COOPERATIVE TRANSMISSIONS FOR RANDOM ACCESS WIRELESS NETWORKS WITH FREQUENCY SELECTIVE FADING

L. Dong, J. Yu, and A. P. Petropulu

Electrical and Computer Engineering Department, Drexel University, Philadelphia, PA 19104
{ld56, yujie, athina}@cbis.ece.drexel.edu

ABSTRACT

A random access protocol for wireless networks was recently proposed that by exploiting cooperation of network nodes can resolve collisions and thus achieve high throughput. In this paper we propose a multichannel extension of that approach that can lead to throughput improvement and significantly reduce packet delays at high traffic loads, while at the same time can handle frequency selective fading. The channel is divided into separable subchannels and each user can transmit packets over multiple subchannels. We propose schemes for resolving collisions on the various subchannels in a way that minimizes the average processing time for each collided packet. At the physical layer we propose an OFDM approach, where the subchannels are groups of carriers.

I. INTRODUCTION

Random access (RA) schemes have been shown to be the best choice for bursty traffic, e.g., *multimedia traffic*. A well studied RA protocol is the ALOHA protocol, which is efficient when the traffic load is low. However, at a higher traffic load, collisions occur frequently and slots are wasted. A collision resolution approach, namely the network-assisted diversity multiple access (NDMA) was recently proposed in [5] [2]. According to NDMA, the collided packets are saved instead of being discarded. For a K-fold collision, the collided users are required to keep retransmitting their packets for an extra $K-1$ slots following the collision slot. Combining the originally collided packets and their retransmissions, the base station (BS) can recover the packets. NDMA exploits time diversity to resolve the collided packets, thus requiring the channel to change between slots, which is not very realistic. In [3], [4], a cooperative scheme was proposed, where instead of requiring the collided users to retransmit their packets as in NDMA, other network nodes designated as relays transmit the signal that they received during the collision slot. The spatial diversity introduced by the relaying was shown to improve network performance, while the retransmission scheme maintains an evenly distributed network power, all at minimal control overhead.

The cooperative scheme in [3], [4], assumes that the channel is flat fading during each slot. However, the wireless channel is frequency selective (dispersive) in nature. The NDMA approach in dispersive channels was investigated in [6], where Lagrange codes were used to detect the users involved in the collision. For recovery of the user bits, the channel coefficients were assumed constant during the entire collision resolution epoch, which is again an extreme and might not be realistic in high traffic load cases where the collision order is high and thus the collision resolution epoch is long.

In this paper we propose a multichannel extension of [3], [4] that further improves throughput in case of high traffic load, and at the same time can handle frequency selective channels. In particular, we propose a protocol according to which the channel is divided into non-overlapping subchannels. Users transmit their packets over multiple but different subchannels, which are selected in a random fashion. Although multiple packets can be successfully transmitted at the same slot over different subchannels, collisions can still occur. Following a collision slot, and taking into account availability of subchannels and also other issues (fairness, delays), the BS will decide to either deal with the collisions on several subchannels in parallel, or on one subchannel at a time. To resolve a collision over a certain subchannel, the BS will assign a set of relay nodes to forward the signal they heard during the collision slot. To minimize packet processing time, the BS will first consider the subchannel with the highest collision order, and assign as many relays as the maximum number of available subchannels. If the number of available subchannels is higher than the collision order, then the collision can be resolved in one additional slot. Otherwise, more slots will be required. Depending on the availability of subchannels, collision resolution on several subchannels can be carried out in parallel. The contributions of the relays will be independent linear equations involving packets that collided over a particular subchannel. When enough equations have been collected, the packets can be recovered. When all collisions have been resolved, transmission of new packets will resume.

At the physical layer, we propose transmission along the lines of an Orthogonal Frequency Division Multiple Access (OFDMA) system [1]. The subchannels mentioned above are groups of carriers. OFDM systems handle frequency selective fading very effectively, since depending on the bandwidth of each carrier, the channel experienced on each carrier is flat fading.

We show that the proposed approach is particularly effective in the cases where the network load is high. It results in short packet processing time, while at the same time combats a frequency selective channel.

II. THE PROPOSED MEDIA ACCESS PROTOCOL

We consider a slotted small-scale multi-access wireless system which includes a set of J user nodes, denoted by $\mathcal{J} = \{1, 2, \ldots, J\}$ and one BS, denoted by $d \notin \mathcal{J}$. All user nodes

This work has been supported by NSF under grant CNS-0435052.

are designated to solely communicate with the BS. All packets are equal-sized and the transmission time of one packet is within one slot. Each node is equipped with a buffer to save all incoming packets until they are transmitted successfully (then they will be discarded). Upon a successful reception, the BS will send the source node an acknowledgment. In the absence of an acknowledgment, the source will resend its packet in the next available time slot.

The channel is assumed to be in general frequency selective fading, and slot-invariant, i.e., the channel taps are constant over one slot but may vary between different slots. It consists of M separable subchannels, denoted by C_m ($m = 1, \dots, M$), and the total channel capacity is M (packets/slot). Each node is allowed to send out up to M packets in one slot but using different subchannels. We assume that a user cannot hear and transmit on the same subchannel.

In [4] Once the BS detects collision(s), i.e. more than one packet transmitted in the channel, the system enters a *cooperative transmission epoch* (CTE) to resolve the collision. In the CTE, the collided packets are not discarded but are rather saved in the buffer. In the slots following the collision, a set of nodes designated as non-regenerative relays retransmit the signal that they received during the collision slot. Based on the initially collided packets and the signals forwarded by the relays, the BS formulates a multiple-input multiple-output (MIMO) problem, the solution of which yields the original packets.

II-A. Resolving collisions on subchannels

For resolving collisions on suchannels, a straightforward extension of [3] would be to group all collided packets that were transmitted via the various subchannels together and resolve the collision along the lines of [3]. However, by taking advantage of the available channels, we here propose a different approach that achieves faster collision resolution and thus higher throughput.

Let us term the process of resolving packets that collided over C_m as *cooperative transmission process* (CTP_m). During CTP_m, the designated relays will use a set of subchannels indicated to them by the BS. If the relay node is a source node that transmitted over C_m it will retransmit its original packet, otherwise it will retransmit the mixture of signals it received during the collision slot corresponding to C_m.

For the case that collisions occurred on multiple subchannels during the same slot, the BS will decide how to resolve collisions over all subchannels according to some predefined strategy, e.g. a strategy that will achieve the least average processing time for every transmitted packet, fairness, etc.. Here we propose the following strategy that achieves the *least processing time*. Suppose that $K(n)$ nodes intend to send packets in the n-th slot. Since each node does not have information on who is active, it will randomly pick subchannels for transmission. Let $K_m(n)$ denote the number of nodes who select subchannel C_m. It holds that $K(n) = \sum_{m=1}^{M} K_m(n)$.

The average processing time, for all transmitted packets is

$$\bar{\tau}_n = \frac{1}{K(n)} \sum_{i=1}^{M} K_m(n) \tau_m(n) \quad (1)$$

where $\tau_m(n)$ denotes the processing time (in slots) for each packet that collided on C_m; $\tau_m(n)$ is equal to the duration of CTP_m.

Collisions with higher order carry more weight in the calculation of the average processing time. To achieve the least processing time for every collided packet, we propose to allocate all available and necessary subchannels to resolve collisions over one subchannel at a time, starting from the highest order collision and moving towards the lowest order collision.

If the number of available subchannels is higher than the collision order, the collision can be resolved in only one additional slot. Otherwise, more slots will be required. Depending on the availability of subchannels, collision resolution on several subchannels can be carried out in parallel (i.e., in the same slot).

Control overhead- To indicate the state of each subchannel, in the beginning of every slot, the BS will broadcast an n-bit control message via every C to all nodes. The n-bit message ($n = \lceil log_2(M+1) \rceil$) is designed to convey to the nodes one of the following $(M+1)$ possible states of that subchannel: **State** 1: subchannel reserved for CTP_1;... **State** M: subchannel reserved for CTP_M ; **State** (M+1): C open to new packets.

Example: Consider a system with only two subchannels. Three users collide over C_1 and two users collide over C_2 at the slot n. At the $(n+1)$-th slot, the BS allocates both subchannels for CTP_1 to resolve the collisions that occurred over C_1, and one subchannel at the $(n+2)$-th slot for CTP_2 to resolve collisions that occurred over C_2. With a straightforward extension of [3], the BS would resolve all five collisions together. The waiting times for all 5 packets would be 3 slots (the collision slot + two for collision resolution). Thus, the average waiting time would be 3 slots. On the other hand, according to the proposed method, at the end of the $(n+1)$-th slot, the collisions that occurred over C_1 have been resolved. The collisions that occurred over C_2 are resolved at the end of $(n+2)$-th slot. So the processing time for those three packet over C_1 is 2 slots while the processing time for two packets over C_2 is 3 slots. The average waiting time is $(3 \times 2 + 2 \times 3)/5 = 2.4$ slots. The reduction of waiting time as compared to the straightforward extension of [3] can be more obvious if the number of collided users and number of subchannels are larger.

II-B. Relay selection

For relay node selection we here introduce the following simple scheme that establishes a predetermined order. A counter, denoted by w is maintained by each user and it increases by one for each relaying request. It may increase more than one time in one slot if multiple relays are required in that slot. The node indexed by $r = \text{mod}(w, J) + 1$ is chosen as relay node. For example, three packets collide over one C in slot n, and at that time the value of the counter is w. Two relays are required to resolve the collision and two Cs are reserved in the next time slot for that purpose. Therefore, the nodes $r_1 = \text{mod}(w+1, J) + 1$ and $r_2 = \text{mod}(w+2, J) + 1$ are selected as relays. More complex cases (i.e. more collisions occurring on more than one Cs) can be handled in an analogous manner. According to this approach a relay will not be reused until all relays have been used in CTP_m.

III. THE PHYSICAL LAYER

In a typical OFDM system with F carriers, blocks of F symbols (BPSK, QAM, etc.), i.e., $\mathbf{x}$ $(1 \times F)$, are subjected to an F-point inverse discrete Fourier Transform (IDFT). A cyclic prefix is pre-appended to the block and the augmented block symbol (OFDM block) is sent through the channel, $h(n)$, in a serial fashion. At the receiver, the cyclic prefix portion is discarded and a DFT is performed on the remainder of the block. Assuming that the channel does not change over the duration of a block, the received block after the DFT equals [1]: $\mathbf{y} = \mathbf{x}\mathbf{H}$ where $\mathbf{H}$ is an $F \times F$ diagonal matrix with its (k,k)-th element equal to the k-th carrier gain, $H(k)$, or equivalently to the k-th sample of the F-point DFT of $h(n)$. Thus, the effect of the channel over the k-th carrier is just a multiplication by the carrier gain, $H(k)$. However, the gains of different carriers are different and packet recovery requires estimation of all channel gains.

The channel can be estimated via a block that contains pilots, and assuming that it will not change over the next few blocks, the obtained channel estimate can be used to recover the subsequent information bearing blocks.

Let us now consider an F-carrier OFDM system, where the carriers are divided into groups of N carriers each, i.e., $C_1, \ldots, C_M$ where $M = F/N$. Also, let us focus on CTP$_m$, in which a collision of order K_m that occurred at slot n over subchannel C_m is to be resolved. Let $\mathcal{S}_m(n)$ denote the set of collided nodes.

A packet consists of B of OFDM symbols.

At slot n, the i-th source node ($\in \mathcal{S}_m(n)$) sends a packet over C_m. Let us denote the j-th symbol in that packet (before IDFT and addition of cyclic prefix) by $\mathbf{x}_i^m(n,j)$. We can think of it as a row vector with length F, where the information bearing symbols occupy positions $(m-1)N, (m-1)N+1, ..., (m-1)N+(N-1)$.

The demodulated received signal over C_m at node r ($\in \bar{\mathcal{S}}_m(n)$) and also at the BS (denoted as d) can be expressed by:

$$\mathbf{y}_r^m(n) = \sum_{i \in \mathcal{S}_m(n)} \mathbf{x}_i^m(n)[\mathbf{I}_B \otimes \mathbf{H}_{ir}(n)] + \mathbf{w}_r^m(n) \quad (2)$$

where $\mathbf{x}_i^m(n) = [\mathbf{x}_i^m(n;0), \ldots, \mathbf{x}_i^m(n;B-1)]$ is a row vector containing the packet that was sent by user i over C_m; $\mathbf{H}_{ir}(n)$ is a diagonal matrix with elements $H_{ir}(p;n),\ p = 0, \ldots, F-1$. $H_{ir}(p;n)$ represents the gain of the p-th carrier between nodes i and r, during slot n; $\mathbf{w}_r^m(n)$ is the noise received at the node r ($r \in \{d\} \bigcup \bar{\mathcal{S}}_m(n)$); $\otimes$ denotes Kronecker product and $\mathbf{I}_B$ is an identity matrix of size $B \times B$.

The selected relay node will either retransmit its original packet or will retransmit the mixture that it received signal during the collision slot. Consider the relay r_i that during the collision slot received a signal over subchannel m and forwarded it during slot $n+k$ to the BS via subchannel l. The received signal at the BS is:

$$\mathbf{z}_{r_i,d}^{m,l}(n+k) = \begin{cases} \mathbf{x}_{r_i}^m(n)[\mathbf{I}_B \otimes \mathbf{H}_{r_i d}(n+k)] \\ +\mathbf{w}_d^l(n+k) \\ \qquad r_i \in \mathcal{S}_m(n) \\ c_{r_i}^m(n)\mathbf{y}_{r_i}^m(n)[\mathbf{I}_B \otimes \mathbf{H}_{r_i d}(n+k)] \\ +\mathbf{w}_d^l(n+k) \\ \qquad r_i \in \bar{\mathcal{S}}_m(n) \end{cases} \quad (3)$$

where $c_{r_i}^m(n)$ is used to scale the relay signal to maintain a constant transmission power.

For mathematic simplicity, we assume that among the total $K_m - 1$ relaying nodes, the first t_m nodes are not source nodes (i.e., $r_j^m \notin \mathcal{S}_m(n)$ for $1 \le j \le t_m$) and the remaining $\gamma_m \triangleq (K_m - t_m - 1)$ nodes are source nodes (i.e. $r_j^m \in \mathcal{S}_m(n)$ for $t_m + 1 \le j \le K_m - 1$).

Let us define: 1) the row vector $\mathbf{X}^m$ as the concatenation of row vectors $\mathbf{x}_{i_1}^m(n), \mathbf{x}_{i_2}^m(n), \ldots, \mathbf{x}_{i_{K_m}}^m(n)$, i.e., the packets of the users involved in the collision on C_m. The noise vector $\mathbf{W}^m$ is defined in a similar fashion.

2) the row vector $\mathbf{Z}^m$ as the concatenation of the packets received by the BS in the CTP$_m$, i.e., $[\mathbf{y}_d^m(n), \mathbf{z}_{r_1 d}^{m,l_1}(n+1), \mathbf{z}_{r_2 d}^{m,l_2}(n+2), \ldots]$.

3) the matrix $\mathbf{H}$, whose first block of F columns is a stack of the matrices $\mathbf{H}_{i_1 d}(n), \mathbf{H}_{i_2 d}(n), \ldots, \mathbf{H}_{i_{K_m} d}(n)$ in that order; its j-th block of F columns (for $(j = 2, \ldots, t_m)$) is a stack of the matrices $[c_{r_j}^m(n)\mathbf{H}_{i_1 r_j}(n)\mathbf{H}_{r_j d}(n+j)], \ldots, [c_{r_j}^m(n)\mathbf{H}_{i_{K_m} r_j}(n)\mathbf{H}_{r_j d}(n+j)]$; and its j-th block of F columns for $(j = t_m + 1, \ldots, K_m)$ is a stack of the matrices $\mathbf{0}, \ldots, \mathbf{0}, \mathbf{H}_{i_j d}(n+j), \mathbf{0}, \ldots, \mathbf{0}$, where i_j is one of the collided sources that was picked to serve as a relay at slot $k+j$.

Then it holds:

$$\mathbf{Z}^m = \mathbf{X}^m[\mathbf{I}_B \otimes \mathbf{H}] + \mathbf{W}^m \quad (4)$$

Alternatively, based on the row vector $\mathbf{Z}^m$ we can form a matrix $\tilde{\mathbf{Z}}^m$, whose j-th row contains the j-th OFDM symbols of the received packets, i.e., $[\mathbf{y}_d^m(n,j), \mathbf{z}_{r_1 d}^{m,l_1}(n,j), \ldots]$. Then it holds:

$$\tilde{\mathbf{Z}}^m = \tilde{\mathbf{X}}^m\mathbf{H} + \tilde{\mathbf{W}}^m \quad (5)$$

where the j-th row of matrix $\tilde{\mathbf{X}}$ equals $[\mathbf{x}_{i_1}(n,j), \ldots, \mathbf{x}_{i_{K_m}}(n,j)]$ and the $\tilde{\mathbf{W}}^m$ is similarly defined. Now $\tilde{\mathbf{Z}}^m$, $\tilde{\mathbf{X}}^m$ are of size $B \times FK_m$.

For collision detection we need to include a user ID in the packet of each user with ID's being orthogonal between different users. To maintain orthogonality of user IDs despite the channel, we propose to distribute the symbols of each user's ID as follows. All will be on the same carrier, and will be distributed one in each OFDM block. For example, for some j, the columns $j, j+F, \ldots, j+K_m F$ of matrix $\tilde{\mathbf{X}}^m$ will contain the orthogonal IDs of users $i_1, i_2, \ldots, i_{K_m}$, respectively. After extracting the j-th column of $\tilde{\mathbf{Z}}^m$ and performing cross-correlation with the known user IDs we can determine whether a user is present in the collision by comparing the cross-correlation result to a threshold [4]. For such an approach we need $B > J$, and we also need the channel to stay fairly constant over B OFDM blocks.

For *channel estimation* we can include in each OFDM symbol of each user a number of pilot symbols. If the channel is of length L then it suffices to include L pilots in each OFDM symbol. These pilots will help recover L samples of the channel frequency response, and due to the redundancy in the frequency domain the remaining samples can be subsequently recovered.

Suppose that we place the pilots at positions $k_1, k_2, \ldots, k_L$ in each OFDM block, and let $\mathbf{S}$ be a column selection matrix of size $B \times K_m L$, that selects the columns $k_1, \ldots, k_L, k_1+F, \ldots, k_L+F, k_1+2F, \ldots, k_L+2F, \ldots$. Then,

$$\tilde{\mathbf{Z}}^m\mathbf{S} = (\tilde{\mathbf{X}}^m\mathbf{S})(\mathbf{H}\mathbf{S}) + \tilde{\mathbf{W}}^m\mathbf{S} \quad (6)$$

where now $\tilde{\mathbf{X}}^m\mathbf{S}$ contains pilots only and $\mathbf{HS}$ contains the channel response samples at frequencies $k_1, k_2, \ldots, k_L$. We can obtain a least squares solution of $\mathbf{HS}$ as $[(\tilde{\mathbf{X}}^m\mathbf{S})^H\tilde{\mathbf{X}}^m\mathbf{S}]^{-1}(\tilde{\mathbf{X}}^m\mathbf{S})^H\tilde{\mathbf{Z}}^m\mathbf{S}$. Once the channel matrix $\mathbf{H}$ is estimated, the transmitted bits over C_m can be obtained via a maximum likelihood (ML) or zero forcing (ZF) decoder as in [4].

IV. SIMULATIONS AND DISCUSSIONS

The total number of users is $J = 32$. The users' ID sequences are selected based on a J-th order Hadamard matrix and the IDs are used to estimate the active users involved in collisions. The length of the packets is 4,800 bits and thus each packet contains 2,400 4-QAM symbols. The frequency selective channel has $L = 3$ taps. The channel taps are chosen independently from zero-mean Gaussian processes, which are constant within one slot, but variant between slots. The number of OFDM carriers is 64.

To investigate the network performance under certain traffic load λ, a Bernoulli model is used. The number of trials is 2,000. In each trial, all users are statistically the same, and each user sends out their packets with probability λ/J. Packets received at the BS with bit error rate higher than $P_{\rm e} = 0.02$ are considered lost or corrupted. In our simulation results we used a zero-forcing decoder for signal recovery. The channel matrix is estimated using the users' orthogonal ID or pilots as described in the previous section.

Fig. 1 shows the throughput versus traffic load. Here the throughput is defined as the average number of packets that are successfully transmitted in a time slot under traffic load λ. For comparison, we also presented the throughput for the single-subchannel cooperative protocol based on OFDM: if one node has a packet to send, it will use all OFDM subcarriers, i.e., there is only one group of subchannels. A collision occurs if more than one node sends packets, in which case cooperative retransmissions will follow. We should note that in the single-subchannel case the packet is M times larger (M: number of subchannels in the multi-subchannel scheme). For the purpose of comparison, we consider same packet length and same transmitted information bits in both schemes. The average processing time of both schemes is shown in Fig. 2.

Under low traffic load, since in the proposed multi-channel scheme the subchannels are picked randomly, it is possible that in a certain slot collisions occur in some subchannels, while other subchannels remain idle. This represents bandwidth waste and lowers the throughput. On the other hand, the single-channel scheme utilizes all available bandwidth. This is evident in Fig. 1 in the low traffic load regime. As the traffic load increases, the single-subchannel scheme suffers from high order collisions, which makes packet recovery sensitive and results in high bit error rate. The multi-channel scheme on the other hand splits packets and sends them through multiple subchannels, thus resulting in smaller order collisions. The probability that the available subchannels are not picked is lower under high traffic load. Therefore, under high traffic load, the throughput for the proposed multichannel scheme is higher, as is also evident in Fig. 1. For the same traffic load regime the corresponding processing time is shorter too (see Fig. 2).

V. CONCLUSION

We proposed a new cooperative random access scheme that allows users to transmit and collide over multiple subchannels. We proposed a novel scheme for resolving collisions that minimize packet processing time. The physical layer is along the lines of an OFDM system, which enables our approach to handle frequency selective channels. The proposed scheme was shown to be effective in high traffic load cases.

VI. REFERENCES

[1] J. A. C. Bingham, Multicarrier modulation for data transmission: An idea whose time has come, IEEE Commun. Mag., May 1990.

[2] G. Dimic, N. D. Sidiropoulos, R. Zhang, "Medium access control - physical cross-layer design", *Signal Processing Magazine, IEEE*, Vol. 21, Issue: 5, Pages: 40-50, Sep. 2004.

[3] R. Lin, A. P. Petropulu, "Cooperative Transmission for Random Access Wireless Networks," *IEEE ICASSP-05*, Philadeplhia PA, March 2005.

[4] R. Lin, A. P. Petropulu, "New Wireless Medium Access Protocol Based On Cooperation," *IEEE Trans. Signal Processing*, vol. 53, no 12, pp. 4675-4684, Dec. 2005.

[5] M. K. Tsatsanis, R. Zhang, and S. Banerjee, "Network-assisted diversity for random access wireless networks," *IEEE Trans. Signal Processing,* vol. 48, no. 3, pp. 702-711, Mar. 2000.

[6] R. Zhang and M. K. Tsatsanis, "Network-assisted diversity multiple access in dispersive channels," *IEEE Trans. Commun.,* vol. 50, no. 4, pp. 623-632, Apr. 2002.

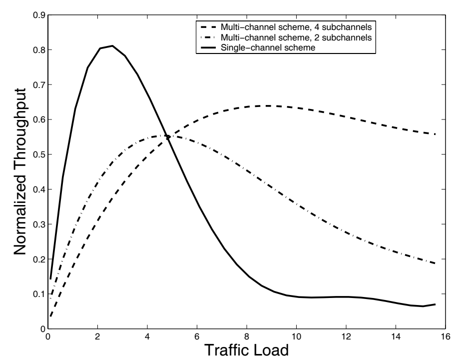

Fig. 1. Normalized throughput v.s. Traffic load for the proposed multi-channel scheme and the single-channel scheme.

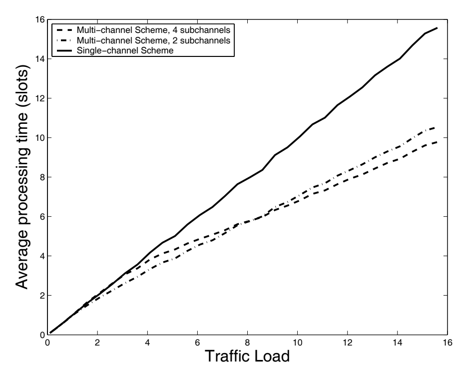

Fig. 2. Normalized average processing time v.s. Traffic load for the proposed multi-channel scheme and the single-channel scheme.