

Efficient packet transmission in wireless ad hoc networks with partially informed nodes

Sara Berri, Samson Lasaulce, Mohammed Said Radjef

▶ **To cite this version:**

Sara Berri, Samson Lasaulce, Mohammed Said Radjef. Efficient packet transmission in wireless ad hoc networks with partially informed nodes. EURASIP Journal on Wireless Communications and Networking, 2019, 2019, 10.1186/s13638-019-1422-4 . hal-02411646

HAL Id: hal-02411646
https://hal.science/hal-02411646

Submitted on 15 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Berri *et al. EURASIP Journal on Wireless Communications and Networking* (2019) 2019:148
https://doi.org/10.1186/s13638-019-1422-4

EURASIP Journal on Wireless Communications and Networking

RESEARCH Open Access

Efficient packet transmission in wireless ad hoc networks with partially informed nodes

Sara Berri[1,2*], Samson Lasaulce[2] and Mohammed Said Radjef[1]

Abstract

One formal way of studying cooperation and incentive mechanisms in wireless ad hoc networks is to use game theory. In this respect, simple interaction models such as the forwarder's dilemma have been proposed and used successfully. However, this type of models is not suited to account for possible fluctuations of the wireless links of the network. Additionally, it does not allow one to study the way a node transmits its own packets. At last, the repeated game models used in the related literature do not allow the important scenario of nodes with partial information (about the link state and nodes actions) to be studied. One of the contributions of the present work is precisely to provide a general approach to integrate all of these aspects. Second, the best performance the nodes can achieve under partial information is fully characterized for a general form of utilities. Third, we derive an equilibrium transmission strategy which allows a node to adapt its transmit power levels and packet forwarding rate to link fluctuations and other node actions. The derived results are illustrated through a detailed numerical analysis for a network model built from a generalized version of the forwarder's dilemma. The analysis shows in particular that the proposed strategy is able to operate in the presence of channel fluctuations and to perform significantly better than the existing transmission mechanisms (e.g., in terms of consumed network energy).

Keywords: Packet transmission, Power control, Game theory, Repeated games, Incentive mechanisms, Wireless ad hoc networks

1 Introduction

In wireless ad hoc networks, nodes are interdependent. One node needs the assistance of neighboring nodes to relay the packets or messages it wants to send to the receiver(s). Therefore, nodes are in the situation where they have to relay packets, but have at the same time to manage the energy they spend for helping other nodes, and therefore exhibiting selfish behavior. To stimulate cooperation, incentive mechanisms have to be implemented [1–7]. The vast majority of incentive mechanisms either rely on the idea of reputation [4, 5, 8] or the use of a credit system [9, 10]. Indeed, to capture the trade-off between a cooperative behavior (which is necessary to convey information through an ad hoc network) and a selfish behavior (which is necessary to manage the node energy), the authors of [5] and [8], and many other papers, exploited a simple but efficient model, which consists in assuming, whatever the size of the network, that the local node interaction only involves two neighboring nodes having a decision-making role; one of the virtues of considering the interaction to be local is the possibility of designing distributed transmission strategies. In the original model, a node has two possible choices, namely, forward or drop the packets it receives from the neighboring node. As shown in [5] and [8], modeling the problem at hand as a game appears to be natural and relevant; in the corresponding game, the node utility function consists of the addition of a data rate term (which is maximized when the other node forwards its packets) and an energy term (which is maximized when the node does not forward the packets of the other node). At the Nash equilibrium of the strategic form static game (called the forwarder's dilemma in the corresponding literature), nodes do not transmit at

*Correspondence: berri.sara2012@gmail.com
[1]Research Unit LaMOS (Modeling and Optimization of Systems), Faculty of Exact Sciences, University of Bejaia, 06000 Bejaia , Algeria
[2]L2S (CNRS-CentraleSupelec-Univ. Paris-Saclay), 91192 Gif-sur-Yvette, France

Springer Open

© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

all. To avoid this situation to happen in a real network, cooperation has to be stimulated by studying the repeated interaction between the nodes [1, 5, 8]. While providing an efficient solution, all the corresponding models still have some limitations, especially regarding the link quality fluctuations and partial knowledge at the nodes; indeed, they do not take into account the quality of the link between the transmitting and the receiving nodes, which may be an important issue since the link quality may strongly fluctuate if it is wireless. The solution in [3, 6], and [8] referred to as ICARUS (hybrId inCentive mechAnism for coopeRation stimUlation in ad hoc networkS) combines the two ideas, namely, reputation and credit system, but it is not suited to scenarios where the actions of the other nodes are not perfectly observed, which results, e.g., in inappropriate punishment (a node is declared selfish while it is cooperative) and therefore in a loss of efficiency. Additionally, in these works, when a node is out of credit, the transmission is blocked and the node cannot send any packet anymore; this might be not practical in some wireless networks where a certain quality of service has to be provided. Also, the authors propose a mechanism to regulate the credit when a node has an excessive number of credits, but the proposed mechanism may be too complex. At last but not least, no result is provided on the strategic stability property, which is important and even necessary to make the network robust against selfish deviations. The purpose of this paper is precisely to overcome the limitations of the aforementioned previous works. More precisely, the contributions of the present paper are as follows.

► The first key technical difference with the closely related works is that the proposed formulation accounts for the possible presence of quality fluctuations of the different links that are involved in the considered local interaction model. In particular, this leads us to a game model which generalizes the existing models since the forwarding game has now a state and the discrete action sets are arbitrary, not just binary; additionally, the node does not only choose the cooperation power but also the power used to send its own packets.

► An important and useful contribution of the paper is to characterize the feasible utility region of the considered problem, by exploiting implementability theorems provided by recent works [11–13]. This problem is known to be non-trivial in the presence of partial information and constitutes a determining element of folk theorems; this difficult problem turns out to be solvable in the proposed reasonable setting (the channel gains are i.i.d. and the observation structure is memoryless). The knowledge of the utility region is very useful since it allows one to measure the efficiency of any distributed algorithm relying on the assumed partial information.

► A third contribution of the paper is that we provide a new transmission strategy whose main features is to be able to deal with the presence of fluctuating link qualities and to be efficient. To design the proposed strategy, we show that the derived utility region can be used in a constructive manner to obtain efficient operating points and propose a new incentive mechanism to ensure that these points are equilibrium points. The proposed incentive mechanism combines the ideas of credit and reputation. To our knowledge, the closest existing incentive mechanism to the one proposed in the present paper is given by ICARUS in [3, 6], and [8]. Here, we go further by dealing with the problem of imperfect observation and that of credit outage or excess. Indeed, the credit evolution law we propose in this paper prevents, by construction, the number of credits from being too large; therefore, one does not need to resort to an additional credit regulation mechanism, which may be too complex.

► In addition to the above analytical contributions, we provide a numerical study which demonstrates the relevance of the proposed approach. Compared to the closest transmission strategies, significant gains are obtained both in terms of packet forwarding rate, network consumed power, and combined utilities. As a sample result, the network-consumed power is shown to be divided by more than two w.r.t. state-of-the art strategies [1, 3, 5, 8].

The remainder of the paper is organized as follows. In Section 2.1, we present the system model; the assumed local interaction model involves two neighboring nodes of an ad hoc network with arbitrary size and generalizes the model introduced in previous works [5, 8]. The associated static game model is also provided in Section 2.1. In Section 2.2, the repeated game formulation of the generalized packet forwarding problem is provided; one salient feature of the proposed model is that partial information is assumed both for the network state and the node actions. In Section 2.3, the feasible utility region of the studied repeated game with partial observation is fully characterized. We also provide an algorithm to determine power control policies that are shown to be globally efficient in Section 3.1. The proposed incentive mechanism and equilibrium transmission strategy are provided in Section 2.4; the proposed transmission strategy allows both the packet forwarding rate and the transmit power to be adapted. A detailed numerical performance analysis is conducted in Section 3.1. Section 4 concludes the paper.

2 Methods/experimental

2.1 System model

The present work concerns wireless ad hoc networks, namely, networks in which a source node needs the assistance of other nodes to communicate with the destination node(s). As well motivated in related papers such as [5]

and [8], we will assume the interaction among nodes to be local, i.e., it only involves neighboring nodes. This means that the network can have an arbitrary size and topology, but a node only considers local interactions to take its decision although it effectively interacts with more nodes. One of the virtues of such an interaction model is to be able *to design distributed transmission strategies* for every node. More specifically, we will assume the model in which local interactions take place in a pairwise manner, which not only allow us to design distributed strategies but also to easily compare the proposed transmission strategy with existing strategies. The key idea of this relevant model is to take advantage of the fact that the network is wireless to simplify the interaction model. For a given node, the dominant interaction will only involve its closest neighbors (see Fig. 1). If several neighboring nodes lie within the radio range of the considered node, then it is assumed to have several pairwise interactions in parallel, as explained in detail in the numerical part.

The nodes are assumed to be non-malicious, i.e., each of them does not aim at damaging the communication of the other. Additionally, they are assumed to operate in an imperfect promiscuous mode, which means that each node imperfectly overhears all packets forwarded by their neighbors. The proposed model generalizes the previous models for at least four reasons. First, the action of a node has two components instead of one: the transmit power used to help the other node, which is denoted by p'_i, and the transmit power used to send its own packets, which is denoted by p_i. Second, the transmit power levels are not assumed to be binary but to lie in a general discrete set[1] $\mathcal{P}_i = \mathcal{P}'_i = \mathcal{P} = \{P_1, P_2, ..., P_L\} = \{P_{\min}, \ldots, P_{\max}\}$, $|\mathcal{P}_i| = |\mathcal{P}'_i| = |\mathcal{P}| = L$. Assuming that the sets are discrete is of practical interest, as there exist wireless communication standards in which the power can only be decreased or increased by step and in which quantized wireless channel state information (CSI) is used (see, e.g., [14, 15]). Similarly, the channel may be quantized to define operating modes (e.g., modulation coding scheme (MCS)) used by the transmitter. Even when the effective channel is continuous, assuming it to be discrete in the model and algorithm part may be very relevant. At last, note that from the limiting performance characterization point of view, the analysis of the continuous case follows from the discrete case but the converse is not true [16]. As a third new feature compared to the related works, the considered model accounts for the possible fluctuations of the quality of each link. With each link, a non-negative scalar is associated, which is called the *channel gain* of the considered link. For a node, the channel gains of the links used to send its own packets and to help the other node are denoted by h_i and h'_i, respectively. These channel gains are assumed to lie in discrete sets (of states): $\mathcal{H}_i = \mathcal{H}'_i = \mathcal{H} = \{h_{\min}, \ldots, h_{\max}\}$ with $|\mathcal{H}_i| = |\mathcal{H}'_i| = |\mathcal{H}| = H$; the realizations of each channel gain will be assumed to be i.i.d.. Technically, continuous channel gains might be assumed. But, as done in the information theory literature for establishing coding theorems, we address the discrete case in the first place, since the continuous case can be obtained by classical arguments (such as assuming standard probability spaces), whereas the converse is not true. Now, from the practical aspect, quantizing the channel gains typically induces a small performance loss compared to the continuous case; one figure assuming a typical scenario illustrates this. The corresponding channel gain model naturally applies to time-selective frequency flat fading single-input single-output channels. If the channel gain is interpreted as the combined effect of path loss and shadowing, our model can also be used to study more general channel models such as multiple-input multiple-output channels. Fourth, the utility function of a node has a more general form than in the forwarder's dilemma. The *instantaneous utility function* for node $i \in \{1, 2\}$ expresses as follows:

$$u_i(a_0, a_1, a_2) = \varphi(\mathrm{SNR}_i) - \alpha(p_i + p'_i), \tag{1}$$

where

- $a_0 = (h_1, h'_1, h_2, h'_2)$ is the *global channel or network state*. The corresponding set will be denoted by $\mathcal{A}_0 = \mathcal{H}_1 \times \mathcal{H}'_1 \times \mathcal{H}_2 \times \mathcal{H}'_2 = \mathcal{H}^4$;
- $a_i = (p_i, p'_i)$ is the *action* of node $i \in \{1, 2\}$;

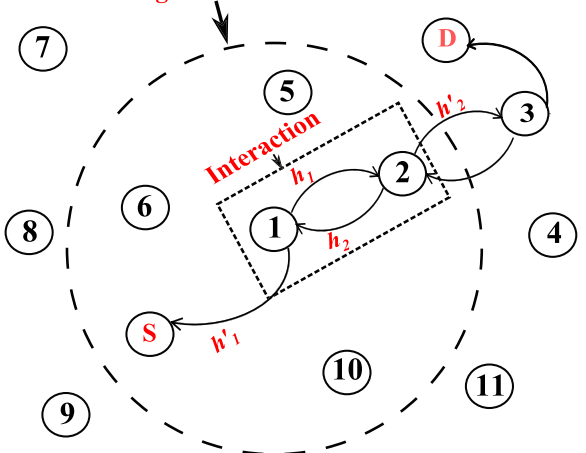


Fig. 1 In this example, the focus is on what node 1 does to allow node S (source) to route its packets to node D (destination). The dashed circle represents the radio range for node 1 and defines its neighbors. To ensure a distributed design, two key elements are exploited: (a) node 1 adopts its transmission behavior to each of its neighbors. Here, node 1 interacts with node 2 (indicated by the dotted box); (b) only the available knowledge of the quality of the most influential links is accounted for (denoted generically by h_1, h'_1, h_2, h'_2)

- The function φ is a communication efficiency function which represents the *packet success rate*. It is assumed to be increasing and lie in $[0,1]$. A typical choice for φ is, for example, $\varphi(x) = \left(1-e^{-x}\right)^{\ell}$, ℓ being the number of symbols per packet (see, e.g., [17–19]) or $\varphi(x) = e^{-\frac{c}{x}}$ with $c = 2^r - 1$, r being the spectral efficiency in bit/s/Hz [20];
- For $i \in \{1,2\}$, the quantity SNR_i is, for node i, the equivalent signal-to-noise ratio (SNR) at the next node after the neighbor. It is assumed to express as follows:

$$\mathrm{SNR}_i = \frac{p_i h_i p'_{-i} h'_{-i}}{\sigma^2}, \tag{2}$$

σ^2 being the noise variance and the *index notation* $-i$ standing for the index of the other node.

Remark. The results derived in Section 2.3 hold for any utility function under the form $u_i(a_0,a_1,a_2)$ (under some assumptions which only concern the observation structure) and not only for the specific choice made above. This choice is made to allow comparisons with existing results (and more specifically with the large set of contributions on the forwarder's dilemma) to be conducted and discussed. □

Remark. The assumed expression of the SNR is also one possible pragmatic choice, but all the analytical results derived in this paper hold for an arbitrary SNR expression of the form $\mathrm{SNR}_i(a_0,a_1,a_2)$; this choice is sufficiently general to study the problem of channel fluctuations which is the main feature to be accounted for. The proposed expression is relevant, e.g., when nodes implement the amplify-and-forward protocol to relay the signals or packets [21]. This simple but reasonable model for the SNR may either be seen as an approximation where the single-hop links dominate the multi-hop links or the talk/listen phases are scheduled appropriately. If another relaying protocol is implemented such as decode-and-forward, other expressions for the equivalent SNR may be used (see, e.g., [22]) without questioning the validity of the analytical results provided in this paper. At last, the parameter $\alpha \geq 0$ in (1) allows one to assign more or less importance to the energy consumption of the node. Indeed, the first term of the utility function represents the benefit of transmitting (i.e., the goodput) while the second term represents the cost of transmitting (i.e., the spent energy). □

The pair of functions (u_1,u_2) defines a strategic-form *static game* (see, e.g., [23]) in which the *players* are nodes 1 and 2 and the *action sets* are respectively $\mathcal{A}_1 = \mathcal{P}^2$ and $\mathcal{A}_2 = \mathcal{P}^2$. This game generalizes the forwarder's dilemma. The latter can be retrieved by assuming that φ is a step function, p'_i is binary, p_i is constant, and all the channel gains are constant. In the next section, we describe mathematically the problem under investigation. It is shown how the problem can be modeled by a repeated game, which is precisely built on the stage or static game:

$$\mathcal{G} = (\mathcal{N}, \{\mathcal{A}_i\}_{i\in\mathcal{N}}, \{u_i\}_{i\in\mathcal{N}}), \tag{3}$$

where $\mathcal{N} = \{1,2\}$.

The unique Nash equilibrium of $\mathcal{G}$ is $p_{i,\mathrm{NE}} = P_{\min}$ and $p_{i',\mathrm{NE}} = P_{\min}$. If the minimum power $P_{\min}$ is taken to be zero, then the situation where the nodes do not transmit at all corresponds to the equilibrium (and thus $(u_1,u_2) = (0,0)$), which clearly shows one of the interests in modeling the packet transmission problem as a repeated game.

2.2 Repeated game formulation of the problem

The problem we want to solve in this paper is as follows. It is assumed that the nodes interact over an infinite number of stages. Over stage $t \in \{1,2,...,T\}$, $T \to \infty$, the channel gains are assumed to be fixed while the realizations of each channel gain are assumed to be i.i.d. from stage to stage. During a stage, a node typically exchanges many packets with its neighbors. At each stage, a node has to make a decision based on the knowledge it has. In full generality, the decision of a node consists in choosing a probability distribution over its set of possible actions. The knowledge of a node is in terms of global channel states and actions chosen by the other node. More precisely, it is assumed that node $i \in \mathcal{N}$ has access to a signal which is associated with the state a_0 and is denoted by $s_i \in \mathcal{S}_i$, $|\mathcal{S}_i| < \infty$. At stage t, the observation $s_i(t) \in \mathcal{S}_i$ therefore corresponds to the image (i.e., the knowledge) that node i has about the global channel state $a_0(t)$. This signal is assumed to be the output of a memoryless observation structure [21][2] whose conditional probability is denoted by $\digamma_i$:

$$\digamma_i(s_i|a_0) = \Pr[S_i = s_i | A_0 = a_0], \tag{4}$$

where capital letters stand for random variables, whereas small letters stand for realizations. Simple examples for s_i are $s_i = h_i$, $s_i = \widehat{h}_i$, $\widehat{h}_i$ being an estimate of h_i, and $s_i = (h_i, h'_i)$; $s_i = a_0 = (h_1, h'_1, h_2, h'_2)$. Now, in terms of observed actions, it is assumed that node $i \in \mathcal{N}$ has imperfect monitoring. In general, node $i \in \mathcal{N}$ has access to a signal $y_i \in \mathcal{Y}_i$, $|\mathcal{Y}_i| < \infty$, which is assumed to be the output of a memoryless observation structure whose conditional probability is denoted by Γ_i:

$$\Gamma_i(y_i|a_0,a_1,a_2) = \Pr[Y_i = y_i | (A_0,A_1,A_2) = (a_0,a_1,a_2)]. \tag{5}$$

The reason why we distinguish between the observations s_i and y_i comes from the assumptions made in terms of causality. Indeed, practically speaking, it is relevant to

assume that a node has access to the past realizations of s_i in the wide sense, namely, to $s_i(1), ..., s_i(t)$ at stage t. However, only the past realizations in the strict sense $y_i(1), ..., y_i(t-1)$ are assumed to be known at stage t. Otherwise, it would mean that a node would have access to the image of its current action and that of the others before choosing the former.

At this point, it is possible to define completely the problem to be solved. The problem can be tackled by using a strategic-form game model, which is denoted by $\overline{\mathcal{G}}$. As for the static game $\mathcal{G}$ on which the repeated game model $\overline{\mathcal{G}}$ is built on, the *set of players* is the set of nodes $\mathcal{N} = \{1, 2\}$. The *transmission strategy* of the node i is denoted by σ_i and consists of a sequence of functions and is defined as follows:

$$\begin{array}{rccl} \sigma_{i,t} : & \mathcal{S}_i^t \times \mathcal{Y}_i^{t-1} & \rightarrow & \Delta\left(\mathcal{P}^2\right) \\ & \left(s_i^t, y_i^{t-1}\right) & \mapsto & \pi_i(t), \end{array} \tag{6}$$

where

- $s_i^t = (s_i(1), \ldots, s_i(t))$, $y_i^{t-1} = (y_i(1), \ldots, y_i(t-1))$.
- $\Delta\left(\mathcal{P}^2\right)$ represents the unit simplex, namely, the set of probability distributions over the set $\mathcal{P}^2$.
- $\pi_i(t)$ is the probability distribution used by the node i at stage t to generate its action $(p_i(t), p_i'(t))$.

The type of strategies we are considering is referred to as a behavior strategy in the game theory literature, which means that at every game stage, the strategy returns a probability distribution. The associated randomness not only allows one to consider strategies which are more general than pure strategies, but also to model effects such as node asynchronicity for packet transmissions. At last, the performance of a node is measured over the long run, and nodes are therefore assumed to implement transmission strategies which aim at maximizing their long-term utilities. The *long-term utility function* of node $i \in \mathcal{N}$ is defined as:

$$\begin{aligned} U_i(\sigma_1, \sigma_2) &= \lim_{T\rightarrow\infty} \sum_{t=1}^{T} \theta_t \mathbb{E}\left[u_i(A_0(t), A_1(t), A_2(t))\right] \\ &= \lim_{T\rightarrow\infty} \sum_{t=1}^{T} \theta_t \sum_{a_0,a_1,a_2} P_t(a_0, a_1, a_2) u_i(a_0, a_1, a_2), \end{aligned} \tag{7}$$

where

- σ_i stands for the transmission strategy of node $i \in \mathcal{N}$.
- It is assumed that the limit in (7) exists.
- θ_t is a sequence of weights which corresponds to a convex combination, that is $0 \leq \theta_t < 1$ and $\sum_{t=1}^{T} \theta_t = 1$. For a repeated game with discount $\theta_t = (1-\delta)\delta^t$ and for a classical infinitely repeated game $\theta_t = \frac{1}{T}$.
- As already mentioned, capital letters stand for random variables, whereas, small letters stand for realizations. Here, $A_0(t)$, $A_1(t)$, and $A_2(t)$ stand for the random processes corresponding to the network state and the node actions.
- The notation P_t stands for the joint probability distribution induced by the strategy profile (σ_1, σ_2) at stage t.

This general model thus encompasses the two well-known models for the sequence of weights which are given by the model with discount and the infinite Cesaro mean. In the model with discount, note that the discount factor may model different phenomena, but in a wireless ad hoc network, the most relevant effect to be modeled seems to be the uncertainty that there will be a subsequent iteration of the stage game, for example, connectivity to an access point can be lost. With this interpretation in mind, the discounting factor represents, for example, the probability that the current round is not the last one , or in terms of mobility, it may also represent the probability that the nodes do not move for the current stage. Therefore, it may model the departure or the death of a node (e.g., due to connectivity loss) for a given routing path. More details about this interpretation can be found in [23] while [24] provides a convincing technical analysis to sustain this probabilistic interpretation.

At this point, we have completely defined the strategic form of the *repeated game* that is the triplet

$$\overline{\mathcal{G}} = \left(\mathcal{N}, \{\Sigma_i\}_{i\in\mathcal{N}}, \{U_i\}_{i\in\mathcal{N}}\right), \tag{8}$$

where Σ_i is the set of all possible transmission strategies for node $i \in \mathcal{N}$.

One of the main objectives of this paper is to exploit the above formulation to find a globally efficient transmission scheme for the nodes. For this purpose, we will characterize long-term utility region for the problem under consideration. It is important to mention that the characterization of the feasible utility region of a dynamic game (which includes repeated games as a special case) with an arbitrary observation structure is still an open problem [25]. Remarkably, as shown recently in [11] and [12], the problem can be solved for an interesting class of problems. It turns out that the problem under investigation belongs to this class provided that the channel gains evolve according to the classical model of block i.i.d. realizations.

In the next section, we show how to exploit [11, 12] to characterize the long-term utility region and construct a practical transmission strategy. In Section 2.4, we will show how to integrate the strategic stability[3] property into this strategy, this property being important to ensure that selfish nodes effectively implement the efficient strategies.

2.3 Long-term utility region characterization

When the number of stages is assumed to be large, the random process associated with the *network state*

$A_0(1), A_0(2), ..., A_0(T)$ is i.i.d, and the observation structure given by $(\urcorner_1, \urcorner_2, \Gamma_1, \Gamma_2)$ is memoryless, some recent results can be exploited to characterize the feasible utility region of the considered repeated game and to derive efficient transmission strategies. The main difficulty to determine the feasible utility region of $\overline{\mathcal{G}}$ is to find the set of possible average correlations between a_0, a_1, and a_2. Formally, the correlation averaged over T stages is defined by:

$$P^T(a_0, a_1, a_2) = \frac{1}{T}\sum_{t=1}^{T} P_t(a_0, a_1, a_2), \tag{9}$$

where P_t is the joint probability at stage t. More precisely, a key notion to characterize the attainable long-term utilities is the notion of implementability, which is given as follows.

Definition 1 *An average correlation Q is said to be implementable if there exists a pair of transmission strategies (σ_1, σ_2) such that the average correlation induced by these transmission strategies verifies:*

$$\begin{aligned} &\forall (a_0, a_1, a_2) \in \mathcal{A}_0 \times \mathcal{A}_1 \times \mathcal{A}_2, \\ &\lim_{T\to\infty} \tfrac{1}{T} \textstyle\sum_{t=1}^{T} P_t(a_0, a_1, a_2) = Q(a_0, a_1, a_2). \end{aligned} \tag{10}$$

Using the above definition, the following key result can be proved.

Proposition 1 *The Pareto frontier of the achievable utility region of $\overline{\mathcal{G}}$ is given by all the points under the form $\left(\mathbb{E}_{Q_\lambda}(u_1), \mathbb{E}_{Q_\lambda}(u_2)\right)$, $\lambda \in [0, 1]$, where Q_λ is a maximum point of*

$$W_\lambda = \lambda \mathbb{E}_Q(u_1) + (1-\lambda)\mathbb{E}_Q(u_2), \tag{11}$$

and each maximum point is taken in the set of probability distributions which factorize as follows:

$$\begin{aligned} Q(a_0, a_1, a_2) = \sum_{v, s_1, s_2} &\rho(a_0) P_V(v) \times \urcorner(s_1, s_2|a_0) \\ &\times P_{A_1|S_1,V}(a_1|s_1, v) P_{A_2|S_2,V}(a_2|s_2, v), \end{aligned} \tag{12}$$

where

- *λ denotes the relative weight assigned to the utility of the first player and can be chosen arbitrarily depending on some prescribed choice, e.g., in terms of fairness or global efficiency.*
- *ρ is the probability distribution of the network state a_0.*
- *$\urcorner$ is the joint conditional probability which defines the assumed observation structure, i.e., a probability which is written as:*[4]

$$\urcorner(s_1, s_2|a_0) = Pr[(S_1, S_2) = (s_1, s_2)|A_0 = a_0]. \tag{13}$$

- *$V \in \mathcal{V}$ is an auxiliary random variable or lottery.*

(See the proof in the Appendix). One interesting comment to be made concerns the presence of the "parameter" or auxiliary variable V. The presence of the auxiliary variable is quite common in information-theoretic performance analyses and in game-theoretic analyses through the notion of external correlation devices (such as those assumed to implement correlated equilibria). Indeed, $V \in \mathcal{V}$ is an auxiliary random variable or lottery which can be proved to improve the performance in general (see [11] for more details). Such a lottery may be implemented by sampling a signal which is available to all the transmitters, e.g., an FM or a GPS signal.

In (22), ρ and $\urcorner$ are given. Thus, W_λ has to be maximized with respect to the triplet $(P_{A_1|S_1,V}, P_{A_2|S_2,V}, P_V)$. In this paper, we restrict our attention to the optimization of $(P_{A_1|S_1,V}, P_{A_2|S_2})$ for a fixed lottery P_V and leave the general case as an extension.

The maximization problem of the functional $W_\lambda(P_{A_1|S_1,V}, P_{A_2|S_2,V})$ with respect to $P_{A_1|S_1,V}$ and $P_{A_2|S_2,V}$ amounts to solving a bilinear program. The corresponding bilinear program can be tackled numerically by using iterative techniques such as the one proposed in [26], but global convergence is not guaranteed, and therefore, some optimality loss may be observed. Two other relevant numerical techniques have also been proposed in [27]. The first technique is based on a cutting plane approach while the second one consists of an enlarging polytope approach. For both techniques, convergence may also be an issue since for the first technique, no convergence result is provided and for the second technique, cycles may appear [28]. To guarantee convergence and manage the computational complexity issue, we propose here another numerical iterative technique, namely, to exploit the sequential best-response dynamics (see, e.g., [29] for a reference in the game theory literature, [23] for application examples in the wireless area, [30] for a specific application to power control over interference channels). Here also, some efficiency loss may be observed, but it will be shown to be relatively small for the quite large set of scenarios we have considered in the numerical performance analysis. The sequential best-response dynamics applied to the considered problem translates into the following algorithm.

Although suboptimal in general (as the available state-of-the art techniques), the proposed technique is of particular interest for at last three reasons. First, convergence is unconditional. It can be proved to be guaranteed, e.g., by induction or by identifying the proposed procedure as an instance of the sequential best-response dynamics for an exact potential game (any game with a common utility is an exact potential game). Second, convergence points are local maximum points, but in all the

simulations performed, those maximums had the virtue of not being too far from the global maximum. At last but not least, it allows us to build a practical transmission strategy which outperforms all the state-of-the art transmission strategies, as explained next. Note that this is necessarily the case when Algorithm 1 is initialized with the state-of-the art transmission strategy under consideration. Although we will not tackle the classical issue of the influence of initialization on the convergence point, it is worth mentioning that many simulations have shown that the impact of the initial point on the performance at convergence is typically small, at least for the utilities under consideration. Therefore, initializing Algorithm 1 with naive strategies such as transmitting at full power $\forall t \in \{1, ..., T\},\ (a_1(t), a_2(t)) = (P_{\max}, P_{\max}, P_{\max}, P_{\max})$ is well suited.

Remark. Algorithm 1 would typically be implemented offline in practice. The purpose of Algorithm 1 is to generate decision functions which are exploited by the proposed transmission strategy. To implement Algorithm 1, only statistics need to be estimated in practice (namely, the channel distribution ρ and the observation structure conditional distribution $\urcorner$); estimating statistics such as the channel distribution information is known to be a classical issue in the communications literature. □

2.4 Proposed equilibrium transmission strategy

The main purpose of this section is to obtain globally efficient transmission strategies. Here, global efficiency is measured in terms of social welfare, namely, in terms of the sum $U_1 + U_2$. This corresponds to choosing $\lambda = \frac{1}{2}$. This choice is pragmatic and follows to what is often done in the literature; it implicitly means that the network nodes have the same importance. Otherwise, this parameter can always be chosen to operate at the desired point of the utility region. Indeed, social welfare is a well-known measure of efficiency, and it also allows one to build other famous efficiency measure of a distributed network such as the price of anarchy [31]. The proposed approach holds for any other feasible point of the utility region which is characterized in the preceding section.

The transmission strategy we propose comprises three ingredients: (1) a well-chosen operating point of the utility region, (2) the use of reputation [5, 8], and (3) the use of virtual credit [9, 10].

1. The proposed operating point is obtained by applying the sequential best-response dynamics procedure described in Section 2.3 and choosing $\lambda = 0.5$, $|\mathcal{V}| = 1$. Each individual maximization operation provides a probability distribution which is denoted by $\pi_i^\star$. Since W_λ is linear in $P_{A_i|S_i,V}$, the maximization operation returns a point which is one of the vertices of the unit simplex. The corresponding probability distribution has thus a particular form ,namely, that of a decision function under the form $f_i^\star(s_i)$. Therefore, when operating at this point, at stage t, node i chooses its action to be $a_i(t) = f_i^\star(s_i(t))$. This defines for node i a particular choice of a lottery over its possible actions; this lottery is denoted by $\pi_i^\star(t)$ and is the unit simplex of dimension $2L$, i.e., $\Delta(\mathcal{P}^2)$. By convention, the possible actions for node i are ordered according to a *lexicographic ordering*. Having $\pi_i^\star(t) = (1, 0, 0, ..., 0) \in \Delta(\mathcal{P}^2)$ means that action (P_1, P_1) is used with probability 1 (wp1); having $\pi_i^\star(t) = (0, 1, 0, ..., 0) \in \Delta(\mathcal{P}^2)$ means that action (P_1, P_2) is used wp1; ...; having $\pi_i^\star(t) = (0, 0, 0, ..., 0, 1) \in \Delta(\mathcal{P}^2)$ means that action (P_L, P_L) is used wp1.
2. The reputation (see, e.g., [5, 8]) of a neighboring node is evaluated as follows. Over each game stage duration, the nodes exchange a certain number of packets which is denoted by K. This number is typically large. Since each node has access to the realizations of the signal y_i for each packet, it can exploit it to evaluate the reputation of the other node at stage t. In this section, we assume a particular observation structure Γ_1, Γ_2, which is tailored to the considered problem of packet forwarding in ad hoc networks. We assume that the signal y_i is binary: $y_i \in \{\mathrm{D}, \mathrm{F}\}$. Let $\epsilon \in [0, 1]$ be the parameter which represents the *probability of misdetection*. If node i chooses the action $a^{\min} = (P_{\min}, P_{\min})$ (resp. any other action of $\mathcal{P}_i^2$), then with probability $1 - \epsilon$, node $-i$ receives the signal D (resp. F). With probability ϵ, node $-i$ perceives what we define as the action Drop D (resp. Forward F) while the action Forward F (resp. Drop D) has been chosen by node i. Thus, Γ_i takes the following form:

Algorithm 1

1. ***Initialization***. *The arguments of the functional* W_λ *are fixed to an initial value:* $(P_{A_1|S_1,V}, P_{A_2|S_2,V}) = \left(P^{(0)}_{A_1|S_1,V}, P^{(0)}_{A_2|S_2,V}\right)$.
2. ***Iteration***. *At iteration* $n \geq 1$, $P^{(n)}_{A_i|S_i,V}$ *is updated by being chosen in the argmax of* $W_\lambda\left(P_{A_i|S_i,V}, P^{(n-1)}_{A_{-i}|S_{-i},V}\right)$. *If there are several maximum points, choose one of them randomly and according to a uniform law.*
3. ***Stopping criterion***. $\left|W_\lambda\left(P^{(n)}_{A_i|S_i,V}, P^{(n)}_{A_{-i}|S_{-i},V}\right) - W_\lambda\left(P^{(n-1)}_{A_i|S_i,V}, P^{(n-1)}_{A_{-i}|S_{-i},V}\right)\right| < \eta$ *for some* $\eta \geq 0$.

$$\Gamma_i(y_i|x_0,a_1,a_2)=\begin{vmatrix}1-\epsilon & \text{if } y_i=\text{F and } a_{-i}\in\mathcal{A}_i^{\text{F}},\\ \epsilon & \text{otherwise,}\end{vmatrix} \tag{14}$$

where $\mathcal{A}_i^{\text{F}}=\mathcal{P}_i\times\mathcal{P}_i\setminus\{0\}$.
Using these notations, node i can compute the reputation of node $-i$ as follows:

$$R_{-i}(t)=\frac{(1-\epsilon)|\{y_i=\text{F}\}|+\epsilon|\{y_i=\text{D}\}|}{K}, \tag{15}$$

where $|\{y_i=\text{F}\}|$ and $|\{y_i=\text{D}\}|$ are respectively the numbers of occurrences of the action Forward and Drop among the K packets node i has been needing the assistance of node $-i$ to forward its packets. The reputation $R_{-i}(t)$ is one of the tools we use to implement the transmission strategy which is described further. Note that one of the interesting features of the proposed mechanism is that reputation (15) of node $-i$ only exploits local observations (first-hand reputation information); node i does not need any information about the behavior of its neighboring nodes. This contrasts with the closest existing reputation mechanisms such as [5] and [8], for which the reputation estimation procedure exploits information obtained from other nodes (second-hand information). The corresponding information exchange induces additional signaling and additional energy consumption. At last, by using these techniques, selfish nodes may collude and disseminate false reputation values.

3. The idea of virtual credit is assumed to be implemented with a similar approach to previous works [9, 10], namely, we assume that the nodes have an initial amount of credits, impose a cost in terms of spent credits for a node that wants to transmit through a neighbor at a certain frequency or probability, and that are rewarded when they forward their neighbors' packets. The reward and cost assumed in this paper are defined next.

The proposed transmission strategy is as follows. While the node has not enough credit, it adopts a cooperative decision rule, which corresponds to operating at the point we have just described. Otherwise, it adopts a signal-based tit-for-tat decision rule, which has been found to be very useful to implement mutual cooperation [1, 32]. Existing tit-for-tat decision rules such as Generous-Tit-For-Tat (GTFT) or Mend-Tolerance Tit-For-Tat (MTTFT) [1] do not take into account the possible existence of a state for the game and therefore the existence of a signal associated with the realization of the state. Additionally, the proposed tit-for-tat decision rule also takes into account the fact that action monitoring is not perfect. In contrast with the conventional setup assumed to implement tit-for-tat or its variants [1], the action set of a node is not binary. Therefore, we have to give a meaning to tit-for-tat in the considered setup. The proposed meaning is as follows. When node i receives the signal D and node $-i$ has effectively chosen the action Drop, it means that node $-i$ has chosen $a_{-i}=a^{\min}=(P_{\min},P_{\min})$. When node i receives the signal Forward and node $-i$ has effectively chosen the action Forward, it means that node $-i$ has chosen $a_{-i}=a_{-i}^{\star}=f_{-i}^{\star}(s_{-i})$. Implementing tit-for-tat for node i means choosing $a_i=a^{\min}=(P_{\min},P_{\min})$ (which represents the counterpart of the action Drop) if the node $-i$ is believed to have chosen the action $a_{-i}=a^{\min}=(P_{\min},P_{\min})$. On the other hand, node i chooses the best action $a_i^{\star}=f_i^{\star}(s_i)$ when it perceives that node $-i$ has chosen the action $a_{-i}^{\star}=f_{-i}^{\star}(s_{-i})$. Note that, contrarily to the conventional tit-for-tat decision rule, the actions $a_i^{\star}$ and $a_{-i}^{\star}$ differ in general. Denoting by $m_i(t)\geq 0$ the credit of node i at stage t, the proposed strategy expresses formally as follows. We will refer to this transmission strategy as SARA (for State Aware tRAnsmission strategy).

Proposed transmission strategy (SARA).

$$\sigma_{i,t}^{\star}(s_i^t,y_i^{t-1})=\begin{vmatrix}\pi_i^{\star}(t) & \text{if } t=0 \text{ or } m_i(t)<\mu,\\ \widehat{\pi}_{-i}(t-1) & \text{otherwise,}\end{vmatrix} \tag{16}$$

where

- *The virtual credit $m_i(t)$ obeys the following evolution law:*

$$m_i(t)=m_i(t-1)+\beta<\pi_i(t-1),e_k>-\beta v_i(t-1), \tag{17}$$

  $\beta v_i(t-1)$ represents the virtual monetary cost for node i when its packet arrival rate is $v_i(t-1)$, with $\beta\geq 0$; $<;>$ stands for the scalar product; e_k is the kth vector of the canonical basis of $\mathbb{R}^{2L}$, namely, all components equal 0 except the kth component which equals 1. The index k is given by the index of action $a_i^{\star}(t)=f_i^{\star}(s_i(t))$.
- *$\mu\geq 0$ is a fixed parameter which represents the cooperation level of the nodes. A sufficient condition on μ and β to guarantee that the nodes have always enough credits is that $\mu\geq 2\beta$.*
- *The distribution $\widehat{\pi}_{-i}(t-1)$ is constructed as follows:*

$$\widehat{\pi}_{-i}(t-1)=R_{-i}(t-1)\pi_i^{\star}(t)+[1-R_{-i}(t-1)]\,\pi^{\min}, \tag{18}$$

  with $\pi^{\min}=(1,0,0,\ldots,0)\in\mathbb{R}^{2L}$ representing the pure action $a^{\min}=(P_{\min},P_{\min})$.

Comment 1 *The second term of the dynamical equation which defines the credit evolution corresponds to the reward a node obtains when it forwards the packets of the other node. On the other hand, the third term is the cost paid by the node for asking the assistance of the other node to forward. The same weight (namely, β) is applied on both terms to incite the node to cooperate. Additionally, such a choice allows one from preventing cooperative nodes to have an excessive number of credits, thus to avoid using a mechanism such as in [6] and [8] to regulate the number of credits. In [6] and [8], the credit excess occurs because the reward in terms of credits only depends on the node action and the cost only depends on the relaying node action. Thus, when the node is cooperative and the relaying node is selfish, there will be a reward but not a cost. As for the credit system, in practice, it might be implemented either by assuming the existence of an external central trusted entity [10] that stores and manages the node credits, or through a credit counter located in the node and maintained by a tamper-resistant security module. Thus, in practice, the operation of this module would not be altered, because it would be designed so that information be accessible only by specific software containing appropriate security measures.*

Comment 2 *Depending on whether packet forwarding rate maximization or energy minimization is sought, it is possible to tune the triplet of parameters $(m_i(0), \beta, \mu)$ according to what is desired. In this respect, it can be checked that the best packet forwarding rate is obtained by choosing any triplet under the form $(2\beta, \beta, 2\beta)$ for any $\beta > 0$. But the power (which is defined by the quantity average network power (ANP) defined through (21)) is then at its maximum. On the other hand, if the triplet of parameters takes the form $(m_i(0), 0, 0)$, with $m_i(0) \geq 0$, the consumed network power will be at its minimum and the packet forwarding rate will be minimized as well. Other choices for the triplet $(m_i(0), \beta, \mu)$ therefore lead to various tradeoffs in terms of transmission rate and consumed power.*

Comment 3 *The proposed strategy can always be used in practice whether or not it corresponds to an equilibrium point of $\overline{\mathcal{G}}$. However, if the strategic stability property is desired, some conditions have to be added to ensure that it corresponds to an equilibrium. Indeed, effectively operating at an efficient point in the presence of self-interested and autonomous nodes is possible if the latter have no interest in changing their transmission strategy. More formally, a point which possesses the property of strategic stability or Nash equilibrium is defined as follows.*

Definition 2 *A strategy profile $\left(\sigma_1^{NE}, \sigma_2^{NE}\right)$ is a Nash equilibrium point for $\overline{\mathcal{G}}$ if*

$$\forall i \in \mathcal{N}, \forall \sigma_i \in \Sigma_i,\ U_i\left(\sigma_1^{NE}, \sigma_2^{NE}\right) \geq U_i\left(\sigma_i, \sigma_{-i}^{NE}\right). \tag{19}$$

In order to obtain an explicit condition for the proposed strategy to be an equilibrium, we consider, as the closest related works [1, 8], a repeated game with discount. This also allows some effects such as the loss of network connectivity to be captured. Remarkably, for the repeated game model with discount, the subgame[5] perfection property is also available. This is useful in practice since it offers some robustness in terms of node behavior. Indeed, this property makes the equilibrium strategy robust against changes in terms of node behavior which might occur during the transmission process; even if some deviations from equilibrium occurred in the past, players have an interest in coming back to equilibrium. A necessary and sufficient condition for a strategy profile to be subgame perfect equilibrium is given by the following result.

Proposition 2 *Assume that $\forall t \geq 1, \theta_t = (1-\delta)\delta^t$, $0 < \delta < 1$. The strategy profile $(\sigma_1^\star, \sigma_2^\star)$ defined by (16) is a subgame perfect equilibrium of $\overline{\mathcal{G}}$ if and only if:*

$$\delta \geq \max\left\{\frac{c_i}{(1-2\epsilon)r_i}, \frac{c_{-i}}{(1-2\epsilon)r_{-i}}, 0\right\}, \tag{20}$$

where: $c_i = \sum_{a_0}\rho(a_0)\left(u_i^1 - u_i^k\right)$, $r_i = \sum_{a_0}\rho(a_0)\left(u_i^k - u_i^2\right)$. $u_i^1 = u_i\left(a_0, a_1^\star, a^{\min}\right)$, $u_i^k = u_i\left(a_0, a_1^\star, a_2^\star\right)$, $u_i^2 = u_i\left(a_0, a^{\min}, a_2^\star\right)$, and $a_i^\star = f_i^\star(s_i)$.

(See the proof in the Appendix)

Comment 4 *The proposed transmission strategy is compatible with a packet delivery mechanism such as an ACK/NACK mechanism. Indeed, in the definition of the transmission strategies (6), the observed signal y_i may correspond to a binary feedback such as an ACK/NACK feedback. Indeed, y_i corresponds to an image of (a_0, a_1, a_2). Such image might then be a binary version of the receive SNR or SINR (e.g., if the receive SNR is greater than a threshold then the packet is well received and the corresponding feedback signal y_i will be ACK). More generally, a binary feedback of the form Forward/Drop is completely compatible with the presence of ACK/NACK feedback-type mechanisms. Simply, the signal Drop may combine the effects of a selfish behavior and bad channel conditions.*

Comment 5 *This section shows that the proposed transmission strategy has five salient features.*

1. *First of all, in contrast with the related works on the forwarder's dilemma, it is able to deal with the problem of time-varying link qualities.*
2. *Second, it not only deals with the adaptation of the cooperation power p_i' of node i (which is the power to forward the packets of the other node) but also the power to transmit its own packets p_i.*
3. *Third, the proposed strategy is built in a way to exploit the available arbitrary knowledge about the global channel state (a_0) as well as possible. The key observation for this is to exploit the provided utility region characterization. Ideally, the nodes should operate on the Pareto frontier. This is possible if a suited optimal algorithm is used.*
4. *Fourth, the proposed transmission strategy is shown to possess the strategic stability property in games with discount under an explicit sufficient condition on the discount factor. Note that, here again, each node has only imperfect monitoring of the actions chosen by the other node. Additionally, the equilibrium strategy is subgame perfect.*
5. *Fifth, the proposed strategy does not induce any problem of credit outage or excess. Credit outage is avoided only if the conditions $\mu \geq 2\beta$ and (20) are satisfied. Therefore, if there is no credit outage problem, there is no need for assisting distant nodes as required in [6] and [8].*

3 Results and discussion

3.1 Numerical performance analysis

All simulations provided in this section have been obtained by an ad hoc simulator developed under *Matlab*. The simulation setup we consider in this paper is very close to those assumed in the closest works and [3] and [8] in particular. The setup we assume by default is provided in Section 3.1.1. When other values for some parameters are considered, this will be explicitly mentioned. In addition to the simulation setup subsection, the simulation section comprises three subsections. The first subsection (Section 3.1.2) aims at conducting a performance analysis in terms of utility function (1), which captures the trade-off between the transmission rate and the energy spent for transmitting. Section 3.2 focuses on the transmission rate aspect while Section 3.2.1 is dedicated to a performance analysis in terms of consumed network energy.

3.1.1 Simulation setup assumed by default

We consider a network of N nodes. When N is considered to be fixed, it will be taken to be equal to 50. The N nodes are randomly placed (according to a uniform probability distribution) over an area of 1000×1000 m^2; only network topology draws which guarantee every node to have a neighbor (in the sense of its radio range) are kept. The assumed topology corresponds to a random topology since the node locations are drawn from a given spatial distribution law (which is uniform for the simulations). Each node only considers the behavior of its neighbors to choose its own behavior. As assumed in the related literature, if a node has several neighbors, it is assumed to play a given game with each of its neighbors. In fact, averaging the results over the network topology realizations has the advantage of making the conclusions less topology dependent. Provided simulations are averaged over 1200 draws for the network topology. Routes are supposed to be fixed and known. Indeed, the proposed transmission strategy is compatible with any routing algorithm. One node can communicate with another node only if the inter-node distance is less than the radio range, which is taken to be 150 m. When a node has several neighbors, it may be involved in several routing paths; then, it is assumed to play several independent forwarding games in parallel and have a given initial credit $m_i(0)$ for each neighbor. The credits are updated separately based on the corresponding forwarding game. This means that the credits a node receives by cooperating with one of its neighbors can only be used for forwarding via the considered neighbor. As a node without neighbors does not need credits and the nodes do not obtain an initial credit, the problem of credit excess is avoided. By default, 50% of the nodes are assumed to be selfish but the network does not comprise any malicious node. The initial packet forwarding rate for cooperative nodes and selfish nodes are respectively set to 1 and 0.1, respectively. Each source node transmits at a constant bit rate of 2 packets/s. For each draw for the network topology, the simulation is run for 1000 s. This period of time is made of 20 frames of 50 s. A frame corresponds to a game stage and to a given draw for the channel vector h. The fact that channel gains are assumed to fluctuate over time is a way of accounting for mobility; in the simulations, they are assumed to evolve according to a (discrete version of the) Rayleigh fading law. Averaging over network topologies allows one to average the results over the path losses. Each channel gain is thus drawn according to an exponential law, which corresponds to a Rayleigh law for the amplitude; if one denotes by h_i the considered channel gain, we have that $h_i \sim \frac{1}{\overline{h}_i} e^{-\frac{h_i}{\overline{h}_i}}$, where $\overline{h}_i = \mathbb{E}(h_i)$ represents the path loss effects. As mentioned above, the channel gain is discrete and the discrete realizations are obtained by quantizing the realizations given by a Rayleigh distribution. The effect of quantization on the performance is typically small. Simulations, which are provided here, show that the loss induced by implementing Algorithm 1 by using quantized channel gains in the presence of actual channel gains which are continuous is about a few percents for the size of channel gain sets used for the simulations. If d denotes the inter-node distance of the considered pair of nodes, then the path loss is assumed

to depend on the distance according to $\bar{h}_i = \frac{\text{const}}{d^2+\kappa^2}$; $\kappa > 0$ is a distance which is used to avoid numerical divergence in $d = 0$. In practice, κ may typically represent the antenna height. During each frame, 100 packets are exchanged. Table 1 recaps the values chosen for the main network parameters.

Concerning the game parameters, the following choices are made by default. The parameter α is set to 10^{-2}. The receive variance σ^2 is always set to 0.1. The sets of possible power levels are defined by: $\forall i \in \{1,2\}, \mathcal{P}_i = \mathcal{P}'_i$, $L = 10$, $P_{\min} = 0$, and $P_{\max} = 10$ W. The power increment is uniform over a dB scale, starting from the minimal positive power which is taken to be equal to 10 mW. The sets of possible channel gains are defined by: $\forall i \in \{1,2\}, \mathcal{H}_i = \mathcal{H}'_i$: $H = 10$, $h_{\min} = 0.04$, and $h_{\max} = 10$, and the channel gain increment equals $\frac{10-0.04}{10}$. The different means of the channel gains are given by $\left(\bar{h}_i, \bar{h'}_i, \bar{h}_{-i}, \bar{h'}_{-i}\right) = (1,1,1,1)$. The communication efficiency function is chosen as in [20]: $\varphi(x) = e^{-\frac{c}{x}}$ with $c = 2^r - 1$, r being the spectral efficiency in bit/s/Hz [20]. In the simulations provided we always have $r = 1$ bit/s/Hz; one simulation will assume a higher spectral efficiency, namely, $r = 3$ bit/s/Hz.

3.1.2 Utility analysis

Here, to be able to easily represent the utility region for the considered problem and to be able to compare our approach with previous models, we consider two neighboring nodes.

The first question we want to answer is to what extent the ability for a node to properly adapt to the link qualities which have an impact on the weighted utility $w_\lambda = \lambda u_1 + (1-\lambda)u_2$; it is related to its knowledge about these qualities, i.e., the global channel state $a_0 = \left(h_1, h'_1, h_2, h'_2\right)$. To this end, we have represented in Fig. 2 the achievable utility region under various information assumptions. The top curve in solid line represents the Pareto frontier which is obtained when implementing the transmission strategy given by Algorithm 1 when $\forall i \in \mathcal{N}$, $s_i = a_0 = \left(h_1, h'_1, h_2, h'_2\right)$ that is each node has *global CSI*. The disks correspond to the performance of the *centralized transmission strategy*, namely the best performance possible. It is seen that for a typical scenario the proposed algorithm does not involve any optimality loss. The curve with squares is obtained with Algorithm 1 under *local CSI*, i.e., $s_i = (h_i, h'_i)$. Interestingly, the loss for moving from global CSI to local CSI is relatively small. This shows that it is possible to implement a distributed transmission strategy without sacrificing too much the global performance. This result is not obvious since the weighted utility w_λ depends on the whole vector a_0. When *no CSI* is available (i.e., $s_i =$ constant), the incurred loss is more significant. Indeed, the curve with diamonds (which is obtained by choosing for each $\lambda \in [0,1]$ the best action profile in terms of the expected weighted utility[6] (11)) shows that the gain in terms of sum-utility or social welfare when moving from no CSI to global CSI is about 10%. The point marked by a star indicates the operating point for which transmitting at full power $a_i = (P_{\max}, P_{\max}, P_{\max}, P_{\max})$ is optimal under no CSI.

As a second step, we compare the performance of SARA, ICARUS [3, 6, 8] and GTFT [1], which do not take into account the channel fluctuations. The three corresponding equilibrium points are particular points of the achievable or feasible utility region represented by Fig. 3. The outer curve is the achieved utility region of Fig. 2 when Algorithm 1 is implemented under local CSI (it is the same as the curve with squares of Fig. 2). The social optimum corresponds to the point indicated by the small disk. The point marked by a square corresponds to the performance of SARA, whereas the points marked by a star and a diamond respectively represent the equilibrium points obtained when using ICARUS [3, 6, 8] and GTFT [1]. Note that the way the strategies ICARUS and GTFT have been designed is such that they are able to adapt the packet forwarding rates but not the transmit power level. As a consequence, they cannot exploit any available knowledge in terms of CSI, which induces a quite significant performance loss; it is assumed that GTFT and ICARUS use a pair of actions (a_1, a_2) which maximizes the expected sum-utility. The gain obtained by the proposed

Table 1 Simulation settings

Network parameters	Value
Space	1000 m × 1000 m
Number of nodes	$N = 50$
Radio range	150 m
Const	10^3
κ	5 m
Initial credit	$m_i(0) = 35$
Initial credit of ICARUS [8]	$m_i(0) = 220$
Parameter cost	$\beta = 10$
Cooperation degree	$\mu = 20$
Probability of misdetection	$\epsilon = 0$
Packet arrival rate	$\nu = 1$
Simulation time	1000 s
Frame/stage duration	50 s
Generosity parameter of GTFT	0.1
IFN of ICARUS [6, 8]	5
edp_{th} of ICARUS [6, 8]	0.85
a of ICARUS [6, 8]	0.5
b of ICARUS [6, 8]	2.3
Number of topology draws	1200

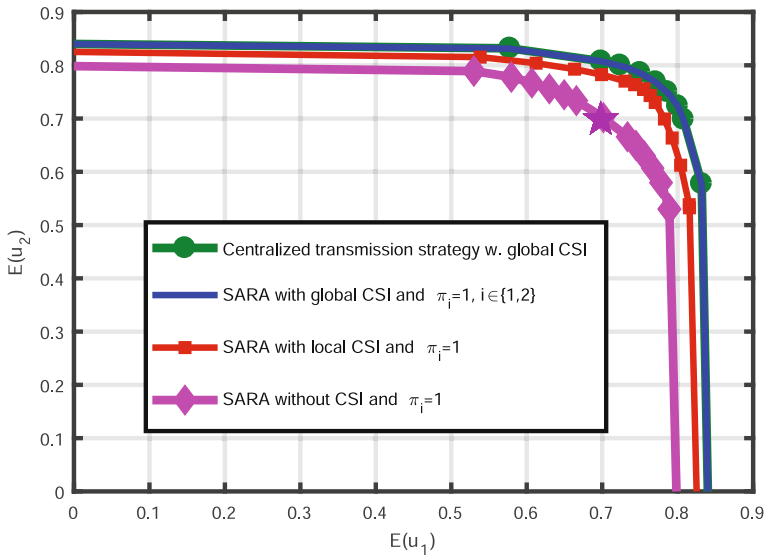


Fig. 2 Achievable utility region under various scenarios of information for the nodes: global CSI, local CSI, and no CSI

transmission strategy comes not only from the fact that the transmit power level can adapt to the wireless link quality fluctuations, but also from the proposed cooperation mechanism. The latter both exploits the idea of virtual credit and reputation, which allows one to obtain a better packet forwarding rate than ICARUS and GTFT. We elaborate more on this aspect in the next subsection. At last, when implementing a transmission strategy built from the one-shot game model given in Section 2.1, the NE of would be obtained, i.e., the operating point would be (0, 0), which is very inefficient.

3.2 Packet forwarding rate analysis

In the previous subsection, we have been assessing the benefits from implementing the proposed transmission strategy in terms of utility. The utility implements a trade-off between the transmission rate and the consumed energy. Here, we want to know how good is the proposed

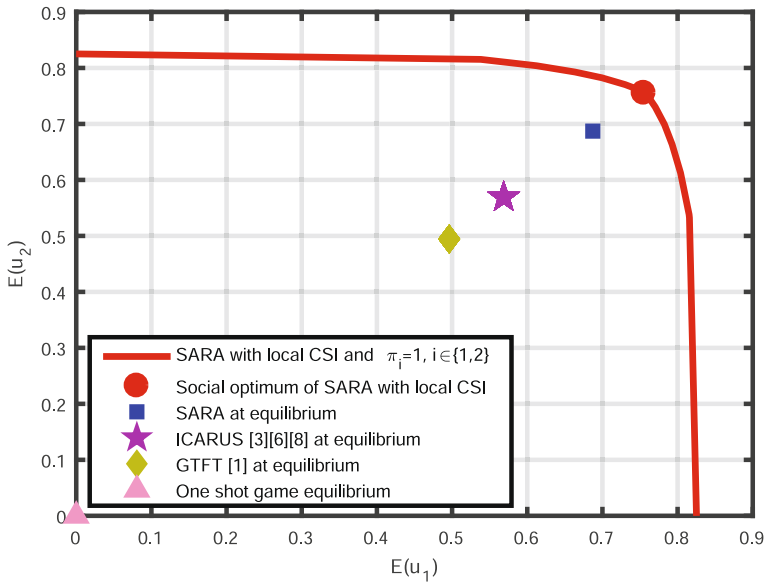


Fig. 3 Achievable utility region with local CSI and the repeated game equilibrium for each strategy: SARA, ICARUS, and GTFT. The one-shot game Nash equilibrium is also represented. The strategies ICARUS and GTFT do not take into account the channel fluctuations

strategy in terms of packet forwarding rate, that is the packet transmission probability.

Figure 4 depicts the evolution of the packet forwarding rate for SARA, ICARUS, and GTFT for a network of 50 nodes. We look at the influence of the fraction of selfish nodes. SARA is very robust to selfishness. Whatever the fraction of selfish nodes, SARA provides a high performance in terms of packet forwarding rate. We see that ICARUS is less efficient than SARA in terms of stimulating cooperation in the presence of selfish nodes, which shows that the proposed punishment mechanism is effectively relevant. The GTFT strategy performance decreases in a significant manner with the number of selfish nodes. For the latter transmission strategy, it is seen that when the network is purely selfish, the operating packet forwarding rate is about 50%; this shows the significant loss induced by using a cooperation scheme which is not very robust to selfishness.

The robustness to observation errors is assessed. More precisely, we want to evaluate the impact of not observing the action Forward or Drop perfectly on the packet forwarding rate. Figure 5 depicts the packet forwarding rate as a function of the probability of misdetection ϵ (see (15)). When $\epsilon > 10\%$, the performance of ICARUS sharply decreases. This is because the retaliation aspect becomes a dominant effect. Nodes punish each other, whereas they should not; this is due to the fact that the estimates of the forwarding probabilities become poor and the ICARUS mechanism is sensitive to estimation errors; illegitimate punishments are implemented, leading to a very inefficient network. On the other hand, observation errors have little influence on SARA because under the equilibrium condition, provided that the credit is less than μ, nodes keep on cooperating. Estimating the forwarding rate does not intervene in the decision process of a node. Note that we have only considered $\epsilon \leq 50\%$. The reason for this is as follows. When $\epsilon > 50\%$, it is always possible, by symmetry, to decrease the effective probability of misdetection to $\epsilon' = 1 - \epsilon$. For this, it suffices to declare the used action to be Forward, whereas the action Drop was observed and vice-versa.

3.2.1 Consumed network energy analysis

Based on the preceding two sections, we know that SARA provides improvements in terms of utility and packet forwarding rate. But the most significant improvements are in fact obtained in terms of consumed energy. Indeed, ICARUS and GTFT have not been designed to account for link quality fluctuations, whereas SARA adapts both the packet forwarding rate and the transmit power level using the parameters assumed by default, except for the path loss $\bar{h}_i = \frac{\text{const}}{d^2+\kappa^2}$, where const=$10^3$ and $\kappa = 5$. In the current formulation of ICARUS and GTFT, the transmit power is fixed (as in Section 3.1.2) according to the best pair of actions in terms of expected sum-utility. In this subsection, the advantage of adapting the power to the quality of the wireless link is clearly observed. Since the consumed network energy is proportional to the network sum-power averaged over time, we will work with the *average network power* (ANP). Here, we consider the

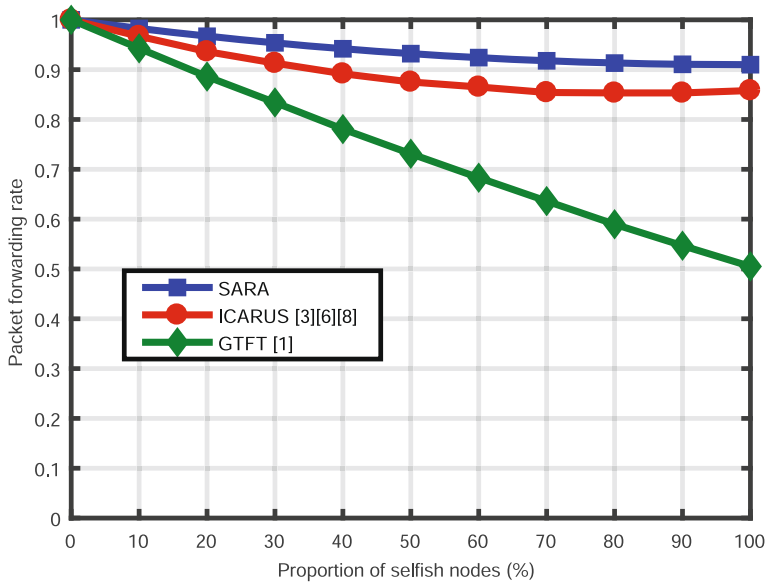


Fig. 4 Packet forwarding rate for different values for the proportion of selfish nodes. The results are averaged over 1200 executions, using the simulation setup assumed by default

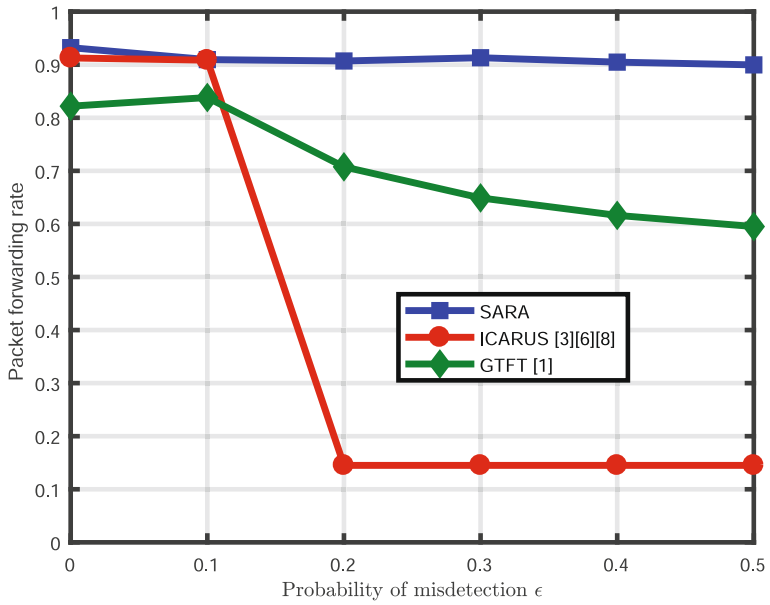


Fig. 5 Packet forwarding rate for different values for the probability of misdetection ϵ. The results are averaged over 1200 executions, using the simulation setup assumed by default

total power which is effectively consumed by the node and not the radiated powers p_i and p'_i (the consumed power therefore includes the circuit power in particular). As explained, for example, in [33, 34], and [35], a reasonable and simple model for relating the radiated power and the consumed power is the affine model: $p_{i,\text{total}} = a(p_i + p'_i) + b$. The parameter b is very important since it corresponds to the power consumed by the node when no packet is transmitted; in [34] and [35], it represents the circuit power, whereas in [33], it represents the node computation power. Here, we assume as in [35] that b is comparable to the $P_{\max}$ and choose the same typical values as in [35], namely, $b = P_{\max} = 1$ W. Eventually, the ANP is obtained by averaging the following quantity $\sum_{i=1}^{N} \{a[p_i(t) + p'_i(t)]\pi_i(t) + b\}$ over all channel and network topology realizations, where N is always the number of nodes in the network and $\pi_i(t)$ the forwarding probability for stage or frame t:

$$\text{ANP} = \frac{1}{T'} \sum_{i=1}^{N} \{a[p_i(t) + p'_i(t)]\pi_i(t) + b\} \tag{21}$$

where T' corresponds to the number of realizations used for averaging. Here, this quantity is averaged over 1200 × 20 stages, the number of network realizations being 1200, and the number of channel realizations being 20. Figure 6 shows how the ANP in dBm scales with the number of nodes for SARA, ICARUS, and GTFT. It is seen that the ANP and therefore the total energy consumed by the network can be divided by more than 2 (the gain is about 4 dB to be more precise) showing the importance of addressing the problem of packet forwarding and power control jointly.

3.2.2 *Impact of quantizing channel gains*

As motivated in Section 2.1, one strong argument for assuming discrete channel gains is that, technically, it corresponds to the most general case; the continuous case follows by using standard information-theoretic arguments. But, from a practical point of view, it matters to assess the loss induced by using an algorithm which exploits quantized channel gains instead of continuous ones. Figure 7 represents the performance in terms of the average utility as a function of cardinality of the set of channel gains. It is seen that the performance obtained by using discrete channel gains in the proposed algorithm, whereas the actual channel gains are continuous is typically small. Here, the simulation setup assumed by default is used.

4 Conclusion

One of the contributions of this work is to generalize the famous and insightful model of the forwarder's dilemma [5, 8] by accounting for channel gain fluctuations. Therefore, the problem of knowledge about global channel state appears, in addition to the problem of imperfect action monitoring when the interaction is repeated. In this paper, we have seen that it is possible to characterize the best performance of the studied system even in the presence of partial information; the corresponding observation structure is arbitrarily provided. The observations

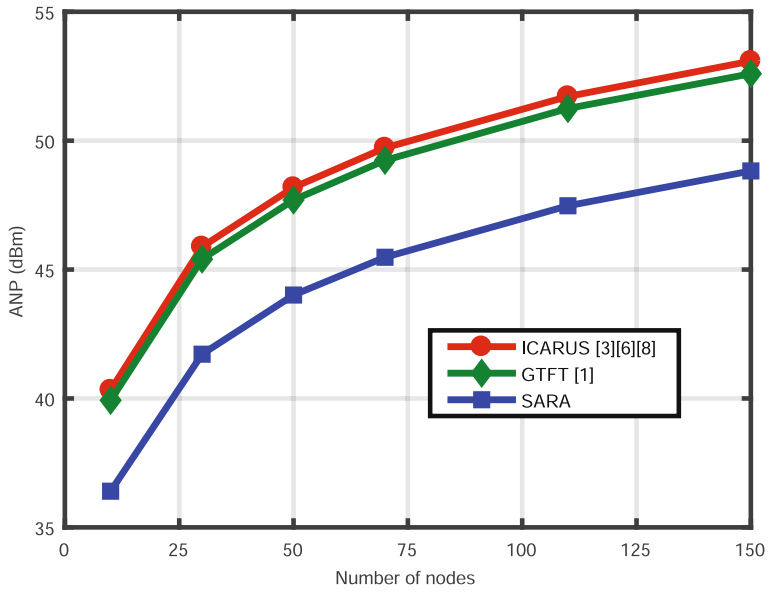


Fig. 6 The figure depicts the average total power (in dBm) or equivalently the total energy consumed by the network against the total number of nodes in the network. It is seen that SARA allows the energy consumed by the network to be divided by more than 2 when compared to the state-of-the art strategies ICARUS and GTFT that do not take into account the channel fluctuations

are generated by discrete observation structures denoted by $\urcorner$ and Γ. In terms of performance, designing power control policies which exploit the available knowledge as well as possible is shown to lead to significant gains. Since, we are in the presence of selfish nodes, we propose a mechanism to stimulate cooperation among nodes. The proposed mechanism is both reputation-based and credit-based. For the reputation aspect, one of the novel features of the proposed strategy is that it generalizes the concept of tit-for-tat to a context where actions are not necessarily binary. For the credit aspect, we propose an evolution law for the credit which is shown to be efficient and robust to selfishness and especially imperfect action monitoring.

From the quantitative aspect, the proposed transmission strategy (referred to as SARA) Pareto dominates ICARUS and GTFT for the utility, the packet forwarding rate, and the energy consumed by the network. Significant gains have been observed; one very convincing result is

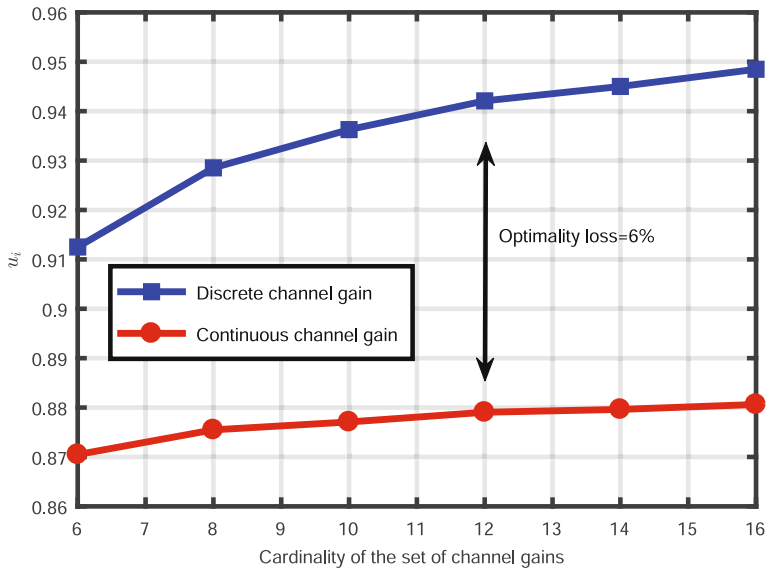


Fig. 7 The average utility for different values for the channel gain set cardinality

that the energy consumed by the network can be divided by 2 when the packet forwarding problem and the power control problem are addressed jointly.

This paper provides the characterization of the best performance in terms of transmission strategy under partial information. Although all performed simulations show that the optimality loss appears to be small, there is no guarantee that the proposed algorithm provides an optimal solution of the optimization problem to be solved to operate on the Pareto frontier or the utility region. Providing such a guarantee would constitute a valuable extension of the present work. Another significant extension would be to relax the i.i.d. assumption on the network state. In this work, the network state corresponds to the global channel state and the i.i.d. assumption is known to be reasonable, but in other setups, where the state represents, e.g., a queue length, a buffer size, or a battery level, the used framework would need to be extended since Markov decision processes would be involved.

Endnotes

[1] The notation |.| stands for the cardinality of the considered set.

[2] The memoryless assumption means that for sequences of realizations of size t (t being arbitrary), $\Pr\left(y_i^t|a_0^t,a_1^t,a_2^t\right)=\Pi_{t'=1}^{t}\Pr(y_i(t')|a_0(t'),a_1(t'),a_2(t'))$.

[3] We will refer to the stability of a point to single deviations as strategic stability.

[4] Note that $\urcorner_1$ and $\urcorner_2$ are directly obtained from $\urcorner$ by a simple marginalization operation.

[5] A subgame of the repeated game is a game that starts at a stage t with a given history.

[6] We therefore assume that the corresponding statistical knowledge is available and exploited.

Appendix 1: Proof of Proposition 1

First, it has to be noticed that long-term utilities are linear images of the implementable distribution. Therefore, characterizing the achievable utility region amounts to characterizing the set of implementable distributions. Note that the set of implementable characterization does not depend on the assumed choice for the infinite sequence of weights $(\theta_t)_{t\geq 1}$, making the result valid for both considered models of repeated games (namely, the classical infinitely repeated game and the model with discount).

Second, to obtain the set of implementable distributions, we exploit the implementability theorem derived in [11]. Therein, it is proved that under the main assumptions of the present paper (namely, the network state is i.i.d. and the observation structure is memoryless), a joint distribution is implementable if and only if it factorizes as in (22). That is, a joint probability distribution or correlation $Q(a_0,a_1,a_2)$ is implementable if and only if it factorizes as:

$$Q(a_0,a_1,a_2)=\sum_{v,s_1,s_2}\rho(a_0)P_V(v)\times\urcorner(s_1,s_2|a_0)\times P_{A_1|S_1,V}(a_1|s_1,v)P_{A_2|S_2,V}(a_2|s_2,v). \quad (22)$$

Third, a key observation to be made now is that if two probability distributions Q_1 and Q_2 are implementable, then the convex combination $\mu Q_1+(1-\mu)Q_2$ is implementable. Indeed, if there is a transmission strategy to implement Q_1 and another to implement Q_2 then by using the first one $\frac{T_1}{T}$ of the time and the second one $\frac{T-T_1}{T}$ of the time, and making T_1 large such that $\frac{T_1}{T}\to\mu$, $\mu Q_1+(1-\mu)Q_2$ becomes implementable. It follows that the long-term utility region is convex. Therefore, the Pareto frontier of the utility region, which characterizes the utility region, can be obtained by maximizing the weighted utility W_λ. This concludes the proof.

Appendix 2: Proof of Proposition 2

We want to prove the following result.

The strategy profile $(\sigma_i^\star,\sigma_{-i}^\star)$ is a subgame perfect equilibrium of $\bar{\mathcal{G}}$ if and only if:

$$\delta\geq\max\left\{\frac{c_i}{(1-2\epsilon)r_i},\frac{c_{-i}}{(1-2\epsilon)r_{-i}},0\right\}, \quad (23)$$

where $c_i=\sum_{a_0}\rho(a_0)\left(u_i^1-u_i^k\right)$, and $r_i=\sum_{a_0}\rho(a_0)\left(u_i^k-u_i^2\right)$.

$u_i^1=u_i\left(a_0,a_1^\star,a^{\min}\right)$, $u_i^k=u_i\left(a_0,a_1^\star,a_2^\star\right)$, $u_i^2=u_i\left(a_0,a^{\min},a_2^\star\right)$, and $a_i^\star=f_i^\star(s_i)$.

As a preliminary, we first review the one-shot deviation "principle" in the context of interest. This "principle" is one of the elements used to prove the desired result.

One-shot deviation principle: For node i, the one-shot deviation principle from strategy σ_i is a strategy $\tilde{\sigma}_i$ writes as:

$$\exists!\ \tau,\ \forall\ t\neq\tau\sigma_{i,t}\left(s_i^t,y_i^{t-1}\right)=\tilde{\sigma}_{i,t}\left(s_i^t,y_i^{t-1}\right). \quad (24)$$

The two strategies $\tilde{\sigma}_i$ and σ_i therefore produce identical actions except at stage τ.

Definition 3 *For node i, the one-shot deviation $\tilde{\sigma}_i$ from strategy σ_i is not profitable if:*

$$U_i(\sigma_i,\sigma_{-i})\geq U_i(\tilde{\sigma}_i,\sigma_{-i}), \quad (25)$$

with $\tilde{\sigma}_i\neq\sigma_i$.

Let us exploit the one-shot deviation principle to prove the result, since it is well known that a strategy profile σ is a subgame perfect equilibrium if and only if there are no profitable one-shot deviations. Assume that for a given

game history, the distributions used by nodes i and $-i$ are respectively π_i and π_{-i}. Following the proposed strategy $\sigma^{\star}_{i,t}$ defined by (16), and by using (15) and (18), one can obtain the distribution of a node i, $\pi_i(t)$, from π_{-i} for each stage t. As defined by relation (18), at each stage t, if $m_i(t) \geq \mu$, node i chooses a distribution $\pi_i(t) = \hat{\pi}_{-i}(t-1)$ stipulating that $a^{min} = (P_{\min}, P_{\min})$ and $a^{\star}_i(t) = f^{\star}_i(s_i(t))$ are the only actions that could be chosen with a positive probability. Thus, it would be sufficient to provide only the kth component of $\pi_i(t)$, which is denoted by $\pi^k_i(t)$. Note that k is the index of action $a^{\star}_i(t) = f^{\star}_i(s_i(t))$. Thus, for $t \geq 1$ we have that:

$$\pi^k_i(t) = \begin{cases} 1, & \text{if } m_i(t) < \mu; \\ (1-2\epsilon)^t \pi_{-i} + \epsilon \sum_{k=0}^{t-1} (1-2\epsilon)^k, & \text{if } \mathrm{mod}(t,2) = 1; \\ (1-2\epsilon)^t \pi_i + \epsilon \sum_{k=0}^{t-1} (1-2\epsilon)^k, & \text{if } \mathrm{mod}(t,2) = 0. \end{cases}$$

Now, we define a one-shot deviation. We consider that node i deviates unilaterally at one stage from the proposed strategy $\sigma^{\star}_{i,t}$, by choosing $\tilde{\sigma}_{i,t}$. If node i deviates, it will be in order to save energy; thus, it chooses $a^{\min}$ with a higher probability than the one provided by the proposed strategy $\sigma^{\star}_{i,t}$. Therefore, we consider that $\tilde{\sigma}_{i,t}$ defines a distribution over the action set as follows (26):

$$\tilde{\pi}_i(t) = \pi_i(t) - d.(\underbrace{-1}_{a^{\min}}, 0, \ldots, \underbrace{1}_{a^{\star}}, 0, \ldots), \tag{26}$$

Using the one-shot deviation $\tilde{\pi}_{-i}(t)$, we have for $t \geq 1$:

$$\tilde{\pi}^k_i(t) = \begin{cases} 1, & \text{if } m_i(t) < \mu; \\ \pi^k_i(t), & \text{if } \mathrm{mod}(t,2) = 0; \\ \pi^k_i(t) - d(1-2\epsilon)^{t-1}. & \text{if } \mathrm{mod}(t,2) = 1. \end{cases}$$

$$\tilde{\pi}^k_{-i}(t) = \begin{cases} 1, & \text{if } m_i(t) < \mu; \\ \pi^k_{-i}(t) - d(1-2\epsilon)^{t-1}, & \text{if } \mathrm{mod}(t,2) = 0; \\ \pi^k_{-i}(t). & \text{if } \mathrm{mod}(t,2) = 1. \end{cases}$$

Now, to accomplish the proof, we need to define the associated expected utilities for each stage provided by $\pi_i(t)$ and $\tilde{\pi}_i(t)$, which are denoted by $u^{\star}_{i,t}$ and $\tilde{u}_{i,t}$, respectively.

$$u^{\star}_{i,t} = \sum_{a_0,a_1,a_2} P_t(a_0,a_1,a_2) u_i(a_0,a_1,a_2), \tag{27}$$

where P_t is the joint probability distribution, and u_i the instantaneous utility (1). Denote by $u^k_i = u_i\left(a_0, a^{\star}_1, a^{\star}_2\right)$, $u^1_i = u_i\left(a_0, a^{\star}_1, a^{\min}\right)$, $u^2_i = u_i\left(a_0, a^{\min}, a^{\star}_2\right)$, and $u^{\min}_i = u_i\left(a_0, a^{\min}, a^{\min}\right)$. $a^{\star}_i = f^{\star}_i(s_i)$, and $a^{\min} = (P_{\min}, P_{\min})$. By means of these notations, we obtain:

$$u^{\star}_{i,t} = \sum_{a_0} \rho(a_0) \left[\pi^k_i(t) \pi^k_{-i}(t) u^k_i + \pi^k_i(t) \left(1 - \pi^k_{-i}(t)\right) u^1_i + \left(1 - \pi^k_i(t)\right) \pi^k_{-i}(t) u^2_i + \left(1 - \pi^k_i(t)\right) \left(1 - \pi^k_{-i}(t)\right) u^{\min}_i \right]. \tag{28}$$

We now define the expected utility of the deviation for each stage, denoted $\tilde{u}_{i,t}$.

$$\tilde{u}_{i,t} = \sum_{a_0} \rho(a_0) \left[\tilde{\pi}^k_i(t) \tilde{\pi}^k_{-i}(t) u^k_i + \tilde{\pi}^k_i(t) \left(1 - \tilde{\pi}^k_{-i}(t)\right) u^1_i + \left(1 - \tilde{\pi}^k_i(t)\right) \tilde{\pi}^k_{-i}(t) u^2_i + \left(1 - \tilde{\pi}^k_i(t)\right) \left(1 - \tilde{\pi}^k_{-i}(t)\right) u^{\min}_i \right]. \tag{29}$$

As the deviation distribution $\tilde{\pi}_i$ depends on the distribution provided by the proposed strategy $\sigma^{\star}_i$, π_i, one can also define $\tilde{u}_{i,t}$ as a function of $u^{\star}_{i,t}$, by using the definitions of $\pi_i(t)$ and $\tilde{\pi}_i(t)$. Hence, we have the following result:

$$\tilde{u}_{i,t} = \begin{cases} u^{\star}_{i,t}, & \text{if } m_i(t) < \mu; \\ u^{\star}_{i,t} - (1-2\epsilon)^{t-1} d \sum_{a_0} \rho(a_0) \breve{U}^1(t) & \text{if } \mathrm{mod}(t,2) = 0; \\ u^{\star}_{i,t} - (1-2\epsilon)^{t-1} d \sum_{a_0} \rho(a_0) \breve{U}^2(t) & \text{if } \mathrm{mod}(t,2) = 1, \end{cases}$$

where $\breve{U}^1(t) = \pi^k_i(t) \left(u^k_i - u^1_i + u^{\min}_i - u^2_i\right) + u^2_i - u^{\min}_i$, and $\breve{U}^2(t) = \pi^k_{-i}(t) \left(u^k_i - u^1_i + u^{\min}_i - u^2_i\right) + u^1_i - u^{\min}_i$. Thus, the deviation utility of node i in the repeated game $\bar{\mathcal{G}}$ is:

$$U_i(\tilde{\sigma}_i, \sigma^{\star}_{-i}) = u^{\star}_{i,0} + \sum_{t=1}^{z} \delta^t \tilde{u}_{i,t} + \sum_{t=z+1}^{\infty} \delta^t u^{\star}_{i,t},$$

where z is the number of stages until the condition $m_i < \mu$ is satisfied. The unilateral deviation from the proposed strategy $\sigma^{\star}_i$ is not profitable if:

$$U_i\left(\tilde{\sigma}_i, \sigma^{\star}_{-i}\right) \leq U_i\left(\sigma^{\star}_i, \sigma^{\star}_{-i}\right). \tag{30}$$

The equilibrium condition could be determined using relation (30). It is defined as follows:

$$u^{\star}_{i,0} + \sum_{t=1}^{z} \delta^t \tilde{u}_{i,t} + \sum_{t=z+1}^{\infty} \delta^t u^{\star}_{i,t} \leq \sum_{t=0}^{\infty} \delta^t u^{\star}_{i,t}.$$

By substituting $\tilde{u}_{i,t}$ by its value, the equilibrium condition writes as:

$$\sum_{t=0}^{t=\frac{z}{2}-1} \delta^{2t} \left((1-2\epsilon)^{2t} d \sum_{a_0} \rho(a_0) \breve{U}^1(2t+1) \right) + \delta \sum_{t=0}^{t=\frac{z}{2}-1} \delta^{2t} \left((1-2\epsilon)^{2t+1} d \sum_{a_0} \rho(a_0) \breve{U}^2(2t+2) \right) \geq 0. \tag{31}$$

We have $\breve{U}^1(2t+1) = \pi^k_i(2t+1) \left(u^k_i - u^1_i + u^{\min}_i - u^2_i\right) + u^2_i - u^{\min}_i$ and $\breve{U}^2(2t+2) = \pi^k_{-i}(2t+2) \left(u^k_i - u^1_i + u^{\min}_i - u^2_i\right) + u^1_i - u^{\min}_i$. We provide results for $\pi^k_i(2t+1) = \pi^k_{-i}(2t+2) = 1$, which implies that relation (31) is satisfied for each $\pi^k_i(2t+1)$ and $\pi^k_{-i}(2t+2)$. With this assumption, the relation (31) becomes:

$$\sum_{a_0}\rho(a_0)\left(u_i^k-u_i^1\right)\sum_{t=0}^{t=\frac{z}{2}-1}\delta^{2t}(1-2\epsilon)^{2t}$$

$$+\sum_{a_0}\rho(a_0)\left(u_i^k-u_i^2\right)\delta\sum_{t=0}^{t=\frac{z}{2}-1}\delta^{2t}(1-2\epsilon)^{2t+1}\geq 0.$$

This is satisfied if and only if:

$$\delta\geq\frac{\sum_{a_0}\rho(a_0)\left(u_i^1-u_i^k\right)\sum_{t=0}^{t=\frac{z}{2}-1}\delta^{2t}(1-2\epsilon)^{2t}}{\sum_{a_0}\rho(a_0)\left(u_i^k-u_i^2\right)(1-2\epsilon)\sum_{t=0}^{t=\frac{z}{2}-1}\delta^{2t}(1-2\epsilon)^{2t}}. \tag{32}$$

The equilibrium condition is thus:

$$\delta\geq\max\left\{\frac{c_i}{(1-2\epsilon)r_i},0\right\}, \tag{33}$$

where $c_i=\sum_{a_0}\rho(a_0)\left(u_i^1-u_i^k\right)$, $r_i=\sum_{a_0}\rho(a_0)\left(u_i^k-u_i^2\right)$, $u_i^1=u_i\left(a_0,a_1^\star,a^{\min}\right)$, $u_i^k=u_i\left(a_0,a_1^\star,a_2^\star\right)$, $u_i^2=u_i\left(a_0,a^{\min},a_2^\star\right)$, and $a_i^\star=f_i^\star(s_i)$.

Thus, the strategy profile $(\sigma_i^\star,\sigma_{-i}^\star)$ is a subgame perfect equilibrium if and only if:

$$\delta\geq\max\left\{\frac{c_i}{(1-2\epsilon)r_i},\frac{c_{-i}}{(1-2\epsilon)r_{-i}},0\right\}. \tag{34}$$

Abbreviations

ANP: Average network power; CSI: Channel state information; GTFT: Generous-Tit-For-Tat; ICARUS: hybrId inCentive mechAnism for coopeRation stimUlation in ad hoc networkS; MTTFT: Mend-Tolerance Tit-For-Tat; NE: Nash equilibrium; SARA: State Aware tRAnsmission strategy; SNR: Signal-to-noise ratio

Acknowledgements

Not applicable.

Funding

Not applicable.

Availability of data and materials

The authors declare that the data supporting the findings of this study are available within the article and its supplementary information files.

Authors' contributions

All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 25 September 2018 Accepted: 3 April 2019
Published online: 03 June 2019

References

1. M. Tan, T. Yang, X. Chen, G. Yang, G. Zhu, P. Holme, J. Zhao, A game-theoretic approach to optimize ad-hoc networks inspired by small-world network topology. Physica A. **494**, 129–139 (2018)
2. L. Feng, Q. Yang, K. Kim, K. Kwak, Dynamic rate allocation and forwarding strategy adaption for wireless networks. IEEE Signal Proc. Lett. **25**(7), 1034–1038 (2018)
3. N. Samian, W. K. G. Seah, in *Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems*. Trust-based scheme for cheating and collusion detection in wireless multihop networks (Computing, Networking and Services, Melbourne, VIC, Australia, 2017)
4. S. Berri, V. Varma, S. Lasaulce, M. S. Radjef, J. Daafouz, in *Proceedings of the 4th International Symposium on Ubiquitous Networking, Lecture Notes in Computer Science*. Studying node cooperation in reputation based packet forwarding within mobile ad hoc networks (Springer LNCS, Casablanca, Maroc, 2017)
5. C. Tang, A. Li, X. Li, When reputation enforces evolutionary cooperation in unreliable MANETs. IEEE Trans. Cybern. **45**(10), 2091–2201 (2015)
6. N. Samian, Z. Ahmad Zukarnain, W. K. G. Seah, A. Abdullah, Z. Mohd Hanapi, Cooperation stimulation mechanisms for wireless multihop networks: a survey. J. Netw. Comput. Appl. **54**, 88–106 (2015)
7. J. M. S. P. J. Kumar, A. Kathirvel, N. Kirubakaran, P. Sivaraman, M. Subramaniam, A unified approach for detecting and eliminating selfish nodes in MANETs using TBUT. EURASIP J. Wirel. Commun. Netw. **2015**(143), 1–11 (2015)
8. D. E. Charilas, K. D. Georgilakis, A. D. Panagopoulos, ICARUS: hybrId inCentive mechAnism for coopeRation stimUlation in ad hoc networkS. Ad Hoc Netw. **10**(6), 976–989 (2012)
9. A. Krzesinski, Promoting cooperation in mobile ad hoc networks. Int. J. Inf. Commun. Technol. Appl. **2**(1), 24–46 (2016)
10. Q. Xu, Z. Su, S. Guo, A game theoretical incentive scheme for relay selection services in mobile social networks. IEEE Trans. Veh. Technol. **65**(8), 6692–6702 (2016)
11. B. Larrousse, S. Lasaulce, M. Wigger, in *Proceedings of the IEEE Information Theory Workshop (ITW)*. Coordinating partially-informed agents over state-dependent networks (IEEE xplore, Jerusalem, 2015)
12. B. Larrousse, S. Lasaulce, in *Proceedings of the IEEE International Symposium on Information Theory*. Coded power control: performance analysis, (2013)
13. B. Larrousse, S. Lasaulce, M. Bloch, Coordination in distributed networks via coded actions with application to power control. IEEE Trans. Inf. Theory. **64**(5), 3633–3654 (2018)
14. A. Gjendemsj, D. Gesbert, G. E. Oien, S. G. Kiani, Binary power control for sum rate maximization over multiple interfering links. IEEE Trans. Wirel. Commun. **7**(8), 3164–3173 (2008)
15. S. Sesia, I. Toufik, M. Baker, *LTE-the UMTS long term evolution: from theory to practice*. (Wiley Publishing, Hoboken, 2009)
16. M. R. Gray, *Entropy and Information Theory*. (Springer, New York, 2013)
17. D. J. Goodman, N. Mandayam, A power control for wireless data. IEEE Pers. Commun. **7**(2), 48–54 (2000)
18. F. Meshkati, M. Chiang, H. V. Poor, S. C. Schwartz, A game theoretic approach to energy-efficient power control in multicarrier CDMA systems. J. Sel. Areas Commun. **24**(6), 1115–1129 (2006)
19. S. Lasaulce, Y. Hayel, R. El Azouzi, M. Debbah, Introducing hierarchy in energy games. IEEE Trans. Wirel. Commun. **8**(7), 3833–3843 (2009)
20. E. V. Belmega, S. Lasaulce, Energy-efficient precoding for multiple-antenna terminals. IEEE Trans. Signal Process. **59**(1), 329–340 (2011)
21. A. El Gamal, Y. Kim, *Network information theory*. (Cambridge University Press, Cambridge, 2011)
22. B. Djeumou, S. Lasaulce, A. G. Klein, in *Proceedings of the IEEE International Sympsium on Signal Processing and Information Technology (ISSPIT)*. Combining decoded-and-forwarded signals in Gaussian cooperative channels (IEEE xplore, Vancouver, 2006)
23. S. Lasaulce, H. Tembine, *Game theory and learning for wireless networks: fundamentals and applications*. (Academic Press, Elsevier, Amsterdam, 2011), pp. 1–336
24. A. Neyman, S. Sorin, Repeated games with public uncertain duration process. Int. J. Game Theory. **39**(1), 29–52 (2010)
25. M. Maschler, E. Solan, S. Zamir, *Game theory*. (Cambridge University Press, Cambridge, 2013)
26. H. Konno, A cutting plane algorithm for solving bilinear programs. Math. Program. **11**(1), 14–27 (1976)
27. G. Gallo, A. Olkticti, Bilinear programming: an exact algorithm. Math. Program. **12**(1), 173–194 (1977)
28. H. Vaish, C. M. Shetty, The bilinear programming problem. Nav. Res. Logist. Q. **23**(2), 303–309 (1976)
29. D. Fudenberg, J. Tirole, *Game theory*. (MIT Press, Cambridge, 1991)
30. A. Agrawal, S. Lasaulce, O. Beaude, R. Visoz, in *Proceedings of the IEEE Fifth International Conference on Communications and Networking (ComNet)*. A framework for decentralized power control with partial channel state information (IEEE xplore, Hammamet, Tunisia, 2015)

31. C. H. Papadimitriou, in *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC)*. Algorithms, games, and the internet (ACM Digital Library, New York, 2001)
32. S. D. Yi, S. K. Baek, J. K. Choi, Combination with anti-tit-for-tat remedies problems of tit-for-tat. J. Theor. Biol. **412**, 1–7 (2017)
33. S. M. Betz, H. V. Poor, Energy rfficient communications in CDMA networks: a game theoretic analysis considering operating costs. IEEE Trans. Signal Process. **56**(10), 518–5190 (2008)
34. F. Richter, A. J. Fehske, G. Fettweis, in *Proceedings of the IEEE 70th Vehicular Technology Conference Fall*. Energy efficiency aspects of base station deployment strategies for cellular networks (IEEE xplore, Anchorage, 2009)
35. V. Varma, S. Lasaulce, Y. Hayel, S. E. El Ayoubi, A cross-layer approach for energy-efficient distributed interference management. IEEE Trans. Veh. Technol. **64**(7), 3218–3232 (2015)

Submit your manuscript to a SpringerOpen journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- ► Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com