Coopération dans les réseaux ad hoc : Application de la théorie des jeux et de l'évolution dans le cadre d'observabilité imparfaite

Pietro Michiardi

Institut Eurecom
2229, route des Cretes BP 193
06904 Sophia-Antipolis, France
Pietro.Michiardi@eurecom.fr

Résumé Les systèmes de communication basés sur le paradigme ad hoc ont reçu beaucoup d'attention dans le passé et les efforts des chercheurs ont été dévoués surtout à produire des protocoles décentralisés pour garantir le routage des paquets en tenant compte de la mobilité des nœuds du réseau. Dans cet article on questionne l'hypothèse de base assumé pendent la conception de ces protocoles selon laquelle les entités impliquées dans l'exécution de ces algorithmes suivent précisément les règles dictées par les concepteurs du système. Pour faire face aux déviations des nœuds du réseau du comportement préétabli afin d'en traire des bénéfices on analyse les caractéristiques d'un mécanisme de coopération basé sur la réputation présenté récemment dans la littérature et nommé CORE. Une approche basée sur la théorie des jeux répétés est utilisée pour déduire les caractéristiques d'une stratégie dérivée de CORE dans le cadre réaliste d'une observabilité imparfaite. Dans cet article on montre comme notre stratégie catalyse la naissance de la coopération entre les nœuds du réseau en présence des erreurs qui affectent l'évaluation de la réputation associé aux nœuds dûes aux interférences et collisions qui caractérisent la technologie radio 802.11.

1 Introduction

Un réseau ad hoc mobile (MANET) est un réseau maillé temporaire constitué par une collection de noeuds sans fil et mobiles sans l'aide d'une infrastructure préétablie utilisée pour exécuter les fonctions de base de gestion de réseau comme le cheminement et l'expédition de paquets. Dans un tel environnement, il peut être nécessaire qu'un noeud mobile demande l'aide d'autres noeuds pour expédier un paquet à sa destination, à cause de la couverture limitée du champ radio disponible à chaque noeud.

En origine, des applications exploitant les réseaux ad hoc ont été envisagées principalement pour des situations de crise (par exemple, dans les champs de bataille ou pour des opèrations de secours). Dans ces applications, tous les noeuds

du réseau appartiennent à une même autorité (par exemple, une unité militaire ou une équipe de secours) et ont un but commun. Cependant, les technologies sans fil se sont sensiblement améliorées ces dernières années et des dispositifs peu coûteux basés sur la norme 802.11 ont envahi le marché et le déploiement des réseaux ad hoc pour des applications commerciales est devenu réaliste. Des exemples incluent le déploiement des réseaux ad hoc pour l'automobile ou pour la fourniture d'équipements de communication pour les régions éloignées ou les périmètres physiquement délimités (par exemple, centres commerciaux, aéroports, etc.....). Dans ces réseaux, les noeuds n'appartiennent pas à la même d'organisation ni à une même autorité et ils ne poursuivent pas un but commun. En outre, les réseaux commerciaux pourraient être plus grands et avoir une plus longue vie ; de plus ils peuvent être complètement autonomes, signifiant que le réseau fonctionnerait seulement grâce à l'opération des utilisateurs.

Selon le type d'application, il est possible de définir deux catégories principales de réseaux ad hoc : les réseaux contrôlés et les réseaux ouverts. Nous nous référons aux réseaux ad hoc contrôlés quand la phase d'initialisation du réseau peut être appuyée par une infrastructure provisoire et quand une confiance à priori entre les noeuds du réseau est disponible. Des relations de confiance à priori peuvent être établies, par exemple, par une autorité contrôlant les noeuds du réseau ou par une organisation commune entre les utilisateurs. D'autre part, dans un réseau ad hoc ouvert, les noeuds sont entièrement autonomes et dans la majorité des cas ne peuvent pas compter sur une infrastructure préétablie pour l'initialisation et l'opération de réseau. De plus, puisque les noeuds sont actionnés par les utilisateurs qui n'appartiennent pas nécessairement à la même organisation ni ne partagent une même autorité, une relation de confiance à priori entre les noeuds n'est pas disponible. La confiance entre les noeuds doit être établie par des mécanismes spécifiques conçus en fonction du scénario offert par un environnement ouvert.

Dans la littérature, une attention particulière a été consacrée à la conception des protocoles d'acheminement optimaux qui réduisent au minimum la consommation d'énergie ou qui sont bien adaptés à une topologie dynamique. Dans cet article il est suffisant de mentionner les propositions fondamentales qui ont contribué au développement successif des mécanismes d'acheminement plus avances, comme par exemple le protocole DSR [5], et AODV [14].

Le dénominateur commun entre les propositions existantes d'acheminement ad hoc est qu'aucune d'elles n'a pris en compte la possibilité (réelle) d'une déviation du comportement des noeuds par rapport à l'exécution définie par le protocole. Néanmoins, il n'est pas difficile aujourd'hui de constater une hausse de comportement illégitime des composants d'un protocole distribué : par exemple, les protocoles pair-à-pair sont facilement modifiables afin d'exploiter le système sans pour autant y contribuer activement.

Dans le cas spécifique offert par les réseaux MANET ouverts, il est très simple de manipuler un protocole d'acheminement afin d'économiser l'énergie dépensée par un noeud d'une façon « égoïste ». Le besoin de coopération entre les nœuds pour assurer le fonctionnement du réseau est en conflit avec l'intérêt individuel

de chaque nœud visant à ne dépenser de l'énergie (une ressource précieuse, car dans la majorité des cas les dispositifs mobiles sont alimentés par batterie ayant une durée de vie limitée) que pour les flux de trafic qui leur sont destinés ou pour les quels ils sont originés.

Dans cet article on étudie un mécanisme proposé dans la littérature pour faire face au comportement égoïste des noeuds d'un réseau ad hoc. Ce mécanisme, nommé CORE [11], se base sur la notion de réputation. Après avoir résumé le fonctionnement de base de CORE on va se concentrer sur son analyse en utilisant les outils fournis par la théorie des jeux. Notre analyse se base sur un modèle très simple pour représenter le conflit d'intérêt auquel chaque noeud fait face lors d'une prise de décision (notamment coopérer pour acheminer un paquet ou ne pas coopérer). Ce modèle simple est étendu pour tenir compte des contraintes physiques imposées par le mode de fonctionnement sans fils. Afin d'évaluer les performances de CORE, une stratégie dérivée de l'implémentation réelle du protocole est présentée. Cette stratégie va être mise en compétition avec d'autres stratégies disponibles dans la littérature dans le cadre d'une ultérieure extension au modèle de base. Dans ce dernier cas, la théorie des jeux évolutifs est appliquée au modèle d'un réseau ad hoc statique.

2 Le mécanisme CORE

L'étude basée sur une simulation effectuée dans notre laboratoire [10] a prouvé que les performances d'un réseau MANET se dégradent sévèrement en présence d'un simple comportement illégitime des noeuds. Indépendamment des cas spéciaux comme pour les réseaux militaires pour lesquels une confiance à priori existe entre tous les noeuds, les noeuds d'un réseau ad hoc ne peuvent pas être considérés fiables pour l'exécution correcte des fonctions critiques du réseau. Des opérations essentielles peuvent être fortement compromises par les noeuds qui n'exécutent pas correctement leur part des opérations comme le routage, l'expédition de paquets, etc... Un mauvais comportement des noeuds qui affecte ces opérations peut s'étendre de l'égoïsme ou du manque simple de collaboration dûe au besoin d'économie de batterie aux attaques actives comme le déni de service et la subversion du trafic. En raison de leur vulnérabilité accrue, les réseaux ad hoc devraient tenir compte des problèmes de sécurité comme condition de base indépendamment des scénarios d'application et des contre-mesures doivent être intégrées aux mécanismes de base de gestion de réseau dés leurs conceptions.

Une autre conclusion importante de notre étude de simulation est que la dégradation de performance dûe aux noeuds égoïstes s'abstenant éxpédier des paquets est plus significative que l'impact du comportement égoïste simulé par des attaques sur le protocole d'acheminement comme le protocole « dynamic source routing » (DSR). Nous croyons que des résultats semblables qui accentuent la sensibilité inhérente de MANET à l'égoïsme des noeuds peuvent être obtenus avec des fonctions de réseau autres que l'acheminement ou l'expédition de paquets. Les mécanismes de sécurité qui imposent seulement l'exactitude ou l'intégrité des opérations de réseau ne seraient ainsi pas suffisants dans MA-

NET. Une condition de base pour maintenir le réseau opérationnel consiste à imposer la contribution des noeuds aux opérations du réseau en dépit de la tendance contradictoire de chaque noeud vers l'égoïsme, motivée par la pénurie des ressources énergétiques.

Dans cette section nous discutons du mécanisme CORE [11], utilisé pour imposer la coopération entre les nœuds. CORE se base sur une technique de surveillance distribuée. Ce mécanisme de coopération n'empêche pas un noeud de nier la coopération ou de dévier d'un comportement légitime mais s'assure que les entités se conduisant mal soient punies en leur refusant graduellement les services de communication. CORE est suggéré comme mécanisme générique qui peut être intégré avec n'importe quelle fonction de réseau comme l'expédition de paquets, découverte de routes, gestion de réseau, et gestion de la localisation. Dans CORE, chaque entité réseau encourage la collaboration d'autres entités en utilisant une métrique de coopération appelée réputation. La métrique de réputation est calculée sur la basé des données recueillies localement par chaque noeuds et peut se baser optionnellement sur l'information fournie par d'autres nœud du réseau impliqués dans des échanges de messages avec les nœuds surveillés. Basé sur la réputation, un mécanisme de punition est adopté comme système de dissuasion pour empêcher un comportement égoïste en refusant graduellement les services de communication aux entités qui se conduisent mal. La conséquence immédiate d'un réseau MANET qui adopte CORE est que les noeuds légitimes (noeuds qui coopèrent à l'opération de réseau) arrivent à économiser de l'énergie car il ne servent pas ceux qui ont été détectés comme égoïstes. En outre, selon le modèle d'égoïsme adopté pour représenter des noeuds se conduisant mal, il est possible de prouver que CORE fournit une incitation efficace à coopérer.

Pour simplifier la description de CORE, on se base sur la Figure 1, qui ne concerne que la détection et la punition d'un comportement égoïste vis-à-vis de l'expédition des paquets.

CORE est composé de trois modules : l'un dédié à l'analyse du trafic réseau, permis par le mode de fonctionnement dit « promiscuous » des cartes sans fils, un module dédié à l'évaluation de la métrique de réputation (pour chaque voisin) et un module utilisé dans la phase de punition des nœuds. Il faut noter que CORE, étant un mécanisme distribué, est exécuté sur chaque nœud mobile du réseau. Le module d'analyse classifie le trafic réseaux aux alentours de chaque nœud pour vérifier le comportement individuel de chaque voisin. Par définition, un nœud B est voisin du nœud A s'il est joignable avec un seul saut, *c.à.d.* s'il se trouve dans le rayon radio du nœud A. Le module d'analyse génère un flux de comportement associé à chaque voisin : un vecteur booléen représente un bon (avec un 1) ou un mauvais (avec un 0) comportement. Le flux de comportement est utilisé par le module de réputation, qui dans sa version de base, n'est rien d'autre qu'un filtre à « réponse finie » de type passe-bas. La fonction de filtrage est utilisée pour réduire l'impact d'une fausse génération du flux de comportement. Dans la réalité, le mode de fonctionnement « promiscuous » n'est pas robuste et il est sujet aux fautes. CORE attribue un bas niveau de réputation seulement aux nœuds qui montrent les traits caractéristiques d'égoïsme d'une façon persistante. D'autres

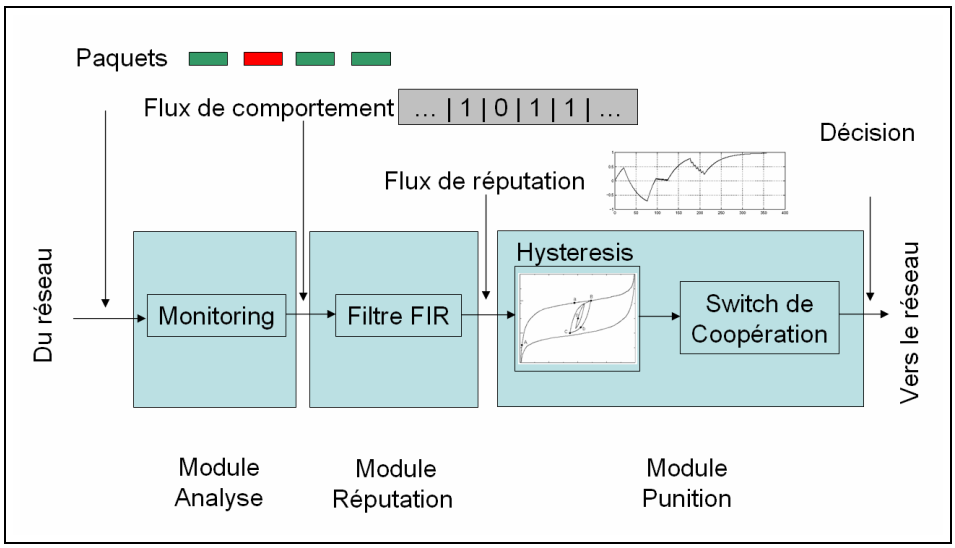

Fig. 1. CORE : principe de fonctionnement

modéles plus complexe peuvent être prévus pour le module de réputation, comme par exemple des filtres passe bande, pour tenir compte de certains types de comportement, ou des filtres dynamiquement accordés. Le module de punition se base sur un seuil (à hystérésis) qui est utilisé pour déclencher le déni de service aux nœuds égoïstes. Dans la Figure 2 trois variations du mécanisme de réputation sont présentées : de droite a gauche, le mécanisme se base 1) sur un flux booléen filtré sur une fenêtre de 5 échantillons, 2) sur un flux fidèle au taux de coopération filtré sur une fenêtre de 5 échantillons, 3) sur la déviation standard correspondant à un taux de coopération de 50%, *c.à.d.* un nœud qui n'achemine que la moitié du trafic, filtré sur une fenêtre de 5 échantillons. Les graphes ont été obtenus par une simulation haut niveau (grâce a MATLAB) du seul mécanisme de réputation, en se basant sur un flux de comportement génère artificiellement. Comme on le peut constater, à parité de comportement (les trois graphes en haut) le niveau de réputation varie en tenant plus ou moins compte des grosses variations en ampleur ou des variations « haute fréquence » comme le montre la forme de la réputation déduite entre 30 et 40 secondes de simulation.

Pour plus de details sur le mechanisme CORE, le lecteur est invité a consulter [11].

3 Modélisation et validation analytique de CORE par la théorie des jeux

Plusieurs mécanismes de coopération ont été proposés par la communauté scientifique dans la tentative de faire face au comportement égoïste des noeuds

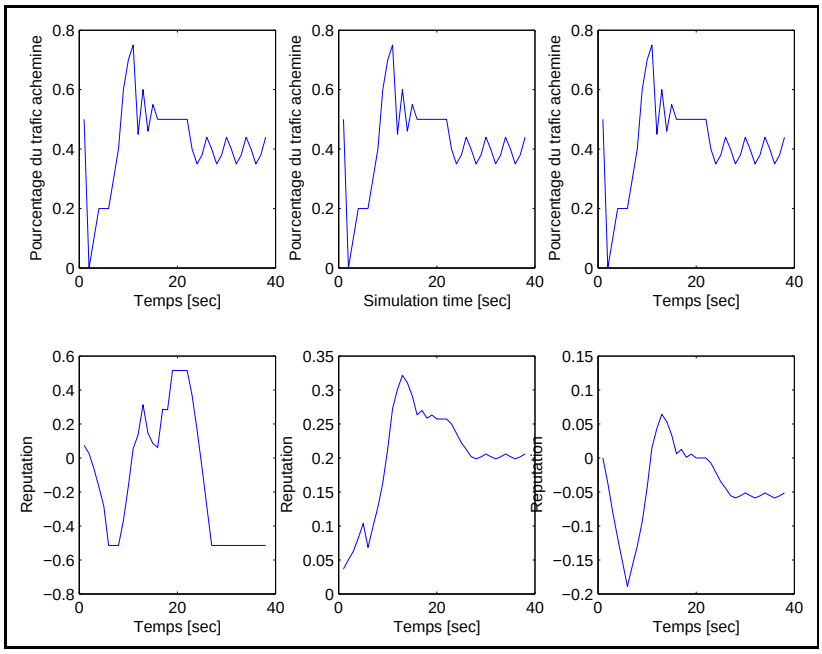

Fig. 2. CORE : évaluation de réputation

présents dans les réseaux ad hoc mobiles (MANET). Comme on le définit par exemple dans [11], un noeud est considéré égoïste quand il ne participe pas à la gestion ordinaire du réseau afin d'économiser de l'énergie. En opposition à la malveillance, l'égoïsme est une menace passive qui ne comporte aucune dégradation intentionnelle de l'opération du réseau au contraire des attaques actives, par exemple, comme la subversion de route, modification des données, *etc...* Dans cette section nous présentons une approche pour évaluer les mécanisme CORE décris dans la section 2. Puisqu'une grande fraction des schémas existants est basée sur des principes apparentés à la modélisation économique, un outil naturel qui s'est présenté pour la validation de tels mécanismes est la théorie des jeux. Dans ce travail, on utilise une méthode basée sur la théorie des jeux non coopérative, ce modèle étant utile pour démontrer les propriétés de base de CORE. En utilisant cette méthode nous adoptions un modèle qui décrit la stratégie d'un noeud égoïste qui doit prendre la décision de coopérer ou de ne pas coopérer avec un noeud voisin aléatoirement choisi. Nous traduisons alors le mécanisme CORE en guise de stratégie qui peut être ainsi comparé à d'autres stratégies bien connues dans la littérature. Sous l'hypothèse généralement utilisée de la « surveillance parfaite », nous démontrons l'équivalence entre CORE et un éventail de stratégies basées sur l'histoire des interactions, comme la stratégie « Tit-for-Tat ». De plus, en adoptant une hypothèse plus réaliste qui tient compte de l'imperfection des observations du comportement des noeuds voi-

sin dûe aux erreurs de communication, le modèle non coopératif met en évidence la supériorité (en termes de stabilité et robustesse) de CORE par rapport à d'autres mécanismes basés sur une histoire d'interaction.

3.1 Modèle du système

L'interaction entre les nœuds d'un réseau MANET et le processus de sélection du niveau de coopération peuvent être décrits en utilisant un simple modèle introduit par Tucker [15], nommée le dilemme du prisonnier (DP). Dans le modèle classique du DP, deux joueurs doivent prendre la décision de coopérer (C) ou de ne pas coopérer (D). Cette décision est prise d'une façon synchrone, sans à priori sur le choix de l'opposant. Si les deux joueurs coopèrent ils reçoivent une prime (R). Si les deux joueurs décident de ne pas coopérer ils reçoivent une punition (P). Dans le cas ou seulement un joueur coopère tandis que l'autre ne coopère pas, les gains vont être T pour le joueur qui n'as pas coopéré et de S pour le joueur qui a coopéré. Souvent, le DP est représenté par une forme canonique dite matricielle, qui exprime les gains en fonction des stratégies adoptées par les joueurs. Autrement dit, la fonction mathématique qui guide le comportement stratégique des joueurs est dite la fonction utilité. Le DP a reçu beaucoup d'at-

		Joueur j	
		C	D
Joueur i	C	(R, R)	(S, T)
	D	(T, S)	(P, P)

		Joueur j	
		C	D
Joueur i	C	$(3, 3)$	$(-2, 4)$
	D	$(4, -2)$	$(0, 0)$

Tab. 1. Forme Matricielle du Dilemme du Prisonnier : (a) canonique, (b) exemple.

tention dans le passé grâce aux amples possibilités d'application, qui recouvrent des domaines tels que l'étude de l'évolution de la coopération en biologie, et bien sûr, qui bien est utile pour l'étude des réseaux. Précisément, le DP appartient à la classe des jeux nommée jeux à deux joueurs, dont la somme des gains n'est pas nulle, avec une sélection de stratégie simultanée. Le dilemme, qui force les jeux dans un état sub-optimal est dicté par l'expression suivante :

$$\begin{array}{l} T > R > P > S \\ R > \frac{S+T}{2} \end{array} \quad (1)$$

Le modèle du DP, s'applique à un réseau MANET statique composé par N nœuds qui représente un terrain de jeux ou deux opposants se rencontrent d'une façon alèatoire. Bien évidement, il s'agit d'un modèle contraint par l'hypothèse de négliger l'impact du routage réseau. D'autres modèles plus compliqués peuvent tenir compte du routage [1] et des contraintes physiques [16] dûes au protocole d'accès au medium partagé (les protocoles MAC). La littérature ne présente pas

de modèles qui tiennent compte de la mobilité. Afin de définir notre modèle, on fait l'hypothèse selon la quelle chaque nœud du réseau est une source du trafic, qui va être acheminée grâce à un protocole de routage (par exemple le protocole DSR). De plus, on assume que deux nœuds qui se rencontrent, car ils sont voisins, vont avoir un besoin mutuel de coopération afin d'acheminer les paquets de chacun. Avant d'envoyer un paquet, suivant le modèle original du DP, chaque nœud doit prendre la décision de coopérer ou pas avec son opposant : en coopèrant un nœud promet d'acheminer un pourcentage du trafic de l'autre nœud. Au lieu de prendre en compte les caractéristiques dûes au modèle énergétique des nœuds, une éventuelle information sur la topologie du réseau et les phénomènes d'interférence, on va se concentrer dans la suite sur une simple similitude pour expliquer le problème. On pourrait imaginer de considérer deux joueurs (les nœuds) ayant une lettre (un ensemble de paquets) à envoyer. Pour chaque lettre envoyée (que ce soit la lettre envoyée par le joueur même ou acheminée pour l'opposant), les joueurs doivent payer le coût d'un timbre (le coût en énergie pour envoyer un paquet). L'exemple donné sur la Table 1, montre que la seule solution possible au jeu, concept qu'on va définire dans la suite, consiste à ne pas dépenser d'énergie, ni pour expédier son propre trafic ni pour acheminer le trafic de l'opposant, car ce choix de stratégie est le seul qui ne comporte pas de perte de gains dans le cas ou l'opposant serait non coopératif.

On pourrait argumenter que le modèle ici présenté est trop simple : au contraire, nous pensons que la perte en réalisme (dûe aux hypothèses qui limitent la prise en compte de phénomènes physiques importants) est grandement compensée par la quantité des résultats dans la littérature du dilemme du prisonnier. De plus, comme il est possible de le voir dans [1], une extension du simple jeu à deux joueurs vers le cas plus réaliste d'un jeu à N joueurs est possible. Il est toutefois important de noter que l'hypothèse d'interaction future, qui dans la théorie des jeux est souvent nommé « l'ombre du futur », est fondamentale pour notre modèle : le trafic dans le réseau est dense, signifiant que tous les nœuds sont source de trafic.

3.2 Version itérée du dilemme du prisonnier

Le simple modèle présenté dans la section précédente peut être enrichi en considérant un jeu qui consiste à répéter un certain nombre de fois le même jeu. La théorie des jeux répétes a été étudiée dans le passé [3]. Comme on l'a constaté dans la section précédente, la seule solution au jeu du DP est que les deux joueurs choisissent de ne pas coopérer. La littérature est riche de définitions de solutions d'un jeu. Dans cet article on va se concentrer sur un concept d'équilibre nommé l'équilibre de Nash. Le seul équilibre de Nash (NE) du DP est de ne pas coopérer. Il est simple de voir que, en supposant fixe le choix de son adversaire, le mieux qu'un joueur puisse faire pour ne pas perdre (*c.à.d.* pour ne pas payer) c'est de choisir la stratégie D. La façon dont les joueurs choisissent leur stratégie est dictée par le principe de rationalité : non seulement les joueurs sont égoïstes, dans le sens qu'ils veulent maximiser leurs profits, mais ils ont à disposition

une puissance de calcul leur permettant de deviner le choix stratégique de l'adversaire. La rationalité des joueurs a été prise en compte seulement récemment comme une limitation de la théorie des jeux appliqués aux problèmes réseaux. Souvent, il est impensable de faire l'hypothèse de rationalité : une hypothèse plus réaliste est proposée, tenant compte des limitations en terme de puissance de calcul des joueurs.

Dans le scénario qu'on considère dans cet article, l'interaction entre deux nœuds (même dans le cas de mobilité) est souvent étendue à plus d'un seul échange. Dans ce contexte, la stratégie choisie par un joueur dans le passé peut avoir une influence sur la décision future de son opposant : le jeu répété comprend les phénomènes de continuité typiques d'un réseau qui se base sur le chemin le plus court entre une source et une destination. Il est donc possible d'observer l'évolution et la naissance de la coopération même si le jeu de base indique que le seul résultat possible est de non coopération, et surtout même dans le cas ou les joueurs sont guidés par l'égoïsme. Il faut noter que dans cette section on considère le jeu répété un nombre infini de fois. Un principe connu sous le nom d « principe d'induction inverse », montre que le simple fait de connaître avec certitude la fin du jeu répété induit un comportement non coopératif. Les joueurs, en partant de la dernière étape du jeu peuvent appliquer le principe de rationalité et ne pas coopérer pour obtenir un gain lors de la dernière itération. Le principe d'induction inverse peut être appliqué à l'avant dernière itération du jeu jusqu'à remonter à la première rencontre, compromettant un résultat positif qui va dans le sens de la naissance d'un comportement coopératif. En pratique, notre modèle se base sur une nuance du concept des jeux répétés indéfiniment (souvent nommés jeux avec horizon infini) : au lieu de répéter un nombre infini de fois le jeu, les joueurs ne connaissent tout simplement pas la fin du jeu. Dans ce contexte l' « ombre du futur » a un poids déterminant pour l'évolution du jeu vers un comportement coopératif.

Pour conclure cette brève introduction aux jeux répétés, il faut caractériser la fonction utilité maximisée par les joueurs : $U_i = \sum_{t=0}^{\infty} \delta^t u_i^t$.

Dans la suite on va donc considérer le joueur i qui maximise la fonction U_i étant la somme de la fonction utilité de chaque itération du jeu de base du DP. Le facteur δ indique le poids d'un gain immédiat par rapport à un gain à long terme. Une simple extension de ce modèle pourrait considérer le facteur δ comme un indice de mobilité des nœuds.

4 Stratégies complexes dans le Dilemme du Prisonnier Itéré

Axelrod et Hamilton [2,**?**] utilisent un tournoi simulé sur ordinateur pour détecter d'une façon numérique les stratégies qui pourraient mener à la naissance de la coopération entre joueurs engagés dans un DP itéré. Dans leur expérience, 14 stratégies complexes et une stratégie complètement aléatoire sont en compétition dès la première itération pour une succession de 200 itérations. Le résultat inattendu de cette expérience est qu'une stratégie très simple séstt

démontrée être celle permettant aux joueurs qui l'adoptent de remporter le maximum de gains. Cette stratégie est nommée la stratégie tit-for-tat (TFT) :

`TIT-FOR-TAT` :
Coopérer dès la première itération, ensuite copier la stratégie de l'adversaire utilisée dans l'itération précédente

Cette stratégie est à l'origine d'un vaste ensemble de stratégies plus complexes, comme on va le démontrer en modélisant CORE comme une stratégie. Pour étudier les caractéristiques d'une stratégie d'un point de vue numérique, deux approches sont possibles :

- Une approche consiste à utiliser une simulation d'un tournoi pair à pair, dans lequel chaque stratégie utilisé dans le jeu fait face aux autres stratégies. Le score final d'une stratégie est la somme (dans un cas simple, sans considerer le facteur δ mentionné dans la Section 3.2) des scores obtenus à chaque itération. A la fin de la simulation, la stratégie gagnante est celle qui a obtenu le score maximal.
- Une seconde approche consiste à exécuter une simulation numérique en utilisant les techniques pour l'étude des systèmes biologiques. Au début de la simulation, les stratégies à étudier sont identiquement distribués sur la population des joueurs. Un tournoi pair à pair est ensuite lancé, ce qui prévoit une rencontre aléatoire entre populations adoptant différentes stratégies. A la fin de chaque round, la partie de population qui utilise une stratégie gagnante sera incrémentée tandis que la population perdante va vers l'extinction. La simulation s'arrête lorsque les populations deviennent stables, c'est a dire dans le cas où le nombre de joueurs qui adoptent une stratégie gagnante est identique. Cette deuxième méthode représente une façon d'estimer la robustesse et la stabilité des stratégies.

Avant d'introduire le modèle de CORE qui définit une stratégie pour le jeu du DP, il est important de décrire en détail la méthodologie de simulation qu'on a choisie pour évaluer les propriétés de stabilité et robustesse de CORE. Dans l'exemple suivant on va considérer trois stratégies, et on va adopter la deuxième approche décrite au début de cette section. Supposons que la population prenant part au jeu du DP itéré utilise, avec une distribution identique sur les joueurs, les stratégies `A`, `B`, et `C`. L'n-ieme itération, que l'on appelle souvent l'n-ieme génération, voit chaque stratégie représentée par une distribution de population : $W_n(A)$ joueurs utilisent la stratégie `A`, $W_n(B)$ utilisent la stratégie `B` et $W_n(C)$ adoptent la stratégie `C`. La représentation matricielle qui guide la stratégie de deux joueurs qui s'affrontent dans le tournoi « pair à pair » est celle présentée dans la Table 1. Le gain des joueurs qui utilisent la stratégie `A` quand ils s'opposent à la stratégie `B` est représenté par $V(A|B)$. Pour chaque simulation on assume une population fixe et constante Π, ce qui constitue une autre limitation du modèle par rapport à la réalité offerte par les réseaux MANET. Notre modèle, aussi comme les modèles disponibles dans la littérature, ne permet pas de tenir compte de l'évolution de la population typique d'un système ouvert. Pour ce qui concerne la taille totale de la population prenant part au jeu itéré, l'expression

suivante est valable :

$$\forall i \in [1, \infty[, \Pi = W_i(A) + W_i(B) + W_i(C) \tag{2}$$

L'évaluation du score obtenu par chaque joueur qui adopte une déterminé stratégie à l'itération n est donc :

$$\begin{array}{l} g_n(A) = W_n(A)V(A|A) + W_n(B)V(A|B) + W_n(C)V(A|C) - V(A|A) \\ g_n(B) = W_n(A)V(B|A) + W_n(B)V(B|B) + W_n(C)V(B|C) - V(B|B) \\ g_n(C) = W_n(A)V(C|A) + W_n(B)V(C|B) + W_n(C)V(C|C) - V(C|C) \end{array} \tag{3}$$

En total, le gain attribué à chaque stratégie est :

$$t(n) = W_n(A)g_n(A) + W_n(B)g_n(B) + W_n(C)g_n(C) \tag{4}$$

On peut en déduire que chaque sub-population à l'itération $n+1$ va donc être :

$$\begin{array}{l} W_{n+1}(A) = \frac{\Pi W_n(A)g_n(A)}{t(n)} \\ W_{n+1}(B) = \frac{\Pi W_n(B)g_n(B)}{t(n)} \\ W_{n+1}(C) = \frac{\Pi W_n(C)g_n(C)}{t(n)} \end{array} \tag{5}$$

En ayant spécifié la dynamique des expérimentations, on va noter les caractéristiques souhaitables définies par Axelrod [3] lors de ses travaux sur le dilemme du prisonnier itéré. Une bonne stratégie doit :

- Ne pas être la première à ne pas coopérer ;
- Réagir aux changements rapidement ;
- Etre flexible et « pardonner » les comportements non coopératifs ;
- Etre simple, soit lors de son implémentation que lors de son exécution

La stratégie TFT, qui satisfait les critères mentionnées par Axelrod, a été considérée comme l'une des meilleurs stratégies pour favoriser la naissance de coopération entre joueurs visant a maximiser leur gains. Dans la suite, on va examiner les conditions dans lesquelles la stratégie TFT ne réalise plus le meilleur score par rapport à d'autres stratégies comme celle dérivée du mécanisme CORE.

4.1 La stratégie CORE : modélisation du mécanisme par une stratégie complexe dans le DP Itéré

Dans le contexte spécifique qui vise à étudier de la naissance de la coopération entre joueurs, on a décrit les étapes théoriques qui nous mènent à utiliser la théorie de l'évolution appliquée aux jeux en passant par une introduction au dilemme du prisonnier et á sa version itérée. Dans cette section nous focalisons notre attention sur la modélisation du mécanisme CORE comme stratégie utilisée dans un simulateur (introduit et expliqué dans [9]) de jeu évolutif. Notre but est aussi celui de comparer les stratégies développées dans la littérature de la théorie des jeux et connues pour être les « meilleures » catalyseurs de coopération avec la stratégie CORE. Pour plus de détails sur le fonctionnement de CORE, le lecteur intéressé peut se référer à [11]. Ici on montre que la stratégie CORE

peut être considérée comme équivalente de la stratégie TFT sous certaines hypothèses (par exemple quand le nombre d'observations utilisées pour évaluer la réputation associé à un nœud est égale à 1). De plus, on montre que la stratégie CORE est la meilleure entre un groupe de stratégies connues pour être très efficaces afin de stimuler la coopération, quand les conditions qui guident le choix stratégique des joueurs ne sont pas idéales. Précisément, dans le contexte d'une observabilité parfaite du choix stratégique de son opposant, CORE et TFT sont équivalentes d'un point de vue stabilité et robustesse, tandis que dans le contexte plus réel d'observabilité imparfaite CORE est nettement meilleur que TFT. L'intérêt d'étudier les caractéristiques d'une stratégie dans le cadre d'observabilité imparfaite vient de la simple remarque que, dans la réalité, tous les mécanismes se basant sur le monitorage de l'activité des nœuds voisins dans un réseau MANET sont très sensibles aux bruits, interférences, collisions, obstacles etc., qui sont typiques d'un environnement se basant sur la technologie sans fils 802.11.

La stratégie CORE, peut être définie comme suit :

`CORE`
- Coopérer dès la première itération
- A chaque itération, observer le B derniers choix stratégiques de son opposant et construire le vecteur $\overrightarrow{b} = (b_1, ..., b_k, ..., b_B)$ où chaque élément est égal a 1 pour une coopération (C) et à -1 pour une non coopération (D) ;
- Calculer la réputation associée à son opposant comme : $reputation = \frac{1}{B} \sum_k b_k$;
- Si $reputation \geq 0$ alors choisir de coopérer (C) autrement ne pas coopérer (D).

Il faut noter que la stratégie CORE décrite ici n'est qu'une simplification du vrai mécanisme CORE, qui se base comme c'est expliqué dans la section 1 sur le filtrage d'un flux booléen, au lieu d'effectuer un calcul sur un vecteur. La représentation simplifiée nous permet d'observer immédiatement la similitude entre CORE et TFT quand le nombre d'observations utilisé pour évaluer la réputation (B) est égale à 1. Quand $B = 1$, la stratégie CORE est exactement décrite par la stratégie TFT, et elle hérite des propriétés définies par Axelrod. Dans cet article on ne va pas présenter le processus analytique qui montre que CORE est une stratégie d'équilibre car il s'agit d'un travail en cours de développement. En règle générale, la littérature de la théorie des jeux fournit une vaste ensemble de théorèmes pour déterminer les conditions et l'existence d'un point d'équilibre, soit pour les jeux non répétés que pour les jeux itéré. Un sujet de recherche très actuel est l'étude algorithmique pour calculer exactement le point d'équilibre d'un système modélisé grâce à la théorie des jeux. Dans la suite, on discute les résultats de notre étude basée sur des simulation numériques.

4.2 Simulations sous l'hypothèse d'observabilité parfaite

Dans cette section on analyse les résultats d'une simulation qui implique quatre stratégies : `tit-for-tat`, `CORE`, `all-C` (toujours coopérer) et `all-D` (ne jamais coopérer). Comme on le décrit dans la section 4 la première itération voit 100 joueurs pour chaque stratégie. Dans la figure 3 il est possible d'observer qu'après seulement 5 itérations la stratégie `all-D` disparait complètement tandis que les autres stratégies présentent le même type d'évolution. Ceci implique que les stratégies gagnantes ont obtenu le même gain dans chaque tournoi à deux joueurs, donc elle peuvent être considèrées comme équivalentes d'un point de vue évolutif. Il faut noter que la stratégie `all-C` n'est pas considérée comme une stratégie canonique car le choix stratégique ne dépend pas des actions passées des opposants.

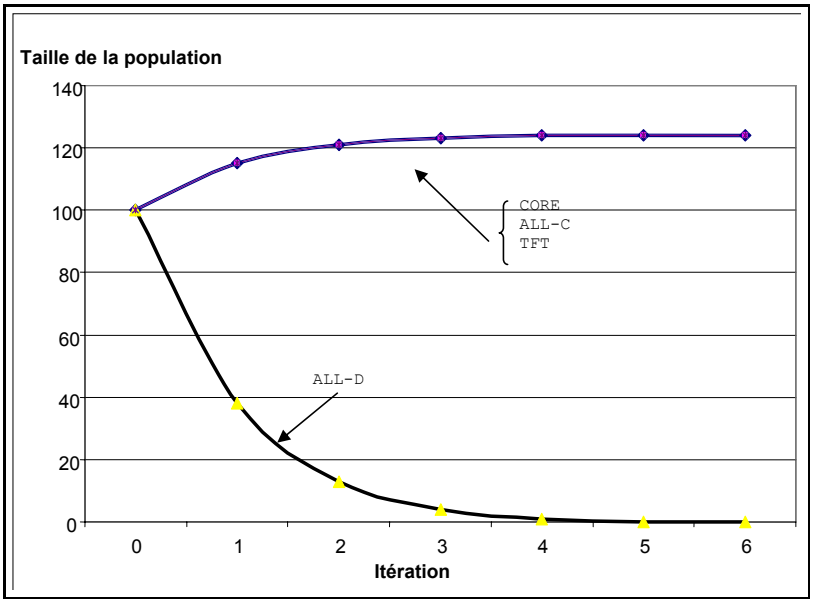

Fig. 3. Simulation de l'évolution des stratégies sous l'hypothèse d'observabilité **parfaite**

4.3 Simulations sous l'hypothèse d'observabilité imparfaite

La majorité des travaux effectués dans le domaine du dilemme du prisonnier itéré a été concentrée sur l'étude de l'équilibre et de l'évolution des systèmes en absence de « bruit ». Cela implique qu'il n'existe pas de possibilité d'erreur d'évaluation de la stratégie choisie par un opposant dans les itérations

passées. Cette hypothèse n'est pas nécessairement valable dans le cadre d'une modélisation d'un système réel : dans le cas spécifique d'un réseau MANET, il est impératif de prendre en compte les erreurs commises par le mécanisme dit « watchdog » implémenté pour fournir à CORE (et à une vaste panoplie d'autres mécanismes à base de réputation) les informations sur le comportement des nœuds voisins. Le lecteur intéresse pourrais par exemple étudier les travaux de Marti [8] pour connaître les problèmes liés a l'utilisation du « watchdog ».

Il y a différents moyens qui peuvent être utilisés pour introduire du bruit dans une simulation :

- Mauvaise implémentation, dans le cas ou un joueur se « tromperait » au moment de l'actuation d'une stratégie
- Mauvaise perception, dans le cas ou un joueur se « tromperait » au moment de l'observation du choix de son opposant

Dans cet article on se concentre sur le cas d'une mauvaise perception car on estime que ce type de problème est le plus semblable aux problèmes introduits par le mécanisme du « watchdog ». Dans la littérature, Kahn and Murnighan [7] on trouve que dans les expériences menées sur le jeu du DP, la coopération entre joueurs est d'autant plus probable quand les joueurs sont sûrs du gain obtenu par les autre joueurs. Les expériences menées par Miller [13] qui utilise la théorie des algorithmes génétiques appliquée au DP montrent que la coopération entre joueurs atteit un maximum lorsque les expériences sont exécutées dans un environnement sans « bruit », tandis que la coopération montrée par les joueurs décroit rapidement avec une hausse du bruit dans le système. Axelrod a aussi proposé des idées pour stimuler la coopération entre joueurs dans le cas d'un système sujet au bruit : techniques qui incluent le groupage de stratégies similaires, l'apprentissage, modifications dynamiques de la forme matricielle qui décris le jeux. Hoffman [6] a étudié aussi les implications dûes à une mauvaise implémentation des stratégie, en simulant un " tremblement de main " lors de la décision finale de coopérer ou pas.

En particulier, on remarque une forte sensibilité au bruit pour les stratégies dépendent d'une façon très simple du passé, de l'histoire qui décrit l'interaction entre joueurs. Par exemple, dans un tournoi entre joueurs qui adoptent la stratégie TFT, une simple erreur de perception peut impliquer une divergence qui mène a la non-coopération : il suffit qu'un joueur se trompe en considérant comme non coopératif un comportement coopératif pour que, dans la prochaine itération du jeu répété, son opposant choisisse de ne pas coopérer en raison d'un comportement non coopératif erroné de son opposant. Les études présentées dans [4] confirment qu'une stratégie capable de s'adapter à un environnement bruyant en « pardonnant » certains choix stratégiques non coopératifs pourrait mener à la coopération entre les joueurs.

Comme on le trouve dans la section 4 on a étudié les résultats d'une simulation qui implique 5 stratégies en présence de bruit dans le système. Chaque point du graphe de la Figure 4 est le résultat de 10 simulations quand le bruit et égal a 10% : une fois sur dix une erreur de perception affecte les décisions des joueurs. 100 joueurs pour chaque stratégie se confrontent dès la première

itération du tournoi. Pour une description des stratégie `spiteful`, `gradual` et `soft-majo`, le lecteur intéressé peut consulter [12]. Comme on peut l'observer dans la Figure 4, la stratégie CORE est celle qui évolue le plus rapidement, en gagnant de plus en plus de parties de population. La raison pour laquelle CORE réalise des meilleurs résultats en présence du bruit est dûe a l'utilisation de la réputation : la valeur de la réputation étant basée sur plus d'une observation, ne comporte pas de déviation inattendue même en présence de bruit car, grâce à la propréteur de lissage de la fonction utilisée dans [11] pour évaluer la réputation, des variations à haute fréquence dans le vecteur (ou flux) qui représente le choix stratégique d'un opposant ont une influence minime.

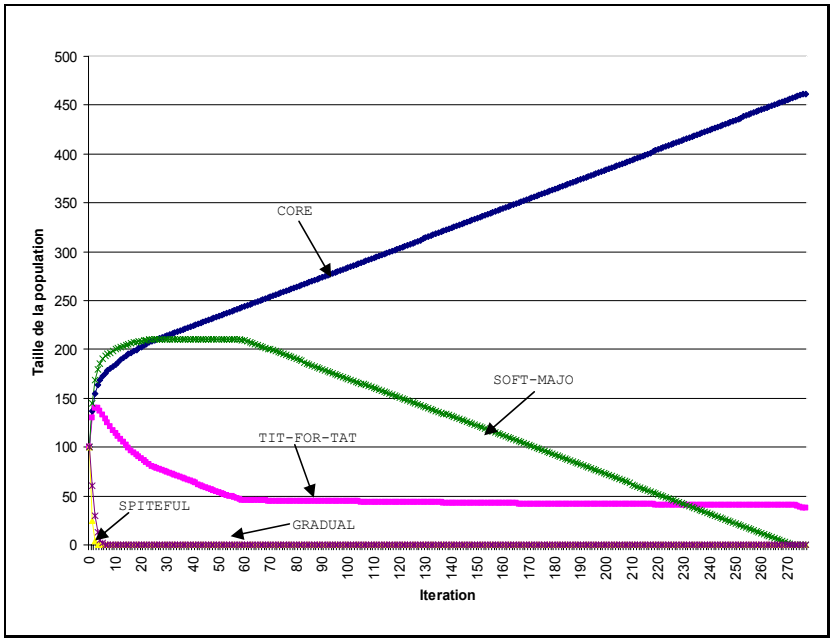

Fig. 4. Simulation de l'évolution des stratégies sous l'hypothèse d'observabilité **imparfaite**

5 Conclusion

Dans cet article on a présenté un mécanisme à base de réputation nommé CORE qui est utilisé pour stimuler la coopération entre le nœuds d'un réseau mobile ad hoc. Le mécanisme CORE est ensuite modélisé grâce à la théorie des jeux et à l'application d'une version itérée du jeu du dilemme du prisonnier. Le modèle analytique est étendu pour étudier, grâce a des techniques utilisées

dans la théorie de l'évolution, la stabilité e la robustesse d'une stratégie directement déduite du mécanisme CORE. Des simulations ayant le but d'analyser les caractéristiques de la stratégie CORE et d'autres stratégies étudiés dans la littérature montrent que CORE est équivalent à la très célèbre stratégie tit-for-tat. De plus, dans le cas ou un élément de réalisme est introduit dans le modèle, pour tenir compte des problèmes créés par un environnement qui exploite la technologie sans fils 802.11, la stratégie CORE montres des caractéristiques de robustesse et de stabilité qui n'ont pas d'égal parmi les autres stratégies considérées dans la littérature. Dans le futur, on se propose d'approfondir l'etude du modèle d'un réseau MANET et des mécanismes de réputation pour tenir compte de la topologie du réseau, des différentes capacités en terme de puissance de calcul et de durée de vie des dispositifs formant le réseau, et du caractère distribué du mécanisme pour évaluer la réputation.

Références

1. Altman E., Kherani A., Michiardi A. et Molva R. (2005) Non cooperative forwarding in Ad hoc Networks. In : *Proceedings of IFIP Networking Conference*, Waterloo, Canada.
2. Axelrod R. (1984) *The Evolution of Cooperation*, Basic Books, New York.
3. Axelrod R. (1987) The evolution of strategies in the iterated prisoner's dilemma, *Journal of Genetic Algorithms and Simulated Annealing*, pp. 32–41.
4. Carraro C. et Siniscalco D. (1993) Strategies for the international protection of the environment, *Journal of Public Economics*, Vol. 52, pp. 309–328.
5. Johnson D. B., Maltz D. A. et Broch J. (2001) DSR : The Dynamic Source Routing Protocol for Multi-Hop Wireless Ad Hoc Networks In : *Ad Hoc Networking*, edited by Charles E. Perkins, Chapter 5, pp. 139-172, Addison-Wesley.
6. Hoffman R. (2000) Twenty years on : The evolution of cooperation revisited, *Journal of Artificial Societies and Simulation*, Vol. 3 Elsevier Science Ltd.
7. Kahn L. M. et Murnighan J. K. (1993) Conjecture, uncertainty, and cooperation in prisoner's dilemma games : Some experimental evidence, *Journal of Economic Behaviour and Organisation*, Vol. 22, pp. 91–117, Elsevier Science Ltd.
8. Marti S., Giuli T. J., Lai K. et Baker M. (2000) Mitigating routing misbehavior in mobile ad hoc networks, In : *Proceedings of the 6th IEEE/ACM International Conference on Mobile Computing and Networking*".
9. Mathieu P., Beaufils B. et Delahaye J. P. (2001) Iterated Prisoner's Dilemma Simulation Software, disponible sur `http://www.lifl.fr/IPD`.
10. Michiardi P. et Molva R. (2001) Simulation-based Analysis of Security Exposures in Mobile Ad Hoc Networks, In : *Proceedings of the European Wireless Conference*, Florence, Italy.
11. Michiardi P. et Molva R. (2002) Core : A COllaborative REputation mechanism to enforce node cooperation in Mobile Ad Hoc Networks, In : *Proceedings of IFIP Communications and Multimedia Security Conference (CMS)*, Portoroz, Slovenia.
12. Michiardi P. et Molva P. (2004) Identity based hash chains for message authentication, Rapport de recherche Eurecom numéro RR-04-11.

13. Miller J. (1989) The coevolution of automata in the repeated prisoner's dilemma, Rapport de recherche 89-003, Santa Fe Institute.
14. Perkins C. E. et Royer E. M. (1999) Ad Hoc On-Demand Distance Vector (AODV) Routing. In : *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications (WMCSA).*
15. Poundstone W. (1993) *Prisoner's Dilemma*, Oxford University Press,
16. Urpi A., Bonuccelli M. et Giodano S. (2003) Modelling Cooperation in Mobile Ad Hoc Networks : A Formal Description of Selfishness, In : *Proceedings of the Workshop : Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WIOPT)*, Sophia Antipolis, France.