Utility-Optimal Random Access: Optimal Performance Without Frequent Explicit Message Passing

Journal:	*IEEE Transactions on Wireless Communications*
Manuscript ID:	Letter-TW-Dec-07-1446
Manuscript Type:	Original Transactions Letter
Date Submitted by the Author:	24-Dec-2007
Complete List of Authors:	Mohsenian Rad, Hamed; University of British Columbia, Electrical and Computer Engineering Huang, Jianwei; Chinese University of Hong Kong, Information Engineering Chiang, Mung; Princeton University, Electrical Engineering Wong, Vincent; University of British Columbia, Electrical and Computer Engineering
Keyword:	medium access control < Multimedia, Networks and Systems, wireless ad-hoc networks, network utility maximization

Utility-Optimal Random Access: Optimal Performance Without Frequent Explicit Message Passing

Hamed Mohsenian-Rad, Jianwei Huang, Mung Chiang, Vincent W.S. Wong

Abstract

In this paper, we propose a distributed random medium access control (MAC) algorithm for wireless ad hoc networks based on the framework of network utility maximization (NUM). Compared with the related algorithms proposed in the literature, our algorithm achieves the optimal network performance without frequent explicit message passing among wireless users. This is of critical importance in practice, since any explicit message passing among wireless users will lead to further contentions in the network and reduce the network performance. We prove the convergence of our proposed algorithm under the assumption that the users estimate the required information through local observation of the shared wireless medium with asymptotically converging estimation errors. This includes the important case where the underlining communication channel is lossy and thus not every transmission can be correctly decoded. When the channel is perfect, our algorithm converges to the global optimal solution of the NUM problem. Simulation results show the optimality and fast convergence of our algorithm, and better efficiency-fairness tradeoff compared with the IEEE 802.11 distributed coordination function.

I. Introduction

In the existing contention-based medium access control (MAC) protocols, there is a tradeoff between system performance (e.g., throughput and fairness) and the amount of explicit message passing required among wireless users. One example is the IEEE 802.11 distributed coordination function (DCF), where users do not explicitly exchange any message related to their transmission probabilities[1] and adapt their transmission probabilities only based on the binary implicit feedback from the network (e.g., collision or not). This typically leads to low throughput and unfair resource allocation [1]. On the other hand, several MAC algorithms (e.g., [2]–[4]) have been designed based on the framework of network utility maximization (NUM) which lead to

H. Mohsenian-Rad and V.W.S. Wong are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada, emails:{*hamed, vincentw*}*@ece.ubc.ca*. J. Huang is with the Information Engineering Department, Chinese University of Hong Kong, Hong Kong, email: *jwhuang@ie.cuhk.edu.hk*. M. Chiang is with the Department of Electrical Engineering, Princeton University, Princeton, USA, email: *chiangm@princeton.edu*.

[1]In this paper, we use “messages” to denote control signals that are explicitly related to users’ transmission probabilities. IEEE 802.11 DCF does not have any explicit message passing, although it has various other control signals (e.g., RTS/CTS/ACK).

the optimal system performance without taking the signalling overhead into account. However, these algorithms require extensive frequent message passing among users. Considering the fact that any message transmission leads to additional contention in a random access network, this paper aims to address the following question: *is it possible to design a MAC algorithm that can achieve the optimal performance without frequent explicit message passing?*

We provide a positive answer to the above question in some special but important cases, based on the NUM-based MAC algorithms we proposed in [5]. Compared with the previous algorithms (e.g., [2]–[4]), the algorithms in [5] support a wider range of utility functions, converge faster, and allow fully asynchronous operations among users. However, frequent explicit message passings are still needed in [5]. In this paper, we show that in the simple case of a single-cell interference topology (e.g., as in wireless personal and local area networks), we can completely eliminate the need for frequent message passing. Users will be able to estimate the required information through local observation of the channel contention history. We prove the convergence of our algorithm under various channel conditions. If the channel is perfect and the estimations are asymptotically accurate, then the optimality of the algorithm is also guaranteed. The estimation techniques we use here are related to [6], [7]. However, our estimation model is more elaborate and captures more information (i.e., each user's transmission probability). Simulation results show that our algorithm is robust to changes in user populations and channel conditions. These encouraging results provide important insights and useful hints to design fully distributed utility optimal MAC algorithms without frequent explicit message passing for more general topologies.

The rest of this paper is organized as follows. The system model is described in Section II. Our algorithm is presented in Section III. Convergence and optimality of the algorithm are proved in Section IV. Simulation results are shown in Section V. We conclude the paper in Section VI.

II. System Model

Consider a single-hop wireless ad-hoc network with $\mathcal{N} = \{1, \ldots, N\}$ as the set of wireless links. Each link, together with its dedicated transmitter and receiver nodes, is called a *user*. A sample network with 3 users is shown in Fig. 1. We assume that each user's receiver node can hear every other user's transmissions. Thus, each user interferes with all other users. This models some important wireless networks including *wireless personal area networks* where wireless devices interact with each other (e.g., in an office) and indoor *wireless local area* *networks* where the nodes interact with each other and an access point (e.g., in a large conference room). Time is divided into equal-length slots. At each slot, user i transmits with probability $p_i \in \mathcal{P}_i = [P_i^{\min}, P_i^{\max}]$, with $0 < P_i^{\min} < P_i^{\max} < 1$. A transmission is successful only if it is the only transmission in the current slot. Let r_i denote the average data rate for user i. We have [8]:

$$r_i(\boldsymbol{p}) = \gamma_i p_i \prod_{j \in \mathcal{N} \setminus \{i\}} (1 - p_j), \qquad \forall\, i \in \mathcal{N}, \tag{1}$$

where $\boldsymbol{p} = (p_i, \forall i \in \mathcal{L})$ is the vector of all users' transmission probabilities and γ_i denotes the fixed peak data rate for user i. Each link $i \in \mathcal{L}$ maintains a *utility* which is an increasing and concave function of r_i and indicates link i's level of satisfaction on its average data rate. The utility of link i is denoted by $u_i(r_i(\boldsymbol{p}))$ which is also a function of $\boldsymbol{p}$. We are interested in finding the value of $\boldsymbol{p}$ that solves the following *network utility maximization* (NUM) problem [9]:

$$\max_{\boldsymbol{p} \in \mathcal{P}} \; \sum_{i \in \mathcal{N}} u_i(r_i(\boldsymbol{p})), \tag{NUM}$$

where $\mathcal{P} = \{\boldsymbol{p} : p_i \in \mathcal{P}_i, \forall i \in \mathcal{N}\}$, and the utility functions are α-fair [10]. That is, $u_i(r_i(\boldsymbol{p})) = (1-\alpha)^{-1} r_i(\boldsymbol{p})^{1-\alpha}$ if $\alpha \in (0,1) \cup (1,\infty)$, and $u_i(r_i(\boldsymbol{p})) = \log r_i(\boldsymbol{p})$, if $\alpha = 1$. In [5], we have shown that the α-fair utility functions can model a wide range of efficient and fair allocations.

III. Algorithm with No Frequent Explicit Message Passing

1) Local Optimization: For each user i, consider the following *local* optimization problem:

$$\max_{p_i \in \mathcal{P}_i} \; \sum_{j \in \mathcal{N}} u_j(r_j(p_i, \boldsymbol{p}_{-i})), \tag{LOCAL-NUM}$$

where $\boldsymbol{p_{-i}} = (p_j,\ \forall j \in \mathcal{N} \setminus \{i\})$ denotes the transmission probabilities of all users *other than* user i. To solve Problem (LOCAL-NUM), user i will choose p_i to maximize the *total* network utility, *assuming* that none of the other users change their transmission probabilities.

Theorem 1: For each user $i \in \mathcal{N}$, the unique global optimal solution of Problem (LOCAL-NUM) is $p_i^*(\boldsymbol{p}_i) = f_i(\boldsymbol{p}_i)$, where the mapping function $f_i(\boldsymbol{p}_i)$ is defined as

$$f_i(\boldsymbol{p}_{-i}) = \left[1 / \left(1 + \sqrt[\alpha]{v_i(\boldsymbol{p}_{-i})} \,\right) \right]_{P_i^{\min}}^{P_i^{\max}}. \tag{2}$$

Here $[x]_b^a = \max\left[\min\left[x, a\right], b\right]$ and $v_i(\boldsymbol{p}_{-i}) = \gamma_i^{\alpha-1} \sum_{j \in \mathcal{N} \setminus \{i\}} (1/\gamma_j)^{\alpha-1} (1/p_j - 1)^{\alpha-1}$. The proof of Theorem 1 is similar to that of [5, Theorem 1] and is omitted for brevity. It is clear that if user i wants to compute (2), the only information it needs from other users is $v_i(\boldsymbol{p}_{-i})$. If each user i can *estimate* the value of

$$m_j = (1/\gamma_j)^{\alpha-1} (1/p_j - 1)^{\alpha-1}, \qquad \forall\, j \in \mathcal{N}, \tag{3}$$

then it can compute $v_i(\boldsymbol{p}_{-i}) = \gamma_i^{\alpha-1} \sum_{j \in \mathcal{N} \setminus \{i\}} m_j$ and set $p_i = f_i(\boldsymbol{p}_{-i})$. Notice that for each $j \in \mathcal{N}$, m_j is bounded between $M^{\min}$ and $M^{\max}$. If $\alpha \geq 1$, then $M^{\min} = (1/\gamma^{\max})^{\alpha-1} (1/P^{\max} - 1)^{\alpha-1}$

and $M^{\max} = (1/\gamma^{\min})^{\alpha-1}(1/P^{\min}-1)^{\alpha-1}$ where $P^{\min} = \min_{i\in\mathcal{N}} P_i^{\min}$, $P^{\max} = \max_{i\in\mathcal{N}} P_i^{\max}$, $\gamma^{\min} = \min_{i\in\mathcal{N}} \gamma_i$, and $\gamma^{\max} = \max_{i\in\mathcal{N}} \gamma_i$. If $\alpha < 1$, then $M^{\min} = (1/\gamma^{\min})^{\alpha-1}(1/P^{\min}-1)^{\alpha-1}$ and $M^{\max} = (1/\gamma^{\max})^{\alpha-1}(1/P^{\max}-1)^{\alpha-1}$. As shown in [5, Section IV-A], if each user i updates its transmission probability p_i accordingly to (2), then the whole system will converge to the optimal solution of Problem (NUM). The key question is how to obtain the values of m_j for all $j \neq i$. Next, we show how this can be done through local observations of the shared channel.

2) Learning from Contention History: From (3), we see that only the values of γ_j and p_j are required to calculate the value of m_j. Notice that α is the same for all users. The value of the peak rate γ_j depends on the channel gain between the transmitter and receiver of user j; thus, it can only be measured by user i and then announced to the whole network once user i joins the network. The remaining task is to determine how to obtain the value of p_j.

From user i's viewpoint, any time slot falls into one of the following possible states: *idle* (no user transmits), *busy* (at least one other user transmits), *success* (user i transmits successfully), and *failure* (user i transmits but it fails). Let p_i^{idle}, p_i^{busy}, p_i^{succ}, and p_i^{fail} denote the probabilities of experiencing these four states, respectively. Also let $p_{i,j}^{\text{err}}$ denote the *packet error rate* of the channel from the transmitter node of user j to the receiver node of user i. We have:

$$p_i^{\text{idle}} = \prod_{j\in\mathcal{N}}(1-p_j), \tag{4}$$

$$p_i^{\text{busy}} = (1-p_i) - p_i^{\text{idle}}, \tag{5}$$

$$p_i^{\text{succ}} = (p_i/(1-p_i))\, p_i^{\text{idle}}\,(1-p_{i,i}^{\text{err}}), \tag{6}$$

$$p_i^{\text{fail}} = p_i - (p_i/(1-p_i))\, p_i^{\text{idle}}\,(1-p_{i,i}^{\text{err}}). \tag{7}$$

Note that p_i^{idle} and p_i^{busy} are independent of $p_{i,i}^{\text{err}}$. By knowing the value of $p_{i,i}^{\text{err}}$ and estimating any subset (or all) of the probabilities in (4)-(7), user i can only estimate the value of $\prod_{j\in\mathcal{N}}(1-p_j)$. However, user i needs more information to calculate the value of p_j for all $j \neq i$.

Recall that, at a busy slot seen by user $i \in \mathcal{N}$, at least one other user transmits. Since users can hear each other, user i may successfully *decode* the transmission of user $j \neq i$ with probability

$$p_{i,j}^{\text{decd}} = p_j(\textstyle\prod_{l\in\mathcal{N}\setminus\{j\}}(1-p_l))(1-p_{i,j}^{\text{err}}) = (p_j/(1-p_j))\left(\prod_{l\in\mathcal{N}}(1-p_l)\right)(1-p_{i,j}^{\text{err}}). \tag{8}$$

Let $n_{i,j}^{\text{decd}}$ denote the number of slots between any two consecutive successful decoding of transmissions of user j by user i. We have:

$$p_{i,j}^{\text{decd}} = 1/(1+\bar{n}_{i,j}^{\text{decd}}), \tag{9}$$

where $\bar{n}_{i,j}^{\text{decd}}$ is the mean value of $n_{i,j}^{\text{decd}}$ and can be *locally* estimated by user i through observation of the channel contention history. Notice that in practice, the transmitted signal by user j can be decoded by the network interface of user i's receiver node; however, as its destination MAC address is not the same as the one in user i, the packet is simply discarded. Now, user i needs to obtain the sender's MAC address from the packet header before discarding the packet.

Similarly, let n_i^{idle} denote the number of non-idle slots that user i observes between any two consecutive idle time slots. User i can estimate p_i^{idle} as follows [6]:

$$p_i^{\text{idle}} = 1/(1 + \bar{n}_i^{\text{idle}}), \tag{10}$$

where $\bar{n}_i^{\text{idle}}$ is the mean value of n_i^{idle}. Substituting (4), (9), and (10) into (8), for each $j \in \mathcal{N}\backslash\{i\}$,

$$1/p_j - 1 = \left((1 + \bar{n}_{i,j}^{\text{decd}})/(1 + \bar{n}_i^{\text{idle}})\right)\ (1 - p_{i,j}^{\text{err}}). \tag{11}$$

Let T_{idle}^i and $T_{j,\text{idle}}^i$ denote the set of time slots at which user i observes an idle slot and decodes the transmissions of user $j \neq i$, respectively. We estimate $\bar{n}_i^{\text{idle}}$ and $\bar{n}_{i,j}^{\text{decd}}$ iteratively as follows:

$$\bar{n}_i^{\text{idle}}(t+1) \;=\; (1 - \rho_i(t))\bar{n}_i^{\text{idle}}(t) + \rho_i(t) n_i^{\text{idle}}(t) I\{t \in T_{\text{idle}}^i\}, \tag{12}$$

$$\bar{n}_{i,j}^{\text{decd}}(t+1) \;=\; (1 - \varrho_{i,j}(t))\bar{n}_{i,j}^{\text{decd}}(t) + \varrho_{i,j}(t) n_{i,j}^{\text{decd}}(t) I\{t \in T_{j,\text{decd}}^i\}, \tag{13}$$

where $\bar{n}_i^{\text{idle}}(t)$, $n_i^{\text{idle}}(t)$, $\bar{n}_{i,j}^{\text{decd}}(t)$ and $n_{i,j}^{\text{decd}}(t)$ denote the estimation of $\bar{n}^{\text{idle}}$, the measurement of n_i^{idle}, the estimation of $\bar{n}_{i,j}^{\text{decd}}$, and the measurement of $n_{i,j}^{\text{decd}}$ at time slot t, respectively, and $I\{\cdot\}$ is an indication function. Here ρ_i and $\varrho_{i,j}$ are *tapering* stepsizes. Based on the asynchronous stochastic approximation theory [11], we know that the estimation error decreases to zero when users do not change their transmission probabilities.

For each user i and any other user $j \neq i$, given γ_j, $\bar{n}_{i,j}^{\text{decd}}$ and $\bar{n}_i^{\text{idle}}$, we define:

$$m_j^i(t) = (1/\gamma_j)^{\alpha-1} \left((1 + \bar{n}_{i,j}^{\text{decd}}(t))/(1 + \bar{n}_i^{\text{idle}}(t))\right)^{\alpha-1}, \qquad \forall\, j \in \mathcal{N}\backslash\{i\}, \tag{14}$$

where $m_j^i(t)$ denotes the estimation of m_j made by user i at time slot t. In general, we have:

$$m_j^i(t) = \beta_j^i(t)\; m_j(t), \tag{15}$$

where $\beta_j^i(t) > 0$ is the *estimation gain*, which can represent accurate estimation (i.e., $\beta_j^i(t) = 1$), over-estimation (i.e., $\beta_i^i(t) > 1$) or under-estimation (i.e., $\beta_i^i(t) < 1$). From (11), if the estimations on $\bar{n}_{i,j}^{\text{decd}}$ and $\bar{n}_i^{\text{idle}}$ are accurate and the channel is perfect (with zero packet error rate), then $\beta(t) = 1$ and we have $m_j^i(t) = m_j(t)$ for all $j \in \mathcal{N}\backslash\{i\}$. Notice that if the value of the existing packet error rate $p_{i,j}^{\text{err}}$ is known (e.g., via measurements at the physical layer), then we can redefine $m_j^i(t) = (1/\gamma_j)^{\alpha-1} \left((1 + \bar{n}_{i,j}^{\text{decd}}(t))/(1 + \bar{n}_i^{\text{idle}}(t))\right)^{\alpha-1}/(1 - p_{i,j}^{\text{err}})^{\alpha-1}$ and obtain a more accurate estimation by canceling out the effect of channel imperfections. However, in this paper, we consider the general case and assume that the packet error rates are not known by users.

For each user $i \in \mathcal{N}$ and for all $j \in \mathcal{N}\backslash\{i\}$, we set $T^i_{j,m}$ such that as time goes by, the minimum difference between any two consecutive time slots in the union of sets $\{T^i_{j,m}, \forall j \in \mathcal{N}\backslash\{i\}\}$ increases. This implies that for each j, we update m_j less frequently to be able to collect more samples of n_i^{idle} and $n_{i,j}^{\text{decd}}$. Thus, the estimations of mean values $\bar{n}_i^{\text{idle}}$ and $\bar{n}_{i,j}^{\text{decd}}$ improve gradually and become asymptotically accurate. We also reset the tapering stepsizes ρ_i and $\varrho_{i,j}$ to 1 after each $t \in T^i_{j,m}$ so that the errors in previous estimations do not affect new estimations. Based on these assumptions, there exists a $\beta^i_j > 0$ such that $\lim_{t\to\infty} \beta^i_j(t) = \beta^i_j$. From (14) and (15),

$$\beta^i_j = 1/(1-p^{\text{err}}_{i,j})^{\alpha-1}, \qquad \forall i,j \in \mathcal{N},\ i \neq j.$$

If the channel is perfect, then $\beta^i_j = 1$ and all estimations are *asymptotically accurate*. For a lossy channel, if $\alpha < 1$, then $\beta^i_j < 1$ and m^i_j is *asymptotically under-estimated* for all $j \neq i$. On the other hand, if $\alpha > 1$, then $\beta^i_j > 1$ and m^i_j is *asymptotically over-estimated*.

3) Distributed Algorithm: Our proposed distributed MAC algorithm with no explicit message passing (except when each user joins or leaves the network) is shown in Algorithm 1. In this algorithm, each user $i \in \mathcal{N}$ continuously updates $\bar{n}_i^{\text{idle}}$ and $\bar{\boldsymbol{n}}_i^{\text{decd}} = (\bar{n}_{i,j}^{\text{decd}},\ \forall j \in \mathcal{N}\backslash\{i\})$ based on its local observations from the shared channel to estimate $\boldsymbol{m}_i = (m^i_j,\ \forall j \in \mathcal{N}\backslash\{i\})$. Then, it chooses p_i according to (2) with $v_i = \sum_{j\in\mathcal{N}\backslash\{i\}} m^i_j$. Sets $T_{i,p}$ and $T_{i,m}$ are two unbounded sets of time slots at which user i updates p_i and $\boldsymbol{m}_i$, respectively. Notice that the updates are *asynchronous* across different users which includes synchronous updates as a special case.

IV. Convergence and Optimality

For each $i\in\mathcal{N}$, and at any time $t \in T_{i,p}$, Algorithm 1 updates

$$p_i(t+1) = f'_i(\boldsymbol{p}_{-i}, t) = \left[1/\left(1 + \sqrt[\alpha]{v'_i(\boldsymbol{p}_{-i}, t)}\ \right)\right]^{P_i^{\max}}_{P_i^{\min}},$$

where $v'_i(\boldsymbol{p}_{-i}, t) = \sum_{j\in\mathcal{N}\backslash\{i\}} (\gamma_i/\gamma_j)^{\alpha-1} (1/p_j - 1)^{\alpha-1} \beta^i_j(t)$. For any $t \geq 0$, we define $f'(\boldsymbol{p}, t) = (f'_i(\boldsymbol{p}_{-i}, t), \forall\ i \in \mathcal{N})$. Notice that $f'(\boldsymbol{p}, t)$ is a time-varying vector mapping. Since $\beta^i_j(t)$ approaches β^i_j as $t \to \infty$ for all $i, j \in \mathcal{N}$, the sequence of mapping $\{f'(\boldsymbol{p}, t)\}$ converges to a unique mapping $f'(\boldsymbol{p}, \infty)$ as $t \to \infty$. That is, for any $\boldsymbol{p} \in \mathcal{P}$ and any $\epsilon' > 0$, there exists $t_{\epsilon'} \geq 0$ such that $\|f'(\boldsymbol{p}, t) - f'(\boldsymbol{p}, \infty)\| < \epsilon'$ for all $t \geq t_{\epsilon'}$.

Theorem 2: Assume there exists $t'_0 \geq 0$ such that for all $t \geq t'_0$ and any $\boldsymbol{p} \in \mathcal{P}$, we have:

$$\frac{\beta^{\max}(t)}{\beta^{\min}(t)} \left(\frac{|1-\alpha|}{\alpha}\ \Psi\ \Phi(V'^{\min}, V'^{\max})\right)^2 \left(\frac{\gamma^{\max}}{\gamma^{\min}}\Gamma\right)^{|1-\alpha|} < 1, \tag{16}$$

where $\beta^{\min}(t) = \min_{i,j\in\mathcal{N}} \beta_j^i(t)$, $\beta^{\max}(t) = \max_{i,j\in\mathcal{N}} \beta_j^i(t)$, $\Psi = \max\left\{\frac{1}{P^{\min}(1-P^{\min})}, \frac{1}{P^{\max}(1-P^{\max})}\right\}$,

$$\Gamma = \frac{P^{\max}(1-P^{\min})}{P^{\min}(1-P^{\max})}, \qquad \text{and} \qquad \Phi(V'^{\min}, V'^{\max}) = \begin{cases} \frac{(V'^{\max})^{1/\alpha}}{(1+(V'^{\max})^{1/\alpha})^2}, & \text{if } V'^{\max} \leq 1, \\ \frac{(V'^{\min})^{1/\alpha}}{(1+(V'^{\min})^{1/\alpha})^2}, & \text{if } V'^{\min} \geq 1, \\ 0.25, & \text{otherwise}. \end{cases} \tag{17}$$

Then, Algorithm 1 globally and asynchronously converges to the unique fixed point of $f'(\boldsymbol{p}, \infty)$. Notice that $V'^{\min}$ and $V'^{\max}$ are the lower and upper bounds on $v_i'(\boldsymbol{p}, t)$ for each $i \in \mathcal{N}$ and at any time t. If $\alpha \geq 1$, then $V'^{\min} = (N-1)M^{\min}(\gamma^{\min})^{\alpha-1}$ and $V'^{\max} = (N-1)M^{\max}(\gamma^{\max})^{\alpha-1}$. If $\alpha < 1$, then $V'^{\min} = (N-1)M^{\min}(\gamma^{\max})^{\alpha-1}$ and $V'^{\max} = (N-1)M^{\max}(\gamma^{\min})^{\alpha-1}$.

The proof of Theorem 2 is given in the Appendix. Notice that, at any time $t \geq 0$,

$$(M^{\min}/M^{\max}) \leq \beta^{\min}(t) \leq \beta^{\max}(t) \leq (M^{\max}/M^{\min}). \tag{18}$$

We notice that all the terms in (16), except Φ, are bounded and independent of the number of users N. Thus, Φ can be arbitrarily close to 0 if N is large enough. This results in the following:

Corollary 1: For any choice of system parameters, there exists an integer $\hat{N} > 0$, such that Algorithm 1 globally and asynchronously converges to the unique fixed point of mapping $f'(\boldsymbol{p}, \infty)$, if the number of users $N > \hat{N}$, i.e., there are enough users competing for the channel.

Theorem 2 is general and does not depend on the exact values of the estimation errors as $t \to \infty$; however, the performance at the asymptotic fixed point still depends on the accuracy of the estimations. The following Theorem can be shown for perfect channel case.

Theorem 3: If the channel is perfect such that $\lim_{t\to\infty} \beta^{\min}(t) = \lim_{t\to\infty} \beta^{\max}(t) = 1$, then the unique fixed point of Algorithm 1 is the unique global optimal solution of Problem (NUM).

The proof of Theorem (3) is similar to that of [5, Theorem 4] and is omitted. Notice that since $\lim_{t\to\infty} \beta_j^i(t) = 1$, we have $f'(\boldsymbol{p}, \infty) = f(\boldsymbol{p}) = (f_i(\boldsymbol{p}), \forall i \in \mathcal{N})$ where $f_i(\boldsymbol{p})$ is as in (2).

From Theorems 2 and 3, if the channel is perfect and (16) holds, Algorithm 1 asynchronously converges to the unique global optimal solution of non-convex Problem (NUM). If the channel is not perfect, although the algorithm still converges, optimality is not always guaranteed.

V. Simulation Results

To evaluate the performance of our proposed distributed algorithm, we develop a discrete-event simulator that implements Algorithms 1 and the IEEE 802.11 DCF access method.

We first consider a network with $N = 4$, $P^{\min} = 0.01$, and $P^{\max} = 0.99$. We set γ_1 = 6, γ_2 = 18, γ_3 = 36, and γ_4 = 54, all in Mbps. Utility parameter $\alpha = 0.5 < 1$. Notice that none of the previous NUM-based MAC algorithms (e.g, [2]–[4]) support α-fair utility functions with $\alpha \in (0, 1)$ because of non-convexity (see [5, Sections II and IV-A]). Each slot is 20 μs (as in 802.11a) and the simulation time is $20s$. We assume that from time $t = 0$ to $t = 10s$, the channel is perfect and $N = 4$. Then, from $t = 10s$ to $t = 20s$, the channel is lossy and $N = 3$ (i.e., user 4 leaves the network). Packet error rates are randomly selected between 0 and 0.01 (i.e., the maximum allowed packet error rate in 802.11a) at $t = 10s$ and then become fixed until $t = 20s$. Results are shown in Fig. 2. We see that Algorithm 1 converges to a small neighborhood of the optimal values very fast. It is also robust to the change of user population and channel conditions. Similar results have also been obtained for $\alpha \geq 1$.

It is well-known that 802.11 DCF has a *short-term fairness* problem, due to binary exponential backoff. Next, we compare 802.11 DCF with Algorithm 1 in terms of both system throughput and Jain's fairness index [12]. The short-term fairness is obtained using sliding windows with size of 200 slots. There are $N = 10$ users in the network and their fixed peak rates are randomly selected between 6 and 54 Mbps. Simulation time is $100s$. The results when α varies between 0.5 to 5 are shown in Fig. 3. We see that, parameter α acts as a *knob* to control the tradeoff between efficiency and fairness. By increasing α we can make the system more fair but less efficient (and vice versa). If $\alpha = 0.5$, then throughput is 29.7% higher than DCF (see Fig. 3(a)). Besides, for any choice of $\alpha \in [0.5, 5]$, the fairness is much better than DCF (Fig. 3(b)).

VI. Conclusion

In this paper, we designed a distributed contention-based MAC algorithm to solve a network utility maximization (NUM) without frequent explicit message passing among users. Our algorithm is fully asynchronous problem, enjoys fast convergence, and supports a wider range of utility functions compared to previously proposed NUM-based MAC algorithms. Simulation results show that our algorithm achieves a better efficiency-fairness trade-off compared with the IEEE 802.11 DCF. It is also robust to the changes of user population and channel conditions.

This work represents a first step towards building practical and utility-optimal random access protocols. Results can be extended in several directions. For example, it is possible to extend our work to a general interference model where each user may not interfere with every other user. The proof of convergence in Theorem 2 will still be valid after slight modifications. However, the performance may not be optimal due to the well-known hidden terminal problem.

APPENDIX

For any $\boldsymbol{p} \in \mathcal{P}$ and $t \geq t_0'$, the Jacobian $J(\boldsymbol{p}, t)$ is defined as an $N \times N$ matrix whose entry in row i and column j is $\partial f_i(\boldsymbol{p}, t)/\partial p_j$. We can show that,

$$\|J'(\boldsymbol{p}, t)\|_\infty \leq \ (|1-\alpha|/\alpha)\ \Psi\ \Phi(V'^{\min}, V'^{\max}), \tag{19}$$

$$\|J'(\boldsymbol{p}, t)\|_1 \ \leq \ (|1-\alpha|/\alpha)\ \left(\beta^{\max}(t)/\beta^{\min}(t)\right)\ \left((\gamma^{\max}/\gamma^{\min})\,\Gamma\right)^{1-\alpha}\ \Psi\ \Phi(V'^{\min}, V'^{\max}). \tag{20}$$

Let $\tilde{\boldsymbol{p}}, \hat{\boldsymbol{p}} \in \mathcal{P}$. From (16), (19), (20), and by Cauchy Schwarz inequality we have [13, pp. 635]:
$$\|f'(\tilde{\boldsymbol{p}}, t) - f'(\hat{\boldsymbol{p}}, t)\|_2 \leq \|J'(\boldsymbol{p}, t)\|_2\ \|\tilde{\boldsymbol{p}} - \hat{\boldsymbol{p}}\|_2 \leq \sqrt{\|J'(\boldsymbol{p}, t)\|_\infty \|J'(\boldsymbol{p}, t)\|_1}\ \|\tilde{\boldsymbol{p}} - \hat{\boldsymbol{p}}\|_2 < \|\tilde{\boldsymbol{p}} - \hat{\boldsymbol{p}}\|_2,$$
where $\boldsymbol{p}$ is any convex combination of $\tilde{\boldsymbol{p}}$ and $\hat{\boldsymbol{p}}$. Thus, for any $t \geq t_0'$, vector function $f'(\boldsymbol{p}, t)$ is a contraction mapping [13, pp. 181] and has a unique fixed point [13, pp. 183], denoted by $\boldsymbol{p}_t^*$. We also denote the unique fixed point of mapping $f'(\boldsymbol{p}, \infty)$ by $\boldsymbol{p}_\infty^*$. Thus,

$$\|f'(\boldsymbol{p}, t) - \boldsymbol{p}_t^*\|_2 \leq \eta_t\, \|\boldsymbol{p} - \boldsymbol{p}_t^*\|_2 \leq \eta\xi, \tag{21}$$

where $\eta_t = \|J'(\boldsymbol{p}, t)\|$, $\eta = \max_{t > t_0'} \eta_t$, and $\xi = \|\boldsymbol{p} - \boldsymbol{p}_t^*\|_2$. Note that $\eta < 1$, and ξ is bounded. Since $f'(\boldsymbol{p}, t)$ is continuous at $\boldsymbol{p}_t^*$ and $\lim_{t\to\infty} f'(\boldsymbol{p}, t) = f'(\boldsymbol{p}, \infty)$, we have $\lim_{t\to\infty} \boldsymbol{p}_t^* = \boldsymbol{p}_\infty^*$. In other words, $\forall \epsilon > 0$, $\exists t_0 \geq t_0'$, such that $\forall t \geq t_0$,

$$\|\boldsymbol{p}_t^* - \boldsymbol{p}_\infty^*\|_2 \leq \epsilon. \tag{22}$$

Together with (21), we have $\|f'(\boldsymbol{p}, t) - \boldsymbol{p}_\infty^*\|_2 \leq \|f'(\boldsymbol{p}, t) - \boldsymbol{p}_t^*\|_2 + \|\boldsymbol{p}_t^* - \boldsymbol{p}_\infty^*\|_2 \leq \eta\xi + \epsilon$. Similarly,

$$\begin{aligned}\|f'(f'(\boldsymbol{p}, t), t+1) - \boldsymbol{p}_\infty^*\|_2 &\leq \left\|f'(f'(\boldsymbol{p}, t), t+1) - \boldsymbol{p}_{t+1}^*\right\|_2 + \left\|\boldsymbol{p}_{t+1}^* - \boldsymbol{p}_\infty^*\right\|_2 \\ &\leq \eta\left(\|f'(\boldsymbol{p}, \boldsymbol{t}) - \boldsymbol{p}_\infty^*\|_2 + \left\|\boldsymbol{p}_{t+1}^* - \boldsymbol{p}_\infty^*\right\|_2\right) + \epsilon \leq \eta(\eta\xi + \epsilon + \epsilon) + \epsilon = \eta(\eta\xi + 2\epsilon) + \epsilon. \end{aligned} \tag{23}$$

For any $k \geq 0$, we recursively define $f'^k(\boldsymbol{p}, t) = f'(f'^{k-1}(\boldsymbol{p}, t), t+k-1)$ where $f'^0 = \boldsymbol{p}$. From (23), and by mathematical induction, we can show that for any $k \geq 0$,
$$\left\|f'^k(\boldsymbol{p}, t) - \boldsymbol{p}_\infty^*\right\|_2 \leq \eta^k\xi + \frac{2\left(1-\eta^k\right)}{1-\eta}\epsilon - \epsilon < \eta^k\xi + \frac{1+\eta}{1-\eta}\epsilon.$$
For any $\varepsilon > 0$, there exist k_ε such that if $k \geq k_\varepsilon$, then $\eta^k\xi \leq \frac{\varepsilon}{2}$. By choosing $\epsilon = \frac{1-\eta}{1+\eta}\frac{\varepsilon}{2}$,

$$\left\|f'^k(\boldsymbol{p}, t) - \boldsymbol{p}_\infty^*\right\|_\infty \leq \left\|f'^k(\boldsymbol{p}, t) - \boldsymbol{p}_\infty^*\right\|_2 < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \tag{24}$$

For all $t \geq t_0$, define $\epsilon_t' = \max_{k\geq 0}\left\|\boldsymbol{p}_{k+t}^* - \boldsymbol{p}_\infty^*\right\|_\infty$, $\varepsilon_t' = \max_{k\geq 0, \boldsymbol{p}' \in \mathcal{P}} \|f'^{k+t-t_0}(\boldsymbol{p}', t_0) - \boldsymbol{p}_\infty^*\|_\infty$, and

$$\varepsilon_t = \begin{cases} \max\left[\varepsilon_t',\ \frac{2(1+\eta)}{1-\eta}\ \epsilon_t'\right], & \text{if } t < t_0 + C, \\ \max\left[\varepsilon_t',\ \frac{2(1+\eta)}{1-\eta}\ \epsilon_t', \chi(C)\,\varepsilon_{t-C}\right], & \text{otherwise}, \end{cases} \tag{25}$$

where function $\chi(C) = \frac{1}{2}(\frac{3+\eta}{1+\eta}\,\eta^C + 1)$, integer constant $C = \lceil \log(\frac{1+\eta}{3+\eta})/\log(\eta)\rceil + 1$, and $\lceil\cdot\rceil$ denotes the ceiling function. From (22) and (24), $\{\epsilon_t'\}$ and $\{\varepsilon_t\}$ are infinite decreasing sequences and converge to zero as $t \to \infty$. Construct a new *time sequence* $\{\bar{t}_l\}$ where $\bar{t}_l = t_0 + lC$ for all integer $l \geq 0$. Since $\chi(C) < 1$, sequence $\{\varepsilon_t\}$ is also decreasing and in particular, we have $\lim_{l\to\infty} \varepsilon_{\bar{t}_l} = 0$. For each $l \geq 0$, define $\mathcal{P}_{\bar{t}_l} = \{\boldsymbol{p} : \|\boldsymbol{p} - \boldsymbol{p}^*_\infty\|_\infty \leq \varepsilon_{\bar{t}_l}\}$. It is clear that $\boldsymbol{p}^*_\infty \in \mathcal{P}_{\bar{t}_l}$ and $\mathcal{P}_{\bar{t}_{l+1}} \subseteq \mathcal{P}_{\bar{t}_l}$ for all $l \geq 0$. Furthermore, $\mathcal{P}_{\bar{t}_{l+l'}} \subset \mathcal{P}_{\bar{t}_l}$ for some finite l'. For any $\boldsymbol{p} \in \mathcal{P}_{\bar{t}_l}$,

$$\left\|\boldsymbol{p} - \boldsymbol{p}^*_{\bar{t}_l}\right\|_\infty \leq \|\boldsymbol{p} - \boldsymbol{p}^*_\infty\|_\infty + \left\|\boldsymbol{p}^*_{\bar{t}_l} - \boldsymbol{p}^*_\infty\right\|_\infty \leq \varepsilon_{\bar{t}_l} + \epsilon'_t.$$

From (24), we know that $\left\|f'^k(\boldsymbol{p}, \bar{t}_l) - \boldsymbol{p}^*_\infty\right\|_\infty < \eta^k \left(\varepsilon_{\bar{t}_l} + \epsilon'_{\bar{t}_l}\right) + \frac{1+\eta}{1-\eta}\epsilon'_{\bar{t}_l}$. If $\varepsilon_{\bar{t}_l} = \varepsilon'_{\bar{t}_l}$, then $\epsilon'_{\bar{t}_l} \leq \frac{1-\eta}{1+\eta}\frac{\varepsilon_{\bar{t}_l}}{2}$. On the other hand, if $\varepsilon_{\bar{t}_l} = \frac{2(1+\eta)}{1-\eta}\epsilon'_{\bar{t}_l}$, or if $\varepsilon_{\bar{t}_l} = \chi(C)\varepsilon_{\bar{t}_l - C}$, then $\varepsilon'_{\bar{t}_l} \leq \varepsilon_{\bar{t}_l}$ and $\epsilon'_{\bar{t}_l} \leq \frac{1-\eta}{1+\eta}\frac{\varepsilon_{\bar{t}_l}}{2}$. Thus, for all three possibilities in (25), we have

$$\|f'^C(\boldsymbol{p}, \bar{t}_l) - \boldsymbol{p}^*_\infty\|_\infty < \eta^C \left(\varepsilon_{\bar{t}_l} + \frac{1-\eta}{1+\eta}\frac{\varepsilon_{\bar{t}_l}}{2}\right) + \frac{1+\eta}{1-\eta}\frac{1-\eta}{1+\eta}\frac{\varepsilon_{\bar{t}_l}}{2} = \chi(C)\,\varepsilon_{\bar{t}_l} \leq \varepsilon_{\bar{t}_{l+1}}.$$

Thus, $\forall \boldsymbol{p} \in \mathcal{P}_{\bar{t}_l}$, $f'^C(\boldsymbol{p}, \bar{t}_l) \in \mathcal{P}_{\bar{t}_{l+1}}$. Since *synchronous convergence* and *box* conditions hold, Algorithm 1 globally and asynchronously converges to the unique fixed point $\boldsymbol{p}^*_\infty$ [13, pp. 431].

References

[1] J. Lee, A. Tang, J. Huang, M. Chiang, and A. Calderbank, "Reverse engineering MAC: A game-theoretic model," *IEEE J. on Selected Areas in Communications*, vol. 6, pp. 2741–2751, Jul. 2007.

[2] J. Lee, M. Chiang, and R. Calderbank, "Utility-optimal random-access control," *IEEE Trans. on Wireless Communications*, vol. 25, pp. 1135–1147, Aug. 2007.

[3] X. Wang and K. Kar, "Cross-layer rate control for end-to-end proportional fairness in wireless networks with random access," *IEEE J. on Selected Areas in Communications*, vol. 24, pp. 1548–1559, Aug. 2006.

[4] L. Chen, S. Low, and J. Doyle, "Joint congestion control and media access control design for ad hoc wireless networks," in *Proc. of IEEE INFOCOM*, Miami, FL, Mar. 2005.

[5] H. Mohsenian Rad, J. Huang, M. Chiang, and V. W. S. Wong, "Utility-optimal random access: Reduced complexity, fast convergence, and robust performance," *Submitted to IEEE Trans. Wireless Communications*, Dec. 2007, http://www.ece.ubc.ca/∼hamed/uora.pdf.

[6] L. Chen, S. H. Low, and J. C. Doyle, "Random access game and medium access control design," Submitted to *IEEE/ACM Trans. on Networking*, Dec. 2006.

[7] F. Calì, M. Conti, and E. Gregori, "Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit," *IEEE/ACM Trans. on Networking*, vol. 8, no. 6, pp. 785–799, 2000.

[8] D. P. Bertsekas and R. Gallager, *Data Communications*, 2nd ed. Prentice Hall, 1992.

[9] F. Kelly, "Charging and rate control for elastic traffic," *European Trans. on Telecommunication*, vol. 8, pp. 33–37, 1997.

[10] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. on Networking*, vol. 8, pp. 556–567, Oct. 2000.

[11] V. S. Borkar, "Asynchronous Stochastic Approximation," *SIAM J. on Control and Optim.*, vol. 36, pp. 840–851, May 1998.

[12] R. Jain, W. Hawe, and D. Chiu, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," Tech. Rep. DEC-TR-301, Sept. 1984.

[13] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice Hall, 1989.

Algorithm 1 *Executed by each user* $i \in \mathcal{N}$.

```
Allocate memory for $p_i$ and $\boldsymbol{m}^i = (m_1, \cdots, m_N)$.
Allocate memory for $\bar{n}_i^{\text{decd}}$ and $\bar{\boldsymbol{n}}_i^{\text{decd}} = (\bar{n}_{i,1}^{\text{decd}}, \cdots, \bar{n}_{i,N}^{\text{decd}})$.
Randomly choose $p_i \in \left[P_i^{\min}, P_i^{\max}\right]$.
Randomly choose $m_j^i \in \left[M^{\min}, M^{\max}\right]$ for all $j \in \mathcal{N}$.
Choose $\bar{n}_i^{\text{idle}} = 1$ and $\bar{n}_{i,j}^{\text{decd}} = 1$ for all $j \in \mathcal{N}$.
Broadcast the fixed data rate $\gamma_i$ to all other users.
repeat
    Transmit with probability $p_i$.
    Update $\bar{n}_i^{\text{idle}}$ and $\bar{\boldsymbol{n}}_i^{\text{decd}}$ according to Eqs. (12) and (13).
    if $t \in T_{i,p}$ then
        Update $p_i = \left[1 / \left(1 + \sqrt[\alpha]{{\gamma_i}^{\alpha-1} \sum_{j \in \mathcal{N} \setminus \{i\}} m_j^i}\right)\right]_{P_i^{\min}}^{P_i^{\max}}$.
    end if
    if $t \in T_{j,m}^i$ then
        Update $m_j^i$ according to Eq. (14).
    end if
until the user decides to leave the network.
Broadcast termination message.
```

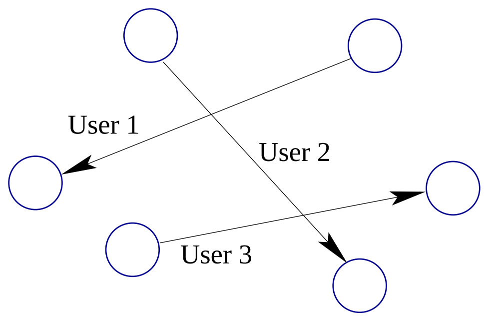

Fig. 1. A single-hop wireless ad-hoc network with $N = 3$ users. Each user includes a wireless link and its dedicated transmitter and receiver nodes.

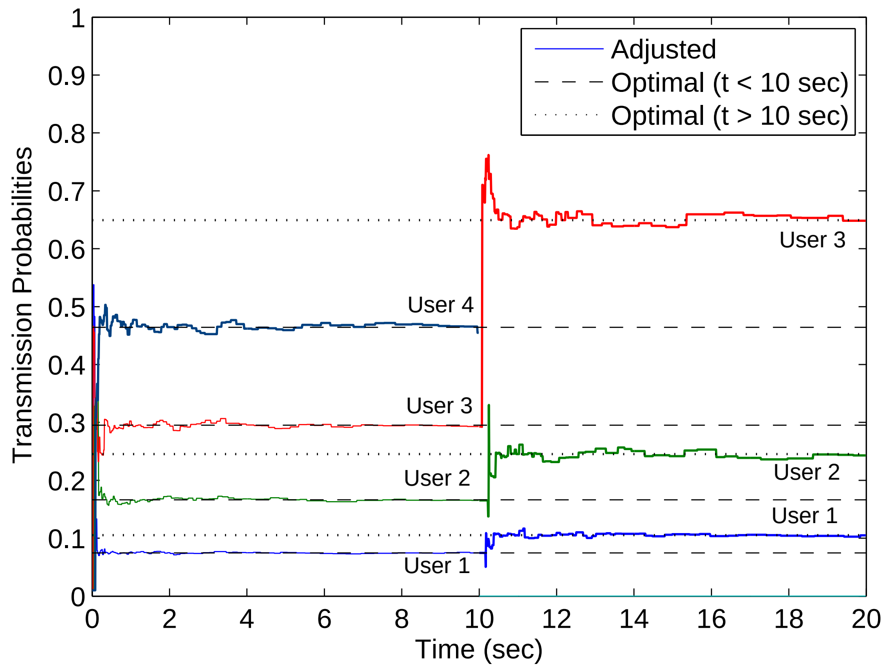

Fig. 2. Simulation results for Algorithm 1 when $\alpha = 0.6$. The number of users and the features of the communication channel change after $t = 10s$. The optimal transmission probabilities before $t = 10s$ (i.e., dashed lines) and after $t = 10s$ (i.e., dotted lines) are accurate and obtained using [5, Algorithm 1].

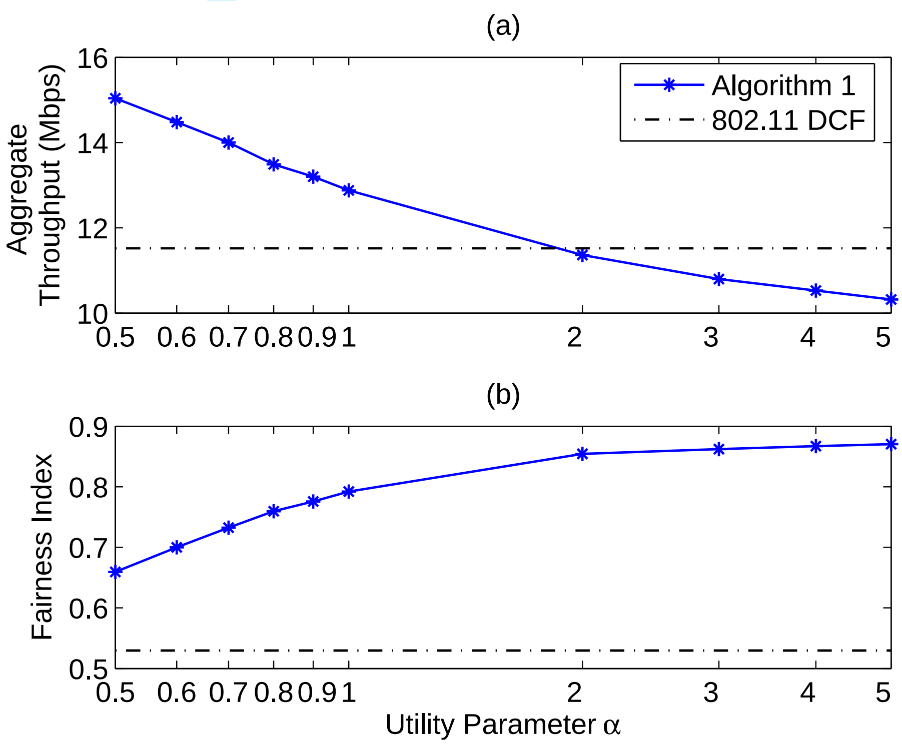

Fig. 3. Comparison between Algorithm 1 and 802.11 DCF when the number of users $N = 10$.