Power-trading in wireless communications: a cooperative networking business model

Sithamparanathan, Kandeepan; Jayaweera, Sudharman; Fedrizzi, Riccardo

Sithamparanathan, K., Jayaweera, S., & Fedrizzi, R. (2012). Power-trading in wireless communications: a cooperative networking business model. IEEE Transactions on Wireless Communications, 11(5), 1872–1880. https://doi.org/10.1109/TWC.2012.030812.111350
Document Version: Accepted Manuscript

Published Version: https://doi.org/10.1109/TWC.2012.030812.111350

Power-Trading in Wireless Communications: A Cooperative Networking Business Model

Sithamparanathan Kandeepan *Senior Member, IEEE*, Sudharman K. Jayaweera *Senior Member, IEEE* and Riccardo Fedrizzi

Abstract

Managing the power resource in battery operated wireless devices is very crucial for extending the lifetime, here we propose the concept of power trading in wireless communications. We present a business model using sealed bid procurement auction based game theory for power-trading in cooperative wireless communication with quality of service (QoS) constraints. We formulate the problem as an auction in a buyer's market sequentially/repeatedly played with a single source and a multiple relay network. The source, in-need of cooperation of a relay due to lack of battery power to communicate with the destination, broadcasts a cooperation-request specifying its QoS requirements. The QoS that we consider here are the bit error rate and the total delay associated with relaying the source data. The relays respond with their bids in terms of Euros/bit, and the source selects the best relay based on the bids. The relays compete with each other to win the game and profit from power trading. Each relay updates its pricing index via reinforcement learning to win the game during successive bidding intervals of the repeated game. Based on this model our results show that the relay node with the best features such as a better wireless channel and a better geographical position with respect to the source and destination nodes has a better chance of winning the game, and hence giving rise to a dominant strategy. More importantly, we show that the gains from the wireless channels can be converted into economic profits which is an attractive feature of the proposed business model for power trading.

Index Terms

Power trading, auction game, power efficiency, economic cost, price power profile.

Sithamparanathan Kandeepan is with the School of Electrical and Computer Engineering, RMIT University, Melbourne and the CREATE-NET Research Centre, Italy, kandeepan@ieee.org, Riccardo Fedrizzi is with the CREATE-NET Research Centre, Italy, Sudharman. K. Jayaweera is with the Department of Electrical Engineering, University of New Mexico, USA, jayaweera@ece.unm.edu.

The research was partly funded by the European Commission under the C2POWER project (EU-FP7-ICT-248577), and S. K. Jayaweera was supported by the US National Science foundation (NSF) under the grant CCF-0830545.

I. INTRODUCTION

The current 3G systems have a tendency to rapidly dissipate power in the mobile devices due to the power hungry applications. This has brought up concerns in the industrial sector considering such devices depend on batteries for their power. The future 4G devices on the other hand are expected to be always connected supporting higher data rates with multiple radios that require more and more power as well. In order to improve the power efficiency of wireless communications cooperative wireless networks [1] have been proposed in the literature at the expense of additional complexity and overheads. Here we consider such a cooperative communication network as a means to **trade power** in wireless and mobile communications for local power efficiency.

We present a business model to trade power between a single source node and several relay nodes considering procurement auction based game theoretic model using a first price auction [2]. The source node which is in need of help of a relay node to communicate with its destination due to insufficient power broadcasts a cooperation request to its neighbors. The neighboring nodes then respond with a sealed-bid (price offer) for the cooperation based on a price-power profile as explained in the paper that is deemed to be locally power efficient to the source node. Since we only consider a single source node the auction becomes a buyer's market. The relay nodes compete with each other to win the cooperation for the source node to profit from it hence forming a procurement auction game. The relay nodes also perform reinforcement learning (RL) in order maximize their profits during the successive bidding processes for a discrete sequential/repeated game. In our business model, we also consider the quality of service requested by the source node to the relay nodes in terms of the maximum bit error probability and the maximum acceptable delay in relaying the data. We show that the gains from the wireless channels which are obtained at no additional cost[1] can be directly converted into economic profits which is a highly desirable feature of the proposed business model for power trading.

While the authors in [3] have presented a relay selection scheme for multiple buyer/multiple

[1]Wireless channel gains can be treated as natural gains coming from the natural/man-made environments of the wireless channels perceived with no additional cost in the proposed power trading business model. Thus, making economic profit out of such natural gains is a very attractive feature of this proposed power trading business model

seller scenario to maximize the information rate for the source nodes, this work does not consider RL nor the QoS constraints and more importantly does not deal with power efficiency. By means of RL a relay node could learn the pricing offers from the other relays in the past and bid its future values subsequently to maximize the economic profit, which is one of the key contributions presented in this paper. The authors in [4] study when and how to opt for cooperation for power efficiency considering network fairness and efficiency; again no RL nor the concept of power trading were proposed here.

Moreover, the authors in [5] consider a game theoretic approach to present a pricing function for noncooperative power control game in CDMA networks, where the proposed pricing function is proportional to the signal to interference plus noise ratio (SNR). In [6], the authors consider pricing for transmit power levels in order to achieve Pareto improvement in the noncooperative power control game. They consider a pricing function proportional to the transmit power. The authors in [7] on the other hand present a spectrum trading framework that considers pricing for multiple-buyer multiple-seller based dynamic spectrum access networks, which can possibly be adopted for power trading. Again, none of the above work consider reinforcement learning and the QoS constraints such as the delay.

The rest of the paper is organized as follows. In Section II we present the considered wireless communications problem with the proposed cooperative and business models. The network model is presented in Section III followed by the power trading game formulation in Section IV considering the BER constraint. In Section V the RL model is presented followed by the inclusion of the delay constraint into the power trading game in Section VI. In Section VII we present the convergence of the game followed by the analysis of the overhead cost in Section VIII. Simulation results are presented in Section IX with some final concluding remarks in Section X.

II. The Cooperation and Business Models

The reference scenario considered here for power trading is motivated by the ICT-EU-FP7 funded 'C2POWER' project [12]. In this project multi-radio terminals are considered in cooperative networks for power efficient communications. The source node in order to save power to transmit to the access point (AP), especially when the battery level is low, tries to cooperate with the neighboring node(s) as shown in Figure-1, thus increasing its life time. Such a scenario forms a strong power trading application in wireless and mobile communications attracting the cellular network operators and other stakeholders.

Based on the above scenario we consider a two-hop cooperative system with N relay nodes $R_i, i = 1, 2..N$ and a single source(S)-destination(D) pair, as described in Figure 2. For such a network model one could adopt various cooperation and business models. In our cooperation model S broadcasts a cooperation-request with the destination details and the required QoS in terms of the maximum (acceptable) bit error probability ξ and the maximum allowable delay ζ_d for data relaying. The relay nodes then privately respond to the request with the price quotations in terms of α_i Euros/bit with the offered QoS. The determination/computation of the price α_i by the respective relay nodes is described later in this section and in detail in Section IV. The source node then selects the best option (relay selection) to minimize its transmit power P_S^t and the economic cost based on its own policies and affordability. After selecting the best relay S cooperates with it to communicate with the destination as shown in Figure 2. The selected R_i receives the information from S, decodes, and relays to the destination node using the decode-and-forward cooperative protocol [1]. Note that the proposed framework can easily be adapted to any other cooperative communications protocol, such as amplify-and-forward.

We next describe the business model adopted for the power trading in the network. We first present the power trading model based on the minimum required bit error rate ξ and then extend the game model considering the delay constraint ζ_d. Based on the required bit error probability ξ, R_i calculates its own cost in terms of the required power $P_{R_i}^t$ to relay to the destination node. The transmit power from the relay to destination $P_{R_i}^t$ would depend on the transmit power from the S to relay P_S^t to meet the the bit error rate (BER) ξ as well as various other parameters such as locations of S, R_i, D nodes, and the quality of the wireless channels etc. The transmit power requirement at R_i is then converted to an equivalent economic cost, priced as α_i (Euros/bit).

In our model, we assume that for all R_i's and S, there exists a fixed common base-cost of Λ (Euros/bit/Watt) known to all the nodes in the network. This assumption can be waived considering different base-costs Λ_i for each R_i and Λ_s for S. However, for the sake of simplicity, in our model we consider the fixed common base-cost in analyzing the proposed business model (i.e. the game). Thus, the price α_i quoted to S by R_i is given by

$$\alpha_i(P_S^t) = \Lambda \kappa_i P_{R_i}^t(P_S^t) \tag{1}$$

where α_i and $P_{R_i}^t$ are functions of P_S^t, κ_i is a real number always greater than one known as the pricing index which is decided by R_i and is only known to R_i. When $\kappa_i = 1$, the i^{th} relay node gains no profit by helping S but simply covers the cost incurred due to the committed power for the cooperation. When $\kappa_i > 1$ an economic profit of $\Lambda(\kappa_i - 1)P_{R_i}^t$ (Euros/bit) is made by the respective relay node. In our model, every R_i will have a minimum value for κ_i denoted as $\kappa_i^{\min}$ only known to R_i such that $1 < \kappa_i^{\min} \leq \kappa_i$ which gains a minimum profit of $\Lambda(\kappa_i^{\min} - 1)P_{R_i}^t$ for the cooperation (if it wins). The determination of $\kappa_i^{\min}$ is modeled as a function of the residual battery charge amount q_i^{res} of R_i, and is given by,

$$\kappa_i^{\min} = 1 + a \exp(-b q_i^{\text{res}}) \tag{2}$$

where $a, b \in \mathbb{R}^+$. Note that since $P_{R_i}^t$ is a function of P_S^t (i.e. in order to meet the minimum BER requirement), the price quotation α_i is also a function of P_S^t which we denote as the *price-power profile*. The price-power profile (α_i vs P_S^t) is then used to calculate the bid to be sent to S by the respective R_i given by the pair $\{\hat{P}_S^t, \hat{\alpha}_i\}$ where $\hat{P}_S^t$ and $\hat{\alpha}_i$ are some points in the domains of P_S^t and α_i respectively. The determination of the bid values $\{\hat{P}_S^t, \hat{\alpha}_i\}$ to maximize the profit for R_i and minimize the cost for S is discussed in Section IV. Figure 3 depicts an example of the price-power profile corresponding to a particular R_i with guaranteed QoS.

Based on our business model it is possible that one node (or a few nodes) can be advantaged by its (their) geographical location(s) with respect to S and destination nodes, especially when the environmental parameters are unchanged. In this case the advantaged R_i, by means of RL (as discussed later), can unfairly increase the price for the cooperation bringing out dominance in the game. The competition from the other players however would avoid such an unfair act from the dominant R_i and in case of no competition from the other relays it will be controlled by the buyer (the source node) which would simply decline such offers. This is an example of a Nash equilibrium for the considered game. Furthermore, if the maximum power that could be committed by S is insufficient to guarantee the QoS even with cooperation then we have a no-play situation in the game.

It is important to note here that our objective is not to achieve power efficiency at the network level but only at the nodal level at S. In other words, even though when the cooperation proves to be power efficient to S, the total power spent for the cooperative communication considering $P^t_{R_i}$ and P^t_S may be more than the power required for the non-cooperative case where S communicates with D directly. We also restate that the motivation of our work is to support the source node in need of help due to low battery power level as described in Section II.

A. *Business Plans for Power Trading*

The business providers (such as the Portuguese Telecom (PTI) involved in the C2POWER project [12]) can make use of the power trading application presented in this paper and derive their own business plans to trade power between two nodes. The business plans can vary depending on whether the nodes belong to the same network or different nodes, in the latter case economic incentives transacted between the nodes become complicated. An interesting feature that could be adopted in the business plan could be the 'friendly-mode' and 'business-mode' features. In the friendly-mode feature the cooperating R_i can choose to help a friendly node without being motivated by any economic gains, and in the business-mode feature the cooperative node will have a motivation to make economic profit as discussed in this paper. The user therefore could choose between the two options for power trading.

III. The Communication Network Model

In this section we present the network model adopted here, we also note that the proposed power trading business model and the concept of pricing for cooperation is not restricted to the adopted network model here (with specific modulation-channels etc.) but could be extended to any network in general. The network model also assumes context awareness meaning that many parameters of its own are known to the nodes by means of cognitive learning processes as mentioned in Section II, hence giving us a cognitive-cooperative network model. We only present the cooperative network model here and assume that cognition exists in the network and provides the necessary information. For the cooperative network in Figure 2 the received signals for the communications from $S \rightarrow D$, $S \rightarrow R_i$, and $R_i \rightarrow D$ can be expressed in the form of,

$$r(t) = \frac{1}{L(d)} h(t)s(t) + \upsilon(t) \tag{3}$$

where, ignoring the subscripts, s is the transmitted signal, h is the small scale fading channel given by a complex Gaussian process or equivalently the envelope of the channel $|h|$ is given by a Rayleigh process with a slowly varying mean square average of $E[h^2] = \overline{h^2}$, υ is the additive noise at the receiver with a double-sided power spectral density of $N_0/2$, and $L(d)$ is the mean pathloss due to a given transmitter-receiver (T-R) separation of d in meters. The pathloss component $L(d)$ for all the links are given by the pathloss exponent γ as [9],

$$L(d) = L(d_0)(d/d_0)^{\gamma} \tag{4}$$

where, $L(d_0)$ is the pathloss at a reference distance d_0 which is given by the free-space pathloss [9] $L(d_0) = (4\pi f d_0/c)^2$, where f is the carrier frequency and $c = 3 \times 10^8$ is the speed of light. With the assumption of context awareness, the above mentioned parameters such as the T-R separations d, the channel power levels $\overline{h^2}$, the pathloss exponents γ are all known to the respective nodes (but not to the other nodes).

We consider BPSK communication system as an example for all the links considered in equation (3) in order to illustrate the power trading business model with a guaranteed BER. The bit error probability under Rayleigh fading with AWGN thus for all the links is given by [9],

$$\Pi = 0.5 \left(1 - \sqrt{\Gamma/(1+\Gamma)}\right) \tag{5}$$

where $\Gamma = E_b\overline{h^2}/N_0$ is the average SNR, E_b is the bit energy given by $E_b = P^t/(\Delta L(d))$, and Δ(bits/s) is the data rate of the respective links. Note that in the above equations we ignore the subscripts of S, R_i or D because they are true for all the links considered here. Furthermore, if $\bar{\Pi}$ is the overall BER of the concatenated communication from $S \to R_i \to D$ and Π_S and Π_R are the BER of $S \to R_i$ and $R_i \to D$ links respectively then $\overline{\Pi}$ is given by $\overline{\Pi} = \Pi_R(1-\Pi_S)+\Pi_S(1-\Pi_R)$. By recognizing that $2\Pi_R\Pi_S$ is much less than Π_S and Π_R, the overall BER $\overline{\Pi}$ reduces to

$$\overline{\Pi} \approx \Pi_S + \Pi_R \tag{6}$$

IV. The Power Trading Game

First we present the power trading game considering the bit error rate constraint ξ and then we extend the game model for the delay constraint later in Section VI. We consider the strategic form of the game considering the procurement auction model defined by $G = \langle I, A, \Omega \rangle$, where $I = \{1, 2 \cdots, N+1\}$ is the set of players in the game (i.e. all R_i and S), $A = A_1 \times A_2 \cdots \times A_N \times A_S$ is the cartesian product of the sets of actions available to each player with A_i being the action set of player i, and $\Omega = \{\Omega_1, \Omega_2, \cdots, \Omega_N, \Omega_S\}$ is the set of reward functions. We also note that in our notations we refer to i for all the parameters related to R_i and refer to $i-$ for all the parameters related to the relay nodes other than R_i.

Repeated Game: Since we assume multiple cooperation requests (sufficiently large number of times) from S, the game is thus played every time S requests for the cooperation. Hence giving us a repeated discrete game, in which the bidders can learn from the previously deployed strategies Ψ_i corresponding to the action $a_i \in A_i$ and revise their bids successively with a revised strategy if improvement in the bids are foreseen, as discussed in the next section.

First Price Auction: The first price auction is considered here, in which the relay node with the bid corresponding to the minimum cost for S wins the game and cooperates with S with the quoted minimum bid.

The Rewards: The payoff or the reward $\Omega_i \in \Omega$ for R_i for the cooperation (power sale) can be directly related to the economic profit described subsequently to (1). The reward Ω_i is only known to R_i and not known to the other relays R_{i-} or the source node S. If R_i wins the cooperation for S then the reward is $\Omega_i = \Lambda(\kappa_i - 1)P_{R_i}^t$, and if R_i does not win the cooperation then the reward is $\Omega_i = 0$, as depicted in Figure 4. In order to maximize the reward R_i needs to maximize κ_i or in other words needs to increase κ_i as described below.

$$\hat{\kappa_i} = \arg\max_{\kappa_i}\{|\alpha_i - \Lambda P^t_{R_i}|\} = \arg\max_{\kappa_i}\{\Lambda P^t_{R_i}(\kappa_i - 1)\} \tag{7}$$

The above maximization is performed by R_i by simply increasing κ_i by means of RL as described in the next section. The reward for S on the other hand is given by $\Omega_S = \Lambda P_0 - C_i$, where P_0 is the required transmit power for the $S \to D$ direct transmission link and ΛP_0 is the economic cost associated with it with C_i being the cost for cooperating with R_i as described subsequently.

Relay Selection: The economical cost (price to pay) for S to cooperate with R_i is given by

$$C_i = \Lambda P^t_S + \alpha_i \tag{8}$$

where, P^t_S is the required transmit power for the relay communication link $S \to R_i$. Based on the source node's local policy it can choose R_i which minimizes power or price (at the expense of each other), here we consider minimizing the price. That is, S selects the relay R_i and chooses the corresponding P^t_S by maximizing the economical gain (the reward) Ω_S, given by,

$$\hat{P}^t_S = \max_{i,P^t_S}\{\Lambda P_0 - C_i; C_i \leq \Lambda P_0, P^t_S \leq P_0\} \tag{9}$$

Alternatively, we could say that S minimizes C_i with the conditions $C_i \leq \Lambda P_0$ and $P^t_S \leq P_0$. It should be noted that S chooses $\hat{P}^t_S$ based on (9) only when the conditions $C_i \leq \Lambda P_0$ and $P^t_S \leq P_0$ are satisfied, if not the cooperation is deemed to be more expensive or infeasible, respectively, and therefore S rejects the offers from R_i.

Bid Calculations: We now describe how R_i calculates the price-power profile $\alpha_i = \Lambda\kappa_i P^t_{R_i}$. Considering the BER requirement of ξ at S, using (6) and the BER expression in (5), node R_i calculates the relationship between $P^t_{R_i}$ (required transmit power at R_i) and P^t_S (the corresponding required transmit power at node S), given by,

$$P^t_{R_i} = \frac{\lambda^2}{\overline{h_2^2}(1-\lambda^2)} L(d_2) N_0 \Delta_R \tag{10}$$

where $\overline{h_2^2}$ is the channel power, d_2 is the T-R separation, Δ_R is the data rate for the transmission from $R_i \to D$, and

$$\lambda = 1 - 2\left(\xi - 0.5\left(1 - \sqrt{\frac{\Gamma_1}{1+\Gamma_1}}\right)\right) \tag{11}$$

where $\Gamma_1 = \frac{\overline{h_1^2} P_S^t}{L(d_1) N_0 \Delta_S}$ is the average SNR, $\overline{h_1^2}$ is the channel power, Δ_S is the data rate, and $L(d_1)$ is the T-R separation associated with the link $S \rightarrow R_i$. The relay node R_i then uses (10) and (11) together with (1) to generate the price-power profile α_i vs P_S^t. It is important to note here that R_i (most of the battery powered wireless devices in general) would have a maximum transmit power limit (typically around 15dBm to 20dBm for commercial products), which in turn determines the start points of the price-power profiles α_i described in Figure 3. If S accepts the deal with R_i then the relay node could either keep the same price for the successive times or increase the price by increasing κ_i. On the other hand, if S does not accept the offer initially, or successively, then R_i may choose to decrease the price by decreasing κ_i. Note that, R_i cannot decrease κ_i below $\kappa_i^{\min}$, and in such situations R_i has to simply wait until the environmental conditions change (e.g. improvement in the channel gain etc.) to provide a better offer to S.

Considering the strategy adopted by S as in (9), which is assumed to be known to R_i, the relay would then also adopt the same strategy to choose $\{\hat{P}_S^t, \hat{\alpha}_i\}$ given by (9) for its bids from the computed price-power profile and chooses $\hat{\alpha}_i = \Lambda \kappa_i P_{R_i}^t(\hat{P}_S^t)$ correspondingly. The value of κ_i for the initial bid may be chosen to be a random value between $\kappa_i^{\min}$ and $\kappa_i^{\min} + 1$ to maximize its reward, and by means of RL, as discussed in the next section, κ_i would be changed subsequently based on the relative values of C_i and C_{i-} considering the decisions made by S. Once the bid values are computed relays will place their bids to S by means of a response message to the cooperation request from S, and S will collect all the bids and initiate the cooperation with the relay corresponding to the minimum C_i.

The Rules of the game: The game is governed by a set of rules followed by all the players and in a real scenario can be imposed by means of policies and protocols. (1) S cannot negotiate the price but can only decide on accepting or declining the offer given by R_i. (2) Any relay R_i can re-quote their prices to S after RL by changing their pricing index κ_i in the successive cooperation requests until R_i cannot afford to provide a better price unless the environmental conditions improve. 3) If two or more R_i offer the same bid with the lowest price then S will select one of them randomly.

V. Reinforcement Learning

Since the power trading game is played sequentially, depending on whether the bid from R_i was accepted or not by S, the relay R_i can learn the relative values of C_{i-} with respect to its own C_i (i.e. whether $C_i \gtrless C_{i-}$) by means of RL. Note that C_{i-} are not known to R_i and hence this game resembles a classical closed-bid auction model. Based on RL during the subsequent bids, if R_i wins the cooperation, it knows that $C_i < C_{i-}$ and hence κ_i is subsequently updated to increase the reward as

$$\kappa_i(t+1) = \kappa_i(t) + \omega_i^+; \quad \omega_i^+ \in \mathbb{R}^+ \tag{12}$$

The value of κ_i is increased as in (12) in a greedy manner until the cooperation is lost (at a point where $C_i > C_{i-}$) at which point R_i will reduce κ_i as described below. The corresponding κ_i value (when R_i loses the cooperation) is recorded as $\kappa_i^{\max}$ and R_i does not exceed this value in the future unless the environmental conditions are changed. Note that $\kappa_i^{\max}$ needs to be updated when the environmental conditions change. If R_i does not win the cooperation for S, or if it loses the cooperation after winning, then it knows that $C_i > C_{i-}$ and hence κ_i is subsequently updated as,

$$\kappa_i(t+1) = \kappa_i(t) - \omega_i^-; \quad \omega_i^- \in \mathbb{R}^+ \tag{13}$$

whilst meeting the condition $\kappa_i \geq \kappa_i^{\min}$. In general, the updates of the pricing index κ_i is given by $\kappa_i(t+1) = \kappa_i(t) + \omega_i$ for R_i, where $\omega_i \in \{+\omega_i^+, -\omega_i^-\}$. Note that κ_i is lower bounded by $\kappa_i \geq \kappa_i^{\min} \geq 1$ and the updating process for κ_i is stopped when it reduces to $\kappa_i = \kappa_i^{\min}$. Likewise, when κ_i increases continuously (when R_i keeps on winning), κ_i will be upper bounded by $\kappa_i^{\max}$. Note that a strict upper bound for κ_i corresponds to the condition where $C_i < \Lambda P_0$.

VI. Delay-Constrained Systems with Guaranteed BER

In the previous sections we formalized the sequentially played auction game for power trading considering the BER constraint. Here, we also include the delay constraint into the power trading business model for S with delay sensitive traffic. The delay that we consider here is the delay in relaying the entire data from S by R_i to the destination node. The delay here is therefore characterized by the traffic that R_i is handling and whether the relay node agrees to provide primary access or secondary access (opportunistic access) to its channel resources for S. In the case of R_i providing primary access to S the delay could be well quantified and it would be rather a straight forward process to say if the delay constraints can be met or not, that is a guaranteed service could be provided to meet the delay constraint. On the other hand if R_i provides secondary access to S, where S can only access R_i opportunistically, then R_i needs to notify S with a parameter quantifying the delay associated with data relaying.

In the work that we present here we consider the secondary access mechanism for S to access the relay (i.e. opportunistic relay access [10]) and quantify the delay QoS by the probability of meeting the delay-threshold ζ_d based on the relay node's traffic. Based on the characteristics of its own traffic at R_i it can come up with a value for the success probability in meeting the delay threshold ζ_d for a given data length say D_p bits at a data rate of Δ bps. As an example by assuming a simplistic Poisson-exponential traffic model at R_i with a Poisson arrival rate of μ (s^{-1}) and a mean exponential hold time of $\hat{\tau}$, the probability of meeting the delay threshold Υ_i at the relay node i can be computed as [10],

$$\Upsilon_i = \Gamma(K_0, \mu(\zeta_d)) - \Gamma\left(K_0, \mu\zeta_d\left(1 - \frac{1}{\bar{\tau}\mu K_0}\right)\right)\left(\frac{\bar{\tau}\mu K_0}{\bar{\tau}\mu K_0 - 1}\right)^{K_0} \exp\left(-\frac{\mu\zeta_d}{\bar{\tau}\mu K_0}\right) \tag{14}$$

where, $K_0 = \lfloor \frac{D_p/\Delta}{1/\mu - \beta - T_H} \rfloor$, β is the transmission switching delay, T_H is the packet overhead delay, and $\Gamma(a, b) = \frac{1}{\Gamma(a)} \int_0^b t^{a-1} \exp(-t) dt$ is the normalized lower incomplete Gamma function with $\Gamma(a)$ being the standard Gamma function. Therefore, by knowing its own traffic characteristics the relay nodes can quantify the delay QoS similar to the example shown above to S during the bidding process. Even though the example provided above considers the Poisson-exponential traffic model it can be used to get an average estimate for the probability of success in meeting the delay constraint in general. A point to note here is that R_i supporting higher data rates can benefit more from this model since K_0 would reduce for higher data rates giving a better Υ_i value for the same traffic rate μ. The bid to S containing the delay metric then would be a triple given by $\{\hat{P}_S^t, \hat{\alpha}_i, \Upsilon_i\}$.

The strategy adopted by R_i for making the bid $\{\hat{P}_S^t, \hat{\alpha}_i, \Upsilon_i\}$ will also be changed when the delay constraint is considered. The delay metric Υ_i does not affect the selection of the parameters from the power delay profile such as the values of $\hat{P}_S^t$ and $\hat{\alpha}_i$, however it would impact the minimum pricing index $\kappa_i^{\min}$ that relates to the rewards Ω_i. In Section II we defined the pricing index κ_i to be a function of the residual battery charge of R_i, now by considering the delay constraints for the relay transmissions as described in this section κ_i could be modified by incorporating the delay metric Υ_i. The new pricing index therefore is defined letting $a = \Upsilon_i$ in (2), given by,

$$\kappa_i^{\min} = 1 + \Upsilon_i \exp(-b q_i^{\text{res}}) \tag{15}$$

In the above expression $\kappa_i^{\min}$ reduces when the traffic load is higher at R_i (i.e. when Υ_i reduces), at the same time and when there is no traffic load at R_i(i.e. when $\Upsilon_i = 1$) then $\kappa_i^{\min}$ only depends on the residual battery charge of R_i. The relay node R_i will then use (7) to maximize its reward as discussed previously.

The strategy adopted by S will also change considering the delay constraint. Below we provide two strategies that S can adopt to choose the winner considering the BER as well as the delay constraints. The first strategy is to simply check if the delay metric from the bids Υ_i meet a predefined value Υ_0 set by S and then choose R_i in a similar manner considering (9), given by,

$$\hat{P}_S^t = \max_{i, P_S^t} \{\Lambda P_0 - C_i; C_i \leq \Lambda P_0, P_S^t \leq P_0, \Upsilon_i \geq \Upsilon_0\} \tag{16}$$

Using the above strategy it is possible that S will not choose any of the relays for cooperation if the delay constraint is not met even though the BER constraint is satisfied. The second strategy that could be adopted by S is to include the delay constraint metric Υ_i directly in the cost function C_i. Considering the delay constraint we can redefine the cost function for S as,

$$C_i = \frac{1}{(\Upsilon_i - \epsilon)} (\Lambda P_S^t + \alpha_i) \tag{17}$$

where $\epsilon > 0$ is a small real number to avoid very large values for C_i to perform the optimization process, moreover the bid from R_i is rejected if $\Upsilon_i \leq \epsilon$ (i.e. if C_i becomes negative). This way S could also maintain the delay constraint by setting $\epsilon = \Upsilon_0$ and at the same time opt for R_i with a smaller delay. Then by using the optimization strategy in (9) S will choose the corresponding relay for the cooperation by using (17).

VII. The Convergence, Nash Equilibrium and Stability

To prove the convergence of the proposed power trading game, though one could provide a rigorous mathematical proof, we argue that the cost function C_i has a global minimum that assures an equilibrium point. For a wireless channel that does not vary over the duration of our interest let us first consider the cost function in (16) first. The cost function $C_i = \Lambda P_S^t + \Lambda \kappa_i P_{R_i}^t$ can be represented as a sum of two functions given by $f = \Lambda P_S^t$ and $g = \Lambda \kappa_i P_{R_i}^t$, in the domain P_S^t. The function f is convex for finite and positive values of Λ because it is linear with P_S^t. Next we consider the function g. For non-zero, positive and finite values of the environmental parameters the BER in (5) for the link $S \rightarrow R_i$, denoted by Π_S, is a convex function[2] in P_S^t. Therefore, the BER for the $R \rightarrow D$ link, denoted by Π_R given by $\Pi_R = \xi - \Pi_S$ (Appendix-1, for finite and positive ξ) is a concave function in P_S^t. From (5), the relay transmit power $P_{R_i}^t = (1 - 2\Pi_R)^2/[1 - (1 - 2\Pi_R)^2]$ is a decreasing convex function in Π_R, and hence we can show that [11] $P_{R_i}^t(\Pi_R(P_S^t))$ is a convex function in P_S^t. Therefore, g is convex in P_S^t for finite positive values of κ_i. The cost function $C_i = f + g$ is therefore a summation of two convex functions and hence C_i itself is convex in P_S^t, which gives rise to a unique (global) minimum for a given range of P_S^t. Since C_i is continuous and convex, and the action sets are compact there exists a Nash equilibrium. It is important to note here that the proposed cost functions in (16) and (17) are constrained by the conditions $C_i \leq \Lambda P_0$ and$P_S^t \leq P_0$ as indicated in the expressions. Therefore, the values of κ_i are bounded by such constraints forcing towards an equilibrium point.

For the cost function in (17), a similar argument as above can also be presented since $(\Upsilon_i - \epsilon)$ does not depend on P_S^t and hence could be treated as a constant giving us $C_i = \frac{1}{\Upsilon_i - \epsilon}(f + g)$.

[2]BER curves that are functions of transmit power are convex in general.

Since $f+g$ is convex and Υ_i is a constant in the domain P_S^t the cost function C_i (as in (17)) is also convex, and hence Nash equilibrium exists as per our previous argument.

The Nash equilibrium can be proved for a general model with an arbitrary modulation technique as long as the relay transmit power $P_{R_i}^t$ is a convex function in P_S^t for a BER constraint of $\xi = \Pi_R + \Pi_S$, which proves that the cost function C_i is convex. It can be easily argued that $P_{R_i}^t$ in general is convex in P_S^t since one needs to be traded off with the other (similar to the curves shown in figure-3) to satisfy $\xi = \Pi_R + \Pi_S$ for most of the modulation schemes.

A. *Sequential Games in Rapidly Varying Channels with Reinforcement Learning*

As in any general learning techniques, the proposed reinforcement learning technique would only become useful for sufficient number of games played sequentially when the channel environment is stable. Therefore, for the channels with smaller coherent times the reinforcement learning feature would not be of much benefit depending on the values of ω_i, the frequency of cooperation requests and the coherence time. Nevertheless, for rapidly changing channels, the stability of the system is still maintained based on the arguments presented in Section VII (Nash equilibrium), and hence the power trading game could be still played. In other words the Nash equilibrium exists for every single cooperation request, the equilibrium point however changes when κ_i is changed by R_i when RL is used, or the equilibrium point changes when the channel environment changes. This is further understood in the numerical results section later.

VIII. Overhead-Cost for Power Trading

The proposed business model for power trading requires a bidding process for every cooperation request for the repeated games. A properly defined protocol suite is therefore required in order to have power-trading game as an application with the appropriate management and billing functionalities. The design of such a protocol suite is an ongoing work and is beyond the scope of this paper/topic. However, in this paper, we consider the associated costs related to the overheads. During the bidding process S broadcasts a *cooperation request message*, sends a *cooperation* *agreement message* once a relay node is selected and finally sends a *cooperation termination message* once the relaying is completed. These messages will also include the necessary management and billing information for the cooperation. Note that we consider the control messages are received without errors such that retransmissions of the same are not required. The overall overhead is minimized and the cooperation is made as simple as possible in order to gain maximum from the power trading application. Therefore an energy efficient protocol suite is much needed to have successful power trading.

Suppose the total number of overhead bits for the overhead messages is Q_{OH} (bits) and the total information/data bits to be relayed is Q_{data} (bits) then it is reasonable to assume that $Q_{OH} << Q_{data}$. Correspondingly, if C_{OH} is the total economic cost associated with the overhead messages then the cost functions given in (8) and (17) can be modified to include C_{OH}. Note that the overhead cost C_{OH} is a fixed cost and is common for every single cooperation for all the relays. C_{OH} can be well quantified in practice since all the parameters associated with the overhead transmissions such as transmit power level, packet length, data rate etc. are known to S once the protocol suite is designed. Therefore, considering the overhead costs C_{OH}, equations (8) and (17) can be respectively redefined as,

$$C_i = \Lambda P_S^t Q_{data} + \alpha_i Q_{data} + C_{OH} \tag{18}$$

$$C_i = \frac{1}{(\Upsilon_i - \epsilon)}(\Lambda P_S^t Q_{data} + \alpha_i Q_{data} + C_{OH}) \tag{19}$$

Note that the above given cost functions consider the total cost for sending an entire message of length Q_{data} as oppose to the previous cost functions in (8) and (17) which consider only the cost associated with sending a single bit. The redefined reward for S in this case, given by $\Omega_S = \Lambda P_0 Q_{data} - C_i$, depends on C_{OH} and thus it turns out that a minimum number of data bits needs to be relayed in order for S to get a profit. Let us consider the cost function C_i in (18), to profit from the cooperation (i.e. for $\Omega_S > 0$), S needs to relay at least $Q_{data}^{\min}$ bits given by,

$$Q_{data}^{\min} = C_{OH}/[\Lambda P_0 - (\Lambda P_S^t + \alpha_i)] \tag{20}$$

From the above equation we observe that $Q_{data}^{\min}$ reduces with reducing overhead cost (numerator of (20)) and increasing economic profit when considering no overhead cost (denominator of (20)). In reality, with high bandwidth applications in the current era, S generally should have sufficient amount of data Q_{data} compared to $Q_{data}^{\min}$ in order to have positive rewards. Quantifying $Q_{data}^{\min}$ requires C_{OH} to be quantified which in turn needs the protocol suite to be defined for the power trading application. Therefore, we omit any further analysis with overheads and assume that sufficiently large information data is available to be relayed.

The relay nodes on the other hand will incorporate the overhead costs associated with the bidding and relaying process within the price index κ_i. If κ_{OH} corresponds to the price index value associated with the overheads on the relay node's side then the minimum price index could be modified to include the overhead costs given by.

$$\kappa_i^{\min} = 1 + \Upsilon_i \exp(-bq_i^{\text{res}}) + \kappa_{OH} \tag{21}$$

Moreover, the Nash equilibrium and the convergence of the power trading game does not change since C_{OH} is independent of P_S^t and α_i, in other words C_{OH} is treated as a constant parameter within C_i as given in equations (18) and (19). Therefore the Nash equilibrium, as proven before, holds true for the business models with cooperation overheads.

IX. NUMERICAL EXAMPLE

Numerical and simulation results are provided to explain the power trading game and its convergence under stable channel conditions. We consider sufficiently large information data to be sent from S and hence eliminate the requirement to consider overheads in our numerical examples, in other words we do not consider the cost/reward functions presented in Section VIII in our numerical/simulation analysis since it requires the quantification of C_{OH}. We select the following simulation parameters considering that $\Upsilon_i > \Upsilon_0,\ \forall i$ for the strategy given in (16): $\Lambda = 1, N = 3, \xi = 1e-3, \Delta_S = 400\text{kbps}, \overline{h^2}$ for all the channels are $1, f_S = 1800\text{MHz}, \Delta_{R_i} = 1000\text{kbps}, f_R = 1800\text{MHz}, N_0 = -163\text{dBw/Hz}, d_{SR_i} = [\sqrt{5}, \sqrt{80}, \sqrt{45}], d_{R_iD} = [\sqrt{113}, \sqrt{26}, \sqrt{45}]$ and $d_{SD} =$

$\sqrt{162}$. Figures 5 and 6 show the cost function (equation (8)) and the price-power profiles (equation (1)) associated with the three relays considering various environmental conditions. In both the figures R_2 dominates the strategy set due to its relative position and the channel gains, and therefore having a lower cost respect to the other two relays. The figures also depict the power requirement P_0 for the link $S \rightarrow D$ and the corresponding economic cost ΛP_0 for the source for a minimum BER of ξ. Therefore, based on these results S selects R_2 for the cooperation which provides better power efficiency with a lower economic cost. This clearly proves that S and R_2 benefit from the trade due to the cooperation using the natural gains in the communication channels.

The steady state of the repeated power trading game for an unchanged radio environment (channel) with RL is analyzed next. Figure 7 shows the results of the repeated game played for power trading between the relays. The figure shows the dynamic adaptation of the pricing index κ_i by the relays and the corresponding rewards Ω_i (scaled by $1000\Omega_i + 3$ to display in the same figure) during the successive games. As observed, the dominant relay R_2 wins the game initially and then tries to maximize its profit by increasing κ_2, and at one point (during the 11^{th} sequence) the cost associated with R_2 exceeds the cost of R_3 (i.e. $C_2 > C_3$) and thus R_3 wins the game. At this point R_2 updates its $\kappa_2^{\max}$ and starts to reduce κ_2 and wins the game successively. The outcomes of the game thereafter remain the same until the environmental conditions change. Once the environmental conditions change the game goes through another transient stage and converges again. Figure 8 shows the convergence of the game by observing the reward vs price-index trajectories. As we observe; R_1 does not win at any time due to its weak strategy; R_3 wins the game once (due to the greediness of R_2); R_2 wins the game initially due to its dominant strategy, increases its profit subsequently, loses once to R_3, and then after reducing its pricing index it keeps winning successively until the environmental conditions change.

From the results we see that the RL process leads to the second price auction model result by rewarding the winner with the second best bidder's reward (in this case with a small difference due to the finite values of ω). Therefore, one could argue that the second price auction model could be adopted here directly without RL which would maximize the reward for the winner. However, in this case it would not provide the chance for the losers of the game to win during the subsequent cooperations without RL. Therefore RL plays an important role to adapt the bids and maximize the rewards.

Finally, we present the simulation results for a time varying channel with varying channel gains. Figure 9 presents the simulation results for the rewards Ω_i for the case of time varying channels. In the figure we observe two time durations where the channel is almost static and the corresponding advantaged R_i uses RL technique to maximize the reward. At the same time the figure also shows a time duration where the channel changes rapidly and hence the relays are unable to learn about the other relays' bids due to insufficient training duration in the learning process, however the advantaged node at every instance during this time wins the game. A point to note here is that, as mentioned in Section VII, we observe from the figure that the game converges to an equilibrium point (i.e. not more than one advantaged player wins the game) at any instance of the time, the equilibrium point changes for different cooperation requests when κ_i or the channels are varied, and the equilibrium point is unchanged for a static (slowly varying) channel when κ_i is unchanged.

X. Conclusions

A power trading business model with QoS constraints for wireless communications was presented using cooperative communications and game theory with the first price auction model. The business model considers the bit error rate and the total delay as QoS constraints, and an example was provided for BPSK communication in Rayleigh fading channel and for a Poisson-exponential traffic model at the relays. The same business model however could be easily adopted for a generic communication model and a traffic model. Reinforcement learning was used by the competing relay nodes in order to maximize the reward for the repeated games. Our results showed that the relay node with a better wireless channel and a geographic position (hence having a dominant strategy) wins the game, and is able to maximize its reward using reinforcement learning for successive cooperations with sufficient learning time. The Nash equilibrium was also proved for the proposed game model even when reinforcement learning is considered at the relay nodes. We also showed that the natural gains from the wireless communication channels can be converted into economic profits based on the proposed business model which we consider is a very attractive feature. The overhead cost related to the bidding process was also considered and it turns out the the source node requires to send a minimum data length through the selected relay in order to gain from power trading. The design of the protocol suite for this application is an ongoing task which would then allow us to quantify the true overhead cost. Future research directions in this topic are to consider multihop scenarios with multiple source terminals where the relay nodes would then have the choice to choose the source node that would maximize its reward, giving us a multiple-buyer multiple-seller auction model.

REFERENCES

[1] K. J. Ray Liu, et.al, *Cooperative Communications and Networking*, Cambridge University Press, 2009.

[2] Drew Fudenberg, '*Game Theory*', MIT Press, 1991, Cambridge.

[3] B. Wang, et. al, 'Distributed Relay Selection and Power Control for Multiuser Cooperative Communication Networks Using Stackelberg Game', *IEEE Transactions on Mobile Computing*, Vol.8, No.7, July 2009, pp. 975-990.

[4] J. Huang, Z. Han, M. Chiang, and H. V. Poor, 'Auction-Based Resource Allocation for Cooperative Communications, '*IEEE Journal on Selected Areas in Communications*', Vol.26, No.7, September 2008, pp. 1226-1237.

[5] A. Ghasemi, K. Faez, A New Pricing Function for Power Control Game in Wireless Data Networks, in Proc. of IEEE Vehicular Technology Conference, Fall, 2006, 25-28 Sept, Montreal.

[6] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, 'Efficient Power Control via Pricing in Wireless Data Networks', *IEEE Transactions on Communications*, Vol.50, No2, February 2002, pp. 291-303.

[7] D. Niyato, et. al, "Dynamics of Multiple-Seller and Multiple-Buyer Spectrum Trading in Cognitive Radio Networks: A Game-Theoretic Modeling Approach", *IEEE Transactions on Mobile Computing*,Vol 8,No 8,August 09, pp. 1009-1022.

[8] J. Rodriguez,..S. Kandeepan, et. al, Cognitive radio and cooperative strategies for power saving in multi-standard wireless devices, In Proc. of Future Network and Mobile Summit, 16-18, June, 2010, Florence.

[9] T. S. Rappaport, *Wireless Communications: Principles and Practice* 2Ed, Prentice Hall, 2002.

[10] S. Kandeepan et. al, 'Delay Analysis of Cooperative Communication with Opportunistic Relay Access', in Proc. of IEEE Vehicular Technology Conference, 15-18 May 2011, Budapest.

[11] Mathematics Tutorial, '*University of Toronto*', http://www.economics.utoronto.ca/osborne/MathTutorial/QCCF.HTM

[12] URL of C2POWER Project http://www.ict-c2power.eu

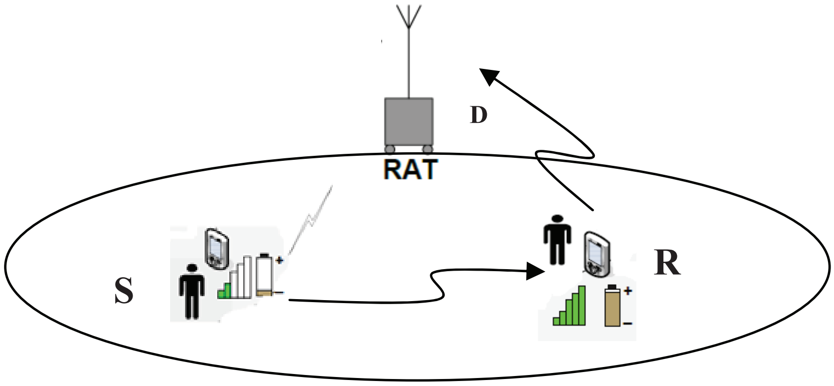

Fig. 1. Cooperative scenario for power trading

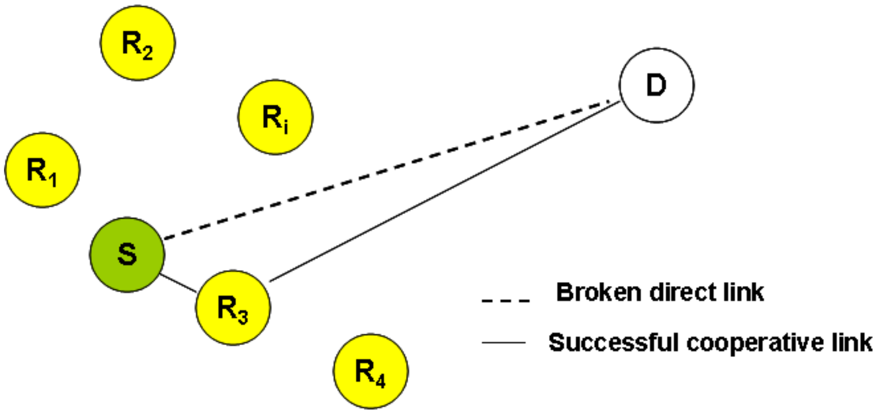

Fig. 2. Power trading game with a set of relays R_i

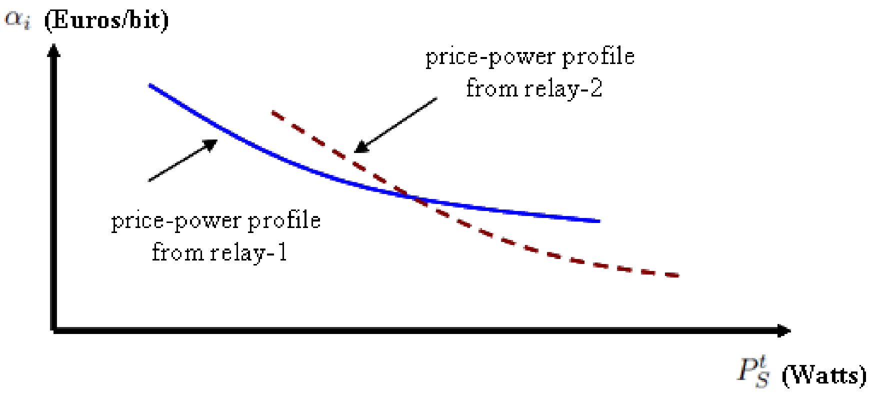

Fig. 3. The price-power profiles obtained by the relays for guaranteed QoS

R_{i-}

R_i	$S \rightarrow R_i$	$S \rightarrow R_{i-}$
$S \rightarrow R_i$	$\Lambda(\kappa_i - 1)P_R^t$	0
$S \rightarrow R_{i-}$	0	$\Lambda(\kappa_{i-} - 1)P_R^t$

Fig. 4. Expected payoffs/rewards for the relay nodes from power trading

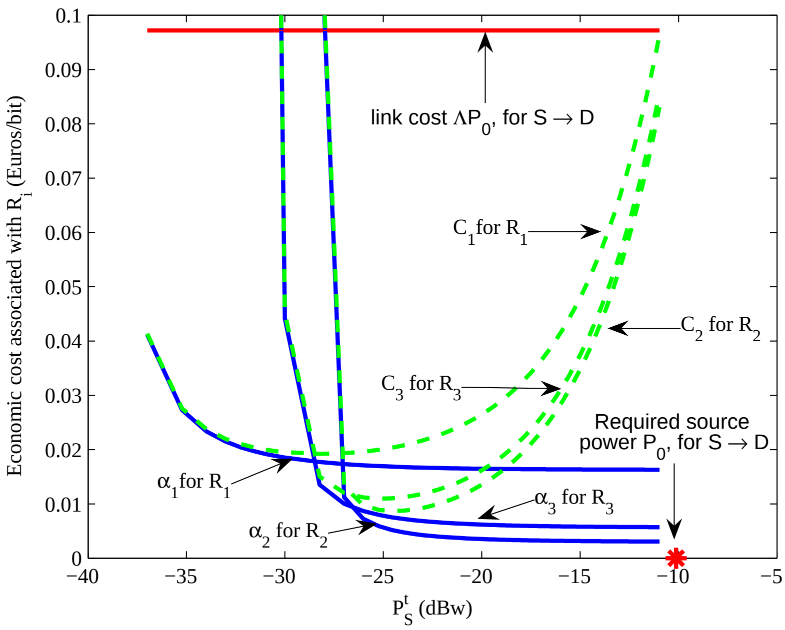

Fig. 5. The cost function C_i and the price-power profiles α_i for the network with $\kappa_i = 1, \gamma_{SR} = 1.8, \gamma_{RD} = 2.3, \gamma_{SD} = 2.7$

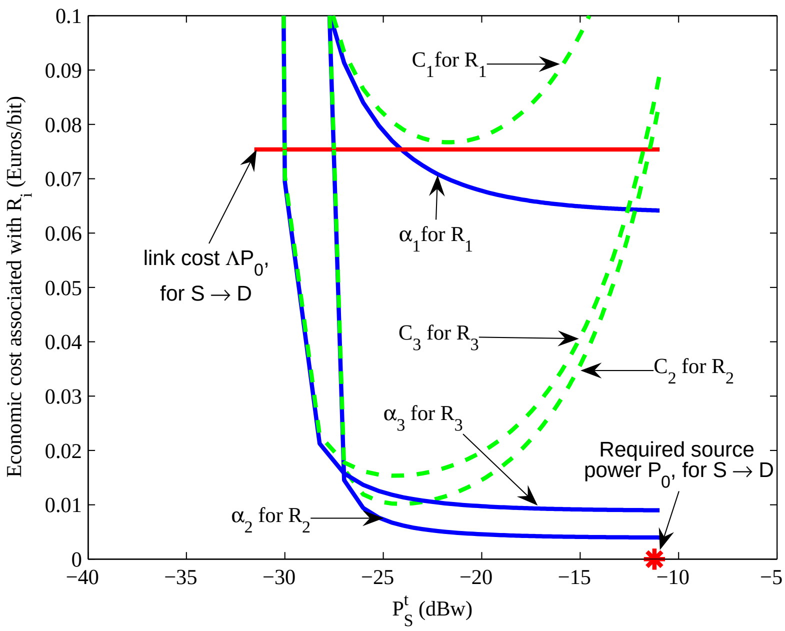

Fig. 6. The cost function C_i and the price-power profiles α_i for the network with $\kappa_i = [1.2, 1.1, 1.3], \gamma_{SR_i} = [3.8, 1.8, 1.8], \gamma_{R_iD} = [2.8, 2.4, 2.4], \gamma_{SD} = 3.1$

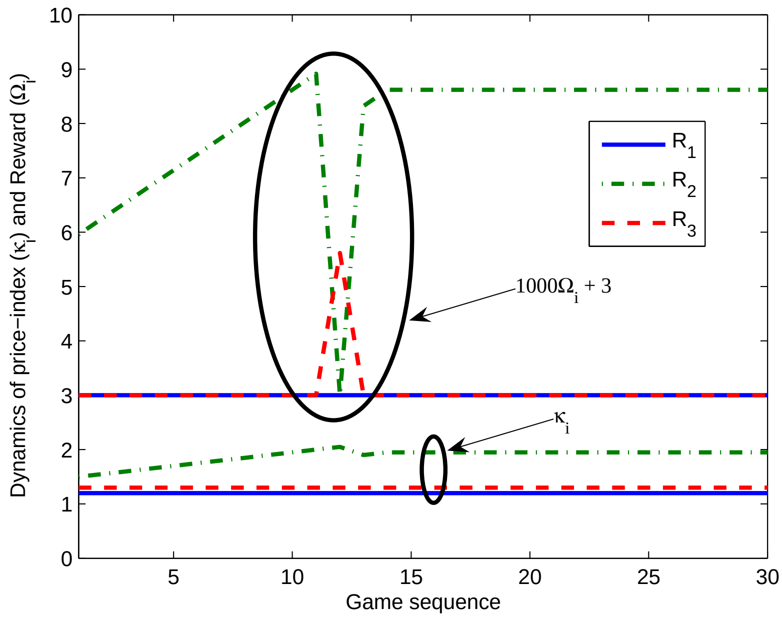

Fig. 7. Variations in cost C_i and price index κ_i for repeated cooperations for similar network parameters as in Figure 6, and $\omega_i \in \{0.05, 0.15\}$.

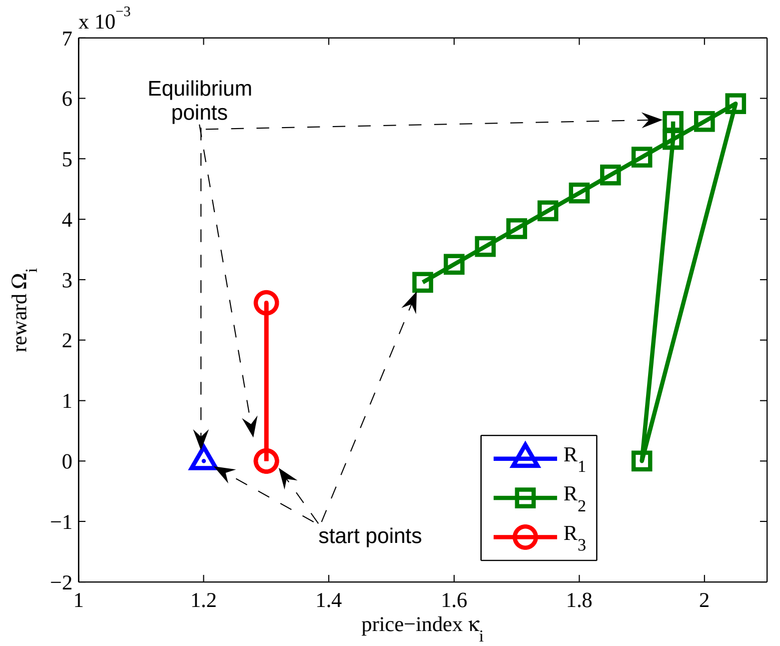

Fig. 8. Reward vs Price-Index trajectories for the power trading game with similar network parameters as in Figure 6 and $\omega_i \in \{0.05, 0.15\}$.

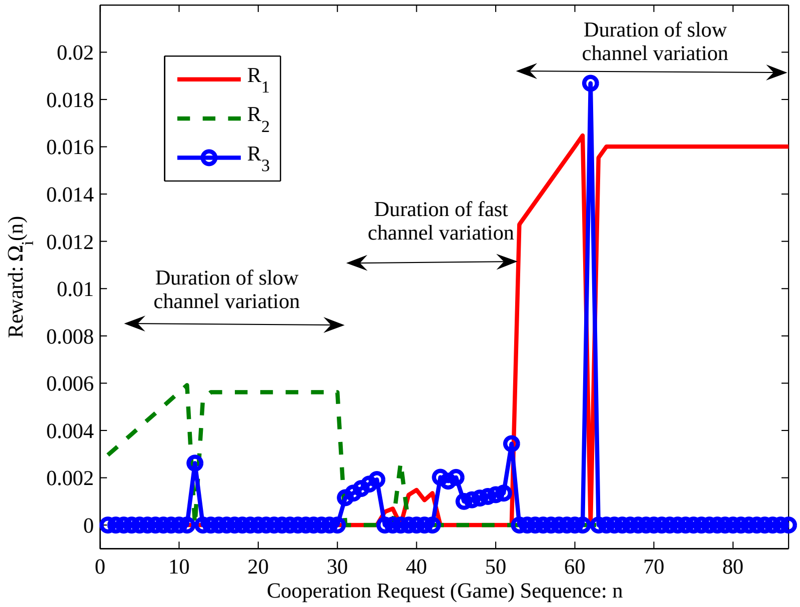

Fig. 9. Variations in the rewards Ω_i for repeated cooperations for time varying wireless channels and $\omega_i \in \{0.05, 0.15\}$, note that the variation in the channel is implemented by randomly varying γ_{SR_i} and γ_{R_iD}.