Jeux stochastiques et contrôle de puissance distribué

François MÉRIAUX[1], Maël LE TREUST[1], Samson LASAULCE[1], Michel KIEFFER[12]

[1]L2S - CNRS - SUPELEC - Université Paris-Sud
3 rue Joliot-Curie F-91192 Gif-sur-Yvette, France

[2]LTCI - CNRS - Telecom ParisTech
46 rue Barrault F-75013 Paris, France

francois.meriaux@lss.supelec.fr, mael.letreust@lss.supelec.fr
samson.lasaulce@lss.supelec.fr, michel.kieffer@lss.supelec.fr

Résumé – Les émetteurs d'un canal à accès multiple sont supposés choisir eux-mêmes leur stratégie de contrôle de puissance de manière à être efficaces énergétiquement. Nous montrons que le concept de jeux stochastiques permet de concevoir des stratégies de contrôle à la fois distribuées, efficaces globalement et ne nécessitant qu'une connaissance partielle du système de communication. La région de tous les points d'utilité d'équilibre est établie et une stratégie pratique de contrôle de puissance de l'émetteur, reposant sur le partage temporel légitime, est proposée.

Abstract – Transmitters of a multiple access channel are assumed to freely choose their power control strategy in order to be energy-efficient. We show that in a stochastic game framework, we can develop energy-efficient distributed control strategies which only require partial knowledge of the entire system. Achievable utility equilibrium region is characterized and based on time-sharing, an explicit power control strategy is proposed.

1 Introduction

Dans un système de communication sans fil où plusieurs émetteurs voient leur signaux interférer en réception, la disparité des dynamiques de puissance des composantes du signal reçu pose généralement problème au récepteur. Et ce, notamment lorsque le récepteur doit décoder plusieurs de ces composantes. Le contrôle de puissance à l'émission vise précisément à compenser cette forte disparité. Dans cet article, nous nous intéressons à un scénario d'importance croissante, celui des systèmes distribués. Dans ce cadre, l'émetteur décide de sa politique de contrôle de puissance en vue de maximiser sa propre métrique de performance. La métrique retenue, appelée utilité, est l'efficacité énergétique (en bit par Joule). Ce cadre est exactement celui introduit par Goodman et al. dans [4]. Les auteurs de [4] ont remarqué que la théorie des jeux, théorie dont l'essence même est d'étudier des preneurs de décisions dont les actions sont inter-dépendantes, est un outil pertinent pour analyser ce problème. Leur modèle, à savoir un modèle de jeu en un coup joué pour chaque paquet de données émis (les joueurs étant les émetteurs et l'action d'un joueur consistant à choisir son niveau de puissance), conduit à une stratégie de contrôle pratique (reposant sur une connaissance limitée du système) mais inefficace globalement. Plus précisément, on peut démontrer qu'il existe une politique de contrôle qui Pareto-domine leur solution, c'est-à-dire pour laquelle tous les émetteurs font mieux en termes d'utilité. Les auteurs de [8] ont démontré qu'un modèle de jeu répété [7] permet d'avoir une modélisation plus fine du problème, modélisation qui conduit à des solutions plus efficaces globalement. L'idée fondamentale et nouvelle en contrôle de puissance, et que nous exploitons dans cet article, est qu'il ne faut pas supposer le contrôle de puissance indépendant d'un paquet à l'autre, et ceci même si les réalisations des gains des canaux sont indépendantes. Un modèle de jeu dynamique tel que le jeu répété permet de tenir compte du fait que les joueurs interagissent plusieurs fois et ceci conduit à créer une corrélation entre les niveaux de puissances choisis par un joueur au cours du temps, et nous le répétons, même pour des canaux dits i.i.d. La contribution de cet article est de généraliser les travaux de [8] en relaxant une hypothèse de normalisation de l'utilité individuelle par le gain de canal. Pour faire cela, nous utilisons un modèle de jeux stochastiques [6], ce qui nous amène à supprimer la sous-optimalité en termes de performances induite par la normalisation nécessaire au modèle de jeu répété. Les travaux de [3, 5] sont alors utilisés pour obtenir un *Folk* théorème qui caractérise la région des utilités atteignables de ce jeu stochastique. Nous présentons également une stratégie de contrôle de puissance explicite pour ce jeu.

Dans le paragraphe 2, nous détaillons le modèle du jeu stochastique que nous considérons. Au paragraphe 3, nous présentons les résultats analytiques obtenus en ce qui concerne la région des utilités atteignables ainsi que les résultats d'équilibre et de performance de la stratégie de Sélection des Meilleurs Utilisateurs (SMU). Dans le paragraphe 4 sont présentés les résultats de simulation obtenus pour comparer la stratégie SMU à d'autres stratégies de contrôle de puissance.

2 Modélisation du problème par un jeu stochastique

Nous considérons un canal à accès multiple, décentralisé au sens du contrôle de puissance, pour lequel K utilisateurs transmettent vers un récepteur sur des intervalles de temps (durée d'un paquet), que nous appellerons étapes du jeu répété, sur lesquels les canaux sont supposés statiques. À chaque étape, les canaux sélectifs en temps mais non sélectifs en fréquence, notés h_i, sont tirés de manière indépendante sur un ensemble admissible : $|h_i|^2 \in [\eta_i^{min}, \eta_i^{max}] = \Gamma_i$. Nous supposons vérifiée l'hypothèse de réciprocité des canaux montants et descendants. De plus, nous supposons que les terminaux sont capables d'estimer avec une erreur négligeable leur canaux montants (via un mécanisme de séquences d'apprentissage, une boucle de retour, etc). Le signal reçu peut s'écrire :

$$Y = \sum_{i=1}^{K} h_i X_i + Z \tag{1}$$

avec $\mathbb{E}|X_i|^2 = p_i$ et $Z \sim \mathcal{N}(0, \sigma^2)$. Dans un contexte où le récepteur décode le signal de chaque émetteur séparément et où il n'y a pas de mécanisme tel que la formation de voie [9] pour atténuer les interférences, pour chaque utilisateur $i \in \mathcal{K} = \{1, 2, ..., K\}$, le rapport signal sur interférence plus bruit (RSIB) est donné par :

$$\text{RSIB}_i = \gamma_i = \frac{p_i |h_i|^2}{\sum_{j \neq i} p_j |h_j|^2 + \sigma^2} \tag{2}$$

Nous pouvons maintenant définir le jeu stochastique qui modélise l'interaction entre les émetteurs qui choisissent leur niveau de puissance au cours du temps.

Définition 1 (Jeu stochastique) *Un jeu stochastique avec observation parfaite est défini par l'uplet :*

$$\mathcal{G} = (\mathcal{K}, (\mathcal{T}_i)_{i \in \mathcal{K}}, (v_i)_{i \in \mathcal{K}}, (\Gamma_i)_{i \in \mathcal{K}}, \pi, \Theta), \tag{3}$$

avec $\mathcal{K}$ l'ensemble des joueurs, $\mathcal{T}_i$ l'ensemble des stratégies pour le joueur i, v_i la fonction d'utilité du joueur i sur le long terme, Γ_i l'intervalle des états de canaux accessibles au joueur i, π la probabilité de transition sur les états et Θ l'espace des observations.

La stratégie et l'utilité sur le long terme du joueur i sont définies comme suit.

Définition 2 (Stratégie des joueurs) *La stratégie du joueur $i \in \mathcal{K}$ est une séquence de fonctions $(\tau_{i,t})_{t \geq 1}$ avec*

$$\tau_{i,t} : \left| \begin{array}{ccc} \Theta^t & \rightarrow & \mathcal{A}_i \\ \underline{h}_t & \mapsto & p_i(t). \end{array} \right. \tag{4}$$

À l'histoire $\underline{h}_t = (\theta(1)..., \theta(t-1), \underline{\eta}(t)) \in \Theta^t$ (observations passées et état présent), on associe une action $p_i(t) \in \mathcal{A}_i$.

La stratégie du joueur i est notée τ_i et le vecteur de statégies $\underline{\tau} = (\tau_1, ..., \tau_K)$ est nommé stratégie jointe. Une stratégie jointe $\underline{\tau}$ entraîne une unique séquence d'actions $(\underline{p}(t))_{t \geq 1}$.

Définition 3 (Utilité des joueurs) *Soit $\underline{\tau}$ une stratégie jointe. L'utilité du joueur $i \in \mathcal{K}$ sachant que l'état initial du canal est $\underline{\eta}(1)$ est définie par*

$$v_i(\underline{\tau}, \underline{\eta}(1)) = \sum_{t \geq 1} \lambda (1-\lambda)^{t-1} \mathbb{E}_{\underline{\tau}, \pi} \left[u_i(\underline{p}(t), \underline{\eta}(t)) | \underline{\eta}(1) \right] \tag{5}$$

avec $u_i(p_1, ..., p_K) = \frac{R_i f(\text{RSIB}_i)}{p_i}$ [bit/J], l'utilité instantanée du joueur i telle que définie dans [4]. R_i est le débit d'émission du joueur i, f est la fonction d'efficacité, elle prend ses valeurs entre 0 et 1. Le paramètre λ est appelé facteur d'escompte. Il peut être interprété comme une probabilité d'arrêt ou le fait que les joueurs apprécient différemment leurs gains à court terme et leurs gains à long terme.

3 Résultats analytiques

3.1 Folk Théorème

Théorème 4 (Folk) *Soit F l'ensemble des utilités atteignables et individuellement rationnelles. Sous l'hypothèse que les joueurs disposent du même signal public, alors pour tout profil d'utilité $\underline{u} \in F$, il existe λ_0 tel que pour tout $\lambda < \lambda_0$, il existe une stratégie d'équilibre public et parfait du jeu stochastique dont l'utilité à long terme vaut $\underline{u} \in F$.*

Il faut noter qu'une telle caractérisation de la région d'utilités atteignables est très puissante. En effet, la technique classique pour obtenir la région d'utilités atteignables consisterait à déterminer toutes les stratégies possibles pour les joueurs puis de calculer les utilités correspondantes. Dans un jeu très simple où chaque joueur n'aurait le choix qu'entre deux niveaux de puissance à chaque étape, il faudrait considérer 2^N stratégies possibles, avec N le nombre d'étapes du jeu. D'après [2], le *Folk* théorème nous autorise à considérer uniquement les stratégies dites de Markov sans perte d'optimalité, le nombre de stratégies à étudier se réduit donc à $2^{|\Gamma|}$ avec $|\Gamma|$ le nombre d'états de canaux.

3.2 Stratégie de Sélections des Meilleurs Utilisateurs

Obtenir une région d'utilités atteignables est une chose, mais il reste à définir formellement des stratégies efficaces dans cette région. C'est ce que nous proposons de faire avec l'introduction d'une stratégie dite de *Sélection des Meilleurs Utilisteurs.*

La stratégie proposée est basée sur le point de fonctionnement présenté dans [8]:

$$\forall i \in \mathcal{K},\ \tilde{p_i}(t) = \frac{\sigma^2}{\eta_i(t)} \frac{\tilde{\gamma}_K}{1-(K-1)\tilde{\gamma}_K} \tag{6}$$

où $\tilde{\gamma}_K$ est l'unique solution non nulle de

$$x(1-(K-1)x)f'(x) - f(x) = 0. \tag{7}$$

Contrairement au cas du jeu répété où les gains des canaux sont constants, quand ces derniers varient à chaque étape, la stratégie consistant à ce que chaque joueur émette au point de fonctionnement (6) n'est plus optimale. Il se trouve qu'on obtient de meilleurs résultats en termes de bien-être social si on réduit l'ensemble des joueurs émettant au point de fonctionnement. Cette approche est intitulée stratégie de *Sélection des Meilleurs Utilisateurs*, elle est caratérisée de la manière suivante.

A chaque étape t du jeu, le récepteur fixe $\mathcal{K}'^t \subset \mathcal{K}$, l'ensemble optimal de joueurs émettant au point de fonctionnement (6) pour maximiser la somme des utilités instantannées des joueurs. Pour chaque joueur $i \in \mathcal{K}$:

- Si $i \in \mathcal{K}'^t$, il lui est recommandé d'émettre au point de fonctionnement (6) à l'étape t.
- Si $i \notin \mathcal{K}'^t$, il lui est demandé de ne pas émettre à cette étape.

Il faut bien noter que le comportement des joueurs n'est pas imposé, le récepteur envoie seulement des recommandations aux joueurs. Pour assurer que cette stratégie soit un équilibre, un mécanisme de punition est établi: si un joueur dévie de la stratégie, les autres joueurs jouent l'équilibre de Nash en un coup pour le restant du jeu. L'équilibre de la stratégie est assuré si le maximum (en termes d'utilité) que peut gagner un joueur en déviant à une étape du jeu est inférieur à ce qu'il va perdre en étant puni par les autres joueurs jusqu'à la fin du jeu. Nous obtenons alors la condition d'équilibre suivante :

Théorème 5 (Équilibre de la stratégie) *La stratégie SMU est un équilibre du jeu stochastique si* $\forall i \in \mathcal{K}$

$$\lambda \leq \frac{\mathbb{E}[u_i(\underline{p}^{smu}, \underline{\eta})] - \mathbb{E}[u_i(\underline{p}^{*}, \underline{\eta})]}{\frac{R\eta_{\max}}{\sigma^2}\frac{f(\beta^*)}{\beta^*} + \mathbb{E}[u_i(\underline{p}^{smu}, \underline{\eta})] - \mathbb{E}[u_i(\underline{p}^{*}, \underline{\eta})]} \tag{8}$$

avec $\underline{p}^{smu}$ le profil de puissance résultant de l'application de la stratégie SMU et $\underline{p}^*$ et β^* respectivement le profil de puissance et le RSIB correspondant à l'équilibre de Nash en un coup.

La complexité de calcul nécessaire à l'éxecution de cette stratégie est faible puisqu'on peut prouver qu'à débit d'émission égal, la sélection optimale de k joueurs pour émettre au point de fonctionnement (6) est l'ensemble des k joueurs avec les meilleurs gains de canaux. Ainsi dans un jeu à K joueur, le récepteur doit comparer K combinaisons de joueurs et non 2^K.

4 Résultats numériques

Pour l'obtention de résultats numériques, nous utilisons la fonction d'efficacité $f(\gamma) = e^{-\frac{a}{\gamma}}$ avec $a = 2^R - 1$. Cette fonction est introduite dans [1].

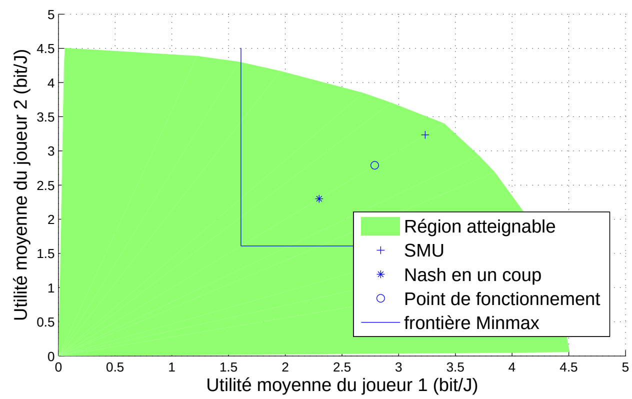

FIG. 1: *Région atteignable et utilités moyennes de diverses stratégies pour un jeu à 2 joueurs.*

La figure 1 illustre la région atteignable pour un jeu à 2 joueurs et 2 états de canaux (avec $\frac{\eta_{\max}}{\eta_{\min}} = 4$) en considérant toutes les stratégies possibles. La frontière *min-max* délimite la région d'équilibre. Les utilités moyennes de SMU, du point de fonctionnement et de l'équilibre de Nash en un coup sont également représentées à l'intérieur de cette région. Notons que que la stratégie SMU Pareto-domine les autres stratégies considérées.

La simulation présentée en figure 2 compare les utilités instantannées moyennes de quatre mécanismes de contrôle de puissance en fonction du nombre d'émetteurs. Pour cette simulation, on considère un nombre fini de gains de canal. La loi d'évolution des gains des canaux suit la propriété de Markov, c'est-à-dire qu'il existe une matrice de probabilité de transtion entre l'état des canaux à l'instant t et l'état des canaux à l'instant $t+1$. Cette marice ainsi que les états de gains de canal accessibles sont les mêmes pour tous les joueurs. A travers l'étude de ces quatres mécanismes, nous étudions les performances atteignables en fonction du caractère centralisé ou décentralisé du mécanisme ainsi que de la quantité d'information disponible sur le système. Ces mécanismes sont les suivants :

- Une version centralisée de SMU, dans laquelle le récepteur choisit qui émet à chaque tour et impose la puissance d'émission en connaissant les gains des canaux à l'instant t. Dans le modèle considéré, les émetteurs appliquent à l'instant $t+1$ la puissance d'émission décidée à l'instant t. Ce retard se justifie par un temps de transmission entre le récepteur et les émetteurs.
- SMU, pour lequel le récepteur décide uniquement

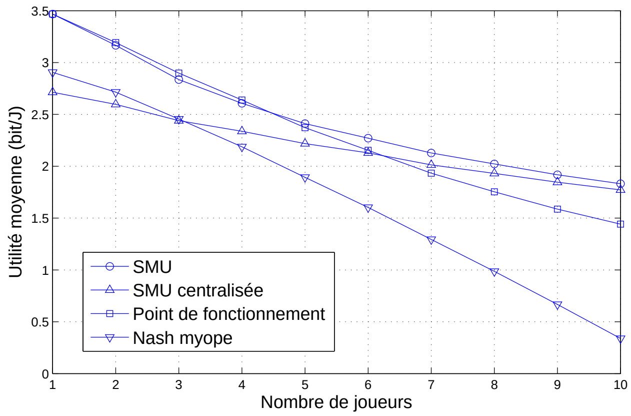

FIG. 2: *Utilités moyennes de quatre mécanismes de contrôle de puissance en fonction du nombre d'émetteurs.*

l'ensemble des émetteurs conseillé à chaque tour du jeu. Chaque émetteur connaissant le gain de son canal et le nombres des autres émetteurs qui vont transmettre avec lui, il fixe lui-même sa puissance d'émission. De la même manière que précédemment, on prend en compte le retard de transmission entre le récepteur et les émeteurs. L'ensemble des joueurs qui émettent à l'instant $t+1$ est donc décidé par le récepteur à l'instant t.

– La stratégie reposant sur le point de fonctionnement développée dans [8]. L'approche est encore plus décentralisée puisque tous les émetteurs fixent leur puissance à chaque tour en connaissant le gain de leur canal et le nombre de joueurssans recommandation de la part du récepteur.
– Un équilibre de Nash "myope". Dans ce cas, les émetteurs n'ont aucune information sur le système mis à part l'espérance du gain de leur canal et le nombre de joueurs. Ils se contentent donc de jouer l'équilibre de Nash statique.

Il est intéressant de noter que SMU offre de meilleures performances que les trois autres mécanismes. En ce qui concerne l'approche centralisée, le fait que la puissance d'émission soit connue des émetteurs avec un temps de retard par rapport à l'état des gains des canaux est un véritable handicap qui n'est compensé que pour un nombre suffisant d'émetteurs.

5 Conclusion et perspectives

Dans un réseau sans fil distribué où les émetteurs sont des agents égoistes libres de choisir leur puissance d'émission pour chaque paquet, les interactions à long terme méritent d'être étudiées. Le cadre des jeux stochastiques permet de prendre en compte le caractère répété de ces interactions ainsi que les variations des gains des canaux d'un paquet au suivant. Cette approche nous permet notamment de caractériser la région des utilités atteignables. Il apparaît qu'étant données les interactions sur à long terme entre les émetteurs, ces derniers peuvent avoir intérêt à ne pas émettre certains paquets si leurs conditions de canal sont trop mauvaises. Cela nous mène à établir une stratégie de contrôle de puissance fondée sur le partage temporel qui se montre performante en termes d'efficacité énergétique.

Les perspectives de ce travail sont d'intégrer dans le contrôle de puissance plusieurs aspects visant à mieux prendre en compte les caractéristiques des flux d'information dans des réseaux réels: la possibilité de tolérer un retard sur l'émission d'un paquet (*delay tolerant networks*); la possibilité d'avoir un flux de paquets sporadique; le fait que la taille mémoire de stockage des paquets à l'émetteur est finie.

Références

[1] E. V. Belmega, S. Lasaulce, and M. Debbah. Power allocation games for MIMO multiple access channels with coordination. *Trans. Wireless. Comm.*, 8(6):3182–3192, 2009.

[2] P. K. Dutta. A folk theorem for stochastic games. *Journal of Economic Theory*, 66(1):1 – 32, 1995.

[3] D. Fudenberg and Y. Yamamoto. The folk theorem for irreducible stochastic games with imperfect public monitoring. *Journal of Economic Theory*, In Press, Corrected Proof, 2011.

[4] D. J. Goodman and N. B. Mandayam. Power control for wireless data. *IEEE Person. Comm.*, 7:48–54, 2000.

[5] J. Hörner, T. Sugaya, S. Takahashi, and N. Vieille. Recursive methods in discounted stochastic games : An algorithm for delta approaching 1 and a folk theorem. Cowles Foundation Discussion Papers 1742, Cowles Foundation for Research in Economics, Yale University, December 2009.

[6] L. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences of the United States of America*, 39(10):1095–1100, 1953.

[7] S. Sorin. *Repeated Games with Complete Information, in Hanbook of Game Theory with Economic Applications*, volume 1. Elsevier Science Publishers, 1992.

[8] M. Le Treust and S. Lasaulce. A repeated game formulation of energy-efficient decentralized power control. *IEEE Trans. on Wireless Commun.*, 2010.

[9] B. D. Van Veen and K. M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE Signal Processing Magazine*, 5(2):4–24, April 1988.