TRANSACTIONS ON EMERGING TELECOMMUNICATIONS TECHNOLOGIES
Trans. Emerging Tel. Tech. 2012; **23**:378–392
Published online 28 February 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/ett.1546

RESEARCH ARTICLE

A utility minimization approach for energy-aware cooperative content distribution with fairness constraints

Elias Yaacoub[1]*, Lina Al-Kanj[2], Zaher Dawy[2], Sanaa Sharafeddine[3], Fethi Filali[1] and Adnan Abu-Dayya[1]

[1] QU Wireless Innovations Center (QUWIC), Qatar Science and Technology Park, Doha, Qatar
[2] Department of Electrical and Computer Engineering, American University of Beirut, Beirut, Lebanon
[3] Department of Computer Science and Mathematics, Lebanese American University, Beirut, Lebanon

ABSTRACT

Cooperation between mobile terminals (MTs) with the objective of energy minimization is studied. The purpose is to distribute a content of common interest to collaborating MTs while ensuring a reduced energy consumption. To reach this goal, the content is sent to selected MTs on a long-range link. Then, it is forwarded to other MTs on short-range mobile-to-mobile links. The problem is formulated as an optimization problem, and the optimal solution is shown to consist of sending the content to a single MT on the long-range link and of having that MT distribute it on the short-range links. This leads to an unfair energy consumption for the selected MT. Thus, to ensure fairness in energy consumption, a low complexity utility minimization algorithm is proposed. Using the appropriate utilities, the algorithm can be used to implement the optimal greedy energy minimization solution or to ensure different degrees of fairness in energy consumption. Practical constraints concerning the centralised and distributed implementations of the proposed algorithm are also discussed. Simulation results show that significant energy savings can be achieved with the proposed approach compared with the non-collaborative case. In addition, a tradeoff between fairness and energy savings is achieved depending on the utility selected. Copyright © 2012 John Wiley & Sons, Ltd.

KEY WORDS

energy minimization; mobile-to-mobile cooperation; short-range communications; proportional fairness; min–max fairness

***Correspondence**

Elias Yaacoub, QU Wireless Innovations Center (QUWIC), Qatar Science and Technology Park, P.O.Box 5825 Doha, Qatar.
E-mail: eliasy@quwic.com

Received 13 July 2011; Revised 12 October 2011; Accepted 12 December 2011

1. INTRODUCTION

The increase in power demand of future mobile terminals (MTs) because of the high throughput and low latency requirements of emerging multimedia services is one of the major challenges towards the development of next generation 4G wireless networks. In fact, studies show that the high energy consumption of battery-operated MTs will be one of the main limiting factors for future wireless communication systems. Emerging multimedia applications that require the MTs' wireless interfaces to be active for long periods while downloading large data sizes require batteries with longer lifetime than what existing battery technologies can provide. To tackle this limitation, mechanisms to reduce energy consumption appear extensively in the literature, for example, see [1–4]. These mechanisms mainly rely on the fact that MTs with multiple wireless interfaces are becoming common in next generation wireless networks. This results in a heterogeneous network architecture. An interesting scenario that is attracting a lot of research interest is the case where MTs support multiple radio access technologies (RAT), and the best RAT to serve an MT is selected according to certain criteria, for example, as in [5]. Another scenario consists of having a heterogeneous network architecture with MTs that actively use two wireless interfaces: one to communicate with the base station (BS) or access point over a long-range (LR) wireless technology (e.g., UMTS/HSPA, WiMAX, or LTE) and one to communicate with other MTs over a short-range (SR) wireless technology [e.g. Bluetooth or wireless local area network (WLAN)]. Hence, the throughput and power limitations of a given wireless technology can be overcome by allowing cooperation among MTs over other wireless interfaces [6, 7]. It is this latter scenario that is investigated in this paper.

Cooperative wireless networks proved to have a lot of advantages in terms of increasing the network throughput [8–13], extending the network coverage [9, 14], decreasing the end-user communication cost [15, 16], decreasing the file download time [10–12] and decreasing energy consumption at MTs [17, 18]. The integrated cellular and ad hoc relaying architecture integrates an ad hoc component into a cellular system by placing stationary special-purpose relay nodes to help improve network throughput [19, 20]. The integrated cellular and ad hoc multicast architecture presents an integrated cellular and ad hoc multicast to increase the cellular multicast throughput through the use of ad hoc relays that are MTs themselves [8]. In the unified cellular and ad hoc network architecture [9], the MTs use their WLAN interface to increase the coverage of a wireless wide area network and to enhance the network throughput. The authors in [10] present a cooperative mobile-to-mobile (M2M) file dissemination architecture over a UMTS wireless interface to increase the network throughput and decrease the file download time. In [15], an MT is assumed to be connected to several wireless networks with different characteristics in terms of bandwidth, packet loss probability and transmission cost. Because of the complexity of the optimization framework, a near-optimal solution shows a reduction in end-user cost while meeting the distortion and delay constraints. In [13], an optimization framework for cooperative relay node selection in heterogeneous wireless communication networks is presented, and a suboptimal cooperative relay node selection algorithm is proposed. The considered heterogeneous network scenario assumes that BSs communicate with MTs over a UMTS wireless interface, whereas MTs communicate with each other over a WiMAX wireless interface. In [21], data substream distribution and energy consumption are studied in an optimised way using integer linear programming. The energy minimization problem is solved using a mathematical solver, e.g., CPLEX, for both unicasting and multicasting on the SR, assuming the same energy consumption on the LR and SR links. The benefits of collaboration for content distribution have been even investigated in wired networks, for example, in [22], where broadband access sharing is studied.

The advantages of cooperative wireless networks are particularly important for cooperative M2M video streaming. For example, the Cooperating ad Hoc networking to sUpport Messaging [18, 23, 24] and Collaborative Streaming among Mobiles [16] architectures assume that all users are interested in the same video that is divided into multiple descriptions. In the CHUM architecture, each MT randomly selects and pulls a video description through an LR cellular link and multicasts it to all members in its cooperation group that is formed in an ad hoc manner. In [1], a cooperative network architecture composed of an LR link technology and an SR link technology is presented to reduce energy consumption among MTs during real-time video streaming. Results show promising opportunities to decrease the total energy consumed by increasing the number of collaborative MTs. In [25], preliminary experimental analysis for a collaborative video streaming architecture using test bed implementation is presented. A group of MTs interested in the same video are connected to a WLAN access point through which they pull one of the video descriptions that they share with other MTs using their Bluetooth interface. This collaborative scheme proved to be more energy efficient than pulling all the video over the WLAN interface. A more comprehensive study is conducted in COMBINE [26] where experimental results are presented for a test bed composed of a general packet radio service LR interface and a WLAN SR interface.

The previous references do not present the energy minimization solution in closed form, neither do they include fairness consideration as part of the problem formulation or solution approach. In this paper, we present the optimal solution for energy minimization in content distribution with M2M collaboration in a single cluster of cooperating MTs. The problem is formulated in a general setup with different possible wireless technologies on the LR and SR. Scenarios with multicasting and unicasting on the SR links are studied. The optimal solution is shown to be unfair because it consists of sending all the data to a single MT on the LR. To add fairness to the energy minimization problem, we present a low complexity algorithm that performs utility minimization. With an appropriate choice of the utility, the algorithm can achieve the unfair greedy energy minimization solution. However, the main purpose of the algorithm is to be used in conjunction with utilities leading to fairness in energy consumption. The algorithm can be used with utilities ensuring min–max fairness or corresponding to a game theoretical formulation where MTs are assumed to play a bargaining game to reach the Nash bargaining solution (NBS).

This paper is organised as follows. The system model is presented in Section 2. The network energy minimization formulation and optimal solution are presented in Section 3. In Section 4, utility minimization is formulated, and a low complexity utility minimization algorithm is presented. Different utilities that can be used with the proposed algorithm are discussed in Section 5. Several simulation scenarios are studied and analysed in Section 6. Practical considerations concerning MT grouping into cooperating clusters in addition to feedback overhead are discussed in Section 7. Future extensions of this work are presented in Section 8. Finally, conclusions are drawn in Section 9.

2. SYSTEM MODEL

The system model adopted in this work is depicted in Figure 1. The design consists of a number K of cooperating MTs in the range of a BS. The BS is connected via wired LAN to the server that holds the content. Terminals

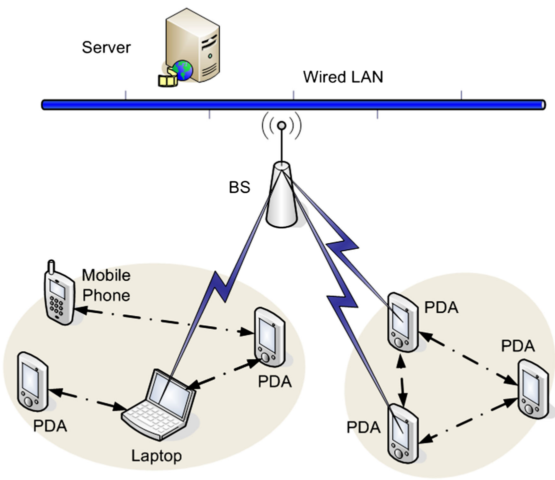

Figure 1. General system model.

can communicate with each other over SR links. Unicasting is considered on the LR. This allows the BS to transmit at the rate supported by each MT instead of forcing all MTs to receive at the rate of the MT having the worst LR channel gain, as is the case in LR multicasting. On the SR, both unicasting and multicasting are investigated in this paper.

In a traditional setup, the server either separately streams the complete content to each requesting MT or multicasts the content once to all requesting MTs. In both cases, the communication interface of each MT remains active for the whole reception duration, which depends on the length of the content and the transmission rate. This results in high energy consumption because of the required processing during data reception.

In this work, we assume the establishment of an M2M network between the MTs over SR wireless links that are more energy efficient than the LR wireless link. In this scheme, the content is divided into N parts. If K MTs are requesting the content, then each will be receiving a subset of the N data parts from the server. Over the SR wireless links, each MT receives the remaining data subsets from the other cooperating MTs in the M2M network. Being exchanged over an energy efficient SR wireless technology, the SR exchanged subsets require lower reception power at the communication interface of the MTs. However, an additional overhead in this case is that each MT needs to spend additional energy to transmit its received data subset to the other cooperating MTs.

In this work, we consider a single cluster of connected MTs where each MT can communicate with every other MT in the cluster. Furthermore, a low mobility scenario is adopted. Thus, it can be assumed that the channel conditions on the MT–MT links and BS–MT links remain approximately constant during the distribution of a single file. Furthermore, we assume that all collaborating MTs exchanging a certain file remain available during the exchange; that is, an MT does not leave the cluster while it is in the midst of a collaborative content exchange process.

2.1. Channel model

The channels on the LR and SR links are assumed to be orthogonal and are modelled by pathloss, shadowing and fading. Thus, the received power P_r can be linked to the transmitted power P_t by a pathloss model as in [27]:

$$\frac{P_r}{P_t}(\mathrm{dB}) = \underbrace{10\log_{10}\kappa - 10\nu\cdot\log_{10} d}_{\text{distance based pathloss}} + \underbrace{h_{\mathrm{dB}} + f_{\mathrm{dB}}(a)}_{\text{random variables}} \tag{1}$$

where κ is a unitless constant that depends on the antenna characteristics and the average channel attenuation, ν is the path loss exponent, d is the distance where the received power is calculated, h is a Gaussian random variable representing shadowing or slow fading having a zero mean and a variance $\sigma^2_{h_{\mathrm{dB}}}$ and f is a random variable representing Rayleigh fading with a Rayleigh parameter a.

2.2. Data rates

Given for each MT: the transmit power P_t the sender is transmitting with the pathloss, shadowing and fading on the channel, and the thermal noise power σ^2, the received signal-to-noise ratio (SNR) γ can be calculated following $\gamma = \frac{P_r}{\sigma^2}$. Given the target bit error rate P_e and the SNR, the bit rates on the LR and SR links can be calculated according to the following:

$$R = B\cdot\log_2(1+\beta\gamma) \tag{2}$$

In (2), B is the passband bandwidth of the channel, and β is called the SNR gap. It indicates the difference between the SNR needed to achieve a certain data transmission rate for a practical multilevel quadrature amplitude modulation system and the theoretical Shannon limit [27, 28]. It is given by $\beta = \dfrac{-1.5}{\ln(5P_e)}$.

2.3. Parameters and variables

The parameters that affect the energy consumption in the chosen system model are the following:

- K: the number of requesting MTs.
- S_T: the size of the content to be sent in one transmission interval. This depends on the content server transmission rate.
- N: the number of parts the content is divided into. Thus, the size of one part is S_T/N.
- $R_{\mathrm{L},k}$: transmission rate on the LR links from the BS to MT k.
- $R_{\mathrm{S},kj}$: transmission rate on the SR links from MT k to MT j.
- $P_{\mathrm{L,Rx}}$: power consumed by the MT during reception on the LR link.
- $P_{\mathrm{S,Rx}}$: power consumed by the MT during reception on the SR links.

Trans. Emerging Tel. Tech. **23**:378–392 (2012) © 2012 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

- $P_{S,Tx,kj}$: power consumed by MT k while transmitting to MT j on the SR links.
- Decision variables: the decision variables are x_k, with x_k an integer variable that determines the number of parts received by MT k over the LR link.

It should be noted that $P_{L,Rx}$, $P_{S,Rx}$ and $P_{S,Tx,kj}$ correspond to the power consumed by the MT, that is, energy drained per second from its battery, during reception and transmission, respectively. They are not to be confused with P_r and P_t; the respective receive and transmit powers over the air measured at the antenna. It should be noted that $P_{S,Tx,kj}$ can be expressed as

$$P_{S,Tx,kj} = P_{S,Tx,0} + P_{t,kj} \tag{3}$$

where $P_{S,Tx,0}$ corresponds to the power consumed by the circuitry of the MTs during transmission on the SR links, and $P_{t,kj}$ corresponds to the power transmitted over the air on the SR links from MT k to MT j.

Power consumption of the BS is not considered because the interest in this paper is in the battery life of the MTs. This can be justified by the fact that most BSs rely on power line cables and not on batteries and thus do not have as stringent power limitations as the MTs.

3. OPTIMAL ENERGY MINIMIZATION

In this section, the energy minimization problem is formulated, and the optimal solution is presented. Unicasting and multicasting are considered on the SR, both with rate adaptive and power adaptive transmissions.

3.1. Energy minimization with unicasting

Considering K requesting MTs interested in downloading a content from a server on the internet in a cooperative manner, and assuming that the content is divided into N parts of equal size and importance, the time t_k required to send x_k parts over a link with rate R_k is

$$t_k = \frac{x_k \cdot S_T}{N \cdot R_k} \tag{4}$$

Multiplying the power drained from the MT battery by the the time needed, then the expression of the energy consumed can be obtained. Consequently, the energy consumed by MT k is

$$E_k = \frac{x_k \cdot S_T}{N \cdot R_{L,k}} P_{L,Rx} + \frac{x_k \cdot S_T}{N} \sum_{j=1, j \neq k}^{K} \frac{P_{S,Tx,kj}}{R_{S,kj}} + \frac{S_T}{N} P_{S,Rx} \sum_{j=1, j \neq k}^{K} \frac{x_j}{R_{S,jk}} \tag{5}$$

In (5), the first term corresponds to the energy consumed by MT k for receiving x_k parts over the LR, the second term corresponds to sending the data received by MT k on the LR to the other MTs on the SR and the last term corresponds to the energy consumed by MT k while receiving the data parts from the other MTs on the SR, that is, receiving x_j data parts from each MT j on the SR.

The total energy consumed by the requesting MTs is

$$E_{coop} = \sum_{k=1}^{K} E_k \tag{6}$$

By substituting (5) in (6), the total energy consumed by the requesting MTs is

$$E_{coop} = \frac{S_T}{N} \cdot P_{L,Rx} \sum_{k=1}^{K} \frac{x_k}{R_{L,k}} + \frac{S_T}{N} \cdot \sum_{k=1}^{K} \sum_{j=1, j \neq k}^{K} \left(\frac{x_k \cdot P_{S,Tx,kj}}{R_{S,kj}} + \frac{x_j \cdot P_{S,Rx}}{R_{S,jk}} \right) \tag{7}$$

After exchanging the dummy indices in the double summation for the last term of (7), we can write

$$E_{coop} = \frac{S_T}{N} \cdot P_{L,Rx} \sum_{k=1}^{K} \frac{x_k}{R_{L,k}} + \frac{S_T}{N} \cdot \sum_{k=1}^{K} \sum_{j=1, j \neq k}^{K} \frac{x_k \cdot (P_{S,Tx,kj} + P_{S,Rx})}{R_{S,kj}} \tag{8}$$

The objective is to minimise the total energy consumption of the MTs when all of them are cooperating. A constraint that guarantees that the whole content is transmitted on the LR should be added. The optimization problem can be formulated as follows:

$$\min_{\mathbf{x}} \; E_{coop} = \frac{S_T}{N} \cdot P_{L,Rx} \sum_{k=1}^{K} \frac{x_k}{R_{L,k}} + \frac{S_T}{N} \cdot \sum_{k=1}^{K} \sum_{j=1, j \neq k}^{K} \frac{x_k \cdot (P_{S,Tx,kj} + P_{S,Rx})}{R_{S,kj}} \tag{9}$$

such that

$$\sum_{k=1}^{K} x_k = N \tag{10}$$

$$\mathbf{x} \in Z_{+}^{K} \tag{11}$$

The problem in (9) is a linear integer optimization problem because the objective function E_{coop} is linear, and the constraints are also linear as shown by (10) that guarantees that the whole content is transmitted on the LR and by (11) that guarantees that the decision variable is integer and positive. Such an integer linear programme can be solved using an integer linear programming solver, that is, as in [21]. However, we prove that the energy is minimised when the data is sent to an MT k^*, and we determine the expression of k^*.

The energy expression in (5) corresponds to the energy consumed by MT k to transmit the parts allocated to it on the LR in addition to receiving the remaining parts allocated to the other MTs.

However, the total energy consumed in the network in order to distribute the x_k parts allocated to MT k is given by

$$E_k^{(\text{dist})} = \frac{x_k \cdot S_T}{N \cdot R_{\text{L},k}} P_{\text{L,Rx}} + \frac{x_k \cdot S_T}{N} \sum_{j=1, j\neq k}^{K} \frac{P_{\text{S,Tx},kj}}{R_{\text{S},kj}} + \frac{S_T}{N} P_{\text{S,Rx}} \sum_{j=1, j\neq k}^{K} \frac{x_k}{R_{\text{S},kj}} \tag{12}$$

where the first term corresponds to the energy consumed by MT k on the LR, the second term corresponds to the energy consumed by MT k to transmit the data on the SR and the last term corresponds to the energy consumed by all other MTs to receive the data transmitted by MT k on the SR. The definition in (12) is used to prove the following theorem:

Theorem 1. *To minimise the energy consumed when distributing a single data part, this part should be sent to an MT k^* satisfying*

$$k^* = \arg\min_k \frac{S_T}{N} \cdot \left(\frac{P_{\text{L,Rx}}}{R_{\text{L},k}} + \sum_{j=1, j\neq k}^{K} \frac{P_{\text{S,Tx},kj}}{R_{\text{S},kj}} + \sum_{j=1, j\neq k}^{K} \frac{P_{\text{S,Rx}}}{R_{\text{S},kj}} \right) \tag{13}$$

Proof. The energy to distribute a data part by sending it to MT k on the LR corresponds to (12) divided by x_k; that is, the energy to distribute a data part by sending it to MT k is given by (12), with x_k replaced by 1. In this case, it is clear that minimum energy is consumed in the network by sending the data part to the MT that satisfies (13). ☐

Lemma 1. *The total energy consumed to distribute all parts in the network is equal to the total energy consumed by the MTs, that is, $E_{coop} = \sum_{k=1}^{K} E_k = \sum_{k=1}^{K} E_k^{(\text{dist})}$.*

Proof. In fact, the energy consumed to distribute the parts allocated to MT k is given by (12). Thus, the energy consumed to distribute all data parts is the energy required to distribute the data parts allocated to all MTs. This is obtained by taking the summation over k of the expression given in (12), which leads directly to the expression in (8) corresponding to the total energy consumed by the MTs. ☐

Theorem 2. *The optimal energy minimization solution consists of sending all the data parts to a single MT k^* on the LR, and MT k^* is in charge of distributing the whole content on the SR.*

Proof. From Theorem 1, the energy consumed to distribute the first data part is minimised by sending it to MT k^* satisfying (13). Similarly, the energy to distribute another data part, say the second part, is minimised by sending that part to MT k^* satisfying (13) since (13) corresponds to the minimum energy consumed by distributing any data part. Hence, by recurrence, the third, fourth, ... and Nth data parts should be sent to MT k^* satisfying (13) to minimise the energy consumed in the network while distributing the data. From Lemma 1, it is known that the total energy for content distribution corresponds to the sum of the energies consumed by the MTs in the network. Consequently, from Theorem 1 and Lemma 1, it is straightforward to conclude that the optimal solution minimizing the energy consumption in the network consists of sending all data parts on the LR to MT k^* satisfying (13), and MT k^* is in charge of distributing the data on the SR links. ☐

The aforementioned formulation and solution considered the general unicasting case. In the succeeding subsections, this formulation is customised to special cases of practical interest: adaptive rate control and adaptive power control.

3.1.1. Rate adaptive transmission.

$P_{\text{S,Tx},kj}$ is the power consumed by MT k while transmitting to MT j on the SR links with a transmit power $P_{t,kj}$ as given by (3). In the case of adaptive rate control, the MT transmit power is constant, that is, $P_{t,kj} = P_t$ and hence, $P_{\text{S,Tx},kj} = P_{\text{S,Tx}}$. Consequently, the rate $R_{\text{S},kj}$ on the SR link between MTs k and j is the rate achievable with the transmit power P_t. It is varied adaptively depending on the channel conditions between MTs k and j. High data rates result in low energy per bit consumption, thus leading to a gain in total energy consumption. For example, the WLAN technologies apply rate control [29].

In the adaptive rate control scenario, the energy consumed in the network (8) becomes

$$E_{\text{coop}} = \frac{S_T}{N} \cdot P_{\text{L,Rx}} \sum_{k=1}^{K} \frac{x_k}{R_{\text{L},k}} + \frac{S_T}{N} (P_{\text{S,Tx}} + P_{\text{S,Rx}}) \cdot \sum_{k=1}^{K} \sum_{j=1, j\neq k}^{K} \cdot \frac{x_k}{R_{\text{S},kj}} \tag{14}$$

The expression of the optimal solution (13) in the case of adaptive rate control is given in Table I.

3.1.2. Power Adaptive Transmission.

In the case of adaptive power control, the MTs communicate at a constant rate on the SR $R_{\text{S},kj} = R_{\text{S}}$. The transmit power $P_{t,kj}$ is varied adaptively depending on

 Trans. Emerging Tel. Tech. **23**:378–392 (2012) © 2012 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

Table I. Optimal solution in the different scenarios.

Parameter	Value
Unicasting - rate adaptive	$k^* = \arg\min_k \frac{S_T}{N} \cdot \left(\frac{P_{\mathrm{L,Rx}}}{R_{\mathrm{L},k}} + (P_{\mathrm{S,Tx}} + P_{\mathrm{S,Rx}}) \sum_{j=1, j\neq k}^{K} \frac{1}{R_{\mathrm{S},kj}} \right)$
Unicasting - power adaptive	$k^* = \arg\min_k \frac{S_T}{N} \cdot \left(\frac{P_{\mathrm{L,Rx}}}{R_{\mathrm{L},k}} + \frac{\left(\sum_{j=1, j\neq k}^{K} P_{\mathrm{S,Tx},kj} \right) + (K-1) P_{\mathrm{S,Rx}}}{R_{\mathrm{S}}} \right)$
Multicasting - rate adaptive	$k^* = \arg\min_k \frac{S_T}{N} \cdot \left(\frac{P_{\mathrm{L,Rx}}}{R_{\mathrm{L},k}} + \frac{P_{\mathrm{S,Tx}} + (K-1) P_{\mathrm{S,Rx}}}{\min_j R_{\mathrm{S},kj}} \right)$
Multicasting - power adaptive	$k^* = \arg\min_k \frac{S_T}{N} \cdot \left(\frac{P_{\mathrm{L,Rx}}}{R_{\mathrm{L},k}} + \frac{\max_j P_{\mathrm{S,Tx},kj} + (K-1) P_{\mathrm{S,Rx}}}{R_{\mathrm{S}}} \right)$

the channel conditions between MTs k and j to achieve the target data rate R_{S}. MTs that are in proximity of each other will communicate with lower power than MTs that are further apart. This will result in a reduction of consumed energy. Some technologies such as Bluetooth apply power control [30].

In the adaptive power control scenario, the energy consumed in the network (8) becomes

$$E_{\mathrm{coop}} = \frac{S_T}{N} \cdot P_{\mathrm{L,Rx}} \sum_{k=1}^{K} \frac{x_k}{R_{\mathrm{L},k}} + \frac{S_T}{N} \cdot \sum_{k=1}^{K} \sum_{j=1, j\neq k}^{K} \frac{x_k \cdot (P_{\mathrm{S,Tx},kj} + P_{\mathrm{S,Rx}})}{R_{\mathrm{S}}} \tag{15}$$

The expression of the optimal solution (13) in the case of adaptive power control is given in Table I.

3.2. Energy minimization with multicasting

The results of Section 3.1 correspond to unicasting, where each MT transmits the data on the SR to each other MT individually. With multicasting, the transmitting MT sends the data once to all the other cooperating MTs.

In the case of multicasting with adaptive rate control, every MT transmits with a data rate that is equal to the minimum one among its neighbours so that all neighbours can receive the data with high reliability. Thus, $R_{\mathrm{S},kj} = \min_{j'} R_{\mathrm{S},kj'}$, $j' = 1, \ldots, K$ and $j' \neq k$. In this case, the energy is expressed as

$$E_{\mathrm{coop}} = \frac{S_T}{N} \cdot P_{\mathrm{L,Rx}} \sum_{k=1}^{K} \frac{x_k}{R_{\mathrm{L},k}} + \frac{S_T}{N} (P_{\mathrm{S,Tx}} + (K-1) P_{\mathrm{S,Rx}}) \sum_{k=1}^{K} \frac{x_k}{\min_j R_{\mathrm{S},kj}} \tag{16}$$

In the case of multicasting with adaptive power control, every MT transmits with a power high enough such that all its neighbours are able to achieve the rate R_{S}. Thus, $P_{\mathrm{S,Tx},kj} = \max_{j'} P_{\mathrm{S,Tx},kj'}$, $j' = 1, \ldots, K$ and $j' \neq k$. In this case, the energy consumed is expressed as

$$E_{\mathrm{coop}} = \frac{S_T}{N} \cdot P_{\mathrm{L,Rx}} \sum_{k=1}^{K} \frac{x_k}{R_{\mathrm{L},k}} + \frac{S_T}{N} \cdot \sum_{k=1}^{K} \frac{x_k \cdot \max_j P_{\mathrm{S,Tx},kj}}{R_{\mathrm{S}}} + \frac{S_T}{N} \cdot (K-1) P_{\mathrm{S,Rx}} \sum_{k=1}^{K} \frac{x_k}{R_{\mathrm{S}}} \tag{17}$$

Following the same approach as in Section 3.1, it can be proven that the optimal solution consists of sending all the data to a single MT on the LR link and that MT can be determined in closed form. The closed form results for multicasting with rate adaptive and power adaptive transmissions are shown in Table I.

3.3. Energy consumption without cooperation

The total energy consumption spent when no cooperation takes place is

$$E_{\mathrm{No\text{-}coop}} = S_T \cdot P_{Rx} \sum_{k=1}^{K} \frac{1}{R_{L,k}} \tag{18}$$

This corresponds to the case where each MT receives the whole content on the LR links. The normalised energy consumption η can be calculated as follows:

$$\eta = \frac{E_{\mathrm{coop}}}{E_{\mathrm{No\text{-}coop}}} \tag{19}$$

DOI: 10.1002/ett

The value of η indicates whether the cooperation is beneficial in terms of energy consumption or not; if $\eta < 1$, then the cooperation results in a gain of energy consumption while $\eta > 1$ reflects a non-beneficial cooperation.

4. LOW COMPLEXITY ALGORITHM FOR UTILITY MINIMIZATION

The results of Section 3 show that the optimal solution is to send all content from the BS to a single MT k^*, and k^* is in charge of distributing the content to all other MTs over the SR links. Although this solution minimises the total energy consumption in the network, it is unfair to k^* because it has to spend more energy than it actually needs to receive all the content on the LR and then transmit it to the other MTs on the SR.

To deal with this problem, we propose an approach on the basis of utility minimization. The utility of each MT, U_k, is a function of its energy E_k. By an appropriate choice of the utility function, we aim to ensure more fairness in the content distribution process. In other words, the purpose of minimizing a function of the energy instead of the energy itself is to allow MTs other than k^* to take part in the SR content distribution process, while still achieving significant gains compared with the non-cooperative scenario.

4.1. Utility minimization formulation

The utility minimization problem can be formulated as follows:

$$\min_{\mathbf{x}} \; U_{\text{coop}} = \sum_{k=1}^{K} U_k(E_k) \tag{20}$$

such that

$$\sum_{k=1}^{K} x_k = N \tag{21}$$

$$\mathbf{x} \in Z_+^K \tag{22}$$

where U_{coop} is the total network utility. When the utility is set to the energy itself, that is, $U_k(E_k) = E_k$, then the utility minimization becomes a greedy energy minimization as in Section 3, and the problem becomes an integer linear programme. However, depending on the utility function, the derivation of an optimal solution for the problem might not be straightforward. For example, considering a logarithmic utility, the utility function will be concave, but the problem cannot be solved using convex optimization techniques. In fact, it will not be a convex problem because the optimization variables are integers.

To deal with this issue, we propose a low complexity suboptimal algorithm that can be used with a wide variety of utilities. With an appropriate choice of the utility, the algorithm can achieve fairness in the content distribution process or can achieve the optimal energy minimization results of Section 3.

4.2. Proposed algorithm

The proposed algorithm consists of allocating part n to MT k in a way to minimise the difference

$$\Lambda_{n,k} = U_k(E_k|\mathcal{I}_{N,k} \cup \{n\}) - U_k(E_k|\mathcal{I}_{N,k}) \tag{23}$$

where the marginal utility, $\Lambda_{n,k}$, represents the increase in the utility function U_k when part n is allocated to MT k, compared with the utility of MT k before the allocation of n. In addition, $\mathcal{I}_{N,k}$ denotes the set of parts allocated to MT k among the N available parts, such that $|\mathcal{I}_{N,k}| = x_k$, where $|\cdot|$ denotes set cardinality. The algorithm is described as follows:

- Consider the set of available parts $\mathcal{I}_N \subseteq \{1, 2, \ldots, N\}$. At the start of the algorithm, $\mathcal{I}_N = \{1, 2, \ldots, N\}$.
- **Step 1:** Find the MT that has the lowest marginal utility defined in (23) among all MTs when the first available part in $\mathcal{I}_N$ is allocated to it. In other words, for each part n, find the MT k^* such that

$$k^* = \arg\min_k \Lambda_{n,k} \tag{24}$$

- **Step 2:** Allocate part n to MT k^*: $\mathcal{I}_{N,k^*} = \mathcal{I}_{N,k^*} \cup \{n\}$
- **Step 3:** Delete part n from the set of available parts:

$$\mathcal{I}_N = \mathcal{I}_N - \{n\} \tag{25}$$

- Repeat Steps 1, 2 and 3 until all parts are allocated.

After the allocation of data parts to MTs, each MT sends the parts it received to the other MTs on the SR using a scenario depending on the utility selected, that is, either via multicasting or unicasting, using either rate adaptive or power adaptive transmissions.

4.3. Complexity analysis

The proposed algorithm allocates each data part after performing a linear search on the MTs to find the MT that minimises the marginal utility. Consequently, the total complexity of the algorithmx is $\mathcal{O}(NK)$, that is, the algorithm has linear complexity in the number of MTs and in the number of data parts, and thus could be easily implemented in real time.

5. UTILITY SELECTION

The proposed algorithm can be applied with a wide range of utility functions, thus being able to achieve various objectives, with each objective represented by a certain utility function.

Trans. Emerging Tel. Tech. **23**:378–392 (2012) © 2012 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

5.1. Greedy energy minimization

The energy in (12) corresponds to the energy consumed by MT k to receive x_k parts on the LR and to transmit these x_k parts on the SR, in addition to the energy consumed by all other MTs to receive these x_k parts. The proposed algorithm reaches the optimal energy minimizing solution when the utility used is equal to the energy consumed to distribute the parts, that is, when $U_k = E_k^{(\text{dist})}$ defined in (12).

In fact, the marginal utility in (23), when $U_k = E_k^{(\text{dist})}$, corresponds to the energy consumed when distributing one data part. Thus, $\Lambda_{n=1,k}$ corresponds to the energy consumed to distribute part $n = 1$ when it is allocated to MT k for distribution. The algorithm finds the MT that minimises $\Lambda_{n=1,k}$, that is, the MT that minimises (12) when $x_k = 1$. Hence, this MT clearly satisfies (13) and thus corresponds to the MT k^*.

For part $n = 2$, $\Lambda_{n=2,k}$ corresponds to the energy to distribute one data part when allocated to MT k if $k \neq k^*$, that is, $\Lambda_{n=2,k} = E_k^{(\text{dist})}$ with $x_k = 1$ for $k \neq k^*$. When $k = k^*$, $\Lambda_{n=2,k^*}$ corresponds to the difference between $E_{k^*}^{(\text{dist})}$ with $x_{k^*} = 2$ and $E_{k^*}^{(\text{dist})}$ with $x_{k^*} = 1$. Equivalently, this corresponds to the energy consumed to distribute one additional data part allocated to MT k^*. Hence, the MT that minimises $\Lambda_{n=2,k}$ is again the MT that minimises (12) when $x_k = 1$. Consequently, the second data part is also sent to MT k^* satisfying (13). By a similar reasoning up to $n = N$, it can be easily shown that all data parts are allocated to the same MT k^*.

The previous discussion was based on the general case represented by (13). Following the same approach, the same results can be proven for all the scenarios listed in Table I.

5.2. Proportional fairness: bargaining game model

The results of Section 3 showed that the optimal solution is unfair: all the content is sent to a single MT, and that MT is responsible for transmitting the whole content to the other cooperating MTs. Although this solution is optimal in terms of minimizing the total consumed energy, it is unfair towards the selected MT, whose energy consumption would exceed its consumption in the non-cooperative scenario.

In this section, to ensure a more fair allocation of data parts, we model the problem as a bargaining game. We consider that each MT is a player (in this section, both terms MT and player are used interchangeably) who wants to maximise its payoff, considered to be its energy savings, or equivalently, who wants to minimise its energy consumption. Cooperation is assumed between players. Consequently, players should share the resources in an optimal way, that is, a way they cannot jointly improve on. The resources to be shared are the N data parts that the content is divided into. Allocating the shared resources in a way to maximise the players' payoffs is equivalent to allocating the N data parts to MTs in a way to minimise each MT's energy, given the shares allocated to the other MTs. With each MT wanting to selfishly minimise its consumed energy, the MTs engage in a 'bargaining' process. It is a well-known result in game theory that the solution to the cooperative bargaining problem maximises the Nash product N_P [31]:

$$N_P = \prod_{k=1}^{K} (W_k(y_k) - F_k) \tag{26}$$

where y_k represents the fraction of resources allocated to player k, $W_k(y_k)$ corresponds to the payoff of player k when y_k is allocated to it and F_k is the payoff of player k in the case where no agreement is reached in the bargaining problem. In the energy minimization problem, the objective of each player is to minimise its consumed energy or, equivalently, maximise its energy savings, and thus has a payoff of $(E_{k,\text{no–coop}} - E_{k,\text{coop}})$. In case no agreement is reached, each MT obtains its data on the LR link and thus consumes $E_{k,\text{no–coop}}$, which leads to a payoff (or energy savings) of zero. Hence, the optimization problem becomes

$$\max \prod_{k=1}^{K} (E_{k,\text{no–coop}} - E_{k,\text{coop}}) \tag{27}$$

Because the logarithm is a continuous strictly increasing function, solving the problem in (27) is equivalent to finding the solution of the following problem:

$$\begin{aligned} &\ln\left(\max\left(\prod\nolimits_{k=1}^{K} (E_{k,\text{no–coop}} - E_{k,\text{coop}})\right)\right) \\ &= \max \ln\left(\prod\nolimits_{k=1}^{K} (E_{k,\text{no–coop}} - E_{k,\text{coop}})\right) \\ &= \max \sum\nolimits_{k=1}^{K} \ln\left(E_{k,\text{no–coop}} - E_{k,\text{coop}}\right) \end{aligned} \tag{28}$$

Maximising the sum in (28) is equivalent to maximising the product in (27) and is easier to implement numerically. This approach represents a notion of 'proportional fairness' in energy because it has analogies with proportional fair (PF) scheduling, a well-known resource allocation approach in wireless communications systems. PF scheduling is known to correspond to a sum of the logarithms of the user rates and represents the NBS equivalent in the rate maximisation problem [32].

Hence, when each MT wants to selfishly maximise its energy savings using the bargaining model of this section, we set the utility to $U_k = -\ln(E_{k,\text{no–coop}} - E_{k,\text{coop}})$. In this case, minimizing $\sum_{k=1}^{K} U_k$ is equivalent to maximising $\sum_{k=1}^{K} \ln(E_{k,\text{no–coop}} - E_{k,\text{coop}})$, which in turn is equivalent to maximising the product $\prod_{k=1}^{K} (E_{k,\text{no–coop}} - E_{k,\text{coop}})$, that is, the Nash product.

5.3. Altruistic utilities

Selecting the utility to be equal to the consumed energy, that is, $U_k = E_k^{(\text{dist})}$, the algorithm performs a greedy minimization of the total energy consumed in the network. This solution is shown to be unfair to one of the MTs to which all the data is forwarded on the LR link so that it distributes it on the SR links. However, setting the utility of each MT as the total energy in the network, that is, $U_k = E_{\text{coop}}$, the algorithm will lead to the same energy minimization approach as when setting $U_k = E_k^{(\text{dist})}$ because of the use of the marginal utility in (23). But, in the case when $U_k = E_{\text{coop}}$, the MT utility is forced to be equal to the network utility, and thus each MT is led to act altruistically by seeing a benefit to the whole network as its own benefit, although this solution is actually unfair to one of the MTs. Assuming the utilities can be hardwired in the mobile devices, this approach can be followed in a distributed scenario to reach the minimum energy consumption in the network, even in the framework of a bargaining game. In fact, with $U_k = E_{\text{coop}}$, maximising the Nash product is equivalent to maximising

$$\prod_{k=1}^{K} \left(E_{\text{no–coop}} - E_{\text{coop}}\right) = \left(E_{\text{no–coop}} - E_{\text{coop}}\right)^K \quad (29)$$

which is equivalent to maximising $(E_{\text{no–coop}} - E_{\text{coop}})$ or minimizing E_{coop}, thus retrieving the greedy minimization of the total consumed energy in the network through the game theoretical formulation itself.

5.4. Min–max utilities

In this section, utilities that attempt to minimise the energy consumption of the MT having maximum energy consumption are presented. We refer to them as min–max utilities. They are derived by analogy to the widely investigated problem of rate maximisation with fairness constraints, that is, [33, 34]. In the case of rate maximisation with fairness, their counterparts are referred to as max–min utilities because the objective is to maximise the minimum data rate in the network. A vector $\mathbf{R}$ of MT data rates is max–min fair if and only if, for each k, an increase in R_k leads to a decrease in R_j for some j with $R_j < R_k$ [33]. Max–min utilities lead to more fairness by increasing the priority of MTs having lower rates [34]. It was shown that max–min fairness can be achieved by utilities of the form [34]

$$U_k(R_k) = -\frac{R_k^{-\alpha}}{\alpha}, \alpha > 0 \quad (30)$$

where the parameter α determines the degree of fairness. Max–min fairness is attained when $\alpha \to \infty$ [34].

Because in the case of energy minimization the objective is the opposite, that is, to minimise the maximum energy, the minus sign in (30) is removed, and the utility used with the algorithm of Section 4.2 can be expressed as

$$U_k(E_k) = \frac{E_k^{-\alpha}}{\alpha}, \alpha > 0 \quad (31)$$

6. RESULTS AND ANALYSIS

In this section, simulation results are presented and analysed. We consider a file of size $S_T = 1$ Mbits, subdivided into $N = 25$ parts, to be transmitted to all requesting MTs. The main simulation parameters are presented in Table II. Channel parameters are obtained from [35], whereas energy consumption parameters are taken as in [36], where measurements are made with 3G communications on the LR, and 802.11b on the SR using the rate adaptive approach.

MTs are assumed to be uniformly distributed in a rectangular area of size 20m × 20m, whose origin is at a distance $d_{\text{LR}} = 400$ m from the BS. This corresponds to a scenario where, for example, a group of mobile users are in close proximity in a cafe or park that is 400 m away from the cellular BS. For the min–max utility, we set $\alpha = 10$. The BS transmit power is considered to be 37 dBm, and the MT transmit power is set to 15 dBm. The simulations are performed using Matlab, and the results are averaged over 10 000 iterations.

6.1. Average results

In this section, results averaged over the positions and channel variations of the MTs are presented. The results are shown in Figure 2 for unicasting and multicasting using different utility functions.

Figure 2 shows that the optimal solution consisting of greedy minimization of the total energy in the network leads to better results than the min–max and PF methods. However, the use of the PF and min–max approaches still allows achieving significant savings compared with the non-cooperative scenario because the value of η is significantly smaller than one, especially with the min–max approach whose performance is close to the optimal scenario.

Comparing multicasting with unicasting, it is clear in Figure 2 that multicasting gives better results than

Table II. Simulation parameters.

Parameter	Value
κ	-128.1 dB
ν	3.76
$\sigma_{h_{\text{dB}}}$	8 dB
$P_{\text{S,Tx}}$	1.425 Joules/s
$P_{\text{S,Rx}}$	0.925 Joules/s
$P_{\text{L,Rx}}$	1.8 Joules/s

Trans. Emerging Tel. Tech. **23**:378–392 (2012) © 2012 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

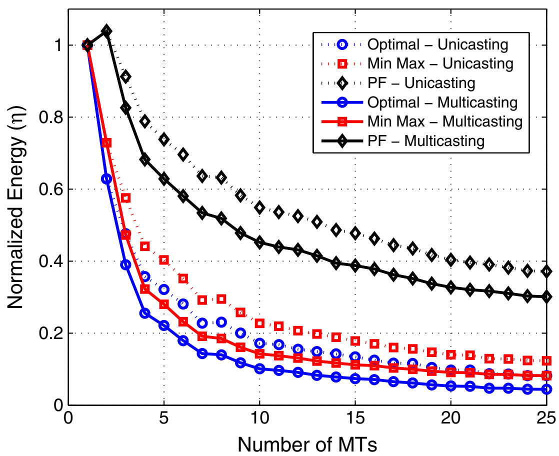

Figure 2. Normalised energy consumption versus the number of mobile terminals (MTs).

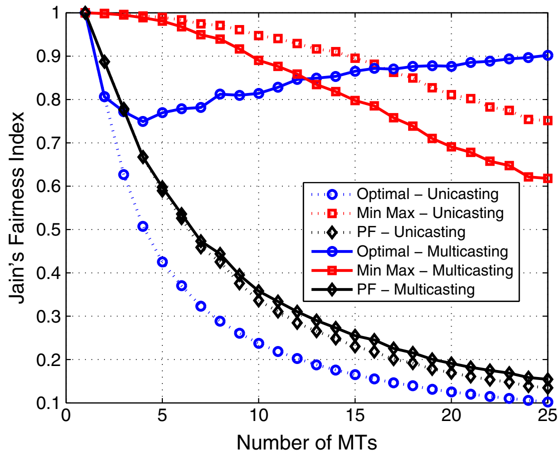

Figure 3. Jain's Fairness index versus the number of mobile terminals (MTs).

unicasting, although in multicasting, every MT is transmitting with a bit rate that is the lowest among its neighbours; however, the transmission is done only once. In unicasting, the MT is spending more energy to transmit the data to everyone of its neighbours. Thus, multicasting is more energy efficient than unicasting when the MTs are close enough to form a single cooperating group on the SR.

Results with $d_{\mathrm{LR}} = 1000$ m and MTs distributed within a 50m $\times$ 50m SR area were generated but are not presented here because of space limitations. Similar conclusions to those obtained from Figure 2 were reached. However, the plots were shifted downwards when d_{LR} increased with the same SR area. The plots were shifted upwards while still being far from the $\eta = 1$ threshold when the SR area increased while keeping the same d_{LR}. Hence, when the distance to the BS is decreased and/or the size of the SR area is increased, the plots of Figure 2 will be shifted upwards depending on the decrease of the LR distance and/or the increase of the SR area.

The results of Figure 2 do not show a measure of the fairness obtained by using each of the compared methods. Figure 3 shows the results of the Jain's fairness index. The Jain's fairness index was derived in [37]. It is widely used to assess fairness in resource allocation in wireless communications, mainly in terms of achievable data rates [38]. In this paper, the interest is in the consumed energy and the fairness achieved during content distribution between the different MTs. The application of Jain's fairness index to energy consumption can be written as

$$J = \frac{\left(\sum_{k=1}^{K} E_k\right)^2}{K \sum_{k=1}^{K} E_k^2} \tag{32}$$

A completely fair solution (equal energy consumption by all MTs) will lead to having $J = 1$. The most unfair solution will lead to a value of $J = 1/K$ and corresponds to the situation where all the energy consumption occurs at a single MT.

With unicasting, the results of the PF and min–max utilities outperform the greedy optimal solution in terms of fairness. However, the min–max utility leads to significantly more fairness than the PF approach because it gives more priority to reducing the energy of the MTs having the highest energy consumption. It is interesting to note that in the case of multicasting, the optimal greedy approach outperforms the min–max approach when the number of collaborating MTs exceeds 12. This is explained by the fact that multicasting highly depends on the rate of the MT having the worst channel conditions. When the number of MTs increases, the probability of finding an MT with bad channel conditions increases. Hence, the MT k^* selected for transmission should transmit at the lowest achievable rate between cooperating MTs. We denote by MT k^- the MT having the lowest rate on the SR link with MT k^*. Thus, the same reception time is spent for all MTs, which leads to an equal energy consumption $E_k = E_{k^-}$ for all MTs other than k^*. For a relatively large number of MTs, although k^* consumes more energy, the contribution of its consumption in the total energy would be masked by the consumption of the other MTs because in this case, $E_{\mathrm{coop}} = E_{k^*} + (K-1)E_{k^-}$.

However, the results are still unfair for MT k^*, although the Jain's fairness index does not appropriately capture this fact. This is investigated in Section 6.2 where the results of a single snapshot are presented.

6.2. Snapshot Results

In this section, we consider $K = 5$ MTs located at a constant distance $d_{\mathrm{LR}} = 400$ m from the BS, with a 5 m separation between an MT and its neighbour MTs. We present a snapshot result corresponding to a single fading realisation. Multicasting is considered as an example. Similar conclusions apply to unicasting.

Figure 4 shows the energy consumed by each of the five MTs at a given snapshot, in addition to the number of parts allocated to each MT. It can be clearly seen from Figure 4 that the optimal greedy approach allocates all resources to a single user. Figure 4 also shows that the energy for the last MT (corresponding to MT $k^* = 5$) with the optimal greedy solution exceeds its consumed energy in the non-cooperative case. However, the energy for all other MTs is reduced. The PF and min–max methods lead to energy savings for all MTs. With the PF approach, MT 5 consumes the least energy and is allocated the lowest number of data parts to distribute, although it is the MT to which all parts are allocated with the greedy approach. This is in line with the NBS because MT 5 has the most favourable conditions and in case of no cooperation, can obtain the required data with minimal energy consumption compared to the other MTs. Hence, MT 5 stands in a good 'bargaining position' that keeps its consumption minimal in the cooperative case. With the min–max approach, the energy consumption and data allocation are almost equal among all MTs, which significantly enhances fairness.

7. PRACTICAL CONSIDERATIONS

In this section, we discuss some practical aspects related to the proposed approach and suggest some ideas to address them.

7.1. Grouping of mobile terminals into cooperating clusters

The problem formulated in Section 3 and the algorithm presented in Section 4 are applicable for MTs in a single cooperative group or cluster. In such a scenario, we refer to MT k^* as the cluster head. In practice, MTs distributed throughout the coverage area of a given BS might be interested in the same content, for example users subscribed in a live news service. In some network scenarios, these MTs might be too spread to form a single cooperative group with efficient SR communications. For example, there might exist some MTs k and j such that $R_{S,kj}$ is too low when (2) is used for rate calculations. It might even be equal to zero when discrete modulation and coding schemes (e.g. BPSK, QPSK, QAM) are used instead of the continuous rate expression given in (2).

In this case, MTs can be grouped into several cooperative clusters, and the proposed methods of Sections 3 and 4 can be applied to each cluster independently. A possible clustering approach is presented next. We denote by C_i the set of MTs in cluster i and by $\bar{C}_i$ the set of MTs not in this cluster:

- **Step 1:** Group all MTs into a single cooperating cluster C_1.
- **Step 2:** Find the cluster head k_1^* satisfying:

$$k_1^* = \arg\min_{k \in C_1} \left(\frac{P_{L,Rx}}{R_{L,k}} + \sum_{j \in C_1, j \neq k} \frac{(P_{S,Tx,kj} + P_{S,Rx})}{R_{S,kj}} \right) \tag{33}$$

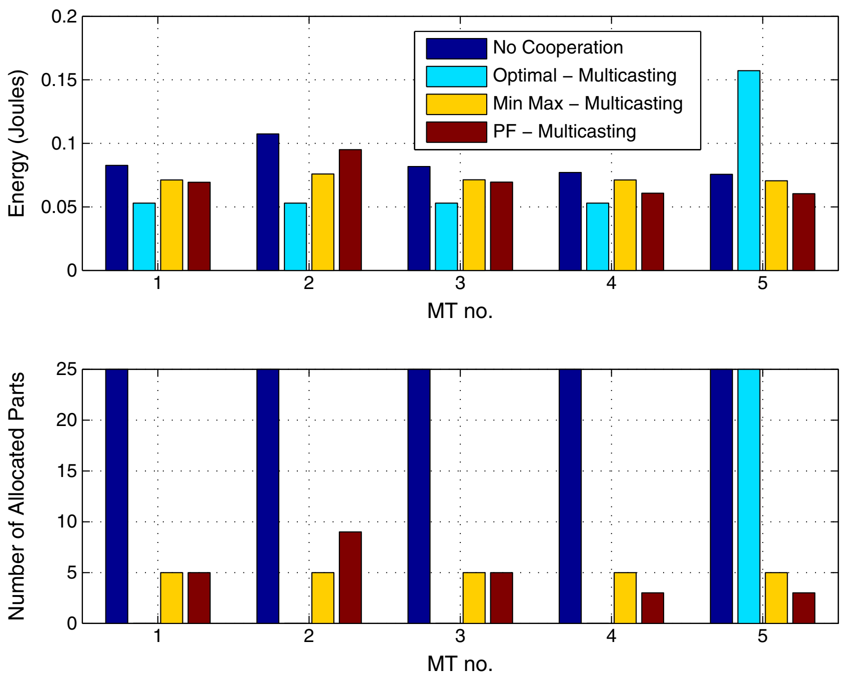

Figure 4. Energy consumption (upper part) and number of allocated data parts (lower part): snapshot results with multicasting.

Trans. Emerging Tel. Tech. **23**:378–392 (2012) © 2012 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

MTs that have $R_{S,k_1^* j} \approx 0$ are kept outside this cluster. They are part of the set $\bar{C}_1$.

- **Step 3:** For each MT j in C_1, remove j from C_1 and place it in $\bar{C}_1$ if

$$\frac{P_{S,Tx,k_1^* j} + P_{S,Rx}}{R_{S,k_1^* j}} \geqslant \frac{P_{L,Rx}}{R_{L,j}} \quad (34)$$

In other words, if it is more energy efficient to send the content to MT j on the LR than on the SR via the cluster head k_1^*, then j is removed from the cluster C_1. Keep MT j in C_1 otherwise.

- **Step 4:** Set $C_2 = \bar{C}_1$.
- **Step 5:** Repeat Steps 2–4 for C_2, that is, find k_2^* and $C_3 = \bar{C}_2$ considering only the MTs in C_2.
- **Step 6:** Repeat the process until all MTs are grouped into clusters or until no improvement can be made. In this case, MTs that are not in any cooperating cluster are standalone MTs, that is, each one of them is considered to form a cluster and receives the content directly from the BS on the LR.

The above clustering method relies on the optimal solution within a single cluster k_i^*. Specifying clusters based on their cluster heads implies that these cluster heads will be used to distribute the content on the SR in each cluster when the optimal greedy energy minimization solution will be implemented. However, after cluster formation, the fair content distribution methods using the proposed utility minimization algorithm can be used for each cluster. In other words, the aforementioned method is used for clustering but not necessarily for content distribution inside clusters: the cluster head does not have to be the only MT to receive on the LR and distribute on the SR within a given cluster. The proposed utility minimization algorithm can be applied independently in each of the obtained clusters C_i after implementing the aforementioned clustering method, and hence all MTs, not just the cluster head, can be involved in the content distribution process.

7.2. Centralised versus distributed implementation

Both the optimal greedy approach and the proposed algorithm can be applied in a centralised way by the BS. In this case, the BS is assumed to be aware of the channel state information (CSI), and hence of the achievable rates $R_{S,kj}$ on the SR links in addition to the CSI and rates $R_{L,k}$ on the LR links. In a low mobility scenario, which is common for the common content distribution applications, this can be achieved by a training phase that precedes the actual content distribution phase. The BS can know the CSI on the LR via feedback from the MTs, which is common in state-of-the-art wireless communication systems. On the SR, MTs can take turns in broadcasting pilot signals. Thus, each MT can estimate its CSI, and hence the rate $R_{S,kj}$, with every other MT by measuring the received strength of the pilot signals. The SR pilot broadcasting process can be coordinated by the BS to avoid collisions. When each MT gets a CSI estimate on its SR links with the other MTs, it can feedback this information to the BS on the LR link. After this training phase, the BS can then coordinate the content distribution process using the proposed methods. In a low mobility scenario, the overhead due to the training phase can be considered low because a long time can elapse before the channel conditions change, and the need arises to repeat the process.

In addition to the centralised implementation, the proposed algorithm lends itself to distributed implementation. The CSI of each MT on the LR can be obtained by measuring the pilot signal of the BS. To estimate the CSI on the SR, each MT broadcasts a pilot signal so that the CSI with other MTs can be estimated on the SR. Then, instead of sending this information to the BS, the MTs can exchange this information, along with their LR CSI estimation, via SR broadcast using the same sequence of turns adopted for SR pilot transmission. Efficient CSI quantization methods can be used to approximate the full CSI with a limited number of feedback bits, which can lead to reducing the overhead of information exchange. After this information exchange phase, each MT would have enough information to implement the utility minimization algorithm in a distributed way and determine the distribution of parts among the various MTs. Then, each MT k would request its parts x_k from the BS and then distribute them on the SR.

Clustering can be also performed in a distributed implementation as the MTs form cooperative clusters with other MTs when they can successfully hear their pilot transmission, that is, when $R_{S,kj}$ is high enough to allow efficient communication between MTs. When $R_{S,kj}$ is too low between two MTs k and j, these will automatically be in different clusters. After exchanging CSI information within each cluster, all MTs would be able to determine the cluster head in that cluster if the optimal (unfair) solution will be implemented, or they would be able to implement the utility minimizing algorithm within that cluster. In that case, all MTs, not just the cluster head, will be involved in the content distribution process. Another approach for distributed cluster formation based on coalitional game theory is presented in [39].

8. FUTURE WORK

In this section, some interesting ideas for future research are presented. An interesting topic would be to elaborate more on the clustering scheme presented in Section 7.1 and to try to formulate the content distribution problem with multiple clusters as an optimization problem and derive the optimal solution. Another direction for future research would be to quantify the feedback overhead because of the exchange of CSI, that is, present a mathematical formulation and calculation for the CSI exchange methods described in Section 7.2, for both the centralised and distributed scenarios. This M2M CSI exchange could have the side effect of being useful in other applications, for example localization as in [40].

In addition, it would be interesting to investigate the scenario where certain MTs can leave the cooperative cluster during the content distribution process. This can correspond in practice to a user leaving a certain area or to the battery of a certain mobile device getting depleted. The impact of these disturbances on the content distribution process should be investigated in this case, and efficient distributed algorithms for restructuring the clusters and retransmitting the lost content should be derived. The use of network coding with several MTs transmitting in each cluster would be an interesting approach to follow in this case. Network coding has in fact been investigated in the framework of M2M collaboration in [41] where an SR collaboration method based on genetic algorithms is proposed. The more challenging problem of optimal content distribution with network coding is not yet investigated.

Furthermore, in this paper, the channel conditions were assumed to remain constant during the content distribution process. This assumption does not hold in a high mobility scenario where the channel changes dynamically or in the case of large files where the content distribution process cannot be completed before the channel conditions vary. Therefore, an interesting direction worthy of further investigation is the extension of the results presented in this paper to a high mobility scenario and the derivation of the optimal solution in the case where the channel varies before the content distribution is complete.

9. CONCLUSIONS

Energy minimization in M2M cooperative networks was investigated. Different practical implementation scenarios including unicasting and multicasting with rate adaptive or power adaptive transmissions were considered. The optimal solution consists of sending the data to a single MT on the LR and of having that MT distribute the data to other MTs on SR links. This leads to an unfair energy consumption for the selected MT. To ensure fairness in energy consumption, a low complexity utility minimization algorithm that can be used with various utility functions was presented. With a certain utility, the unfair energy minimizing solution can also be reached by the proposed algorithm. Using the appropriate utilities, the algorithm can lead to different outcomes: a min–max utility allows to ensure fairness in energy consumption, whereas a PF utility favors MTs that have the best channel conditions by allowing them to reduce their energy consumption. Practical constraints involving the formation of cooperative clusters, and the exchange of information to implement the algorithm in centralised and distributed scenarios were also discussed.

ACKNOWLEDGEMENTS

The authors would like to thank the reviewers for their comments that helped in enhancing the quality of the paper. This work was made possible by an NPRP grant from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors.

REFERENCES

1. Fitzek F, Katz M. *Cooperation in Wireless Networks: Principles and Applications*. Springer: Dordrecht, Netherlands, 2006.
2. Oh BJ, Chen CW. Energy efficient H.264 video transmission overwireless ad hoc networks based on adaptive 802.11e EDCA MAC protocol, In *Proceedings of IEEE International Conference on Multimedia & Expo*, June 2008.
3. Mahfoudh S, Minet P. Survey of energy efficient strategies in wireless ad hoc and sensor networks, In *Proceedings of International Conference on Networking*, pages 1–7, April 2008.
4. Vijay S, Sharma SC. EEN: An energy efficient multi-hop ad hoc wireless networks, In *Proceedings of Conference on Emerging Trends in Engineering and Technology*, pages 145–150, July 2008.
5. Sachs J, Magnusson P. Assessment of the access selection gain in multi-radio access networks. *European Transactions on Telecommunications* 2009; **20**(3): 265–279.
6. Hui SY, Yeung KH. Challenges in the migration to 4G mobile systems. *IEEE Communications Magazine* 2003; **41**(12): 54–59.
7. Dohler M, Meddour D-E, Senouci S-M, Saadani A. Cooperation in 4G - hype or ripe? *IEEE Technology and Society Magazine* 2008; **27**(1): 13–17.
8. Bhatia R, Li LE, Luo H, Ramjee R. ICAM: integrated cellular and ad hoc multicast. *IEEE Transactions on Mobile Computing* 2006; **5**(8): 1004–1015.
9. Luo H, Ramjeey R, Sinhaz P, Li LE, Lu S. UCAN: a unified cellular and ad hoc network architecture, In *Proceedings of MobiCom'03*, pages 353–367, September 2003.
10. Popova L, Herpel T, Gerstacker W, Koch W. Cooperative mobile-to-mobile file dissemination in cellular networks within a unified radio interface. *International Journal of Computer and Telecommunications Networking* 2008; **52**(6): 1153–1165.
11. Popova L, Herpel T, Koch W. Improving downlink UMTS capacity by exploiting direct mobile-to-mobile data transfer, In *Proceedings of International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, pages 1–7, April 2007.

Trans. Emerging Tel. Tech. **23**:378–392 (2012) © 2012 John Wiley & Sons, Ltd.
DOI: 10.1002/ett

12. Popova L, Herpel T, Koch W. Efficiency and dependability of direct mobile-to-mobile data transfer for UMTS downlink in multi-service networks, In *IEEE WCNC 2007*, pages 3318–3324, March 2007.
13. Peng M, Wang W. Investigation of cooperative relay node selection in heterogeneous wireless communication systems, In *IEEE ICC 2008*, pages 174–178, May 2008.
14. Chen J, Li S, Chan S-HG, He J. WIANI: wireless infrastructure and ad-hoc network integration, In *IEEE ICC 2005*, pages 3623–3627, May 2005.
15. Peng K-J, Tsai Z. Distortion and cost controlled video streaming in a heterogeneous wireless network environment, In *IEEE PIMRC 2006*, pages 1–5, September 2006.
16. Leung MF, Chan S-HG. Broadcast-based peer-to-peer collaborative video streaming among mobiles. *IEEE Transactions on Broadcasting* 2007; **53**(1): 50–61.
17. Zhang Q, Fitzek FHP, Iversen VB. Design and performance evaluation of cooperative retransmission scheme for reliable multicast services in cellular controlled P2P networks, In *IEEE PIMRC 2007*, pages 1–5, September 2007.
18. Zhu D, Mutka M. Cooperation among peers in an ad hoc network to support an energy efficient IM service. *Journal of Pervasive and Mobile Computing* 2008; **4**(3): 335–359.
19. Wu H, De S, Qiao C, Yanmaz E, Tonguz OK. Hand-off performance of the integrated cellular and ad hoc relaying (iCAR) system. *Journal of ACM Wireless Networks* 2005; **11**(6): 775–785.
20. Zhaoji X, Nan H, Zhiqiang H. Call dropping and blocking probability of the integrated cellular ad hoc relaying system, In *IEEE GlobeCom 2008*, pages 1–6, December 2008.
21. Al-Kanj L, Dawy Z. Optimized energy efficient content distribution over wireless networks with mobile-to-mobile cooperation, In *IEEE ICT 2010*, pages 471–475, April 2010.
22. Manilici V, Wundsam A, Feldmann A, Vidales P. Potential benefit of flow-based routing in multihomed environments. *European Transactions on Telecommunications* 2009; **20**(7): 650–659.
23. Kang S, Mutka M. Mobile peer membership management to support multimedia streaming, In *Proceedings of ICDCS Workshop on Mobile and Wireless Networks*, pages 770–775, May 2003.
24. Kang S, Mutka M. A mobile peer-to-peer approach for multimedia content sharing using 3G/WLAN dual mode channels. *Journal of Wireless Communications and Mobile Computing* 2005; **5**(6): 633–645.
25. Ramadan M, Zein L, Dawy Z. Implementation and evaluation of cooperative video streaming for mobile devices, In *IEEE PIMRC 2008*, pages 1–5, September 2008.
26. Ananthanarayanan G, Padmanabhan V, Ravindranath L. COMBINE: leveraging the power of wireless peers through collaborative downloading, In *Proceedings of MobiSys'07*, pages 286–298, June 2007.
27. Goldsmith A. *Wireless Communications*. Cambridge University Press: New York, NY, USA, 2005.
28. Qiu X, Chawla K. On the performance of adaptive modulation in cellular systems. *IEEE Transactions on Communications* 1999; **47**(6): 884–895.
29. Haratcherev I, Lagendijk R, Langendoen K, Sips H. Hybrid rate control for IEEE 802.11, In *Proceedings of International Symposium on Mobility Management and Wireless Access*, pages 10–18, October 2004.
30. Bluetooth SIG. Specification of the Bluetooth System, Bluetooth Core Specification version 4.0, 2010.
31. Carmichael F. *A Guide to Game Theory*. Pearson Education Limited: Essex, England, 2005.
32. Yaacoub E, Dawy Z. A game theoretical formulation for proportional fairness in LTE uplink scheduling, In *IEEE WCNC 2009*, pages 1–5, April 2009.
33. Le Boudec J-Y. Rate adaptation, Congestion Control and Fairness: A Tutorial, Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland, December 2008.
34. Song G, Li Y. Cross-layer optimization for OFDM wireless networks-part i: theoretical framework. *IEEE Transactions on Wireless Communications* 2005; **4**(2): 614–624.
35. 3rd Generation Partnership Project (3GPP). 3GPP TR 25.814 3GPP TSG RAN Physical Layer Aspects For Evolved UTRA, v7.1.0, 2006.
36. Mahmud K, Inoue M, Murakami H, Hasegawa M, Morikawa H. Energy consumption measurement of wireless interfaces in multi-service user terminals for heterogeneous wireless networks. *IEICE Transactions on Communications* 2005; **E88-B**(3): 1097–1110.
37. Jain R. *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation and Modeling*. Wiley: New York, 1991.
38. Salem M, Adinoyi A, Rahman M, *et al*. An overview of radio resource management in relay-enhanced OFDMA-based networks. *IEEE Communications Surveys and Tutorials* 2010; **12**(3): 422–438.
39. Al-Kanj L, Saad W, Dawy Z. A game theoretic approach for content distribution over wireless networks with

mobile-to-mobile cooperation, In *IEEE PIMRC 2011*, pages 1–5, September 2011.

40. Frattasi S, Monti M. Ad-coop positioning system (acps): positioning for cooperative users in hybrid cellular ad-hoc networks. *European Transactions on Telecommunications* 2008; **19**(8): 923–934.

41. Militano L, Fitzek FHP, Iera A, Molinaro A. Network coding and evolutionary theory for performance enhancement in wireless cooperative clusters. *European Transactions on Telecommunications* 2010; **21**(8): 725–737.

 Trans. Emerging Tel. Tech. **23**:378–392 (2012) © 2012 John Wiley & Sons, Ltd.
DOI: 10.1002/ett