Analysis of coalition formation and cooperation strategies in mobile ad hoc networks ☆

Pietro Michiardi, Refik Molva *

Institut Eurecom, 2229 Route des Crêtes, 06904 Sophia-Antipolis, France

Available online 20 August 2004

Abstract

This paper focuses on the formal assessment of the properties of cooperation enforcement mechanisms used to detect and prevent selfish behavior of nodes forming a mobile ad hoc network. In the first part, we demonstrate the requirement for a cooperation enforcement mechanism using cooperative game theory that allows us to determine a lower bound on the size of coalitions of cooperating nodes. In the second part, using non-cooperative game theory, we compare our cooperation enforcement mechanism CORE to other popular mechanisms. Under the hypothesis of perfect monitoring of node behavior, CORE appears to be equivalent to a wide range of history-based strategies like tit-for-tat. Further, adopting a more realistic assumption taking into account imperfect monitoring due to probable communication errors, the non-cooperative model puts in evidence the superiority of CORE over other history-based schemes.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Game theory; Cooperation; Coalition formation; Ad hoc networks

1. Introduction

Cooperation enforcement mechanisms have been developed recently in the attempt to cope with the selfish behavior of nodes in mobile ad hoc networks (MANET). As defined in [5,6], a node is considered selfish when it does not participate in the basic network operation in order to save energy. As opposed to maliciousness, selfishness is a passive threat that does not involve any intention to damage the operation of networking functions by active attacks like route subversion, tampering with data, etc.

In this paper we present two different approaches to assess the features of our cooperation enforcement mechanism CORE [6,7]. Using

CORE, every node locally rates its neighbors through a monitoring mechanism. The observations collected by the monitoring mechanism are processed to evaluate a reputation value associated to each neighbor. The reputation value is used by CORE in a step-like cooperation policy: only nodes with a reputation that satisfy the requirement of being greater than a defined threshold are served (i.e. data and routing packets are forwarded), while nodes with low reputation values are gradually isolated from the network.

Since a large fraction of existing cooperation enforcement schemes are based on principles akin to decision making and economic modeling, a natural tool that emerged to be suitable for the validation of such mechanisms is game theory.

In the first part of this paper we present a model that takes into account both a node-centric and a network-centric perception of the interactions between nodes that participate in a MANET by using *cooperative game theory*. We first demonstrate the requirement for a cooperation enforcement mechanism in order to promote cooperation between self-interested nodes by showing that in the absence of such a mechanism the best strategy for a node would be to free ride. Moreover, we analyze which would be the size of a coalition of cooperating nodes based on the importance given by a node to the node-centric and network-centric perspective of the game. We finally suggest how the CORE mechanism could be used to stimulate a node to join the coalition of cooperators. The benefit from using cooperative GT derives from the ability of this method to seize the dynamics of large group of players: the strategy chosen by a player does not only depend on a self-interested perception of the game but also takes into account a group-wide policy of the coalition the player belongs to.

Although the "cooperative games" approach appears to be appropriate to model the dynamics of large coalitions of nodes forming a MANET, the main limitation of this method is that it is based on a high-level representation of the reputation mechanism that does not take into account the features of CORE. To overcome this weakness, we present in the second part of this paper an alternative approach based on *non-cooperative game theory* [8,9]. In this second method we use a model that describes the strategy of a self-interested node that has to take the decision whether to cooperate or not with a randomly chosen neighbor. Under this model, the CORE mechanism can be translated into a strategy profile that can be compared to other popular strategies. Under the commonly used hypothesis of *perfect monitoring*, we demonstrate the equivalence between CORE and a wide range of history-based strategies like *tit-for-tat*. Further, by adopting a more realistic assumption that takes into account unreliable observations of nodes' behavior due to communication errors, the non-cooperative model puts in evidence the superiority (in terms of stability and robustness) of CORE over other history-based schemes.

Although the two methods described in this paper focus on CORE as a specific mechanism, some general conclusions can be drawn from this analysis towards the design of cooperation enforcement mechanisms in general.

2. Related work

Recently, much attention has been dedicated to game theoretical models for MANET in general and for cooperation enforcement mechanisms in particular and an increasing number of models have been presented to the community. It is however out of the scope of this paper to propose an extensive state of the art of game theoretical models of cooperation in MANET, thus we will focus on some approaches that we deem related to our setting.

In an interesting approach presented in [2] the authors propose a game theoretical model in which energetic information is taken into account to describe the conflicting interaction between heterogeneous nodes involved in a forwarding game, i.e. a game in which nodes that belong to a path from a source to a destination have to collaboratively relay data packets. The authors study the properties of a well known strategy (generous-tit-for-tat, G-TFT) and demonstrate that under the energy constraints imposed to the nodes, G-TFT promotes cooperation if every node of the network

conforms to it. The model in [2] provides an accurate description of the energetic constraint of a node, which is the main reason for a selfish behavior, but provides only high-level guidelines towards the design of a cooperation enforcement mechanism based on the G-TFT strategy. The main difference between the work presented in this paper and the research conducted in [2] is that in our model we take into account a more realistic scenario where the observations made by a node on her neighbors can be affected by errors. The monitoring mechanism is indeed the key feature of a cooperation strategy based on the observation of the opponent's move (such as G-TFT) and we believe that a more accurate description of how these observations are made is fundamental.

Another interesting work towards the definition of a generic game theoretical framework to study cooperation in MANET has been presented in [3]. The authors propose a model that takes into account both the available energy to a node and the traffic generated and/or directed to that node and helps derive some interesting guidelines towards the definition of a cooperation mechanism. The authors not only analyze some existing cooperation mechanisms including CORE but also propose to use the tit-for-tat (TFT) as a cooperation strategy. Similarly to the work presented in [3], in our paper we are able to accurately describe not only our cooperation strategy CORE but also a wide-range of history-based cooperation strategies (such as TFT). The performance analysis of the TFT strategy presented in [3] is extended in our work and we prove that CORE outperforms all other strategies when the imperfect monitoring assumption is made.

In [4] the authors propose an alternative model for the forwarding behavior of a node that is part of a specific network topology. By using their model, the authors are able to express the equilibrium forwarding strategy of a selfish node as a function of topology and routing (path length) information. They also propose a punishment mechanism that enforces a cooperative behavior among selfish nodes. Although the results obtained in [4] provide a very useful description of the relation between routing, network topology and the cooperative behavior of a node, the pro-

posed punishment mechanism is limited to a specific instance of the considered network topology and does not take into account the imperfect monitoring of the node behavior.

The research presented in [2–4] and in the second part of this paper, is based on non-cooperative game theory: even when multiple players are considered, the strategy selection phase is always driven by a node-centric perception of the game. As a result, the cooperation strategies obtained through the proposed models take into account only the payoffs obtained by a single player. Hence, in the first part of this paper we propose an alternative approach based on a general model using cooperative game theory as a framework to study cooperation as a group initiative rather than a strategy adopted by single players. We believe that the ''cooperative games'' approach provides an appropriate way of describing the dynamics of group formation in MANET but needs further research in order to introduce in the model a more formal description of cooperation enforcement mechanism.

3. Cooperative games approach

In an attempt to explain cooperation and coalition formation, most theoretical models use a two-period structure as introduced in [12,13]. Players must first decide whether or not to join a coalition. In a second step, both the coalition and the remaining agents choose their behavior non-cooperatively. A coalition is stable if no agent has an incentive to leave. [1] Simulations presented in [16–18] have shown that, although there is cooperation, the coalition size is rather small.

In this paper we suggest an approach based on a preference structure as defined by the ERC-theory [11]. This theory explains most of the behavior of agents observed in diverse experiments but deviates little from the traditional utility concept. The utility of an agent is not solely based on the absolute payoff but also on the relative payoff

---

[1] The definition of stability also implies that no agent wants to join the coalition.

compared to the overall payoff to all agents. Given a certain relative payoff share, the utility is strictly increasing in the own absolute payoff of the agent. Given a fixed absolute payoff, the agent is best off when receiving just the equal (fair) share. To both sides of this equal share, i.e. when receiving less or more than the fair amount, utility is lower, even if the absolute payoff does not change. [2]

In Section 3.2, we first study a symmetric *N*-node prisoner's dilemma (PD) game in a non-cooperative setting, in which the agents have only two options available—cooperate or defect. We analyze Nash-equilibrium of the non-cooperative game when agents' preferences can be described by ERC, i.e. players value both their absolute and their relative payoff. In particular, we look at the number of agents who play cooperatively. We show that non-cooperation is always an equilibrium, since—if no other node cooperates—a node would maximize its absolute payoff and receive the equal share by choosing to defect. Additionally, however, there may be Nash-equilibrium in which nodes cooperate: if, for example, the rest of the agents play cooperatively, a player can get the equal share by choosing to cooperate as well. Hence, if it values its relative payoff being close to the equal share more than its absolute payoff, it will choose to complete the grand coalition. Clearly, partial cooperation can also occur, whereby some nodes cooperate while others defect. For such equilibrium, we show that the number of cooperating nodes is rather large: since cooperation leads to a lower absolute payoff, for a node to choose to cooperate, playing cooperatively must move it closer to the equal share than defecting would. As we show, this can only be the case if at least half of the nodes cooperate. This result contrasts with the standard result presented in [16] which states that the coalition size is rather small.

Note, however, that in the prisoner's dilemma, the nodes have only the discrete choice of cooperating or defecting, but with respect to the cooperation enforcement problem, the nodes of an ad hoc network might choose their cooperation level [3] continuously. We therefore introduce in Section 3.3 a symmetric continuous PD-game based on the ERC preference structure. An interesting finding of this analysis is that ERC alone cannot improve upon the non-cooperative Nash-equilibrium with standard preferences in which only the absolute payoff matters to a node.

As a further refinement, we propose the cooperative-games approach consisting in a combination between the ERC preference structure and the two-stage coalition formation method [12]. In contrast to the traditional models from the game theory literature, the ERC preference structure allows coalitions to involve a rather large fraction of players. Furthermore, this model allows for a precise characterization of conditions under which even a grand coalition can be obtained.

Finally, in Section 3.5 we propose a discussion on the relation between a coalition formation process and our cooperation enforcement mechanism CORE, used as an effective complementary tool to impose a specific ERC-type for every node participating in a cooperative setting as an ad hoc network.

3.1. The preference structure

Our analysis relies on a preference structure in which players, along with their own absolute payoff, are motivated (non-monotonously) by the relative payoff share they receive, i.e. how their standing compares to that of others. We use the ERC model presented in [4] and enhance it with a *complete information framework*. Let the (non-negative) payoff to node i be denoted by $y_i, i, \ldots, N$, and the relative share by $\sigma_i = y_i / \sum_j y_j$.

The ERC-utility function is defined as follows: $\alpha_i u(y_i) + \beta_i r(\sigma_i)$ where $\alpha_i, \beta_i \geqslant 0$ and $u(\ )$ is differentiable, strictly increasing and concave, and $r(\ )$ is differentiable, concave and has its maximum in $\sigma_i = 1/N$. Throughout this paper we assume that nodes' disutility from disadvantageous inequality

---

[2] Note that such a preference for equity is self-centered only and is distinct from altruism.

---

[3] The definition of cooperation level will be given in Section 3.5: here it is sufficient to know that cooperation level stands for the fraction of packets (data or routing) that are forwarded by a node of the network playing the cooperation game.

Table 1
Summarizing table defining the game based on ERC theory

Number of players	N
ERC global utility function for player i	$\alpha_i u(y_i) + \beta_i r(\sigma_i)$
ERC-types for player i	$\alpha_i, \beta_i \geqslant 0$
Absolute payoff for player i	y_i
Absolute utility function for player i	$u(y_i)$, differentiable, strictly increasing, concave
Relative payoff for player i	σ_i
Relative utility function for player i	$r(\sigma_i)$, differentiable, concave, maximum in $\sigma_i = 1/N$

is larger if the node is better off than average, i.e. $r(1/N - x) \leqslant r(1/N + x)$, $\forall x \in [0, 1/N]$. The types of nodes are characterized by the relative weights α_i, β_i (Table 1).

3.2. The prisoner's dilemma with a discrete strategy space

In this section we study a simple symmetric N-node prisoner's dilemma where each mobile node can cooperate, 'c', or defect, 'd': this implies that the *strategy set* available to each player is discrete and only two actions are allowed. In terms of the node misbehavior problem, this means that the node either correctly executes the network functions or it does not.

Let the total number of cooperating nodes be denoted by k. For any given k, the payoff to a node is given by $B(k)$ if the node defects (tries to free-ride). If a node plays cooperatively, it must bear some additional costs $C(k)$. Its payoff is therefore given by $B(k) - C(k)$. We assume decreasing marginal benefits for a node if the number of mobile nodes rises, i.e. $B(k)$ is increasing and concave. Furthermore, the *total cost* of cooperation, $kC(k)$, increases in k.

In order to generate the standard incentive structure of a PD game, we make the following assumption.

Assumption 1. PD structure: $B(k + 1) - B(k) < C(k + 1)$.

Assumption 1 implies that playing cooperatively reduces the absolute payoff, given an arbitrary number of 'c'-nodes. To make cooperation more attractive from both the social and the individual point of view, we make the following assumptions:

Assumption 2. (*Socially desirable*)

$$N \cdot B(k + 1) - (k + 1)C(k + 1) \geqslant N \cdot B(k) - kC(k). \tag{1}$$

Assumption 3. (*Individually desirable*)

$$B(k + 1) - C(k + 1) \geqslant B(k) - C(k). \tag{2}$$

Furthermore, we assume that payoffs for both cooperating and defecting nodes are non-negative for all k.

3.2.1. The Nash-equilibrium

In the following section we analyze the Nash-equilibrium in the one shot PD game under the assumption that all the nodes joining an existing network choose simultaneously.

Assume that k nodes, aside from node i, play cooperatively. We want to study the condition under which node i, which is not part of the set of k cooperating nodes, chooses to cooperate; player i chooses to play 'c' if and only if her utility is higher than when playing 'd', i.e.:

$$\alpha_i u[B(k + 1) - C(k + 1)] \\ + \beta_i r\left[\frac{B(k + 1) - C(k + 1)}{N \cdot B(k + 1) - (k + 1)C(k + 1)}\right] \\ \geqslant \alpha_i u[B(k)] + \beta_i r\left[\frac{B(k)}{N \cdot B(k) - kC(k)}\right]. \tag{3}$$

This is equivalent to node i playing 'c' if and only if

$$\frac{\alpha_i}{\beta_i} \leqslant \delta(k) \tag{4}$$

where

$$\delta(k) = \frac{r\left[\frac{B(k+1)-C(k+1)}{N \cdot B(k+1)-(k+1)C(k+1)}\right] - r\left[\frac{B(k)}{N \cdot B(k)-kC(k)}\right]}{u[B(k)] - u[B(k+1) - C(k+1)]}.$$

In order to choose 'c' the node must be overcompensated for the loss in absolute gain by moving closer to the average gain.

The *general conditions for a Nash-equilibrium of a ERC-PD game* [11] of N nodes whereby the number of cooperating nodes is k^* can be used to study expression (4):

$$\frac{\alpha_i}{\beta_i} \leqslant \delta(k^* - 1) \quad \text{for } k^* \text{ nodes (playing 'c'),} \qquad (5)$$

$$\frac{\alpha_i}{\beta_i} \geqslant \delta(k^*) \quad \text{for the remaining } N - k^* \text{ nodes}$$

$$\text{(playing 'd').} \qquad (6)$$

Conditions (5) and (6) can be used to evaluate the number of nodes k^* that may possibly cooperate in a Nash-equilibrium. On one hand, as long as $\delta(k^* - 1) < 0$, there is no chance of having a coalition of size k^* because $\alpha_i/\beta_i > \delta(k^* - 1)$ for all types and condition (5) cannot hold for any node. On the other hand, the conditions for a Nash-equilibrium given by (5) and (6) imply that if $\delta(k^* - 1) > 0$ then there are types $[(\alpha_i/\beta_i)_{i=1,...,N}]$ of nodes such that k^* nodes cooperate and $N - k^*$ nodes free-ride. Note that for a given distribution of ERC-types, $\delta(k^* - 1) > 0$ is a *necessary but not sufficient condition* to get a coalition size of k^*. For a given payoff structure with $\delta(k^* - 1) > 0$, however, there exist ERC-types such that k^* is the equilibrium for any coalition size.

In order to find feasible coalition sizes, we must therefore study conditions under which $\delta(k)$ is positive.

Note that in (4) the denominator of $\delta(k)$ is positive due to Assumption 1. The sign of the numerator, however, depends on the number k of cooperating nodes.

For $k = 0$ the sign of the numerator is negative, since

$$r\left(\frac{B(1) - C(1)}{NB(1) - C(1)}\right) = r\left(1 - \frac{(N-1)B(1)}{NB(1) - C(1)}\right)$$
$$< r\left(\frac{B(0)}{NB(0)}\right) = r\left(\frac{1}{N}\right).$$

For $k = N - 1$ the sign of the numerator is positive, since

$$r\left(\frac{B(N) - C(N)}{NB(N) - NC(N)}\right) = r\left(\frac{1}{N}\right)$$
$$> r\left(\frac{B(N-1)}{NB(N-1) - (N-1)C(N-1)}\right)$$
$$= r\left(1 - \frac{(N-1)B(N-1) + (N-1)C(N-1)}{NB(N-1) - (N-1)C(N-1)}\right).$$

Therefore, $\delta(0) < 0 < \delta(N - 1)$ and no nodes unilaterally cooperate whereas all nodes playing 'c' can establish an equilibrium, provided that all nodes' types (α_i/β_i) are smaller than $\delta(N - 1)$.

In general, there are equilibria where only a certain number k^* of nodes cooperate. The crucial point is to find whether or not the numerator is positive. Remember that we previously assumed that

$$r\left(\frac{1}{N} - x\right) \leqslant r\left(\frac{1}{N} + x\right), \quad \forall x \in \left[0, \frac{1}{N}\right].$$

It is necessary, in order to obtain $\delta(k) > 0$, that a node choosing 'd' further deviates from the equal share $(1/N)$ than by playing 'c', i.e.

$$\frac{1}{N} - \frac{B(k+1) - C(k+1)}{NB(k+1) - (k+1)C(k+1)}$$
$$> \frac{1}{N} - \frac{B(k)}{NB(k) - kC(k)}. \qquad (7)$$

It is possible to show that inequality (7) is satisfied for $k > N/2$. [4]

Assumptions 1 and 2 imply that the condition $\delta(k) > 0$ is necessary (but not sufficient) to state that, for any given vector of types, if a node plays 'c' at the equilibrium, then at least half of the nodes cooperate.

Proposition 1. *For any given payoff structure of the PD game with ERC preferences, there is always an equilibrium in which all nodes defect.*

Proposition 2. *Given Assumption* 1 *and Assumption* 2, *there is a Nash-equilibrium where at least N/2 nodes cooperate.*

Based on Proposition 2, if there is a coalition of cooperating nodes then it is rather large.

---

[4] The proof of this affirmation is given in Appendix A.

3.3. The prisoner's dilemma with a continuous strategy space

In Section 3.2.1, we assumed that nodes only have a discrete option as to whether to cooperate or not. Now, we turn to a prisoner's dilemma game where nodes can continuously choose their cooperation levels. As we will see, ERC alone cannot improve upon the non-cooperative Nash-equilibrium with standard preferences whereby only the absolute payoff matters. However, introducing more structure to the game, i.e. if nodes play a coalition game (Section 3.4), ERC may yield a rather large coalition size or even support the grand coalition.

Let the number of nodes again be denoted by N. We define the cooperation level q_i $(\in [0,1])$ as the fraction of packets (both data and routing packets) that node i forwards to its neighboring nodes or to the destination node. Each node must choose its cooperation level q_i $(i = 1, \ldots, N)$. Cooperation induces some costs $C(q_i)$ that are assumed to be increasing and convex in the cooperation level $(C'(\ ) > 0, C''(\ ) > 0)$. Cooperation also yields some benefit $B(Q)$ in terms of network connectivity and aggregate cooperation effort made available by cooperating nodes, where $Q = \sum_i q_i$ denotes the aggregate cooperation level. Benefits from cooperation are increasing and concave, $B'(\ ) \geqslant 0$, $B''(\ ) < 0$. The payoff to a node is therefore determined by: $B(Q) - C(q_i)$.

3.3.1. The Nash-equilibrium

In order to find the Nash-equilibrium point(s) of the game (if one exists) it is necessary to identify the strategy q_i that corresponds to the singularity point(s) of the global ERC utility function, i.e. finding the roots of the first order derivative of the utility function and make sure that those points are maximum of the utility function. It should be noted that the assumptions made on the convexity of the utility function (see Section 3.1) alleviate the problem of studying the border conditions in the function domain. In the rest of the paper we refer to the *first order condition* when describing the process of finding the Nash-equilibrium.

We analyze the Nash-equilibrium when nodes act simultaneously. Node i chooses q_i to maximize its utility function $\alpha_i u(y_i) + \beta_i r(\sigma_i)$, where

$$y_i = B\left(\sum_{j \neq i} q_j + q_i\right) - C(q_i)$$

$\rightarrow$ Absolute payoff to player i,

$$\sigma_i = \frac{y_i}{\sum_j y_j} \quad \rightarrow \text{ Relative payoff to player } i.$$

Note that the definition used to express the absolute payoff to node i emphasizes the strategy space for node $i(q_i)$ as compared to the strategy space available to the other nodes of the network.

By choosing q_i, each node determines its own cooperation costs and the benefits from cooperation. The choice of q_i also impacts the payoff of the remaining nodes that in turn is fed back to the node's own utility through the relative payoff.

The first order condition is therefore given by

$$\left[\alpha_i u'(\ ) + \beta_i r'(\ )\frac{\sum_{j \neq i} y_j}{\sum_j y_j^2}\right][B'(Q) - C'(q_i)]$$
$$- \beta_i r'(\ )\frac{y_i}{\sum_j y_j^2}(N - 1)B'(Q) = 0.$$

The first order condition can be rewritten as

$$\alpha_i u'(\ )[B'(Q) - C'(q_i)]$$
$$+ \beta_i r'(\ )\left[\frac{\sum_j y_j - y_i}{\sum_j y_j^2}B'(Q) - \frac{\sum_{i \neq j} y_j}{\sum_j y_j^2}C'(q_i)\right] = 0 \tag{8}$$

or

$$\alpha_i u'(\ )[B'(Q) - C'(q_i)]$$
$$+ \beta_i r'(\ )\left[\frac{1 - N\sigma_i}{\sum_j y_j}B'(Q) - \frac{\sum_{j \neq i} \sigma_j}{\sum_j y_j}C'(q_i)\right] = 0. \tag{9}$$

The strategy (q_i) of node i to a given cooperation policy for the rest of the network can be calculated from this first order condition.

Proposition 3. (Continuous game) *In the continuous PD-game based on ERC preferences, the Nash-equilibrium is given by solving the following expression*: $B'(Nq^*) - C'(q^*) = 0$. *It is symmetrical*

as long as at least one node draws utility from its absolute payoff ($\alpha i > 0$). [5]

Introducing ERC preferences, therefore, does not increase the cooperation effort chosen by the nodes *when playing the PD-game with a continuous action set*. It does not even change the equilibrium cooperation levels. In contrast to the (discrete) prisoner's dilemma, ERC does not add any equilibrium in which there is more cooperation effort. The existence of equilibrium in the PD game that mimics cooperative behavior, therefore, only arises in the presence of discrete action sets. Having a continuous decision variable, ERC does not change the set of equilibrium. The reason is that ERC does not establish a preference for being cooperative, but for being similar to other nodes with respect to the payoff.

In this section, however, we used the ERC theory in a classical non-cooperative setting: let us see how the strategy selection of a selfish node change when introducing more structure to the game, i.e. when considering a cooperative-game setting.

3.4. Coalition formation: the cooperative-game approach

As a further refinement, we now propose a cooperative-games approach consisting in a combination of the ERC preference structure and the two-stage coalition formation method as introduced in [13].

Let us assume that all nodes are identical with respect to their payoff function (i.e. they use the same definition of utility function). In a first stage, nodes decide whether or not to join the coalition. By the principle of "rationality", each node is assumed to know the decisions of the other nodes. The cooperation levels (i.e. the strategy) that will be chosen in the second stage depend on whether the nodes take part in the coalition or not. The coalition thereby maximizes its collective benefits and plays against the nodes that don't take part in the coalition, which simultaneously maximize their individual utility.

We first study the case of nodes that have identical ERC-types. We demonstrate that within the coalition formation game, ERC-preferences can enforce cooperation and even result in the grand coalition. We then look at the case of heterogeneous ERC-types. By studying the extreme scenario of nodes that are solely interested either in their absolute payoff or in equity, we will explore the effects of the existence of some equity-oriented nodes in the network.

3.4.1. Coalition of identical ERC-types

We will now solve the coalition formation game backwards, that is, for any coalition size k, we first study the first order conditions for the choice of the cooperation level inside and outside the coalition. Then, in the second step, the equilibrium coalition size is determined by a stability condition. This means that in the equilibrium, k must satisfy the condition that there is no incentive to leave the coalition. [6]

For standard preferences (using ERC-preferences this results in the special case $\beta = 0$), the game theory literature shows that the coalition size is rather small. Using ERC preferences, however, the number of nodes within a coalition can be much higher in equilibrium.

Instead of solving the game in general, we will show that if nodes only value the relative payoff high enough, i.e. α/β is below a certain bound then even the grand coalition can be stable.

The first order condition for nodes outside the coalition (S) is given by (10), whereas the cooperation strategy of nodes that take part in the coalition is chosen by maximizing the utility function of a representative member: indeed all nodes within the coalition S select the same strategy q_s since they are assumed to be of the same type. This implies that all members of the coalition have identical absolute payoff ($y_S = B(Q) - C(q_S)$) and relative payoff ($\sigma_S = y_S/(ky_S + \sum_{j \notin S} y_j)$).

---

[5] The proof of Proposition 3 is given in Appendix B.

[6] The original work introduced in [23] states that the stability condition is such that there is an incentive to neither leave nor join the coalition.

The first order condition is given by

$$[\alpha u'(\ ) + \beta r'(\ )]\frac{\sum_{j \notin S}\sigma_j}{\sum_j y_j}[kB'(Q) - C'(q_S)]$$

$$- \beta r'(\ )\frac{\sigma_S}{\sum_j y_j}(N-k)kB'(Q) = 0, \qquad (10)$$

$$\alpha u'(\ )[kB'(Q) - C'(q_S)]$$

$$+ \beta r'(\ )\left[-\frac{\sum_{j \notin S}\sigma_j}{\sum_j y_j}C'(q_S) + kB'(Q)\frac{1 - N\sigma_S}{\sum_j y_j}\right] = 0. \qquad (11)$$

- For nodes that do not belong to the coalition S we know from Section 3.4 that if $\sigma_j < (>)1/N$ for $j \notin S$ then $B'(Q) > (<)C'(q_j)$.
- For the coalition, we obtain from (10) and (11) that if $\sigma_S < (>)1/N$ then $kB'(Q) > (<)C'(q_S)$. [7] Since $B'(Q) > kB'(Q)$, [8] the first order conditions imply that for nodes within the coalition $\sigma_S \leqslant 1/N$ and thus: $kB'(Q) \geqslant C'(q_S)$. To prove that inside the coalition $\sigma_S \leqslant 1/N$, assume to the contrary that $\sigma_S > 1/N$ and that $\sigma_j < 1/N$ for some nodes j outside the coalition. Inequalities (10) and (11) imply that $C'(q_j) < B'(Q) < kB'(Q) < C'(q_S)$ which contradicts the assumption of increasing and convex cooperation costs.

Inequalities (10) and (11) can be used to show the following proposition:

Proposition 4. (Coalition game) *In the symmetric coalition game for identical ERC preferences (type α/β), the grand coalition is stable if α/β is sufficiently small, i.e. nodes are interested enough in being close to the equal share.*

Note first, that within the grand coalition, the cooperation level satisfies the condition $NB'(Nq^*) = C'(q^*)$, independently of the ERC-types and nodes that receive the equal share.

If node i leaves the coalition ($k = N - 1$), then from the first order conditions we obtain

---

[7] Assuming that $\sigma_S > 1/N$ then $r'(\ ) < 0$ and (11) implies $kB'(Q) < C'(q_S)$.

[8] For $k \geqslant 2$.

$$(N-1)B'[(N-1)q_S + q_i] \geqslant C'(q_S) \geqslant C'(q_i)$$

$$> B'[(N-1)q_S + q_i]. \qquad (12)$$

Let us now look at the cooperation levels that would result if the ERC-type α/β goes to zero. In this case, nodes get more and more interested in getting their equal share, and their cooperation levels will converge: in the limit $\tilde{q} = q_S = q_i$. However, in the limit, inequality (12) still must hold, i.e. $(N-1)B'(Nq) \geqslant C'(q)$.

In the limit the absolute payoff of a node leaving the coalition is smaller than within the grand coalition, whereas the relative payoff is the same, i.e. $NB'(\tilde{Q}) \geqslant C'(\tilde{q})$.

Therefore, as long as α/β is small enough, the absolute payoff remains lower and the utility derived from the relative payoff is also smaller than in the grand coalition. Thus, no node has an incentive to leave the grand coalition if α/β is small enough.

3.4.2. Coalition of heterogeneous ERC-types

When nodes with heterogeneous ERC-types are allowed to take part in the coalition (S), those nodes that have the largest α_i/β_i will have the greatest interest to leave the coalition in order to obtain a larger absolute payoff.

We will now concentrate on the extreme case in which nodes are either interested in their absolute payoff ($\beta_i = 0$) or in equity ($\alpha_i = 0$). The former are referred to as A-nodes, the latter as B-nodes. In total, there are N_a A-nodes and N_b B-nodes; k_a of these A-nodes and k_b B-nodes form the coalition. The cooperation levels are denoted by q_{as}, q_{bs} for nodes inside S, q_{an} and q_{bn} for nodes outside the coalition.

Let us first look at the behavior of B-nodes.

Outside the coalition, any B-nodes can arrive at the equal share by choosing the average cooperation cost level. Thus,

$$C(q_{bn}) = \frac{1}{N_a + k_b}[k_a C(q_{as}) + k_b C(q_{bs})$$

$$+ (N_a - k_a)C(q_{an})]. \qquad (13)$$

A B-node inside the coalition has no incentive to leave if it also receives the equal share:

$$C(q_{bs}) = \frac{1}{N_a - k_b}[k_a C(q_{as}) + (N_b - k_b)C(q_{bn}) + (N_a - k_a)C(q_{an})]. \qquad (14)$$

In equilibrium, all B-nodes choose the same cooperation level, $q_b \cong \hat{q}_{bn} = q_{bs}$ and receive the equal share:

$$C(q_b) = \frac{1}{N_a}[k_a C(q_{as}) + (N_a - k_a)C(q_{an})]. \qquad (15)$$

A-nodes outside the coalition maximize their absolute payoff, $B(Q) - C(q_{an})$. The first order condition is given by

$$B'(Q) = C'(q_{an}). \qquad (16)$$

Within the coalition, the utility of a representative A-type-member is maximized by guaranteeing that the B-members get the equal share, i.e. $C(q_{bs})$. The first order condition for choosing q_{as} is given by

$$B'(Q)\left[k_a + k_b \frac{\partial q_{bs}}{\partial q_{as}}\right] - C'(q_{as})$$
$$= B'(Q)k_a\left[1 + \frac{k_b}{N - k_b}\frac{C'(q_{as})}{C'(q_{bs})}\right] - C'(q_{as}) = 0. \qquad (17)$$

By construction, for any given k_a and k_b, every B-node is indifferent to being either inside or outside the coalition. For a coalition to be stable, an A-node must not have an incentive to leave the coalition. In general, for any k_b there will be a certain number of A-nodes, k_a, that will join the coalition. We have multiple equilibria.

Inequalities (13)–(17) can be used to infer the following results:

Result 5. The larger the total number of equity-oriented nodes (N_b), the higher the incentives for A-nodes to join the coalition. Hence, for a given k_b, the number of cooperating A-nodes k_a increases in N_b.

Result 6. The more B-nodes join the coalition, the smaller the incentive for A-nodes to do so. In equilibrium, k_b and k_a are negatively correlated.

Result 7. The total cooperation level increases with the number of B-types outside the coalition. A joining B-node improves the payoffs only if it does not drive out an A-node.

The rationale of Results 5 and 6 is the following: if an A-node enters the coalition and the coalition increases its cooperation efforts, B-nodes outside the coalition increase their cooperation activities as well and thereby additionally reward the entering node.

If the number of such equity-oriented B-nodes outside the coalition gets larger, this external reward for joining a coalition increases and, therefore, the equilibrium coalition size increases. Analogously, if B-nodes join the coalition, fewer nodes outside the coalition reward the entering A-node by an increase of their cooperation activities. Hence, the incentives for A-nodes to enter the coalition decrease and the number of A-nodes that are inside the coalition in equilibrium gets smaller.

Result 7 reflects the fact that the more nodes cooperate, the higher the efficiency gains are and the closer the aggregate cooperation level is to the efficient one. The impact of A- and B-nodes on the decision of the coalition, however, differs in the following way: a joining A-node is interested in the absolute payoff and, consequently, the re-optimizing coalition increases its cooperation effort because the positive effect on one more node is now taken into account. A joining B-node, however, is not primarily interested in the absolute payoff, but in the equal share. Therefore, the coalition will not increase the total cooperation level that much because the B-node refrains from deviating from the cooperation level of non-cooperating nodes. Consequently, the efficiency gains are larger if an A-node enters the coalition than if a B-node joins. Therefore, B-nodes are welcome inside a coalition only if their entering does not drive out an A-node.

3.5. Discussion: coalition formation process and the cooperation enforcement mechanism CORE

Self-interested, autonomous mobile nodes of an ad hoc network may interact "rationally" to gain and share benefits in stable (temporary) coalitions:

this is to save costs by coordinating activities with other nodes of the network. For this purpose, each node determines the utility of its actions in a given environment by an individual utility function. In Section 3.1 we introduced a more sophisticated model in which not only self-centered preferences are taken into account to derive the individual payoff of an action but also relative information is used in order to find an extended set of possible equilibrium points.

Results obtained with the proposed model are promising: in a dynamic network formed by nodes that follow the definition of utility given by the ERC theory, depending on the node types, it is possible to obtain stable coalitions of a relatively large size and under certain circumstances, even the grand coalition becomes feasible. Node types are determined by the two parameters α and β which represent the key factor of the coalition formation process.

We believe that the reputation technique implemented in CORE can be used as an effective mechanism to impose a specific identical ERC type for every node participating in a cooperative setting as an ad hoc network. Indeed, the reputation measure introduced in [5] is compliant with the incentive structure given by (1) and (2). Cooperation is made attractive from an individual point of view because the cost of participating to the network operation is compensated with a higher reputation value, which is the pre-requisite for a node to establish a communication with other nodes in the network. On the other hand, when the number of cooperating nodes increases, the cost for participation is compensated by a more connected network that in turn increases the benefit of cooperation. Now, if the two parameters α and β are represented as functions of the reputation r_{ni} as defined in [5], then it is possible to enforce a particular value to the α/β ratio. Specifically it is possible to dynamically adjust the α/β ratio in order to be compatible with Proposition 4. Thus, even the grand coalition is stable and every node of the network cooperates bearing the same costs and getting equal benefits by choosing a fair operating point in which no one deviates from the average cooperation level chosen by the coalition.

The relation between α, β and r_{ni} is indirectly proportional: the lower the reputation value (meaning that the past strategy selected by the node has been to reduce the cooperation level) the higher will be factor β and the lower will be factor α thus reducing the α/β ratio, and vice-versa.

The relation between the reputation value and the ERC type of a node becomes more complicated if we allow the presence of nodes with different ERC types: modeling a network that allows different ERC types is interesting when considering mobile nodes with different capabilities such as different battery power and different computational power.

However, in order to provide a formal assessment of the efficiency of the reputation mechanism proposed in CORE it is necessary to evaluate the node model presented in the previous sections in a dynamic setting: the reputation value is computed based on the past strategies selected by the nodes of the network and have an influence on those nodes' future actions. Furthermore any variation on the strategy selection phase of a node has an impact on the strategies selected by neighboring nodes: solutions to the dynamic coalition formation process still have to be examined.

We believe that the research we have conducted so far has given some interesting results and proposes a useful basis to study the coalition formation process of autonomous self-interested mobile nodes by means of reputation mechanisms which is, to the best of our knowledge, a rather unexplored domain. However, we think that it is possible to express the dynamic coalition formation process using a more elegant and simple methodology, which is a key requirement for studying dynamic games. The relatively recent literature on the subject states that the models of coalition formation may be classified into two main categories: utility-based models, as it is largely favored by game theory, and complementary-based models. Up to now, most classic methods and protocols for the formation of stable coalitions among rational agents follow the utility-based approach and cover two main activities which may be interleaved: the generation of coalition structures, that is partitioning or covering the set of agents into coalitions, and the distribution of gained benefit

among the participants to each of the coalitions. The future research direction we will take is to prove that reputation mechanisms in general are compliant to the so called *Coalition Formation Algorithm*. Coalition formation algorithms are those mechanisms that provide a feasible solution to a cooperative game in coalitional structure: there are several solution concepts and we will focus on the so called *Kernel-oriented* solutions [14,15]. Kernel-oriented coalitions are the most suitable for our purpose because the related literature gives precise conditions for a coalition formation algorithm to be kernel-stable with a polynomial complexity, as opposed to other solution/algorithms that are only of theoretical relevance since they have exponential complexity.

4. Non-cooperative games approach

In an alternative approach, we investigated on the characteristics of CORE by modeling the interactions between the nodes of a MANET as a non-cooperative game. In the following sections, we will introduce a specific and well-known game (the prisoner's dilemma, PD) and explain how and why this model is suitable to describe the decision making process that a mobile node would undertake when participating to the ad hoc network operation. Subsequent to the definition of the model that describes the interaction between decision-makers (nodes) involved in the game play, we will extend our analysis to a particular instance of games that goes under the name of repeated games. Repeated (or iterated) games have been exhaustively treated in the game theoretic literature [8–10,22–25] and interesting results concerning the establishment of a cooperative behavior will be presented. In particular, we will focus on the strategy that a player [9] adopts to determine whether to cooperate or not at each of the moves in the iterated game and describe an important strategy known as tit-for-tat (TFT) which has been considered by a lot of game-theorist to be one of

the best strategies not only to promote cooperation but also for the evolution of cooperation (a definition of "evolution of cooperation" will be given in the following sections). We will then describe how the CORE cooperation enforcement mechanism can be translated into a strategy for a player and compare it to the TFT strategy to numerically prove the equivalence between CORE and TFT. By further extending the game theoretic concept applied to the classical iterated PD game we will show how the performances of TFT and its derivates (i.e. generous-TFT, GTFT) degrade as *noise* is introduced in the model. In the following sections we will describe how the introduction of a noise factor allows grasping the undesirable effects of using the *promiscuous mode* operation of a wireless card as a basis for the monitoring mechanism (the watchdog mechanism) and prove that the CORE strategy outperforms all other known strategies both for promoting cooperation and for the evolution of cooperation. The numerical results obtained through a simulation software designed by [23] are stimulating the more difficult task of providing a formal analysis of the CORE strategy, which is part of our future work.

4.1. System model

In order to describe the interaction between nodes of a MANET and the decision making process that results in a cooperative or selfish behavior of the nodes we will use a classical game introduced by Tucker [24, pp. 117–118]. In the classical PD game, two players are both faced with a decision to either cooperate (C) or defect (D). The decision is made simultaneously by the two players with no knowledge of the other player's choice until the choice is made. If both cooperate, they receive some benefit (R). If both defect they receive a specific punishment (P). However, if one defects, and one cooperates, the defecting strategy receives no punishment (T) and the cooperator a punishment (S). The game is often expressed in the canonical form in terms of pay-offs.

The PD game is a much studied problem due to its far-reaching applicability in many domains. In game theory, the prisoner's dilemma can be viewed

---

[9] In this paper we will adopt the word player and node as synonyms.

as a two-players, non-zero-sum, non-cooperative and simultaneous move game. In order to have a dilemma the following expressions must hold:

$$T > R > P > S,$$
$$R > \frac{S + T}{2}. \tag{18}$$

In our model, a MANET formed by N nodes is considered as an N-player playground in which randomly, any two nodes can meet. We suppose that every node of the network has some data traffic to be sent through some source route that is the result of the execution of some routing protocol (as an example the DSR protocol). We also suppose that when any two nodes meet, at some time period t, they both need to send some data packets through each other, i.e. using each other as a relay node. Before the actual process of sending a packet, the two nodes have to take the decision whether to cooperate or defect. By cooperating a node will forward one (or more) data packet for the requesting node, whereas by defecting a node will not relay data packets on behalf of the requesting node. Instead of including an accurate description of energetic costs, topology information, possible interference and path information we will base our model on some basic economic modeling. As an illustrative and intuitive example, let us consider two players (nodes) with some letters (data messages) to send. For each letter leaving a player, a stamp (energy cost for sending one data packet) has to be used. When a letter is forwarded towards its destination the player benefit is (arbitrarily) fixed to 5: of course, the benefit for a successful communication should be higher than the (energetic) cost for sending the letter. So, for example, if two players meet and both have a letter to send, if the decision of a player is to cooperate, she will have to spend two stamps (one for her letter, and one for her opponent's letter) and eventually receive a benefit of 5 if her opponent cooperated, leading to a payoff equals to 3 in case the opponent decided to cooperate and to a payoff equals to -2 in case the opponent decided to defect. This situation can be translated in a payoff matrix which matches the one illustrated in Table 2 of the classical PD game.

Of course, it is arguable that such a simple model can represent a real MANET, but we believe that the limitations imposed by our model are greatly compensated by the consolidated theoretical results available in the literature for the prisoner's dilemma. Furthermore, we plan to extend the model in order to cope with a T-player simultaneous move game where $T < N$ thus taking into account the cooperative strategy of nodes that are part of an entire path from a data source to her selected destination.

However, the key of the model presented in this paper and any further extensions is the ''willingness to communicate'' assumption: during every play of the game (both in the basic PD and in the iterated PD, as we will see in the next section) both players engaged in the decision making process (cooperate or not) are supposed to have some data packets to be sent through the opponent player. As we will see later this assumption is necessary in order to implement a punishment mechanism for a non-cooperating node.

4.2. The iterated prisoner's dilemma

The iterated version of the PD game, and in general repeated games have been extensively studied in the literature and the interested reader could refer to [10] in order to find a basic yet complete introduction to the theory of games, equilibrium concepts and iterated games. [10] In this paper we will not focus on the basic results from game theory applied to the PD (e.g. Nash-equilibrium of the one shot PD game) but we will introduce some concepts that will be used in the rest of the paper.

One surprising feature of many one-shot games (i.e. games that are played only once) including the PD game, is that the Nash-equilibrium is non-cooperative: each player would prefer to fink (defect) rather than to cooperate. However, in a more realistic scenario (e.g. a MANET) a particular one shot game can be played more than once; in fact, a realistic game could even be a correlated series of one shot games. In such iterated games an action

---

[10] In this section we often refer to or use text and examples available on the Jim Ratliff notes. Text and examples will not be quoted just for clarity of presentation.

Table 2
Prisoner's dilemma payoff matrix: (Panel A) general, (Panel B) example

Player i	Player j	
	Cooperate	Defect
Panel A		
Cooperate	(R, R)	(S, T)
Defect	(T, S)	(P, P)
Panel B		
Cooperate	$(3, 3)$	$(-2, 4)$
Defect	$(4, -2)$	$(0, 0)$

chosen by a player early on can affect what other players choose to do later on: repeated games can incorporate a phenomena which we believe are important but not captured when restricting our attention to static, one shot games. In particular, we can strive to explain how cooperative behavior can be established as a result of rational behavior. In this section we will discuss repeated games which are "infinitely repeated". This need not mean that the game never ends, however. We will see that this framework can be appropriate for modeling situations in which the game eventually ends (with probability one) but the players are uncertain about exactly when the last period is (and they always believe there's some chance the game will continue to the next period).

In the following subsections we will introduce in a more formal way some basic concepts related to repeated games and infinitely repeated games. We will then show the definition of a strategy for a player and explain how to verify if a (simple) strategy is an equilibrium for a game. A reader who is familiar with game theory is invited to skip the following two Sections 4.2.1 and 4.2.2.

4.2.1. Repeated games theory

Consider a game G (which we will call the *stage game* or the *constituent game*). Let the player set be $I = \{1, \ldots, n\}$. In our present repeated-game context it will be clarifying to refer to a player's stage game choices as actions rather than strategies. (We will reserve "strategy" for choices in the repeated game). So each player has a *pure-action* space A_i. The space of action profiles is $A = X_{i \in I} A_i$. Each player has a von Neumann–

Morgenstern utility function defined over the outcomes of G, $g_i : A \to \Re$, that in the particular case of the two players PD game takes the form of a payoff matrix as in Table 2.

Let G be played several times (perhaps an infinite number of times) and award each player a payoff which is the (discounted) sum of the payoffs she got in each period from playing G. Then this sequence of stage games is itself a game: a *repeated game*.

Two statements are implicit when we say that in each period we were playing the same stage game: (a) for each player the set of actions available to her in any period in the game G is the same regardless of which period it is and regardless of what actions have taken place in the past and (b) the payoffs to the players from the stage game in any period depend only on the action profile for G which was played in that period, and this stage-game payoff to a player for a given action profile for G is independent of which period it is played. Statements (a) and (b) are saying that the environment for our repeated game is *stationary* (or, alternatively, independent of time and history). *This does not mean the actions themselves must be chosen independently of time or history.*

We will limit our attention here to cases in which the stage game is a one-shot, simultaneous-move game. Then we interpret (a) and (b) above as saying that the payoff matrix is the same in every period. We make the typical "*observable action*" or "*standard private monitoring*" assumption that the play which occurred in each repetition of the stage game is revealed to all the players before the next repetition. Therefore even if the stage game is one of imperfect information (as it is in simultaneous-move games)—so that during the stage game one of the players does not know what the others are doing/have done that period—each player does learn what the others did before another round is played. This allows subsequent choices to be conditioned on the past actions of other players. We will see later in the paper that if we make the assumption of "*imperfect private monitoring*" results can be significantly different.

Before we can talk about equilibrium strategies in repeated games, we need to get precise about

what a strategy in a repeated game is. We will find it useful when studying repeated games to consider the semi-extensive form. This is a representation in which we accept the normal-form description of the stage game but still want to retain the temporal structure of the repeated game.

Let the first period be labeled $t = 0$. The last period, if one exists, is period T, so we have a total of $T + 1$ periods in our game. We allow the case where $T = \infty$, i.e. we can have an infinitely repeated game.

We will refer to the action of the stage game G which player i executes in period t as a_i^t. The action profile played in period t is just the n-tuple of individuals' stage-game actions:

$$a^t = (a_1^t, \ldots, a_n^t). \tag{19}$$

We want to be able to condition the players' stage-game action choices in later periods upon actions taken earlier by other players. To do this we need the concept of a *history*: a description of all the actions taken up through the previous period. We define the history at time t to be

$$h^t = (a^0, a^1, \ldots, a^{t-1}). \tag{20}$$

In other words, the history at time t specifies which stage-game action profile (i.e. combination of individual stage-game actions) was played in each previous period. Note that the specification of h^t includes within it a specification of all previous histories $h^0, h^1, \ldots, h^{t-1}$. For example, the history h^t is just the concatenation of h^{t-1} with the action profile a^{t-1}; i.e. $h^t = (h^{t-1}; a^{t-1})$. The history of the entire game is $h^{T+1} = (a^0, a^1, \ldots, a^T)$. Note also that the set of all possible histories h^t at time t is just

$$A^t = \prod_{j=0}^{t-1} A, \tag{21}$$

the t-fold Cartesian product of the space of stage-game action profiles A.

To condition our strategies on past events, then, is to make them functions of history. So we write player i's period-t stage-game strategy as the function s_i^t, where $a_i^t = s_i^t(h^t)$ is the stage-game action she would play in period t if the previous play had followed the history h^t. A player's stage-game

action in any period and after any history must be drawn from her action space for that period, but because the game is stationary her stage-game action space A_i does not change with time. The period-t stage game strategy profile s^t is

$$s^t = (s_1^t, \ldots, s_n^t). \tag{22}$$

So far we have been referring to stage-game strategies for a particular period. Now we can write, using these stage-game entities as building blocks, a specification for a player's strategy for the repeated game. We write player i's strategy for the repeated game as

$$s_i = (s_i^0, s_i^1, \ldots, s_i^T), \tag{23}$$

i.e. a $(T + 1)$-tuple of history-contingent player-i stage-game strategies. Each s_i^t takes a history $h^t \in A^t$ as its argument. The space S_i of player-i repeated-game strategies is the set of all such $(T + 1)$-tuples of player-i stage game strategies $s_i^t : A^t \to A_i$.

We can write a strategy profile s for the whole repeated game in two ways. We can write it as the n-tuple profile of players' repeated-game strategies:

$$s = (s_1, \ldots, s_n) \tag{24}$$

as defined in (23). Alternatively, we can write the repeated-game strategy profile s as

$$s = (s^0, s^1, \ldots, s^T) \tag{25}$$

i.e. as a collection of stage-game strategy profiles, one for each period, as defined in (22).

Let us see how this repeated game is played out once every player has specified her repeated-game strategy s_i. It is more convenient at this point to view this repeated-game strategy profile as expressed in (25), i.e. as a sequence of $T + 1$ history-dependent stage-game strategy profiles. When the game starts, there is no past play, so the history h^0 is degenerate: every player executes her $a_i^0 = s_i^0$ stage-game strategy from (23). This zeroth period play generates the history $h^1 = (a^0)$, where $a^0 = (a_1^0, \ldots, a_n^0)$. This history is then revealed (or monitored by the players themselves) to the players so that they can condition their period-1 play upon the period-0 play. Each player then chooses her $t = 1$ stage-game strategy $s_i^1(h^1)$.

Consequently, in the $t = 1$ stage game the strategy profile $a^1 = s^1(h^1) = (s^1_1(h^1), \ldots, s^1_n(h^1))$ is played. In order to form the updated history this stage-game strategy profile is then concatenated onto the previous history: $h^2 = (a^0, a^1)$. This new history is revealed to all the players and they each then choose their period-2 stage-game strategy $s^2_i(h^2)$, and so on. We say that h^{T+1} is the path generated by the repeated-game strategy profile s.

Let us now consider the *payoff function of the repeated game*. We can think of the players as receiving their stage-game payoffs period-by-period. Their repeated game payoffs will be an additively separable function of these stage-game payoffs. Right away we see a potential problem: if the game is played an infinite number of times, there is an infinite number of periods and, hence, of stage-game payoffs to be added up. In order that the players' repeated-game payoffs be well defined we must ensure that this infinite sum does not blow up to infinity. We ensure the finiteness of the repeated-game payoffs by introducing *discounting* of future payoffs relative to earlier payoffs. Such discounting can be an expression of time preference and/or uncertainty about the length of the game. We introduce the average discounted payoff as a convenience which normalizes the repeated-game payoffs to be "on the same scale" as the stage game payoffs.

Infinite repetition can be the key for obtaining behavior in the stage games which could not be equilibrium behavior if the game were played once or a known finite number of times. For example, defection in every period by both players is the unique equilibrium in any finite repetition of the PD. [11] When repeated an infinite number of times, however, cooperation in every period is an equilibrium if the players are "sufficiently patient".

When studying infinitely repeated games we are concerned about a player who receives a payoff in each of infinitely many periods. In order to represent her preferences over various infinite payoff streams we want to meaningfully summarize the desirability of such a sequence of payoffs by a single number. A common assumption is that the player wants to maximize a weighted sum of her per-period payoffs, where she weights later periods less than earlier periods. For simplicity this assumption often takes the particular form that the sequence of weights forms a geometric progression: for some fixed $\delta \in (0, 1)$, each weighting factor is δ times the previous weight. δ is called her *discount factor*. If in each period t player i receives the payoff u^t_i, we could summarize the desirability of the payoff stream $u^0_i, u^1_i, \ldots$ by the number

$$\sum_{t=0}^{\infty} \delta^t u^t_i. \tag{26}$$

Such an intertemporal preference structure has the desirable property that the infinite sum of the weighted payoffs will be finite (since the stage-game payoffs are bounded). A player would be indifferent between a payoff of x^t at time t and a payoff of $x^{t+\tau}$ received τ periods later if

$$x^t = \delta^\tau x^{t+\tau}. \tag{27}$$

A useful formula for computing the finite and infinite discounted sums we will use later in this section is

$$\sum_{t=T_1}^{T_2} \delta^t = \frac{\delta^{T_1} - \delta^{T_2+1}}{1-\delta}, \tag{28}$$

which, in particular, is valid for $T_2 = \infty$.

If we adopted the summation (26) as our players' repeated-game utility function, and if a player received the same stage-game payoff v_i in every period, her discounted repeated-game payoff, using (28), would be $v_i/(1 - \delta)$. It is however more convenient to transform the repeated-game payoffs to be "on the same scale" as the stage-game payoffs, by multiplying the discounted payoff sum from (26) by $(1 - \delta)$. So we define the average discounted value of the payoff stream $u^0_i, u^1_i, \ldots$ by

$$(1 - \delta) \sum_{t=0}^{\infty} \delta^t u^t_i. \tag{29}$$

It is often convenient to compute the average discounted value of an infinite payoff stream in terms of a leading finite sum and the sum of a trailing infinite substream. For example, say that the payoffs v^t_i a player receives are some constant

---

[11] See Theorem 4 in "Repeated Games" handouts by Ratliff [10].

payoff v_i' for the first t periods, i.e. $0, 1, 2, \ldots, t-1$, and thereafter she receives a different constant payoff v_i'' in each period $t, t+1, t+2, \ldots$ The average discounted value of this payoff stream is

$$(1-\delta)\sum_{\tau=0}^{\infty}\delta^{\tau}v_i^{\tau} = (1-\delta)\left(\sum_{\tau=0}^{t-1}\delta^{\tau}v_i^{\tau} + \sum_{\tau=t}^{\infty}\delta^{\tau}v_i^{\tau}\right)$$
$$= (1-\delta)\left(\frac{v_i'(1-\delta^t)}{1-\delta} + \frac{v_i''\delta^t}{1-\delta}\right)$$
$$= (1-\delta^t)v_i' + \delta^t v_i''. \tag{30}$$

It is possible to see that the average discounted value of this stream of bivalued stage-game payoffs is a convex combination of the two stage-game payoffs. We can iterate this procedure in order to evaluate the average discounted value of more complicated payoff streams. Another useful example is when a player receives v_i' for the first t periods, then receives v_i'' only in period t and receive v_i''' every period thereafter. The average discounted value of the stream beginning in period t (discounted to period t) is: $(1-\delta)v_i'' + \delta v_i'''$. Substituting this for v_i'' in (30), we find that the average discounted value of this three-valued payoff stream is

$$(1-\delta^t)v_i' + \delta^t[(1-\delta)v_i'' + \delta v_i''']. \tag{31}$$

We have now defined all the formalism needed to examine the equilibrium of a (infinitely) repeated PD game and to verify if a predefined strategy constitutes an equilibrium. The various definitions of equilibrium and the related theorems can be found in [10].

4.2.2. Cooperation in the repeated prisoner's dilemma

In the one-shot PD, the players cannot avoid choosing their dominant strategy defect (see Table 3). In order to make the following analysis simpler, consider the following payoff matrix.

It is easy to verify that conditions (18) hold.

Even when this game is finitely repeated, because the stage game has a unique Nash-equilibrium, the unique subgame-perfect equilibrium has both players defecting in every period. However, when the players are sufficiently patient it is possible to sustain cooperation (i.e. keeping

Table 3
Modified PD payoff matrix

Player i	Player j	
	Cooperate	Defect
Cooperate	$(1,1)$	$(-1,2)$
Defect	$(2,-1)$	$(0,0)$

"Cooperate") in every period as a subgame-perfect equilibrium of the infinitely repeated game. First we will see that such cooperation is a Nash-equilibrium of the repeated game. We will then show that this cooperation is a subgame-perfect equilibrium.

When an infinitely repeated game is played, each player i has a repeated-game strategy s_i, which is a sequence of history-dependent stage-game strategies s_i^t; i.e. $s_i = (s_i^0, s_i^1, \ldots)$, where each $s_i^t : A^t \to A_i$. The n-tuple of individual repeated-game strategies is the repeated-game strategy profile $s = (s_1, s_2, \ldots, s_n)$.

As a fundamental example, let us consider a particular strategy that a player could follow and which is sufficient to sustain cooperation. This strategy is also known as the `spiteful` strategy.

`Spiteful`

- Cooperate in the first period.
- In later periods, cooperate if both players have always cooperated.
- However, if either player has ever defected, defect for the remainder of the game.

More precisely and formally, it is possible to write player i's repeated-game strategy $\bar{s}_i = (\bar{s}_i^0, \bar{s}_i^1, \ldots)$ as the sequence of history-dependent stage-game strategies such that in period t and after history h^t,

$$\bar{s}_i^t(h^t) = \begin{cases} C, & t = 0 \text{ or } h^t = ((C,C)^t), \\ D, & \text{otherwise.} \end{cases} \tag{32}$$

First, we will show that for sufficiently "patient players" the strategy profile $\bar{s} = (\bar{s}_1, \bar{s}_2)$ is a

Nash-equilibrium of the repeated game. Then we will show that for the same required level of patience these strategies are also a subgame-perfect equilibrium.

Now, if both players conform to the alleged equilibrium prescription, they both play "cooperate" at $t = 0$. Therefore at $t = 1$, the history is $h^1 = (C, C)$; so they both play "cooperate" again. Therefore at $t = 2$, the history is $h^2 = ((C, C), (C, C))$, so they both play "cooperate" again. And so on... The path of s is the infinite sequence of cooperative action profiles $((C, C), (C, C), \ldots)$. The repeated-game payoff to each player corresponding to this path is trivial to calculate: they each receive a payoff of 1 in each period, therefore the average discounted value of each player's payoff stream is 1.

Can player i gain from deviating from the repeated-game strategy $\bar{s}_i$ given that player j is faithfully following $\bar{s}_j$? Let t be the period in which player i first deviates. She receives a payoff of 1 in the first t periods $0, 1, \ldots, t-1$. In period t, she plays "defect" while her conforming opponent played "cooperate", yielding player i a payoff of 2 in that period. This defection by player i now triggers an open-loop "defect" always response from player j. Player i's best response to this open-loop strategy is to "defect" in every period herself. Thus she receives zero in every period $t+1, t+2, \ldots$ To calculate the average discounted value of this payoff stream to player i we can refer to (31), and substitute $v_i' = 1$, $v_i'' = 2$, and $v_i''' = 0$. This yields player i's repeated-game payoff when she defects in period t in the most advantageous way to be $1 - \delta^t(2\delta - 1)$. This is weakly less than the equilibrium payoff of 1, for any choice of defection period t, as long as $\delta \geq \frac{1}{2}$. Thus we have defined what we meant by "sufficiently patient": cooperation in this PD game is a Nash-equilibrium of the repeated game as long as $\delta \geq \frac{1}{2}$.

To verify that $\bar{s}$ is a subgame-perfect equilibrium of the repeated prisoners' dilemma it is necessary to check that this strategy profile's restriction to each subgame is a Nash-equilibrium of that subgame. Consider a subgame, beginning in period τ with some history h^τ. What is the restriction of $\bar{s}_i$ to this subgame? Denoting the restriction by $\hat{s}_i$ we have

$$\hat{s}_i^t(\hat{h}^t) = \bar{s}_t^{t+\tau}(h^\tau; \hat{h}^t)$$
$$= \begin{cases} C, & h^\tau = ((C, C)^\tau) \text{ and } \hat{h}^t = ((C, C)^t), \\ D, & \text{otherwise.} \end{cases}$$
$$(33)$$

We can partition the subgames of this game, each identified by a beginning period τ and a history h^τ, into two classes: (A) those in which both players chose "cooperate" in all previous periods, i.e. $h^\tau = ((C, C)^\tau)$, and (B) those in which a defection by either player has previously occurred. For those subgames in class (A), the sequence of restrictions $\hat{s}_i^t(\hat{h}^t)$ from (33) reduces to the sequence of original stage-game strategies $\bar{s}_i^t(h^t)$ from (32), i.e. for all τ and $h^\tau = ((C, C)^\tau)$ we have

$$\hat{s}_i^t(\hat{h}^t) = \begin{cases} C, & h^\tau = ((C, C)^\tau) \text{ and } \hat{h}^t = ((C, C)^t) \\ D, & \text{otherwise} \end{cases}$$
$$= \begin{cases} C, & \hat{h}^t = ((C, C)^t) \\ D, & \text{otherwise} = \bar{s}_i^t(h^t) \end{cases}.$$
$$(34)$$

Because $\bar{s}$ is a Nash-equilibrium strategy profile of the repeated game, for each subgame h^τ in class (A), the restriction $\hat{s}$ is a Nash-equilibrium strategy profile of the subgame when $\delta \geq \frac{1}{2}$.

For any subgame h^τ in class (B), $h^\tau \neq ((C, C)^\tau)$. Therefore the restriction $\hat{s}_i$ of $\bar{s}_i$ specifies $\hat{s}_i^t = D$ for all $t \in \{0, 1, \ldots\}$. In other words, in any subgame reached by some player having "defected" in the past, each player chooses the open-loop strategy "defect always". Therefore the repeated-game strategy profile $\hat{s}$ played in such a subgame is an open-loop sequence of stage-game Nash-equilibria. From Theorem 1 of [10] we know that this is a Nash-equilibrium of the repeated game and hence of this subgame. We have shown that for every subgame the restriction of $\bar{s}$ to that subgame is a Nash-equilibrium of that subgame for $\delta \geq \frac{1}{2}$. Therefore s is a subgame-perfect equilibrium of the infinitely repeated PD when $\delta \geq \frac{1}{2}$.

4.3. Complex strategies in the iterated prisoner's dilemma

In Section 4.2.2, we detailed the analysis of a particular strategy called `spiteful` that was

shown to be an equilibrium strategy (both a Nash-equilibrium for the whole repeated game and a subgame perfect equilibrium) for the prisoner's dilemma. Axelrod and Hamilton [19–21] used a computer tournament to numerically detect strategies that would favor cooperation among players engaged in the iterated PD. In a first round, 14 more or less sophisticated strategies and one totally random strategy competed against each other for the highest average scores in an iterated PD of 200 moves. Unexpectedly, a very simple strategy did outstandingly well:

> TIT-FOR-TAT
> Cooperate on the first period and then copy your opponent's last move for all subsequent periods

This strategy was called Tit-for-tat (TFT) and became the founder of an ever growing amount of successful strategies. To study the behavior of strategies from a numerical point of view, two kinds of computation can be done:

- The first one is a simple round robin tournament, in which each strategy meets all other strategies. Its final score is then the sum (not the discounted sum) of all scores done in each confrontation. At the end, the strategy's strength measurement is given by its range in the tournament.
- The second type of numerical analysis is a simulated ecological evolution, in which at the beginning there is a fixed population including the same quantity of each strategy. A round robin tournament is made and then the population of bad strategies is decreased whereas good strategies obtain new elements. The simulation is repeated until the population has been stabilized, i.e. the population does not change anymore.
A good strategy is then a strategy which stays alive in the population for the longest possible time, and in the biggest possible proportion. This kind of evaluation quotes the *robustness* of strategies.

Before the introduction of CORE as a strategy for the iterated PD, it is important to detail the computation method for ecological evolution, for example involving three strategies. Suppose that, initially, the population is composed of three strategies A, B, C. At generation n each strategy is represented by a certain number of players: $W_n(A)$ using A, $W_n(B)$ using B and $W_n(C)$ using C.

The payoff matrix of two-by-two meeting between A, B and C is computed and is thus known (see Table 2). $V(A|B)$ is the score of A when it meets B, etc. Let us suppose that the total size of the population is fixed and constant. Let is note it Π:

$$\forall i \in [1, \infty[, \quad \Pi = W_i(A) + W_i(B) + W_i(C). \quad (35)$$

The computation of the score (distributed points) of a player using a fixed strategy at generation n is then

$$
\begin{aligned}
g_n(A) &= W_n(A)V(A|A) + W_n(B)V(A|B) \\
&\quad + W_n(C)V(A|C) - V(A|A), \\
g_n(B) &= W_n(A)V(B|A) + W_n(B)V(B|B) \\
&\quad + W_n(C)V(B|C) - V(B|B), \\
g_n(C) &= W_n(A)V(C|A) + W_n(B)V(C|B) \\
&\quad + W_n(C)V(C|C) - V(C|C).
\end{aligned}
\quad (36)
$$

Note that because of the subtractions the computation of g cannot be simplified. The total points distributed to all involved strategies are

$$
\begin{aligned}
t(n) &= W_n(A)g_n(A) + W_n(B)g_n(B) \\
&\quad + W_n(C)g_n(C).
\end{aligned}
\quad (37)
$$

The size of each sub-population at generation $n + 1$ is finally

$$
\begin{aligned}
W_{n+1}(A) &= \frac{\Pi W_n(A)g_n(A)}{t(n)}, \\
W_{n+1}(B) &= \frac{\Pi W_n(B)g_n(B)}{t(n)}, \\
W_{n+1}(C) &= \frac{\Pi W_n(C)g_n(C)}{t(n)}.
\end{aligned}
\quad (38)
$$

All division being rounded to the nearest lower integer.

Classical results on the iterated PD, which have been emphasized by Axelrod in [38] show that to be good a strategy has to

- Not be the first to defect.
- Be reactive.
- Forgive.
- Be simple.

The TFT strategy which satisfies all those criteria, has, since Axelrod's book, been considered to be one of the *best* strategies not only for cooperation but also for evolution of cooperation.

4.4. CORE as a complex strategy for the iterated prisoner's dilemma

It is now important to define the scope of our analysis. After a brief introduction on the theory behind the study of the iterated PD game, we are focusing on the numerical analysis (through a simulation software [24]) of the features presented by some specific strategies that the players of the iterated PD should follow in order to promote cooperation. Furthermore, we want to compare some of the strategies available in the game theoretic literature and known to be the "best" strategies both from a cooperation point of view and from an evolutionary point of view with the strategy derived from the CORE cooperation enforcement mechanism. We suggest the reader to refer to [6] in order to grasp the details and the functioning of CORE.

Our claim is that the CORE strategy can be considered equivalent to the TFT strategy under certain circumstances (namely when the reputation buffer is of size 1). Furthermore, we will show through the evolutionary simulation outlined in Section 4.3, that the CORE strategy outperforms over all the other analyzed strategies when the assumption of "*perfect private monitoring*" is replaced by the "*imperfect private monitoring*" assumption.

The CORE strategy [12] can be defined as follows:

CORE

- Cooperate on the first move.
- In each period, observe the past B opponent's moves and build a vector $\vec{b} = (b_1, \ldots, b_k, \ldots, b_B)$ where each element equals $+1$ for a cooperation and -1 for a defection.
- Evaluate reputation as $\text{reputation} = \frac{1}{B}\sum_k b_k$.
- If reputation $\geqslant 0$ Cooperate else Defect.

We want to show now that the TFT strategy represents a particular case of the CORE strategy. Indeed, if we set $B = 1$ it means that only one observation over the opponent's past moves is taken into account to build the reputation information. This implies that if the opponent cooperated in the last move her reputation will be positive and the player will chose too cooperate. Vice versa, if the last opponent's move was a defection, the reputation would be negative and the response of the player would be to defect. This is exactly what the TFT strategy implies: cooperate on the first move and do what the opponent did in the previous move.

In this paper an analytical result stating that the CORE strategy is an equilibrium strategy will not be presented as the work in this direction is in progress: however we believe that the analysis will be facilitated thanks to the equivalence of the TFT and the CORE strategy.

In the following subsections we present some results obtained through evolutionary simulations using the iterated PD software available in [34]. The CORE strategy has been coded and added up to the list of available strategies in the software.

---

[12] The reader should be informed that in this paper we consider a limited version of the CORE mechanism in which reputation is evaluated through a simple average over the past observations made through the watchdog mechanism. A more faithful definition of the CORE strategy is reserved for our future work.
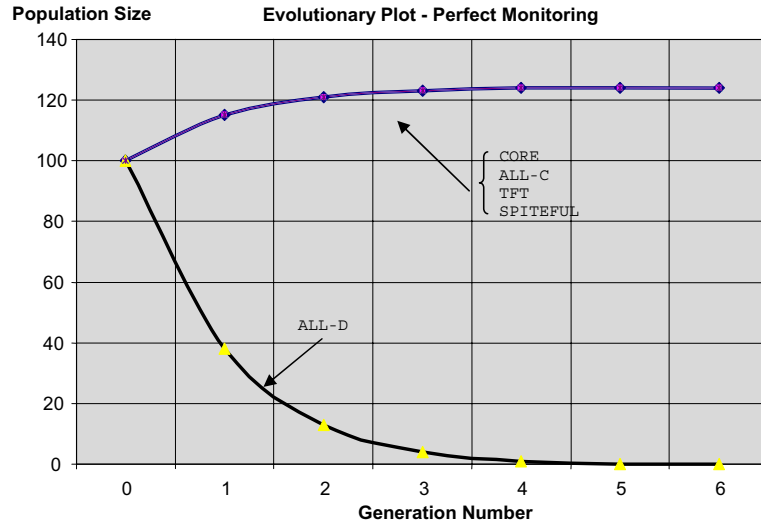
Fig. 1. Evolutionary simulation of complex strategies for the iterated PD with perfect monitoring.

4.4.1. Simulations with the "perfect private monitoring" assumption

We present here the results of the evolutionary simulation involving three strategies when the standard perfect monitoring assumption is made. As described in Section 4.3, suppose that, initially, the population is composed of five strategies `tit-for-tat`, `spiteful`, CORE, `all-C` (cooperate always) and `all-D` (defect always). Initially, each strategy is represented by a certain number of players: 100 players using each of the mentioned strategies. As it is possible to see in Fig. 1, after five generations the `all-D` strategy disappears: the three [13] winning strategies are equivalent for promoting cooperation and, more important, for the evolution of cooperation. This implies that the winning strategies obtained the same payoff in a two-by-two round robin tournament and can be considered equivalent from an evolutionary point of view.

4.4.2. Simulations with the "imperfect private monitoring" assumption

The majority of work in the iterated prisoner's dilemma has focused on games in a noise-free envi-

ronment, i.e. there is no danger of a signal being misinterpreted by the opponent or the message being damaged in transit. This assumption of a noise-free environment is not necessarily valid if one is trying to model real-world scenarios. As a specific example, when considering interactions between two nodes in a MANET where the behavior of a node follows the game theoretical model imposed by the prisoner's dilemma, it would be interesting to consider errors due to the watchdog mechanism. The interested reader should refer to [1] in order to understand intrinsic problems of the watchdog mechanism and the promiscuous mode operation of wireless cards. Specifically, the watchdog mechanism can be thought of as the private monitoring assumption in a two-players iterated prisoner dilemma: it is thanks to the watchdog mechanism (private monitoring) that a node (player) can infer in any period the behavior (opponent's past moves) of her neighbor and decide which actions needs to be taken (strategy).

There are different means that can be chosen to introduce noise to the simulation:

- mis-implementation (when the player makes a mistake implementing its choice),
- *misperception* (when one player misperceives the other player's signal or choice).

---

[13] Note that we are not considering the `all-C` as a winning strategy because of its history independent nature.

In this paper we will concentrate on misperception noise as we believe it significantly linked to the problems introduced by the watchdog mechanism.

Kahn and Murnighan [25] find that in experiments dealing with prisoner's dilemma in noisy environments, cooperation is more likely when players are sure of each other's payoffs. Miller's experiments in genetic algorithms applied to the prisoner's dilemma results in the conclusion that cooperation is at its greatest when there is no noise in the system and that this cooperation decreases as the noise increases [26]. Some ideas to promote cooperation in noisy environments have been posited by Axelrod; these include genetic kinship, clustering of like strategies, recognition, maintaining closeness when recognition capabilities are limited or absent (e.g limpets in nature), increasing the chance of future interactions (certain social organizations, hierarchies in companies etc.), changing the pay-offs, creating social norms where one learns cooperation. Hoffman [27] reports that results are sensitive to the extent to which players make mistakes either in the execution of their own strategy (mis-implementation noise) or in the perception of opponent choices (misperception noise).

In particular, cooperation is vulnerable to noise as it is supported by conditional strategies. For example, in a game between two TFTs, a single error would trigger a series of alternating defection. A number of authors confirm the negative effect of noise of TFT and find that more forgiveness promotes cooperation in noisy environments [28,29].

As described in Section 4.4.1, we executed an evolutionary simulation involving five strategies when misperception noise was taken into account: we decided to set the noise to the value of 10% and we took the average population size over five simulation runs. Hundred players for each of the following strategies competed in a round-robin tournament as described in Section 4.3: `tit-for-tat`, `spiteful`, `CORE`, `gradual` and `soft-majo`. [14] As it is possible to observe in Fig. 2, the `CORE` strategy outperforms and results

to be the most evolutionary stable and robust strategy among all the population (we believe though that exceptions especially constructed in which performances of `CORE` are not so outstanding are possible but are seldom and not easy to obtain).

The reason why `CORE` performs better than the other strategies when the imperfect monitoring assumption is made can be explained as follows: by adopting the `CORE` strategy, a node base her decision of whether to cooperate or not using a certain amount of observations made on the opponents past moves as defined by the B parameter. Thus, the reputation measure evaluated by the node takes into account more than one observation and is less sensible to any misperception noise. Furthermore, in its advanced version (which has not been implemented in the simulations, though), the `CORE` strategy also weights the past B observations giving more relevance to past observations than recent ones. It is intuitive to realize that a transient misbehavior is filtered out by the reputation mechanism that makes `CORE` more flexible and "forgiving" in presence of temporary misbehavior or a momentary high percentage of noise. A specific example of such a situation can be found when thinking of a communication between nodes of a MANET in presence of obstacles or high interference.

In Fig. 3, two populations of 100 members adopt respectively the `CORE` and the TFT strategy. The noise value has been set to 20% and 10 rounds of simulations have been executed in order to take average values of the evolution of population sizes. It is possible to observe that both strategies are evolutionary stable, in the long run; however, `CORE` is the winning strategy as the population size of players adopting it increases at each new generation, as defined in Section 4.3. Furthermore, Fig. 3 shows that the reputation buffer size (B) and both the stability condition and population size are directly related. As B increases, stability is reached at a lower generation number (i.e. earlier) and the population size of players adopting the `CORE` strategy grows faster. We believe however that these interesting results have to be evaluated in an analytical way: the fine-tuning of `CORE` parameters (such as B and the frequency

---

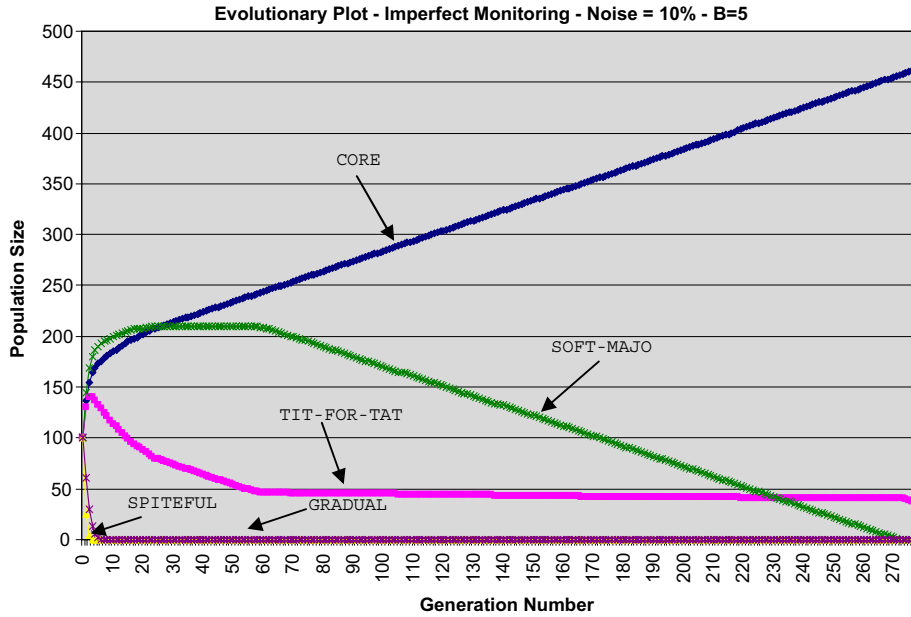[14] The `gradual` and `soft-majo` strategies are described in [23].

Fig. 2. Evolutionary simulation of complex strategies for the iterated PD with noise.
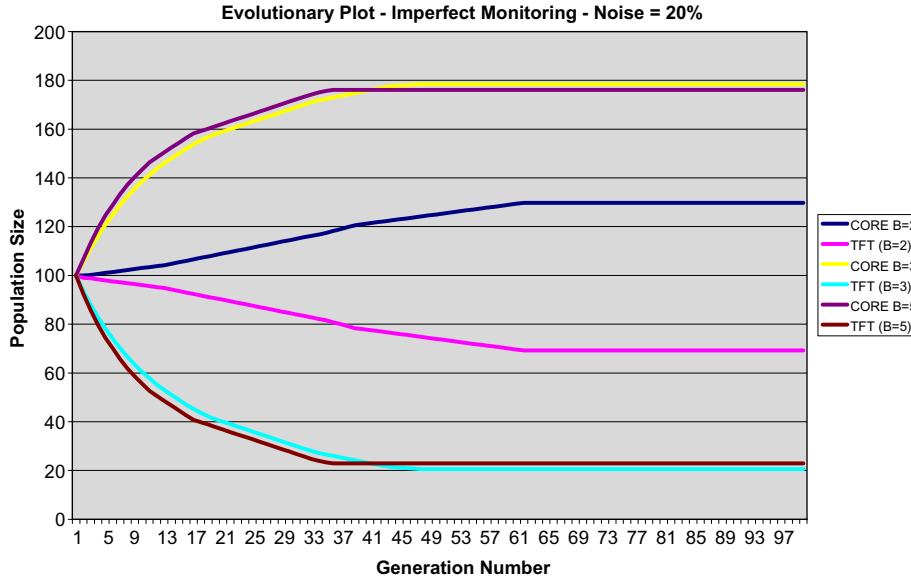

Fig. 3. Evolutionary simulation: CORE vs. TFT with different buffer sizes (*B*).

at which observations are made) would require a laborious empirical study if carried out only by means of evolutionary simulations. We plan to

analyze the CORE strategy in our future work taking as a starting point the analysis of the SPITEFUL strategy presented in Section 4.2.2.

5. Conclusions

In this paper we presented two different approaches based on game theory to assess the features of the CORE cooperation enforcement mechanism. Although the two methods described in this paper focus on CORE as a specific mechanism, some conclusions can be drawn from this analysis towards the design of cooperation enforcement mechanisms in general.

With the ''cooperative approach'' we introduced the concept of coalition of cooperating nodes as a subset of the nodes forming the network that exhibit a cooperative behavior. We then demonstrated that when cooperation is enforced through a mechanism like CORE the guaranteed size of a coalition of cooperating nodes is at least half of the size of the network. Furthermore, we were able to infer the necessary conditions for a node to join the coalition of cooperating nodes, as a function of the ERC-types. The main limitation of this approach derives from an intuitive representation of the CORE mechanism in the GT model. However, we believe that the ''cooperative games'' method is especially suitable to study the dynamics of coalition formation among a large number of nodes thus we plan to further investigate on a faithful representation of the CORE mechanism in this model.

On the other hand, the results provided by the ''non-cooperative'' approach better characterize CORE with respect to other mechanisms in a realistic setting. We were able to demonstrate the equivalence between the ''TFT'' strategy and CORE: precisely TFT can be thought of as a special case of the CORE strategy. Moreover, in order to represent a more realistic scenario for the execution of a (infinitely) repeated game we introduced a noise factor, which affects the players (nodes) perception of the opponents past moves. The ''imperfect private monitoring'' assumption allowed to model in a realistic way the watchdog mechanism used by CORE (and most of the available cooperation enforcement mechanisms) as it is known to be particularly unreliable. We showed through evolutionary simulations that the CORE strategy outperforms all other studied strategies in a noisy environment for its stability and robustness. As a

future research, we plan to extend the system model in order to take into account multiple players involved in the same decision making process in order to overcome the limitations due to a pair-wise interaction.

Appendix A. Proof of Proposition 2

We have to show that $\delta(k) > 0$ for $k > N/2$.

Remember that in (4) the denominator of $\delta(k)$ is positive due to Assumption 1. That is, $\delta(k) > 0$ if the numerator of (4) is positive. Remember also that we assumed $r(1/N - x) \leqslant r(1/N + x), \forall x \in [0, 1/N]$.

The numerator in (4) is positive if $r(cooperate) > r(defect)$. This is the case when Eq. (7) is satisfied.

Let us proceed by showing that $\delta(k) < 0$ for $k < N/2 - 1$.

It is possible to rewrite Eq. (7) as follows:

$$B(k+1)C(k)Nk + B(k)C(k+1)N(k+1-N) + C(k)C(k+1)[Nk - 2k(k+1)] < 0,$$

or

$$B(k+1)C(k)\frac{N}{k+1} + B(k)C(k+1)N\frac{k+1-N}{k(k+1)} + C(k)C(k+1)\left(\frac{N}{k+1} - 2\right) < 0, \qquad (A.1)$$

$$\left[B(k+1)C(k)\frac{N}{k+1} - B(k)C(k+1)\frac{N}{k}\right] + \left[\frac{NB(k)}{k} - C(k)\right]C(k+1)\left(2 - \frac{N}{k+1}\right) < 0. \qquad (A.2)$$

Eq. (A.2) can also be rewritten as

$$\left[\frac{B(k+1)}{B(k)} - \frac{(k+1)C(k+1)}{kC(k)}\right] + (k+1)C(k+1)$$
$$\times \left[\frac{1}{kC(k)} - \frac{1}{NB(k)}\right]\left(2 - \frac{N}{k+1}\right) < 0. \qquad (A.3)$$

Now, from the monotonicity and concavity of $B(\ )$ it follows that $B(k+1)/(k+1) < B(k)/k$.

Furthermore, the total cost of cooperation increases $kC(k)$ in k. Therefore,

$$\frac{B(k+1)}{B(k)} - \frac{(k+1)C(k+1)}{kC(k)} \leqslant \frac{k+1}{k} - 1 = \frac{1}{k}.$$

Since it has also been assumed that payoffs are non-negative, $B(k) \geqslant C(k)$. Thus

$$(k+1)C(k+1)\left[\frac{1}{kC(k)} - \frac{1}{NB(k)}\right]$$
$$\geqslant \frac{(k+1)C(k+1)}{kC(k)}\frac{N-k}{N} \geqslant \frac{N-k}{N}. \qquad (A.4)$$

We therefore obtain

$$\left[\frac{B(k+1)}{B(k)} - \frac{(k+1)C(k+1)}{kC(k)}\right]$$
$$+ (k+1)C(k+1)\left[\frac{1}{kC(k)} - \frac{1}{NB(k)}\right]\left(2 - \frac{N}{k+1}\right)$$
$$\leqslant \frac{1}{k} + \frac{N-k}{N}\left(2 - \frac{N}{k+1}\right)$$
$$= \frac{N(k+1) + 2(N-k)k(k+1) - (N-k)Nk}{Nk(k+1)}.$$

The numerator equals: $-2k^3 + (3N-2)k^2 - N(N-3)k + N$ which can be shown to be negative for $1 \leqslant k < N/2 - 1$, as long as $N > 8$.

Hence for $N > 8$ we have that the general conditions for a Nash-equilibrium of the ERC-PD game $\delta(k^* - 1) > 0$ are satisfied for $k > N/2$.

NOTE: the condition $N > 8$ can be removed if we assume that the total cost of cooperation increases more than the total benefits gained by defecting, i.e.: $(k+1)C(k+1)/kC(k) > NB(k+1)/NB(k)$.

Appendix B. Proof of Proposition 3

We have to show that in the cooperation game for ERC preferences, the Nash-equilibrium is given by solving the expression $B'(Nq^*) - C'(q^*) = 0$. We then show that the Nash-equilibrium point is symmetrical as long as at least one node draws utility from its absolute payoff ($\alpha_i > 0$).

Let us first study the two extreme cases, $\alpha_i = 0$ and $\beta_i = 0$, respectively:

- For $\beta_i = 0$, i.e. a player interested only in her absolute payoff, the first order condition (8 or 9) reduces to: $B'(Q) - C'(q_i) = 0$.
- For $\alpha_i = 0$, the node is solely interested in getting the equal payoff share. Thus, it would choose q_i to satisfy: $NC(q_i) = \sum_j C(q_j)$. Furthermore, when $\alpha_i = 0$, condition (8) also reduces to $B'(Q) - C'(q_i) = 0$.

  Indeed, β_i and $r'(\ )$ are positive by definition and the second summand reduces to

  $$\left[\frac{\sum_j y_j - y_i}{\sum_j y_j^2}B'(Q) - \frac{\sum_{i \neq j} y_j}{\sum_j y_j^2}C'(q_i)\right] = 0,$$

  which can be simplified as $B'(Q) - C'(q_i) = 0$.

For $\alpha_i, \beta_i \neq 0$ the chosen cooperation level is between the levels for those extreme cases: the first order condition must be satisfied for all nodes simultaneously. Since $r'(\sigma_i) = 0$ when $\sigma_i = 1/N$ by assumption, it follows that there is a *symmetric equilibrium* where all nodes choose the same cooperation level, i.e. $\sigma_i = 1/N$ for all types α_i/β_i, for $i = 1, \ldots, N$.

The resulting cooperation level q^* is given by solving the condition: $B'(Nq^*) - C'(q^*) = 0$.

Let us prove by contradiction (*reductio ad absurdum*) that there is an *asymmetric equilibrium*, i.e. some nodes receive less, and others more than the equal share.

In this case, on the one hand, $\sigma_i < 1/N$ implies that $r'(\sigma_i) > 0$, so from Eq. (8), we obtain

$$B'(Q) - C'(q_i) > 0. \qquad (8a)$$

On the other hand, for $\sigma_i > 1/N$ we have $r'(\sigma_i) < 0$, and therefore Eq. (9) implies

$$B'(Q) - C'(q_i) < 0. \qquad (9a)$$

Inequalities (8a) and (9a) imply that a node which gets more than the equal share has larger marginal cooperation costs ($C'(q_i)$) than nodes that receive

less, which contradicts the assumed payoff distribution.

Hence, only *symmetric equilibrium* exists. If $\alpha_i > 0$ for at least one node, we get $B'(Nq) - C'(q) = 0$ from Eq. (8).

References

[1] S. Marti, T. Giuli, K. Lai, M. Baker, Mitigating routing misbehavior in mobile ad hoc networks, in: Proceedings of MOBICOM, 2000.

[2] Vikram Srinivasan, Pavan Nuggehalli, Carla-Fabiana Chiasserini, Ramesh R. Rao, Cooperation in wireless ad hoc networks, in: Proceedings of IEEE Infocom 2003.

[3] A. Urpi, M.A. Bonuccelli, S. Giordano, Modeling cooperation in mobile ad hoc networks: a formal description of selfishness, in: Proceedings of WiOpt 2003.

[4] Eitan Altman, Arzad A. Kherani, Pietro Michiardi, Refik Molva, Non-cooperative forwarding in ad hoc networks, PIMRC 2004, also available as a technical report at INRIA Sophia Anitpolis, submitted for publication.

[5] P. Michiardi, R. Molva, Simulation-based analysis of security exposures in mobile ad hoc networks, European Wireless Conference, 2002.

[6] P. Michiardi, R. Molva, Core: A COllaborative REputation mechanism to enforce node cooperation in mobile ad hoc networks, IFIP Communication and Multimedia Security Conference 2002.

[7] P. Michiardi, R. Molva, A game theoretical approach to evaluate cooperation enforcement mechanisms in mobile ad hoc networks, IEEE/ACM Workshop WiOpt 2003, Sophia Antipolis, France.

[8] M.J. Osborne, A. Rubinstein, A Course in Game Theory, MIT Press, Cambridge, MA, 1997.

[9] D. Fudenberg, J. Tirole, Game Theory, MIT Press, Cambridge, MA, 1991.

[10] J. Ratliff, Game Theory Handouts, Available from <http://virtualperfection.com/gametheory>.

[11] G.E. Bolton, A. Ockenfels, ERC a theory of equity, reciprocity, and competition, The American Economic Review 90 (2000) 166–193.

[12] R.J. Aumann, J.H. Dreze, Cooperative games with coalition structure, International Journal of Game Theory 3 (1974) 217–237.

[13] A. Lange, C. Vogt, Cooperation in international environmental negotiations due to a preference for equity, Journal of Public Economics 87 (2003) 2049–2067.

[14] L.S. Shapley, Utility comparisons and the theory of games, in: T. Guilbau (Ed.), La Decision, Editions du CNRS, Paris, pp. 251–263. Reprinted in: A. Roth (Ed.), The Sahpley Value, Cambridge University Press, Cambridge, 1988, pp. 307–319.

[15] R. van den Brink, G. van der Laan, A class of consistent share functions for games in coalition structure, Tinbergen Institute, 2001.

[16] S. Barrett, International environmental agreements as games, in: R. Pethig (Ed.), Conflicts and Cooperation in Managing Environmental Resources, Springer, Berlin, 1992, pp. 11–37.

[17] C. Carraro, D. Siniscalco, Strategies for the international protection of the environment, Journal of Public Economics 52 (1993) 309–328.

[18] M. Hoel, International environmental conventions: the case of uniform reductions of emissions, Environmental and Resource Economics 2 (1992) 141–159.

[19] R. Axelrod, The Evolution of Cooperation, Basic Books, New York, 1984.

[20] R. Axelrod, The evolution of strategies in the iterated Prisoner's Dilemma, in: L. Davis (Ed.), Genetic Algorithms and the Simulated Annealing, Pitman, London, 1987, pp. 32–41 (Chapter 3).

[21] R. Axelrod, W.D. Hamilton, The Evolution of Cooperation, Science 211 (1981) 1390–1396.

[22] C. O'Riordan, Iterated Prisoner's Dilemma: A review, Technical Report NUIG-IT-260601, National University of Ireland, Galway.

[23] P. Matheiu, B. Beaufils, J.P. Delahaye, Iterated Prisoner's Dilemma Simulation Software, Available from <http://www.lifl.fr/IPD>.

[24] W. Poundstone, Prisoner's Dilemma: John von Neumann, Game Theory, and the Puzzle of the Bomb, Oxford University Press, Oxford, UK, 1993.

[25] L.M. Kahn, J.K. Murnighan, Conjecture, uncertainty, and cooperation in prisoner's dilemma games: Some experimental evidence, Journal of Economic Behaviour & Organisation 22 (1993) 91–117.

[26] J.H. Miller, The co-evolution of automata in the repeated prisoner's dilemma, Journal of Economic Behaviour & Organisation 29 (1996) 87–112.

[27] Robert Hoffman, Twenty years on: The evolution of cooperation revisited, Journal of Artificial Societies and Simulation 3 (2) (2000).

[28] U. Mueller, Optimal retaliation for optimal cooperation, Journal of Conflict Resolution 31 (4) (1988) 692–724.

[29] J. Bendor, R.M. Kramer, S. Stout, When in doubt cooperation in a noisy prisoner's dilemma, Journal of Conflict Resolution 35 (4) (1991) 691–719.

Pietro Michiardi received the Laurea in electronic engineering from the Politecnico di Torino in 2001. He was granted a scholarship by the European Union to take part in a program in advanced telecommunications engineering at the Eurecom Institute where he got a diploma in Multimedia Communications. In January 2000 Pietro joined the Eurecom Institute as a research engineer working on a project for the development of advanced security services for business transactions. Since September 2001 Pietro has been a Ph.D. student at the Eurecom Institute working on routing security

and cooperation enforcement for Mobile Ad Hoc Networks. Pietro Michiardi contributed actively to the definition of new types of security requirements for the ad hoc network paradigm and proposed original security mechanisms that were analyzed using economic principles. The use of game theory to model cooperation in ad hoc networks and to study cooperation enforcement mechanisms was awarded in the IEEE/ACM WiOpt 2003 international workshop on "Modeling and Optimization for Wireless Networks".

Refik Molva is a professor at Institut Eurécom since 1992. He is leading the network security research group that currently focuses on multipoint security protocols, multi-component system security, and security in ad hoc networks. His past projects at Eurécom were on mobile code protection, mobile network security, anonymity and intrusion detection. Beside security, he worked on distributed multimedia applications and was responsible for the BETEUS European project on CSCW over a trans-European ATM network. Prior to joining Eurécom, he worked for five years as a Research Staff Member in the Zurich Research Laboratory of IBM where he was one of the key designers of the Krypto-Knight security system. He also worked as a network security consultant in the IBM Consulting Group in 1997. He is the author of several publications and patents in the area of network security and has been part of several evaluation committees for various national and international bodies including the European Commission.