Decentralized Cross-Layer Optimization for Multichannel Aloha Wireless Networks*

Guowang Miao †
Ye (Geoffrey) Li
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332–0250

Ananthram Swami
Army Research Laboratory
Communication & Information Sciences
Adelphi, MD 20783-1197, USA

***Abstract*—While most existing channel aware Aloha schemes focus on wireless networks where each user intends to send only one traffic flow, and interferes with all the other users, some wireless networks may have more complicated traffic distribution and the transmissions of different users may interfere with different groups of users. In this paper, we consider schemes for the decentralized cross-layer optimization of multichannel random access, in which users are not necessarily within the transmission ranges of all other users, and each user may choose to send packets to or receive packets from different users simultaneously. With cross-layer design, users are configured according to local neighborhood information to adapt to inhomogeneous network characteristics. It is demonstrated by simulation that the proposed scheme significantly outperforms existing channel aware Aloha schemes due to its exploitation of both multiuser diversity and the inhomogeneous characteristics of traffic distribution in the network.**

I. Introduction

Cross-layer optimization is becoming a more and more important research area in wireless communications. Many researches have focused on centralized cross-layer scheduling, where the best performance can be obtained with the help of *channel state information* (CSI) from all active users. However, CSI feedback consumes a large amount of resources in centralized schemes, especially for MIMO and OFDM based networks. Although recent researches [1] are trying to design systems with reduced feedback, the signalling overhead is still large when many users are at high mobility. To reduce CSI feedback, decentralized cross-layer design approaches can be considered. Opportunistic random access schemes have been studied in [2], [3], for example. For opportunistic random access, each user exploits its own CSI for transmission controls. In [2], each user transmits only if its channel fading level is above a certain pre-determined threshold, which is chosen to maximize successful transmission probability. A channel aware multicarrier random access scheme has been proposed in [3], where each user selects some subcarriers with the best channel gains. All these works are for wireless networks where all users transmit to a common receiver. However, this scenario does not fit many wireless communication environments such as sensor networks, mobile ad hoc networks, and so on. Besides, existing policies [2], [3] are designed such that each user has the same transmission probability. Although this guarantees absolute fairness among all users, the network performance is not optimal when the traffic flows are not uniformly distributed in the network or when the link characteristics are not identical. These motivate us to propose a novel scheme in this paper.

* This work was supported by Intel, SABA Program of Motorola and the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-20-0011.
† Corresponding author. Email: gmiao3@gatech.edu.

We will consider wireless networks where users are not necessarily within the transmission ranges of the others. Each user may intend to send packets to different destinations. The proposed scheme successfully exploits both multiuser diversity and the inhomogeneous characteristics of traffic distribution, and can be applied to different types of wireless networks such as wireless sensor networks and mobile ad hoc networks.

II. System Description

Consider multichannel wireless networks with K subchannels. All channels between pairs of users are reciprocal, i.e. when no interference exists, User A can reliably receive signal from User B if and only if User B can reliably receive signal from User A with the same channel gain. Each user has ideal knowledge of its own CSI and applies the same transmission control policy. No communication pair has instantaneous cooperations such as exchange of CSI, subchannel selections, and so forth. A user can not transmit and receive simultaneously on the same subchannel; however, it may transmit over a set of subchannels and receive over a different set of subchannels. A node cannot receive any signals successfully if any of its interfering neighbors is transmitting simultaneously on the same subchannel. Each user may intend to send packets to or receive packets from different users. Each user transmits at a constant data rate. Different traffic flows may have different rates. The transmit power is allocated such that the signal can be reliably received, i.e. the receive power is above a certain level P_r, which may be different for different traffic flows or users. The transmit power is directly given by $P_t = P_r/h$.

The reliable transmission data rate is given by $R(P_r)$, which is assumed to be strictly concave in P_r. Each user is subject to both an average transmit power constraint $\mathbf{E}(P_t) \le \overline{P}$, where $\mathbf{E}$ is the expectation value, and an instantaneous transmit power constraint $P_t \le P_m$. Not all users are necessarily within the transmission ranges of the others. For simplicity, those that can communicate with each other are assumed to experience homogeneous channels, i.e. the channel gains of different links are independent identically distributed with probability density $f(h)$ and distribution $F(h)$.

The following notations are used:

- $\mathcal{V} = \{1, 2, ..., N\}$: set of active users transmitting or receiving packets;
- $\mathcal{E} = \{(i,j)_k | i, j \in \mathcal{V}, i \neq j; k = 1, 2, ..., K\}$: set of transmission links over all K subchannels;
- $\mathcal{L} = \{(i,j) | i, j \in \mathcal{V}, i \neq j$, i and j can receive packets from each other$\}$: set of links available for communication;
- $G(\mathcal{V}, \mathcal{E}, \mathcal{L})$: directed graph denoting the network;
- $\mathcal{N}_i = \{j | \forall j \in \mathcal{V}, s.t. (i,j) \in \mathcal{L}\}$: interfering neighbor set of i;
- $\mathcal{T}_i = \{j | (i,j) \in \mathcal{L}, (i,j)$ is backlogged$\}$: set of users receiving packets from i;
- $\mathcal{R}_j = \{i | (i,j) \in \mathcal{L}, (i,j)$ is backlogged$\}$: set of users sending packets to j.

III. CROSS-LAYER OPTIMIZATION FOR SLOTTED ALOHA

Traditionally, the sender of each backlogged link $(i,j) \in \mathcal{L}$ independently sends a packet in every slot with a certain transmission probability $p_{(i,j)}$ in slotted Aloha. With channel knowledge, the sender will transmit only when the channel is not deeply faded. The MAC layer transmission control is described as follows:

$\forall i, i \in \mathcal{V}$, *user* i *sends packets to user* j *on subchannel* k *when the following conditions are satisfied:* $j \in \mathcal{T}_i$; $\forall a \in \mathcal{T}_i, h_{(i,j)_k} \ge h_{(i,a)_k}$; $h_{(i,j)_k} \ge \overline{H}_{(i,j)_k}$, *in which* $\overline{H}_{(i,j)_k}$ *is the predetermined channel gain threshold for link* $(i,j)_k$.

Proper thresholds $\{\overline{H}_{(i,j)_k} | (i,j)_k \in \mathcal{E}\}$ and the transmission data rates of all traffic flows, i.e. the power allocations, should be determined so that the overall network performance can be optimized from certain perspectives.

A. Cross-Layer Optimization Modeling

According to the above transmission policy and the homogeneity assumption, the probability of a transmission happening on link $(i,j)_k \in \mathcal{E}$ is given by:

$$p_{(i,j)_k} = \frac{1}{|\mathcal{T}_i|}\left(1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k})\right) \tag{1}$$

where $|\cdot|$ denotes the cardinality of the set. The proof of (1) is given in Appendix A. The transmission probability of user i on subchannel k is:

$$p_{i_k} = \sum_{j \in \mathcal{T}_i} p_{(i,j)_k} = \sum_{j \in \mathcal{T}_i} \frac{1}{|\mathcal{T}_i|}\left(1 - F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k})\right) \tag{2}$$

Hence, the throughput on link $(i,j)_k$ is:

$$T_{(i,j)_k} = p_{(i,j)_k}(1 - p_{j_k}) \prod_{a \in \mathcal{N}_j, a \neq i} (1 - p_{a_k}) R(P_{r(i,j)_k}), \tag{3}$$

where $(1 - p_{j_k}) \prod_{a \in \mathcal{N}_j, a \neq i}(1 - p_{a_k})$ is the probability that neither user j nor its neighboring users except user i transmit on subchannel k. $R(P_{r(i,j)_k})$ is the achieved data rate under the given power constraint.

The average consumed power on link $(i,j)_k$ is:

$$\begin{aligned}\mathbf{E}(P_{(i,j)_k}) &= \int_0^\infty \Pr\{H_{(i,j)_k} = h, \text{and} (i,j)_k \text{transmits}\} \frac{P_{r(i,j)_k}}{h}\, \mathrm{d}h \\ &= \int_{\overline{H}_{(i,j)_k}}^\infty f(h) F^{|\mathcal{T}_i|-1}(h) \frac{P_{r(i,j)_k}}{h} \mathrm{d}h \\ &= \frac{1}{|\mathcal{T}_i|} \int_{\overline{H}_{(i,j)_k}}^\infty \frac{P_{r(i,j)_k}}{h} \mathrm{d}F^{|\mathcal{T}_i|}(h)\end{aligned} \tag{4}$$

Hence, the received power is:

$$P_{r(i,j)_k} = \mathbf{E}(P_{(i,j)_k}) \left(\int_{\overline{H}_{(i,j)_k}}^\infty \frac{\mathrm{d}F^{|\mathcal{T}_i|}(h)}{h}\right)^{-1}. \tag{5}$$

Define the configuration of the whole network as $\mathcal{C} = \{\overline{\mathcal{H}}, \mathcal{P}_r\}$, where $\overline{\mathcal{H}} = \{\overline{H}_{(i,j)_k} | (i,j)_k \in \mathcal{E}, j \in \mathcal{T}_i\}$ and $\mathcal{P}_r = \{P_{r(i,j)_k} | (i,j)_k \in \mathcal{E}, j \in \mathcal{T}_i\}$. With the power constraints, the optimal configuration of the whole network, $\mathcal{C}^* = \{\overline{\mathcal{H}}^*, P'^*\}$, that achieves proportional fairness [4] among all subchannels carrying traffic flows is given by:

$$\begin{aligned}&\mathcal{C}^* = \arg \max_{\{\overline{\mathcal{H}}, \mathcal{P}_r\}} \sum_{(i,j)_k \in \mathcal{E}, j \in \mathcal{T}_i} \ln(T_{(i,j)_k}), \\ &\text{such that} \\ &\sum_{j \in \mathcal{T}_i, k=1,...,K} \frac{1}{|\mathcal{T}_i|} \int_{\overline{H}_{(i,j)_k}}^\infty \frac{P_{r(i,j)_k}}{h} \mathrm{d}F^{|\mathcal{T}_i|}(h) \le \overline{P}, \\ &\sum_k \max_{j \in \mathcal{T}_i} \frac{P_{r(i,j)_k}}{\overline{H}_{(i,j)_k}} \le P_m, \forall i, i \in \mathcal{V}.\end{aligned} \tag{6}$$

where the constraints are the average and instantaneous power constraints respectively. The solution to problem (6) involves both the configuration of threshold $\overline{\mathcal{H}}^*$ and the received power level P'^*, i.e. the transmission data rate on each link once they decide to transmit.

IV. DESIGN OF CROSS-LAYER OPTIMIZATION

It is difficult to find the optimal solution to problem (6). Since the objective function of problem (6) is equivalent to:

$$\begin{aligned}\mathcal{C}^* = \arg \max_{\{\overline{\mathcal{H}}, \mathcal{P}_r\}} \sum_{(i,j)_k \in \mathcal{E}, j \in \mathcal{T}_i} \Bigg(&\ln \Big(p_{(i,j)_k}(1 - p_{j_k}) \\ \prod_{a \in \mathcal{N}_j, a \neq i} (1 - p_{a_k})\Big) &+ \ln\left(R(P_{r(i,j)_k})\right)\Bigg),\end{aligned} \tag{7}$$

we decompose the problem into two related problems, and find suboptimal solutions for them. The solution to $\overline{\mathcal{H}}^*$ is given by:

$$\overline{\mathcal{H}}^* = \arg\max_{\overline{\mathcal{H}}} \sum_{(i,j)_k\in\mathcal{E}, j\in\mathcal{T}_i} \Bigg(\ln \big(p_{(i,j)_k}(1-p_{j_k}) \prod_{a\in\mathcal{N}_j, a\neq i} (1-p_{a_k})\big)\Bigg), \tag{8}$$

And the transmission rate on each link, i.e. the received power level P_r^*, is determined by:

$$P_r^* = \arg\max_{P_r} \sum_{(i,j)_k\in\mathcal{E}, j\in\mathcal{T}_i} \ln \big(R(P'_{(i,j)_k})\big),$$

$$\text{such that}$$

$$\sum_{j\in\mathcal{T}_i, k=1,\ldots,K} \frac{1}{|\mathcal{T}_i|}\int_{\overline{H}_{(i,j)_k}}^{\infty} \frac{P_{r(i,j)_k}}{h} \mathrm{d}F^{|\mathcal{T}_i|}(h) \le \overline{P}, \tag{9}$$

$$\sum_k \max_{j\in\mathcal{T}_i} \frac{P_{r(i,j)_k}}{\overline{H}_{(i,j)_k}} \le P_m, \forall i, i\in\mathcal{V}.$$

in which $\{\overline{H}^*_{(i,j)_k}\}$ is the solution given by (8). From problem (6), (8) and (9), we know that problem (8) optimizes the first part of the objective function of (6) without constraints, while (9) optimizes the second part in the objective function of (6) with transmit power constraints.

Theorem 1: The optimal predetermined channel gain threshold for any link $(I,J)_k \in \mathcal{E}$ where $J\in\mathcal{T}_I$, $\overline{H}^*_{(I,J)_k}$, as defined in (8), is given by

$$\overline{H}^*_{(I,J)_k} = F^{-1}\left[(1-\frac{|\mathcal{T}_I|}{|\mathcal{R}_I| + \sum_{j\in\mathcal{N}_I}|\mathcal{R}_j|})^{\frac{1}{|\mathcal{T}_I|}}\right] \tag{10}$$

Theorem 2: Assuming the strict concavity of the data rate function $R(P_r)$, the optimal received power level on any link $(i,j)_k \in \mathcal{E}$ where $j\in\mathcal{T}_i$, $P^*_{r(i,j)_k}$, as defined in (8), is given by

$$P^*_{r(i,j)_k} = \min\Bigg(\frac{\overline{P}}{K}\big(\int_{\overline{H}^*_{(i,j)_k}}^{\infty} \frac{1}{h}\mathrm{d}F^{|\mathcal{T}_i|}(h)\big)^{-1}, P_m\overline{H}_{(i,j)_k}/K\Bigg), \tag{11}$$

in which $\overline{H}^*_{(i,j)_k}$ is given by *Theorem 1*.

Theorem 1 and *Theorem* 2 are proved in Appendix B and Appendix C respectively. Observing (10), we can see that the optimal threshold of user i depends on $|\mathcal{T}_i|$, $|\mathcal{R}_i|$, and $|\mathcal{R}_j|, j\in\mathcal{N}_i$. The first two values constitute local information while $|\mathcal{R}_j|, j\in\mathcal{N}_i$ are information about interfering neighbors. This two-hop knowledge, $|\mathcal{R}_j|, j\in\mathcal{N}_i$, i.e. how may flows each interfering neighbor receives, can be obtained through data broadcasting each time this number changes due to some variation in the network topology. This incurs only trivial signalling overhead. If we assume that each user has no two-hop knowledge, the number of flows each interfering neighbor receives needs to be estimated. Since the transmission of each interfering neighbor $j\in\mathcal{N}_i$ will be detected by User i, i.e. $|\mathcal{T}_j|$ is available, User i can assume $|\mathcal{T}_j|$ to be $|\mathcal{R}_j|$. This is reasonable, since for usual reliable data transmissions, a data flow is always accompanied by an ACK flow in the reverse direction. Hence, instead of (10), the transmission threshold with one-hop knowledge, i.e. local knowledge, is given by:

$$\overline{H}^*_{(I,J)_k} = F^{-1}\left[(1-\frac{|\mathcal{T}_I|}{|\mathcal{R}_I| + \sum_{j\in\mathcal{N}_I}|\mathcal{T}_j|})^{\frac{1}{|\mathcal{T}_I|}}\right]. \tag{12}$$

V. Simulation Results

Consider the wireless network in Fig. 1 as an example. The arrows show the traffic flows in the network. The set of interfering users is $\mathcal{L} = \{(1,3), (1,2), (1,4), (2,3), (2,6), (3,4), (3,5), (3,6), (4,6), (4,5), (5,6), (5,7), (5,8), (6,7), (6,8), (7,8), (7,9), (8,9)\}$. In the following, the network is to operate on both single channel and multichannel architectures. Different schemes will be implemented and detailed performance comparisons provided.

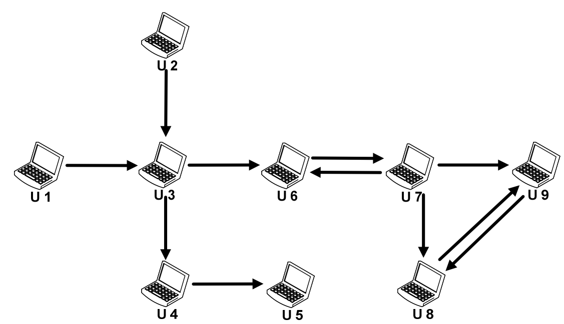

Fig. 1: Network architecture

A. Scenario 1: Single Channel Network

Assume that the network operates with one channel. For simplicity, we assume Rayleigh fading channel and $R(P) = W\log(1+\frac{HP}{Wn_o})$ where W is the channel bandwidth and n_o is the noise power per unit bandwidth. We will compare the performance of the proposed cross-layer transmission policy with the channel-aware Aloha in [2], and the traditional Aloha in [5], which does not consider cross-layer optimizations. For traditional Aloha transmissions, in order to make the comparison meaningful, the same average power constraint and the instantaneous power constraint are enforced. Since there is no cooperation between the MAC and PHY layers, the PHY layer keeps on transmitting except when the channel is deeply faded. In order to satisfy the average power constraint, the transmission threshold is chosen so that the average data rate is maximized, i.e. $\overline{H} = \arg\max_H(1-F(H))R(P_r)$ subject to the instantaneous power constraint, and P_r is given by (5). The threshold is found through a linear search. Fig. 2 shows the logarithm throughput comparison of the whole network when the channel has different average channel gains. The "TwoHop" curve represents the result of the proposed scheme when each user has two-hop information of the neighboring users, while the "OneHop" curve represents the result when each user has only one-hop information. As we can see, with only one-hop knowledge, the system has slight performance degradation as compared with the transmissions when two-hop knowledge is available. Curve "QIN" shows the performance of [2], which assumes that each user knows how many users are in the whole network. Curve "Traditional" shows the result using the traditional optimal Aloha. As shown in Fig. 2, due to the advantage of cross layer design, the proposed scheme considerably outperforms traditional optimal Aloha. Besides, by exploiting the neighborhood information of each user, the proposed method also outperforms the channel-aware Aloha scheme in [2]. This is due to the consideration of the inhomogeneous traffic distribution in the proposed scheme, so that channels are better utilized.

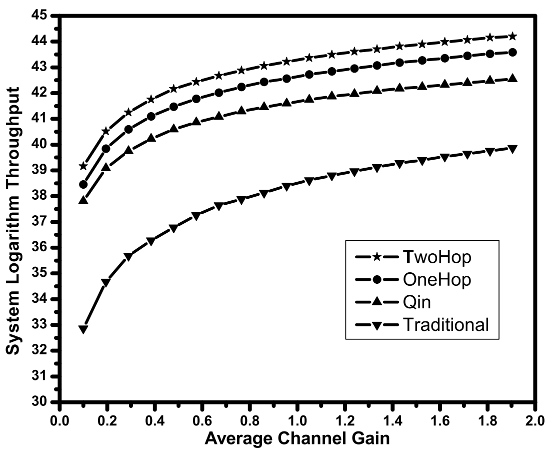

Fig. 2: Throughput comparison with proportional fairness constraint. system parameters: $P_m = 50\text{dBm}, \overline{P} = 43\text{dBm}$, bandwidth $W = 100\text{Hz}$, noise power $n_o = 0.001\text{W/Hz}$.

B. Scenario 2: Multichannel Network

Consider the same wireless network as in scenario 1 except that there are five subchannels. Besides implementing the schemes in scenario 1 for the multichannel environment, we also run the network with the proposed "CAMCRA" in ([3]) for more comparisons. During each transmission slot, "CAMCRA" chooses c subchannels with the c largest significant gains, where $c = \max\left(1, \lfloor\frac{\text{subchannel number}}{\text{user number}}\rfloor\right)$. Then the method in ([2]) is applied on each subchannel given that each user knows how many users are using the subchannel. Since the number of users in each subchannel is a random variable, ([3]) proposes to use $\max\left(1, \frac{\text{user number}}{\text{subchannel number}}\right)$ as an estimate. As shown in Fig. 3, when each user has only one-hop information, the proposed policy performs nearly the same as the "CAMCRA" in [3] for this network configuration. However, when each user has two-hop neighborhood information, the scheme proposed in this paper outperforms the "CAMCRA" in [3]. This is also due to the optimized transmission probability settings and the power allocation for each traffic flow according to the inhomogeneous traffic distribution in the network.

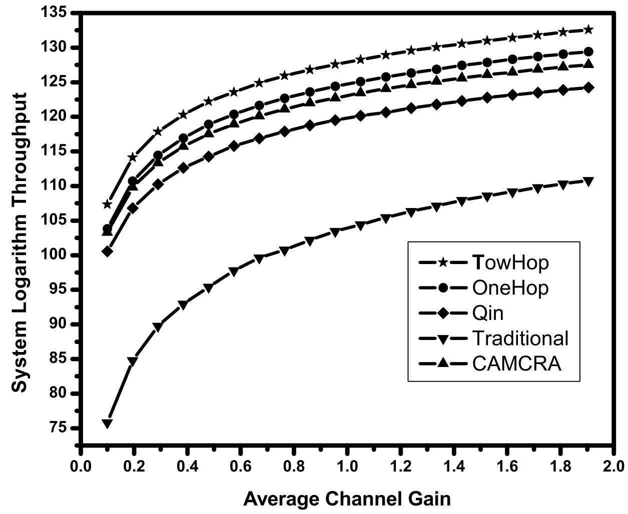

Fig. 3: Five channel network throughput comparison with proportional fairness constraint. system parameters: $P_m = 50\text{dBm}, \overline{P} = 43\text{dBm}$, bandwidth $W = 100\text{Hz}$, noise power $n_o = 0.001\text{W/Hz}$.

VI. Conclusion

We have proposed a joint PHY-MAC layer optimization policy for multichannel Aloha wireless networks in which not all users are necessarily within the transmission ranges of each other, and each user may intend to send packets to different destinations. Assuming ideal information exchange of the MAC and PHY layers, users are configured according to their neighborhood information to adapt to the inhomogeneous network characteristics. The design consists of three steps: (1) neighborhood information collection; (2) transmission control of the MAC layer; and (3) power allocation for each traffic flow on each subchannel. Simulation results show that the proposed scheme outperforms traditional Aloha due to the exploitation of multiuser diversity through cross-layer design. Besides, the proposed scheme also performs much better than existing channel aware Aloha schemes due to the consideration of the inhomogeneous characteristics of traffic distribution in the network. The scheme developed here can be applied to different kinds of wireless networks such as wireless sensor networks and mobile ad hoc networks to fully exploit system capacity.

Appendix A
Proof of Transmission Probability

$$\begin{aligned} p_{(i,j)_k} &= \Pr\left(h_{(i,j)_k} = \max_{a\in\mathcal{T}_i}(h_{(i,a)_k}), h_{(i,j)_k} \geq \overline{H}_{(i,j)_k}\right) \\ &= \Pr\left(h_{(i,j)_k} = \max_{a\in\mathcal{T}_i}(h_{(i,a)_k})\right). \end{aligned}$$

$$\begin{aligned}
&\Pr\left(h_{(i,j)_k} \geq \overline{H}_{(i,j)_k} \middle| h_{(i,j)_k} = \max_{a\in\mathcal{T}_i}(h_{(i,a)_k})\right)\\
= &\frac{1}{|\mathcal{T}_i|}\ \Pr\left(\max_{a\in\mathcal{T}_i}(h_{(i,a)_k}) \geq \overline{H}_{(i,j)_k}\right)\\
= &\frac{1}{|\mathcal{T}_i|}\left(1-\prod_{a\in\mathcal{T}_i}\Pr\left(h_{(i,a)_k} < \overline{H}_{(i,j)_k}\right)\right)\\
= &\frac{1}{|\mathcal{T}_i|}\left(1-F^{|\mathcal{T}_i|}(\overline{H}_{(i,j)_k})\right)
\end{aligned}$$

Appendix B
Proof of Theorem 1

Proof: Denote the objective function as:

$$\begin{aligned}
T(p) = \sum_{(i,j)_k\in\mathcal{E}, j\in\mathcal{T}_i} &\Bigg(\ln\left(p_{(i,j)_k}\right)+\ln\left(1-p_{j_k}\right)+\\
&\sum_{a\in\mathcal{N}_j, a\neq i}\ln\left(1-p_{a_k}\right)\Bigg),
\end{aligned} \tag{13}$$

where $p=\{p_{(i,j)_k}|(i,j)_k\in\mathcal{E}, j\in\mathcal{T}_i\}$, $p_{(i,j)_k}$ and p_{i_k} are given by (1) and (2) respectively. It is easy to see that both $\ln(p_{(i,j)_k})$ and $\ln\left(1-p_{i_k}\right)=\ln\left(1-\sum_{j\in\mathcal{T}_i}p_{(i,j)_k}\right)$ are stricly concave functions of any $p_{(i,j)_k}\in p$. Hence, $T(p)$ is a strictly concave function of p and has a unique global maximum p^* which satisfies $\nabla T(p)|_{p^*}=0$.

$$\begin{aligned}
&\frac{\partial T(p)}{\partial p_{(I,J)_k}} = \frac{\partial}{\partial p_{(I,J)_k}}\sum_{(i,j)_k\in\mathcal{E},j\in\mathcal{T}_i}\Bigg(\ln\left(p_{(i,j)_k}\right)+\\
&\ln\left(1-p_{j_k}\right)+\sum_{a\in\mathcal{N}_j,a\neq i}\ln\left(1-p_{a_k}\right)\Bigg)\\
= &\frac{\partial}{\partial p_{(I,J)_k}}\ln\left(p_{(I,J)_k}\right)+\frac{\partial}{\partial p_{(I,J)_k}}\sum_{(i,I)_k\in\mathcal{E},i\in\mathcal{R}_I}\ln\left(1-p_{I_k}\right)\\
&+\frac{\partial}{\partial p_{(I,J)_k}}\sum_{(i,j)_k\in\mathcal{E},j\in\mathcal{T}_i}\left(\sum_{a\in\mathcal{N}_j}\ln(1-p_{a_k})-\ln(1-p_{i_k})\right)\\
= &\frac{1}{p_{(I,J)_k}}+|\mathcal{R}_I|\frac{\partial}{\partial p_{(I,J)_k}}\ln\left(1-p_{I_k}\right)\\
&+\frac{\partial}{\partial p_{(I,J)_k}}\sum_{(i,j)_k\in\mathcal{E},j\in\mathcal{T}_i}\sum_{I\in\mathcal{N}_j}\ln(1-p_{I_k})\\
&-\frac{\partial}{\partial p_{(I,J)_k}}\sum_{(I,j)_k\in\mathcal{E},j\in\mathcal{T}_I}\ln(1-p_{I_k})\\
= &\frac{1}{p_{(I,J)_k}}+|\mathcal{R}_I|\frac{\partial}{\partial p_{(I,J)_k}}\ln\left(1-p_{I_k}\right)\\
&+\frac{\partial}{\partial p_{(I,J)_k}}\sum_{(i,j)_k\in\mathcal{E},j\in\mathcal{T}_i}\sum_{j\in\mathcal{N}_I}\ln(1-p_{I_k})\\
&-|\mathcal{T}_I|\frac{\partial}{\partial p_{(I,J)_k}}\ln(1-p_{I_k})\\
= &\frac{1}{p_{(I,J)_k}}+(|\mathcal{R}_I|-|\mathcal{T}_I|)\frac{\partial}{\partial p_{(I,J)_k}}\ln\left(1-p_{I_k}\right)\\
&+\frac{\partial}{\partial p_{(I,J)_k}}\sum_{j\in\mathcal{N}_I}|\mathcal{R}_j|\ln(1-p_{I_k})\\
= &\frac{1}{p_{(I,J)_k}}-(|\mathcal{R}_I|-|\mathcal{T}_I|+\sum_{j\in\mathcal{N}_I}|\mathcal{R}_j|)\frac{1}{1-p_{I_k}}
\end{aligned}$$

Since $\frac{\partial T(p)}{\partial p_{(I,J)_k}}\Big|_{p^*}=0$, $p^*_{(I,J)_k}=\frac{1-p^*_{I_k}}{|\mathcal{R}_I|-|\mathcal{T}_I|+\sum_{j\in\mathcal{N}_I}|\mathcal{R}_j|}$. Besides, $p^*_{I_k}=\sum_{j\in\mathcal{T}_I}p^*_{(I,j)_k}=|\mathcal{T}_I|p^*_{(I,J)_k}$. We get:

$$p^*_{(I,J)_k}=\frac{1}{|\mathcal{R}_I|+\sum_{j\in\mathcal{N}_I}|\mathcal{R}_j|} \tag{14}$$

According to (1) and (14), the optimal $\overline{H}^*_{(I,J)_k}$ is given by (10).

Appendix C
Proof of Theorem 2

Proof: According to (10), we can see that $\overline{H}^*_{(i,j)_k}$ is independent of j and k. Hence,the first constraint of (9) is $\sum_{j\in\mathcal{T}_i,k=1,\ldots,K}P_{r(i,j)_k}\leq\overline{P}|\mathcal{T}_i|\left(\int_{\overline{H}^*_{(i,j)_k}}^{\infty}\frac{1}{h}\mathrm{d}F^{|\mathcal{T}_i|}(h)\right)^{-1}$, while the second constraint of equals $\frac{P_{r(i,j)_k}}{\overline{H}_{(i,j)_k}}\leq\frac{P_m}{K}$ since subchannels are symmetric. Since the data rate function $R()$ is assumed to be a strictly concave function, and $\ln$ is a nondecreasing and concave function, it is easy to show that function $\ln(R(x))$ is also strictly concave for all x. Hence, $\sum_{(i,j)_k\in\mathcal{E},j\in\mathcal{T}_i}\ln\left(R(P_{r(i,j)_k})\right)\leq|\mathcal{T}_i|\ln\left(R(\sum_{(i,j)_k\in\mathcal{E},j\in\mathcal{T}_i}P_{r(i,j)_k})/|\mathcal{T}_i|\right)$. The equation holds if and only if $P_{r(i,j_1)_k}=P_{r(i,j_2)_k}$ for any $j_1,j_2\in\mathcal{T}_i$. Then, it is easy to see that (11) satisfies both constraints, and the objective value will be maximized when one constraint takes effect while satisfying the other constraint.

References

[1] G. Miao, and Z. Niu, "Practical Feedback Design based OFDM Link Adaptive Communications over Frequency Selective Channels," in *Proc. IEEE Conf. Commun. (ICC' 2006)*, Istanbul, Turkey, June 2006, pp. 4624-4629.

[2] X. Qin and R. Berry, "Exploiting multiuser diversity for medium access control in wireless networks," in *Proc. IEEE INFOCOM 2003*, San Francisco, CA, Apr. 2003, pp. 1084-1094.

[3] G. Ganesan, G. Song, and Y. (G.) Li, "Asymptotic throughput analysis of distributed multichannel random access schemes," in *Proc. IEEE Conf. Commun. (ICC' 2005)*, Seoul, South Korea, May 2005, pp. 3637-3641.

[4] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol 48, no. 6, pp. 1277-1294, June 2002.

[5] K. Kar, S. Sarkar, and L. Tassiulas, " Achieving proportional fairness using local information in Aloha networks," *IEEE Trans. Autom. Control*, vol. 49, no. 10, pp. 1858-1863, Oct. 2004.