M_{RT})}, \\ T_\epsilon^{BE} &= \frac{1}{\lambda_{BE}} \frac{\alpha(M_{RT})}{1 - \alpha(M_{RT})}. \end{aligned}$$

Nous écrivons le temps de séjour moyen d'appels BE en fonction du seuil L_{BE} , sous la forme finale

$$T^{BE}(L_{BE}) = \frac{\frac{1}{\epsilon}}{\gamma[\sum_{i=0}^{N_{RT}} \tilde{q}(i)\beta_{\epsilon'} + \frac{L_{BE}}{1-L_{BE}} \sum_{i=N_{RT}+1}^{M_{RT}} \tilde{q}(i)] - \bar{\lambda}_{BE}}, \quad (3.22)$$

où $\gamma = \frac{\bar{\mu}_{BE}W}{\Gamma E_{BE}/N_0}$ et $\beta_{\epsilon'} = \frac{\Theta - i\Delta_{RT}^r}{1 - \Theta + i\Delta_{RT}^r}$. L'équation ci-dessus montre l'influence du seuil L_{BE} sur le temps de séjour d'appels BE. Nous rappelons que les nombres d'appels M_{RT} et N_{RT} dépendent aussi du seuil L_{BE} .

Remarque 2 Nous signalons que le choix de l'ordre de l'approximation dans la série de Taylor dépend du degré de précision entre les deux approches la perturbation singulière et l'analyse spectrale. Donc, si le premier coefficient ne donne pas une bonne approximation de la distribution stationnaire, sachant que l'hypothèse considérée, est toujours valide, nous passons au deuxième coefficient et ainsi de suite jusqu'à ce qu'on trouve la bonne approximation.

3.4 Résultats numériques

Nous considérons les paramètres suivants. Le débit minimum d'appels RT est 4.75 *kbps* et le débit maximum est 12.2 *kbps*. L'énergie sur bruit pour transmettre un bit de type RT est $E_{RT}/N_0 = 4.1$ dB, $E_{BE}/N_0 = 4.1$ dB (Holma et Toskala, 2001), $\bar{p}_{BE} = 0.55$, $\bar{p}_{RT} = 0.5$, $R_T = 38$ *kbps* et $\Theta = 1 - 10^{-5}$. Le taux d'arrivée d'appels BE est $\lambda_{BE} = 0.209\epsilon$, leur taux de départ est $\mu_{BE} = 10^{-5}\epsilon$ et le paramètre de la perturbation est $\epsilon = 10^{-3}$. La figure 3.1, présente une comparaison de la probabilité marginale d'appels BE, obtenue par les deux méthodes : l'analyse spectrale (voir la première partie du chapitre 2) et la perturbation singulière. Nous avons considéré dans l'approche de perturbation le premier terme de la série de Taylor comme approximation de la distribution stationnaire. Nous remarquons que ce premier coefficient de Taylor donne une très bonne approximation quand le paramètre de la perturbation est suffisamment petit, par exemple $\epsilon = 10^{-3}$.

Les figures 3.3, 3.4 et 3.2, exposent le temps de séjour d'appels BE en fonction du seuil L_{BE} . Nous rappelons que l'hypothèse de l'applicabilité de la méthode perturbation singulière est que les appels BE se produisent moins fréquemment que les appels RT. Nous observons que le premier coefficient d'approximation dans la série de Taylor donne une parfaite approximation dans la figure 3.2. Nous constatons que l'erreur entre la Méthode exacte (analyse spectrale) et la méthode approximative (perturbation singulière)

devient significative quand le temps de service d'appels RT décroît (voir les figures 3.3 et 3.4). Ce Phénomène est due au fait que lorsque l'on utilise le premier terme de la série de Taylor, on obtient le comportement limite si ϵ tend vers zéro, c'est à dire, que les transitions du nombre d'appels RT se produisent beaucoup plus rapidement que les transitions du nombre d'appels BE. Toutefois, lorsque le temps de service d'appels RT diminue, les appels RT passent plus de temps dans le système, ce qui veut dire que les transitions du nombre d'appels RT devient de moins en moins lent et nous avons donc les mêmes configurations de transition pour les deux appels RT et BE. Or, ce n'est pas le cas dans notre hypothèse où nous avons supposé deux échelles de transitions différentes, une rapide (nombre d'appels RT qui se produisent rapidement) et une lente (nombre d'appels BE qui se produisent lentement). Néanmoins, le fournisseur de service ou l'opérateur peut utiliser cette approche basée sur l'approximation pour évaluer en temps réel avec un calcul rapide des métriques de performances d'appels RT et BE. Ainsi, le fournisseur peut décider dans ce cas entre la meilleure configuration du choix de seuil L_{BE} et le contrôle d'admission qui améliore l'utilisation de la bande passante. En effet, les performances du système dépendent de plusieurs facteurs : les taux d'arrivée d'appels, la durée moyenne d'appel RT, les effets de masques, Ces facteurs changent dynamiquement et le système a besoin de réévaluer les performances à nouveau. Par conséquent, notre approche permet à l'opérateur d'avoir un meilleur CAC et une meilleure utilisation de la bande passante tout en satisfaisant les demandes de qualité de service d'appels RT et BE.

Enfin, l'analyse des résultats nous permet de constater que l'erreur entre la solution exacte (méthode d'analyse spectrale) et la solution approximative (méthode de la perturbation singulière) croît quand la durée moyenne de service ($1/\mu_{RT}$) d'un paquet de type RT augmente. Cet écart d'erreur est dû au fait que la durée de service ($1/\mu_{RT}$) d'un paquet RT augmente alors que les appels en RT restent longtemps dans le système ; ce qui n'est pas le cas dans notre modélisation où nous avons supposé qu'ils varient rapidement dans le système. Il est donc normal que l'approximation soit distincte de la solution exacte.

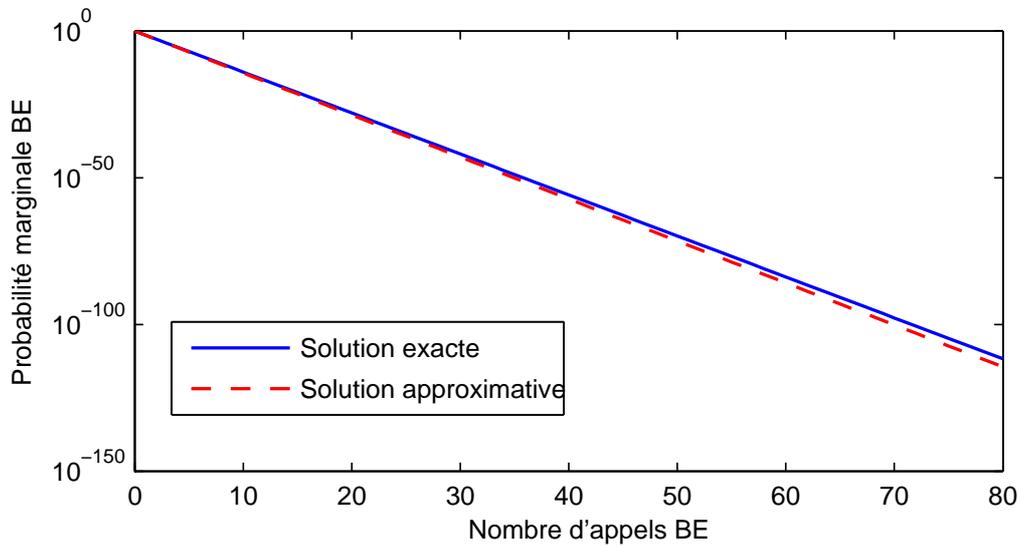


FIGURE 3.1 – Probabilité marginale en fonction du nombre d'appels BE, pour $\epsilon = 10^{-3}$, $\lambda_{BE} = 0.209\epsilon$, $\mu_{BE}^\epsilon = 10^{-5}\epsilon$, $\lambda_{RT} = 0.1557$, $L_{BE} = 0.752$ et $\mu_{RT} = 0.1$. Comparaison entre deux méthodes : Analyse Spectrale (solution exacte) et Perturbation Singulière (solution approximative).

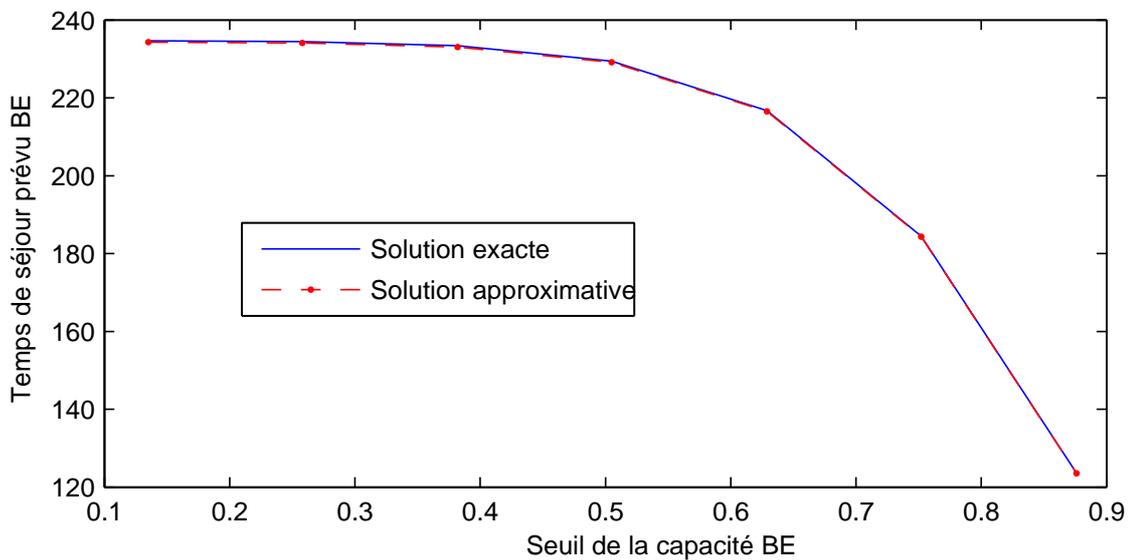


FIGURE 3.2 – Temps de séjour d'appels BE en fonction du seuil L_{BE} , pour $\epsilon = 10^{-3}$, $\lambda_{BE} = 0.209\epsilon$, $\mu_{BE}^\epsilon = 10^{-5}\epsilon$, $\lambda_{RT} = 0.1557$ et $\mu_{RT} = 100$. Comparaison entre deux méthodes : Analyse Spectrale (solution exacte) et Perturbation Singulière (solution approximative).

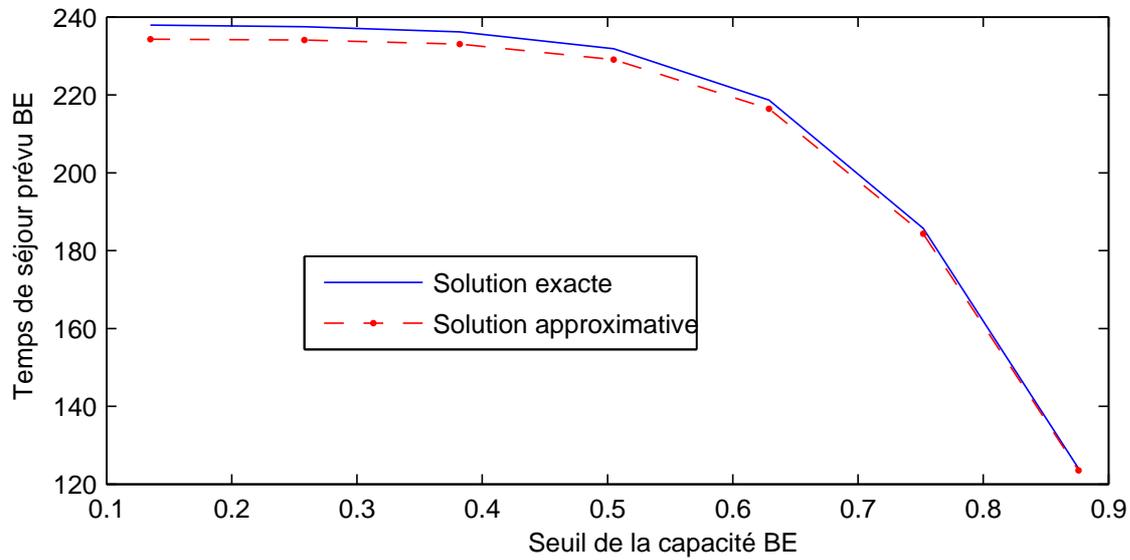


FIGURE 3.3 – Temps de séjour d’appels BE en fonction du seuil L_{BE} , pour $\epsilon = 10^{-3}$, $\lambda_{BE} = 0.209\epsilon$, $\mu_{BE}^\epsilon = 10^{-5}\epsilon$, $\lambda_{RT} = 1.5573$ et $\mu_{RT} = 1$. Comparaison entre deux méthodes : Analyse Spectrale (solution exacte) et Perturbation Singulière (solution approximative).

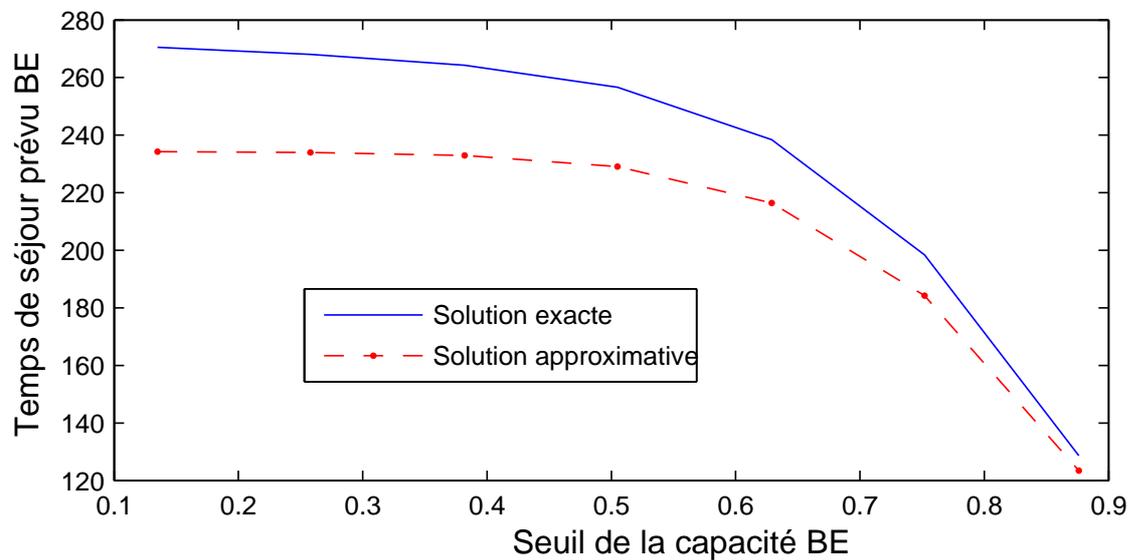


FIGURE 3.4 – Temps de séjour d’appels BE en fonction du seuil L_{BE} , pour $\epsilon = 10^{-3}$, $\lambda_{BE} = 0.209\epsilon$, $\mu_{BE}^\epsilon = 10^{-5}\epsilon$, $\lambda_{RT} = 0.1557$ et $\mu_{RT} = 0.1$. Comparaison entre deux méthodes : Analyse Spectrale (solution exacte) et Perturbation Singulière (solution approximative).

3.5 Conclusion

Dans le chapitre précédent, nous avons fourni une méthode exacte qui donne des formes explicites aux performances du système, en terme de probabilité d'existence et de blocage d'appels, et de temps de séjour. L'avantage de cette méthode est la réduction de la complexité du problème. Les performances du système ne dépendent que du nombre d'appels RT. Cette indépendance en nombre d'appels BE nous permet de ramener un problème avec un nombre arbitraire d'appels BE (voir infini) à un problème relativement simple. Alors quand le nombre d'appels RT augmente, la méthode prend un temps de calcul important et donne quelques erreurs, ce qui nous suggère à utiliser une autre approche complémentaire et rapide au niveau calcul. En effet, dans ce chapitre, nous avons étudié le système WCDMA avec deux types d'appels : temps réel (RT) qui ont des ressources dédiées et non temps réel ou *best effort* (BE) qui partagent la capacité restante du système. Nous avons considéré durant ce travail une capacité minimum réservée aux appels BE ainsi que toute portion de la capacité non utilisée par les appels RT. Notre approche d'analyse du système est basée sur la modélisation du système par une chaîne de Markov à deux dimensions, où la première correspond au nombre d'appels RT et la deuxième correspond au nombre d'appels BE dans le système. Néanmoins, nous avons utilisé une approximation basée sur l'approche de la perturbation singulière pour trouver une approximation simple de la distribution d'états d'équilibre de cette chaîne de Markov. Nous avons montré sous forme numérique que notre approche donne une bonne approximation de la distribution d'états d'équilibre avec des calculs plus simple que dans le cas de l'analyse spectrale. Toutefois, le système peut utiliser cette approche pour évaluer en temps réel les mesures de performances d'appels RT et BE.

