

- [7] X. Li, "Space-time coded multi-transmission among distributed transmitters without perfect synchronization," *IEEE Signal Process. Lett.*, vol. 11, no. 12, pp. 948–951, Dec. 2004.
- [8] Y. Li and X.-G. Xia, "Full diversity distributed space-time codes for asynchronous cooperative communications," in *Proc. IEEE ISIT*, Adelaide, Australia, Sep. 2005, pp. 911–915.
- [9] Y. Li and X.-G. Xia, "A family of distributed space-time trellis codes with asynchronous cooperative diversity," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 790–800, Apr. 2007.
- [10] Y. Shang and X.-G. Xia, "Shift-full-rank matrices and applications in space-time trellis codes for relay networks with asynchronous cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3153–3167, Jul. 2006.
- [11] Y. Mei, Y. Hua, A. Swami, and B. Daneshrad, "Combating synchronization errors in cooperative relays," in *Proc. IEEE ICASSP*, Mar. 2005, vol. 3, pp. 369–372.
- [12] X. Li, F. Ng, and T. Han, "Carrier frequency offset mitigation in asynchronous cooperative OFDM transmissions," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 675–685, Feb. 2008.
- [13] Y. Li, W. Zhang, and X.-G. Xia, "Distributive high-rate full-diversity space-frequency codes for asynchronous cooperative communications," in *Proc. IEEE ISIT*, Seattle, WA, Jul. 2006, pp. 2612–2616.
- [14] Z. Li and X.-G. Xia, "A simple Alamouti space-time transmission scheme for asynchronous cooperative systems," *IEEE Signal Process. Lett.*, vol. 14, no. 11, pp. 804–807, Nov. 2007.
- [15] Z. Li and X.-G. Xia, "An Alamouti coded OFDM transmission for cooperative systems robust to both timing errors and frequency offsets," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1839–1844, May 2008.
- [16] G. S. Rajan and B. S. Rajan, "OFDM based distributed space time coding for asynchronous relay networks," in *Proc. IEEE ICC*, Beijing, China, May 2008, pp. 1118–1122.

A Cooperative Multicast Strategy in Wireless Networks

Binglai Niu, Hai Jiang, *Member, IEEE*, and
H. Vicky Zhao, *Member, IEEE*

Abstract—Multicast is a bandwidth-efficient mechanism to provide wireless services for a group of terminals. Providing reliable wireless multicast is challenging due to packet loss and channel fading. In this research, we investigate the medium-access-control (MAC) layer cooperation to enhance the reliability of wireless multicast. We show that the optimal relaying strategy is a pure threshold policy, i.e., the optimal number of relays is bounded by a threshold. We also propose a simple and efficient algorithm to find the optimal time allocation for the direct and relay transmissions.

Index Terms—Cooperation, multicast, relay, throughput.

I. INTRODUCTION

The past two decades have witnessed a variety of wireless networks [1]–[7]. Recently, multicast applications over various wireless networks have become popular, where information data are

Manuscript received October 16, 2009; revised February 1, 2010; accepted March 15, 2010. Date of publication March 22, 2010; date of current version July 16, 2010. This work was supported in part by the Natural Science and Engineering Research Council of Canada under a Strategic Project Grant and in part by the Alberta Innovates—Technology Futures, Alberta, Canada, under an Alberta Ingenuity New Faculty Award. The review of this paper was coordinated by Prof. O. B. Akan.

The authors are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: binglai@ualberta.ca; hai.jiang@ece.ualberta.ca; vzhao@ece.ualberta.ca).

Digital Object Identifier 10.1109/TVT.2010.2046431

simultaneously delivered to a group of terminals. Examples include Internet Protocol Television over WiMax [8] and multimedia broadcast/multicast service in third-generation networks [9], [10]. However, the erroneous nature and the high packet loss rate of wireless channels make wireless multicast very challenging, and it is of critical importance to design reliable wireless multicast techniques.

Many techniques have been proposed in the literature to combat channel fading in wireless multicast networks, e.g., retransmissions [11], [12], error-correcting codes [13]–[16], and cooperation among terminals [17]–[20]. Among them, cooperative multicast has recently attracted much attention, which is motivated by the following fact. In an ad hoc network or a vehicular ad hoc network [e.g., a group of vehicles is listening to the information from serving base stations (BSs)], when a source tries to multicast to multiple receivers, generally, the receivers undergo different channel fading, i.e., some receivers may have good channel conditions, whereas others do not. To have as many successful receivers as possible, the source may have to use a low transmission rate to make receivers with bad channels correctly receive. On the other hand, cooperative multicast can allow the source to transmit at a higher rate by letting receivers help each other forward packets to exploit the spatial diversity and enhance the reliability of wireless multicast [21]–[32]. In the literature, research on cooperative multicast can roughly be categorized into three directions.

- 1) Network-coding-based cooperation: Network coding has the capability of increasing network throughput by exploiting the shared nature of the wireless medium. Cooperation among terminals is triggered after a batch of broadcast packets has been sent from the source. In [23], it is shown that the optimal transmission scheduling problem in network-coding-based cooperative peer-to-peer repair is NP-hard, and heuristic protocols are given with knowledge of network topology and packet loss information. In [24], random network coding is adopted in WiMAX, and the channel and power allocation in relaying terminals is investigated to optimize multicast rates. It is assumed that the BS has the knowledge of channel state information on each link. In [25], a practical network-coding-based cooperative multicast protocol is investigated, with focus on the one-relay case.
- 2) Space-time coded cooperation: When there are multiple relays, space-time coding can be utilized by the relays to forward signal to other terminals. System performance can be improved, with the cost of computational complexity. In [26], outage capacity is analyzed for two-phase space-time-coded cooperative multicast. It is shown that cooperation can improve outage capacity in a large network with low signal-to-noise ratio (SNR). In [27], the improvement of coverage and latency in space-time-coded cooperative broadcast is studied, and centralized approximation for the optimal cooperative broadcasting problem is also given.
- 3) Power-efficient cooperation: When the size of a network increases, the power consumption in wireless multicast/broadcast will also increase. Therefore, it is important to achieve power efficiency in wireless multicast/broadcast. In [28], the optimal power-allocation problem in a dense network is studied, in which the optimal transmission order and power of the source and relays are found. Low-complexity distributed power-efficient protocols are also proposed. In [31], a maximal lifetime accumulative broadcast algorithm is given, which determines the transmission order and power of the terminals such that the network lifetime is maximized.

Furthermore, the work in [29] proposes a loss-compensation protocol using route diversity, where lost packets are retransmitted to

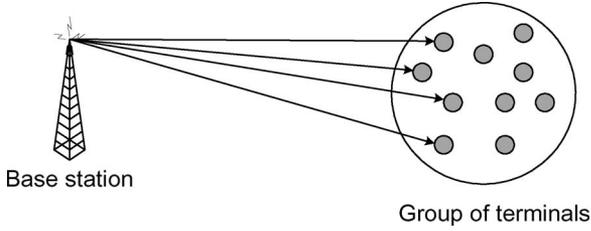


Fig. 1. System model.

terminals near the edges of adjacent cells. In addition, in [32], a cooperative strategy is proposed for video multicast, whereas the work does not consider multipath fading in wireless channels.

In this paper, we study medium-access-control (MAC) layer cooperation in wireless multicast, where the BS sends packets to a group of terminals. The packets to be delivered are delay sensitive (e.g., real-time traffic), such that they are considered useless if not delivered within a period (e.g., a time slot). The motivation of our work is to determine the optimal relaying strategy when there are a number of relays available. The major contribution of our work is twofold. First, we show that, interestingly, the optimal relaying strategy (to maximize the total throughput) is a pure threshold policy, i.e., the optimal number of relays is bounded by a threshold, which can be obtained via numerical methods. Second, to avoid the complexity in exhaustive search for the optimal time allocation between direct and relay transmissions, we propose a simple algorithm that can quickly find the optimal time allocation.

Note that our research focuses on the multicast of one packet, which is delay sensitive, whereas the network-coding-based cooperative multicast usually deals with transmissions of a batch of packets. The delay-sensitive nature of the traffic determines that network coding may not be appropriate since the waiting time for a batch of packets may not be tolerable. On the other hand, our research and the space-time-coded cooperative multicast are complementary to each other. In particular, space-time coding can also be added into our system model for performance improvement, which deserves further investigation.

The rest of this paper is organized as follows. Section II describes the system model and formulates the problem. Section III gives a detailed discussion of the optimum relay selection scheme. Section IV presents a simplified numerical method for optimal time allocation. Section V provides simulation results and performance analysis of the proposed strategy. Section VI gives some further discussion, followed by conclusions in Section VII.

II. SYSTEM MODEL AND PROBLEM FORMULATION

The considered model is a single-hop wireless multicast network shown in Fig. 1, where the BS multicasts to a group of N terminals that are close to each other. The distance between the BS and terminals is much larger than the distance between any two terminals.¹ We consider live multicast applications, which have stringent delay constraint and require that a packet of length L must be delivered to all terminals within time duration T .

Similar to [32], we employ a MAC-layer cooperative multicast strategy, as shown in Fig. 2. It contains two stages that have time duration $T_1 = (1 - \alpha)T$ and $T_2 = \alpha T$, respectively, where $0 \leq \alpha < 1$. In the

¹When the terminals are not close to each other (e.g., when a BS tries to broadcast to all terminals in its coverage), the total coverage can be partitioned into a number of small regions, or the terminals close to each other can self-organize to form a small group. Then, our target system is the particular small groups.

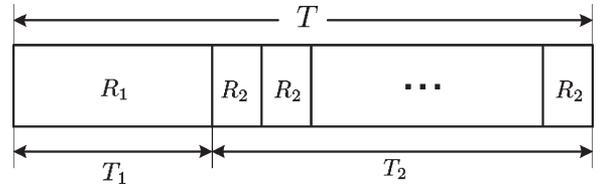


Fig. 2. Cooperative strategy.

first stage, the BS transmits the packet at rate $R_1 = L/T_1$. We assume that the wireless link between the BS and a terminal experiences Rayleigh fading, and the channel gain can be modeled as complex Gaussian with $\mathcal{CN}(0, \sigma_B^2)$.² We further assume that all channel gains from the BS to the terminals are independent and identically distributed (i.i.d.). Similar to [33], we claim that a terminal correctly receives the packet if its instantaneous channel capacity (which is the Shannon's capacity) is larger than or equal to the BS's transmission rate R_1 .³ Thus, given our model, the probability that a terminal successfully receives the packet in stage 1 is $p_1 = \exp((1 - 2^{R_1})N_0/P_B\sigma_B^2) = \exp((1 - 2^{R_1})/\Gamma_1)$, where P_B is the BS's transmission power, N_0 is the variance of the zero-mean additive white Gaussian noise, and $\Gamma_1 = P_B\sigma_B^2/N_0$. Note that p_1 is a function of the time-allocation parameter α and is the same for all terminals. Let s_1 be a random variable denoting the number of terminals that successfully receive the packet in stage 1. We have $\Pr[s_1 = i] = \binom{N}{i} p_1^i (1 - p_1)^{N-i}$ for $0 \leq i \leq N$. These terminals send feedbacks to the BS, indicating their success in receiving the packet. In addition, we assume that the feedback time is negligible.

In stage 2, terminals who successfully receive the packet in stage 1 help forward the packet to others. The channel gain between any two terminals is assumed to be i.i.d. and modeled as complex Gaussian with $\mathcal{CN}(0, \sigma_R^2)$. Note that the non-i.i.d. case is to be discussed in Section VI. Given that i terminals correctly receive the packet in the first stage, since all channel gains between terminals are i.i.d., the BS randomly selects $m_i \leq i$ of them to serve as relays, and the selected relays take turns to broadcast the packet in stage 2. The time duration T_2 in stage 2 is divided into m_i intervals of equal length, each for a relay who transmits the packet at the same rate $R_2 = m_i L/T_2$ using the same power P_R . Note that, when $s_1 = N$ (all terminals correctly receive the packet in stage 1), $m_N = 0$, and there is no need for terminals to forward the packet in stage 2. For a terminal who fails in stage 1, it can correctly receive the packet from a relay if the instantaneous channel capacity (between the terminal and the relay) exceeds the transmission rate R_2 , which happens with probability $p_2 = \exp((1 - 2^{R_2})N_0/P_R\sigma_R^2) = \exp((1 - 2^{R_2})/\Gamma_2)$, where $\Gamma_2 = P_R\sigma_R^2/N_0$. A terminal succeeds in stage 2 if it correctly receives the packet from at least one of the m_i relays, which happens with probability $(1 - (1 - p_2)^{m_i})$. Let s_2 be the number of terminals who fail in stage 1 but successfully receive the packet in stage 2. Given that $s_1 = i$ and m_i relays help forward the packet, s_2 's conditional mean is $\mathbb{E}[s_2|i, m_i] = (N - i)(1 - (1 - p_2)^{m_i})$.

Define $\mathbf{m} \triangleq (m_1, m_2, \dots, m_N)$, and let $s(\alpha, \mathbf{m})$ be the total number of terminals who successfully receive the packet within T (either from the BS or from one of the relays). Following the preceding

²In this research, it is also assumed that the channel gain in a link (from the BS to a terminal or from one terminal to another) does not change within duration T .

³The Shannon's capacity is used, because, in this research, we do not specify any modulation and coding. In a real system, when the modulation and coding are specified, the probability of successful packet reception (p_1 or p_2 in this section) can also be calculated, which depends on the probability density function of the channel fading and the modulation and coding schemes.

analysis, we define the system throughput as the mean of $s(\alpha, \mathbf{m})$, which can be calculated using

$$\begin{aligned} \mathbb{E}[s(\alpha, \mathbf{m})] &= \sum_{i=1}^N [i + (N-i) \cdot (1 - (1-p_2)^{m_i})] \Pr\{s_1 = i\} \\ &= \sum_{i=1}^N \left\{ [i + (N-i) \cdot (1 - (1-p_2)^{m_i})] \right. \\ &\quad \left. \cdot \binom{N}{i} p_1^i (1-p_1)^{N-i} \right\}. \end{aligned} \quad (1)$$

In our work, we select the optimal relay number \mathbf{m} and the optimal time-allocation parameter α to maximize the system's throughput $\mathbb{E}[s(\alpha, \mathbf{m})]$.

III. OPTIMAL RELAY SELECTION

We first fix the time-allocation parameter α and select the optimal relay number \mathbf{m} . From (1), for each $s_1 = i$, the relay-selection parameter m_i only affects the term $(1 - (1-p_2)^{m_i})$. Therefore, for a fixed α , maximization of $\mathbb{E}[s(\alpha, \mathbf{m})]$ is equivalent to maximization of $(1 - (1-p_2)^{m_i})$ or, equivalently, minimization of $(1-p_2)^{m_i}$ for all $1 \leq i \leq N-1$. Define $t \triangleq 1/\Gamma_2$, $q \triangleq R_2/m_i = L/(\alpha T)$, and $p(\alpha, m_i) \triangleq \exp(t(1-2^{q m_i}))$. Thus, $p_2 = \exp((1-2^{R_2})/\Gamma_2) = p(\alpha, m_i)$, and the optimal relay-selection becomes the following optimization problem: For $1 \leq i \leq N-1$

$$\begin{aligned} \text{minimize}_{m_i} \quad & Q(\alpha, m_i) \triangleq (1-p(\alpha, m_i))^{m_i} \\ \text{subject to} \quad & 0 \leq m_i \leq i. \end{aligned} \quad (2)$$

Note that, in (2), m_i is an integer. To gain insights into the optimal relay selection problem, we first consider the minimization of $Q(\alpha, x) \triangleq (1-p(\alpha, x))^x$, where $p(\alpha, x) \triangleq \exp(t(1-2^{q x}))$ and $x \in \mathbb{R}^+$ is a positive real number. The following result can be obtained.

Lemma 1: For a fixed α , $Q(\alpha, x)$ achieves its minimum value at $x = x_0$, where $(\partial Q(\alpha, x)/\partial x)|_{x=x_0} = 0$. In addition, $Q(\alpha, x)$ is a monotonically decreasing function of x when $0 < x \leq x_0$, and it monotonically increases when $x > x_0$.

The proof of Lemma 1 is in Appendix A.

For a given α , let $x_0(\alpha)$ be the root of the equation $(\partial Q(\alpha, x)/\partial x) = 0$, and define

$$m^*(\alpha) = \begin{cases} \lfloor x_0(\alpha) \rfloor, & \text{if } Q(\alpha, \lfloor x_0(\alpha) \rfloor) \leq Q(\alpha, \lceil x_0(\alpha) \rceil) \\ \lceil x_0(\alpha) \rceil, & \text{otherwise.} \end{cases} \quad (3)$$

Here, $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor and ceiling functions, respectively. Based on Lemma 1, for integer m_i , $Q(\alpha, m_i)$ is a monotonically decreasing function of m_i when $m_i \leq m^*(\alpha)$, and $Q(\alpha, m_i)$ monotonically increases if $m_i > m^*(\alpha)$. Therefore, for a given $s_1 = i$, to minimize $Q(\alpha, m_i)$ and to maximize the network throughput, if the number of successful terminals in stage 1 is smaller than $m^*(\alpha)$, then all these successful terminals should be selected as relays; otherwise, the BS should randomly select only $m^*(\alpha)$ of the i successful terminals to serve as relays in stage 2.

To summarize, the optimal m_i can be found by

$$m_i^{\text{opt}} = \begin{cases} \min(i, m^*(\alpha)), & \text{if } i < N \\ 0, & \text{if } i = N \end{cases} \quad (4)$$

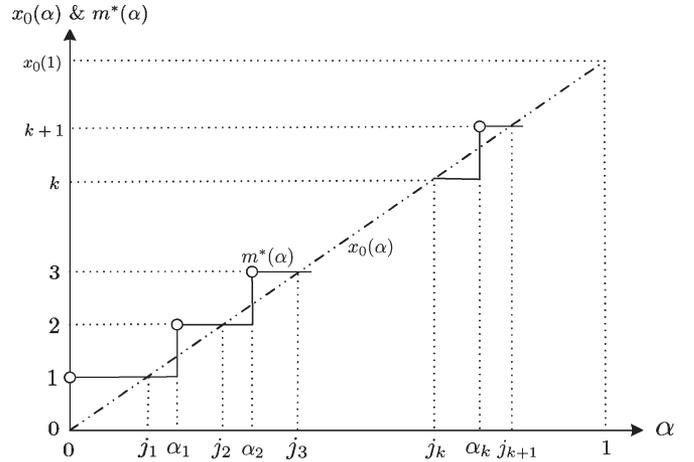


Fig. 3. Illustration of Algorithm 1.

and $m^*(\alpha)$ works as a threshold in determining the optimal number of relays. Interestingly, it can be seen that $m^*(\alpha)$ depends on α but not i and can be numerically found according to (3).

With the optimal relay selection algorithm, the network throughput $\mathbb{E}[s(\alpha, \mathbf{m})]$ in (1) becomes

$$\begin{aligned} \mathbb{E}[s(\alpha)] &= \sum_{i=1}^N \left\{ [i + (N-i) \cdot (1 - (1-p_2(\alpha))^{m_i^{\text{opt}}})] \right. \\ &\quad \left. \cdot \binom{N}{i} p_1(\alpha)^i (1-p_1(\alpha))^{N-i} \right\}. \end{aligned} \quad (5)$$

IV. OPTIMAL TIME ALLOCATION

In this section, based on our optimal relay selection strategy in the previous discussion, we find the optimal time-allocation scheme. Note that $x_0(\alpha)$ and $m^*(\alpha)$ in (3) are functions of α . Therefore, to determine the optimal time allocation, we need to find $m^*(\alpha)$ for all $\alpha \in [0, 1)$, which is computationally expensive.

For any given $\alpha \in [0, 1)$, instead of solving the differential equation $(\partial Q(\alpha, x)/\partial x) = 0$ to obtain $m^*(\alpha)$ based on (3), we propose an efficient algorithm to find $m^*(\alpha)$ with less complexity. The following lemmas are introduced to facilitate the algorithm.

Lemma 2: $x_0(\alpha)$ linearly scales with α . That is, given $\alpha_1 \neq \alpha_2$, let $x_0(\alpha_1)$ and $x_0(\alpha_2)$ be the solutions to $(\partial Q(\alpha_1, x)/\partial x) = 0$ and $(\partial Q(\alpha_2, x)/\partial x) = 0$, respectively. Then, we have $(x_0(\alpha_1)/x_0(\alpha_2)) = \alpha_1/\alpha_2$.

The proof of Lemma 2 is given in Appendix C.

Lemma 2 means that, if we find $x_0(\alpha_0)$ for a specific value α_0 , then for any other $\alpha \in (0, 1)$, we have $x_0(\alpha) = x_0(\alpha_0) \cdot \alpha/\alpha_0$.

Lemma 3: For an integer n , assume that $a, b \in (0, 1)$ satisfy $x_0(a) = n$ and $x_0(b) = n+1$. If $m^*(\alpha_0) = n+1$ for a specific $\alpha_0 \in (a, b)$, then $m^*(\alpha) = n+1$ for any $\alpha \in [\alpha_0, b]$.

The proof of Lemma 3 is given in Appendix D.

For interval $[a, b]$, where $x_0(a) = n$ and $x_0(b) = n+1$ (n is an integer), it is obvious that, when α increases from a to b , there is an α_n that satisfies $Q(\alpha_n, n) = Q(\alpha_n, n+1)$. From Lemma 3, we have $m^*(\alpha) = n$ if $\alpha \in [a, \alpha_n]$ and $m^*(\alpha) = n+1$ for any $\alpha \in (\alpha_n, b]$. Based on this observation, we have the following simplified algorithm to efficiently find the values of $m^*(\alpha)$ for any $\alpha \in [0, 1)$. An illustration is shown in Fig. 3, in which $x_0(\alpha)$ and $m^*(\alpha)$ are represented by the dashed-dotted line and solid line, respectively.

Algorithm 1: Determination of $m^*(\alpha)$ ($0 \leq \alpha < 1$)

- 1: Let $\alpha = 1$, and solve $(\partial Q(\alpha = 1, x)/\partial x) = 0$ to get $x_0(\alpha)|_{\alpha=1}$.
- 2: Draw the figure of $x_0(\alpha) = \alpha \cdot x_0(\alpha)|_{\alpha=1}$.
- 3: Draw lines $x_0(\alpha) = 1, 2, \dots$, which will partition the line of $x_0(\alpha) = \alpha \cdot x_0(\alpha)|_{\alpha=1}$ into K regions with index $0, 1, \dots, K - 1$. Assume that region k is between lines $x_0(\alpha) = k$ and $x_0(\alpha) = k + 1$.
- 4: In region $k = 0$, if $\alpha = 0$, then $m^*(\alpha) = 0$; Otherwise, $m^*(\alpha) = 1$.
- 5: **for** $k = 1 : K - 1$ **do**
- 6: In region k , use numerical methods such as bisection search to find the point of $\alpha = \alpha_k$ such that $Q(\alpha_k, k) = Q(\alpha_k, k + 1)$.
- 7: In region k , $m^*(\alpha) = k$ if $\alpha \in [j_k, \alpha_k]$, and $m^*(\alpha) = k + 1$ if $\alpha \in (\alpha_k, j_{k+1}]$, where $j_k = k/x_0(\alpha)|_{\alpha=1}$.

Step 3 of the algorithm is to partition the line $x_0(\alpha)$ into several regions. In region k , at the left end, we have $x_0(\alpha) = k$, whereas, at the right end, we have $x_0(\alpha) = k + 1$. Therefore, in region k , $m^*(\alpha)$ is either k or $k + 1$. From Lemma 3, when α moves from the left end to the right end in region k , once we have $m^*(\alpha) = k + 1$ at one particular value of α , the value of $m^*(\alpha)$ will keep at value $k + 1$ when α further moves from the particular value to the right end of the region. Therefore, in Step 6, the bisection search method is used to find the particular value of α in region k . In addition, in the region, we have $m^*(\alpha) = k$ and $m^*(\alpha) = k + 1$ at the left and right sides of the particular value, respectively, as shown in Step 7.

The major advantage of the preceding algorithm is that we only need to solve $(\partial Q(\alpha, x)/\partial x) = 0$ once, and then, we can get all values of $m^*(\alpha)$ by fast search. Note that, when the algorithm is not adopted, we have to solve $(\partial Q(\alpha, x)/\partial x) = 0$ for each value of α within $0 \leq \alpha < 1$, which is computationally complex.

Given the preceding efficient method to determine $m^*(\alpha)$ for all $\alpha \in [0, 1)$, the optimal time allocation can be found as follows. First, we uniformly sample α in the range of $[0, 1)$ using a step size $\delta (\ll 1)$. Then, we find $m^*(\alpha)$ for all sampled α values according to Algorithm 1 and calculate the corresponding network throughput $\mathbb{E}[s(\alpha)]$ using (5). Finally, we select the value of α that maximizes $\mathbb{E}[s(\alpha)]$ as the optimal time-allocation parameter. Note that $\alpha = 0$ corresponds to direct transmission without terminal cooperation.

V. PERFORMANCE EVALUATION

In this section, we analyze the performance of the proposed cooperative multicast strategy and compare it with that of the direct multicast scheme without cooperation. We first analyze the performance of our cooperative scheme with different time allocations. Consider a network with $N = 100$ terminals. The packet size L is 8 Kbits, which needs to be received within $T = 10$ ms over a wireless channel with 1-MHz bandwidth. We fix $\Gamma_1 = 0$ dB and $\Gamma_2 = 10$ dB. Fig. 4 shows $\mathbb{E}[s(\alpha)]$, which is the average number of terminals that successfully receive the packet during T with different values of α . It is observed that $\mathbb{E}[s(\alpha)]$ decreases with α at first and then increases to a maximum point and then decreases again. This can be explained as follows. When the time duration for the relay session (i.e., stage 2) is very short (α is close to 0), the success probability p_2 in stage 2 is very small such that almost no terminal can successfully receive from a relay. Thus, a larger α means more waste of time and degrades the performance. As α increases to a certain level, the success probability p_2 becomes larger, and cooperation in stage 2 takes effect. Thus, the overall throughput increases. As α further increases (close to 1), the

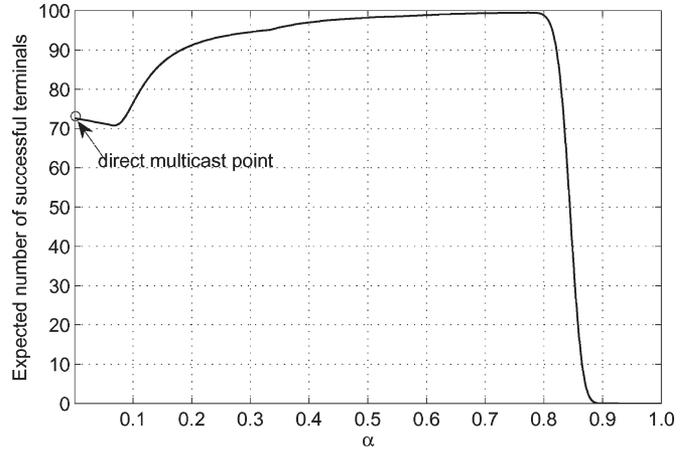


Fig. 4. Expected number of total successful terminals.

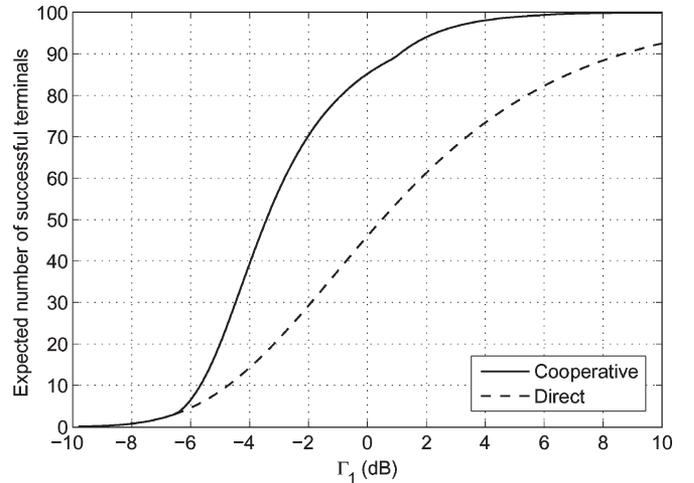


Fig. 5. Comparison of cooperative multicast and direct multicast.

time for BS transmission T_1 is very small (close to 0), which means that the success probability p_1 is very small and that almost no terminal can correctly receive the packet in stage 1. As a result, the throughput dramatically decreases.

Next, we compare our cooperative strategy with direct multicast. The cooperative strategy is designed with the optimal time allocation and optimal relay selection. We vary the values of Γ_1 and Γ_2 while keeping the difference of them at $\Gamma_2 - \Gamma_1 = 10$ dB. Note that Γ_1 and Γ_2 are the average SNR for the channel from the BS to the terminals and between any two terminals, respectively. Fig. 5 shows the average number of successful terminals versus Γ_1 . We observe that, when Γ_1 is smaller than -6.4 dB, the optimal cooperative strategy is just direct multicast. The reason is given as follows. When Γ_1 is very small (which means that the channel from the BS to the terminals is poor), it is not beneficial if we partition the duration T into two stages and let the BS transmit in the first stage, because almost no terminal will successfully receive in the first stage due to poor channel quality. As the value of Γ_1 increases beyond a certain level, cooperation achieves a better performance. Particularly, in Γ_1 's region from -4 to 4 dB, the optimal cooperative strategy achieves significant performance gain, compared with the direct multicast. When Γ_1 further increases (more than 8 dB), both the cooperative strategy and direct multicast can achieve high throughput, and the difference between them becomes insignificant.

TABLE I
OPTIMAL NUMBER OF RELAYS AND MAXIMAL SYSTEM THROUGHPUT VERSUS THE NUMBER k OF SUCCESSFUL TERMINALS IN T_1 IN THE NON-I.I.D. CASE

k		1	2	3	4	5	6	7	8	9
Optimal number of relays	Exhaustive search	1	2	2	2	2	2	2	2	2
	Approximation	1	2	2	2	2	2	2	2	2
Maximal system throughput	Exhaustive search	9.37	9.74	9.78	9.81	9.84	9.87	9.90	9.94	9.97
	Approximation	9.36	9.73	9.76	9.90	9.83	9.86	9.90	9.93	9.97

VI. FURTHER DISCUSSION

Our analysis in the preceding sections is based on an assumption that the channels between terminals are i.i.d., which is referred to as the *i.i.d. case*. It is essential to investigate the impact if the channels between terminals are not identically distributed. Consider the following non-i.i.d. case. The terminals are uniformly distributed within a circle with radius R . The channel between any pair of terminals undergoes path loss with exponent 2 and Rayleigh fading. In the non-i.i.d. case, the probability p_2 in the preceding sections is no longer the same for each pair of terminals and is expressed as $p_2(D, R_2) = \exp(D(1 - 2^{R_2}))$, where $D = N_0/P_R\sigma_R^2$ depends on the channel propagation statistics between any two terminals. To deal with the heterogeneity of p_2 , we use approximation as follows. For the N terminals uniformly distributed within the circle, we can numerically calculate the average value of D , which is denoted as \bar{D} , and further calculate the corresponding value of p_2 , which is denoted as $\bar{p}_2 = p_2(\bar{D}, R_2)$ for a particular value of R_2 . By using \bar{p}_2 to approximate the successful probability between any pair of terminals, we can approximate the system throughput, as shown in (1), by replacing p_2 with \bar{p}_2 . For this approximation, we can further get the optimal relay selection and optimal time allocation.

We use computer simulations to evaluate the accuracy of the approximation. Consider $N = 10$ terminals uniformly distributed within a circle with unit radius. The wireless channel bandwidth is 1 MHz. The BS transmission time is $T_1 = 10$ ms, and the cooperation period is $T_2 = 10$ ms. Packet length $L = 8$ Kbits. The average received SNR for a terminal from BS is $\Gamma_1 = 0$ dB, and the average received SNR between two terminals is 0 dB if the distance between the two terminals is unit. We randomly generate 1000 topologies based on the terminal distribution. When $k \in \{1, 2, \dots, N\}$ terminals successfully receive the packet from BS in T_1 , by exhaustive search (on how many relays are selected and which relays are selected), we can get the optimal number of relays, which achieves the maximal average⁴ system throughput for the 1000 topologies, as shown in Table I. The approximation results are also given in Table I. It can be seen that the approximation matches well with the exhaustively searched optimal results, which demonstrates the accuracy of the approximation for the non-i.i.d. case.

VII. CONCLUSION

In this paper, we have proposed a cooperative multicast strategy for transmissions of delay-sensitive traffic. We have analyzed the optimum relay-selection scheme and shown that the optimal relay number is bounded by a threshold value. We have also proposed a fast algorithm to find the optimal time allocation for BS transmission and terminal cooperation. Numerical results have shown that our proposed strategy can enhance the network performance. This research should provide helpful insights into designs of cooperative multicast strategies over wireless networks.

⁴Note that, here, the average is for 1000 topologies.

APPENDIX A

PROOF OF LEMMA 1

Proof: Taking the first derivative of $\ln Q(\alpha, x) = \ln\{(1 - p(\alpha, x))^{qx}\}$ with respect to x , we have

$$\begin{aligned} \frac{\partial \ln Q(\alpha, x)}{\partial x} &= \frac{1}{Q(\alpha, x)} \frac{\partial Q(\alpha, x)}{\partial x} \\ &= \ln(1 - p(\alpha, x)) + \frac{(-1) \cdot x}{1 - p(\alpha, x)} \frac{\partial p(\alpha, x)}{\partial x} \end{aligned}$$

where

$$\begin{aligned} \frac{\partial p(\alpha, x)}{\partial x} &= \exp(t(1 - 2^{qx})) \cdot t \cdot (-1) \cdot \ln 2 \cdot q \cdot 2^{qx} \\ &= -qt(\ln 2)p(\alpha, x)2^{qx}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \frac{\partial Q(\alpha, x)}{\partial x} &= Q(\alpha, x) \left\{ \ln(1 - p(\alpha, x)) + \frac{(-1) \cdot x}{1 - p(\alpha, x)} \frac{\partial p(\alpha, x)}{\partial x} \right\} \\ &= \frac{Q(\alpha, x)p(\alpha, x)}{1 - p(\alpha, x)} \left\{ \frac{1 - p(\alpha, x)}{p(\alpha, x)} \ln(1 - p(\alpha, x)) \right. \\ &\quad \left. + qt(\ln 2)x2^{qx} \right\} \\ &= \frac{Q(\alpha, x)p(\alpha, x)}{1 - p(\alpha, x)} f(\alpha, x) \end{aligned} \quad (6)$$

where

$$f(\alpha, x) \triangleq \frac{1 - p(\alpha, x)}{p(\alpha, x)} \ln(1 - p(\alpha, x)) + qt(\ln 2)x2^{qx}. \quad (7)$$

Appendix B shows that $f(\alpha, x)$ has two properties.

i)

$$\lim_{x \rightarrow 0} f(\alpha, x) = 0, \quad \lim_{x \rightarrow +\infty} f(\alpha, x) = +\infty. \quad (8)$$

ii) Define x^* as the root of $(\partial f(\alpha, x)/\partial x) = 0$; then, we have

$$\begin{cases} \frac{\partial f(\alpha, x)}{\partial x} < 0, & \text{if } 0 < x < x^* \\ \frac{\partial f(\alpha, x)}{\partial x} = 0, & \text{if } x = x^* \\ \frac{\partial f(\alpha, x)}{\partial x} > 0, & \text{if } x > x^*. \end{cases} \quad (9)$$

From the two properties, it can be seen that $f(\alpha, x) = 0$ has a single root $x_0 \in (0, +\infty)$, and

$$\begin{cases} f(\alpha, x) < 0, & \text{if } 0 < x < x_0 \\ f(\alpha, x) = 0, & \text{if } x = x_0 \\ f(\alpha, x) > 0, & \text{if } x > x_0. \end{cases} \quad (10)$$

For $x \in (0, +\infty)$, we have $p(\alpha, x) \in (0, 1)$ and $Q(\alpha, x) \in (0, 1)$, and thus, $Q(\alpha, x)p(\alpha, x)/(1 - p(\alpha, x)) > 0$. Based on (6), (8), and (10),

we have

$$\begin{cases} \frac{\partial Q(\alpha, x)}{\partial x} < 0, & \text{if } 0 < x < x_0 \\ \frac{\partial Q(\alpha, x)}{\partial x} = 0, & \text{if } x = x_0 \\ \frac{\partial Q(\alpha, x)}{\partial x} > 0, & \text{if } x > x_0. \end{cases} \quad (11)$$

Therefore, $Q(\alpha, x)$ has one global minimum at $x = x_0$, and it is a monotonically decreasing function of x if $x \in (0, x_0)$ and monotonically increases as x increases from x_0 . This completes the proof. ■

APPENDIX B PROOF OF (8) AND (9)

Proof of (8): Note that $p(\alpha, x)$ is monotonically decreasing with x , and

$$\lim_{x \rightarrow 0} p(\alpha, x) = 1, \quad \lim_{x \rightarrow +\infty} p(\alpha, x) = 0.$$

We have

$$\begin{aligned} \lim_{x \rightarrow 0} f(\alpha, x) &= \lim_{p(\alpha, x) \rightarrow 1} \left\{ \frac{1 - p(\alpha, x)}{p(\alpha, x)} \ln(1 - p(\alpha, x)) \right\} \\ &\quad + \lim_{x \rightarrow 0} \{qt(\ln 2)x2^{qx}\} = 0 \\ \lim_{x \rightarrow +\infty} f(\alpha, x) &= \lim_{p(\alpha, x) \rightarrow 0} \left\{ \frac{1 - p(\alpha, x)}{p(\alpha, x)} \ln(1 - p(\alpha, x)) \right\} \\ &\quad + \lim_{x \rightarrow +\infty} \{qt(\ln 2)x2^{qx}\} = +\infty. \end{aligned}$$

Proof of (9): We first show that $(\partial f(\alpha, x)/\partial x)$ is a monotonically increasing function of x for $x \in (0, +\infty)$. We have

$$\begin{aligned} &\frac{\partial f(\alpha, x)}{\partial x} \\ &= \frac{\partial \left\{ \frac{1 - p(\alpha, x)}{p(\alpha, x)} \ln(1 - p(\alpha, x)) \right\}}{\partial x} + \frac{\partial \{qt(\ln 2)x2^{qx}\}}{\partial x} \\ &= qt(\ln 2)2^{qx} \left\{ \frac{1}{p(\alpha, x)} \ln(1 - p(\alpha, x)) + 2 + q \ln 2 \cdot x \right\}. \end{aligned} \quad (12)$$

Let $G(\alpha, x) \triangleq (1/p(\alpha, x)) \ln(1 - p(\alpha, x))$. We have

$$\begin{aligned} &\frac{\partial G(\alpha, x)}{\partial x} \\ &= \left\{ \frac{-1}{p^2(\alpha, x)} \ln(1 - p(\alpha, x)) - \frac{1}{p(\alpha, x)(1 - p(\alpha, x))} \right\} \\ &\quad \cdot \frac{\partial p(\alpha, x)}{\partial x} \\ &= \frac{qt(\ln 2)2^{qx}}{p(\alpha, x)(1 - p(\alpha, x))} \\ &\quad \times \{(1 - p(\alpha, x)) \ln(1 - p(\alpha, x)) + p(\alpha, x)\}. \end{aligned}$$

Let $g(p) \triangleq (1 - p) \ln(1 - p) + p$. We have $g'(p) = -\ln(1 - p) > 0$ for $p \in (0, 1)$, and thus, $g(p)$ is monotonically increasing when $p \in (0, 1)$. Since $\lim_{p \rightarrow 0} g(p) = 0$, we have $g(p) > 0$ for $p \in (0, 1)$. Then, it is obvious that $(\partial G(\alpha, x)/\partial x) > 0$, with $p(\alpha, x) \in (0, 1)$, where $x \in (0, +\infty)$. Thus, $G(\alpha, x)$ is a monotonically in-

creasing function of x for $x \in (0, +\infty)$, and so is the function $T(\alpha, x) \triangleq G(\alpha, x) + 2 + q \ln 2 \cdot x$.

It is easy to prove that

$$\begin{aligned} \lim_{x \rightarrow 0} T(\alpha, x) &= \lim_{p(\alpha, x) \rightarrow 1} \left\{ \frac{1}{p(\alpha, x)} \ln(1 - p(\alpha, x)) \right\} + 2 \\ &= -\infty \\ \lim_{x \rightarrow +\infty} T(\alpha, x) &= \lim_{p(\alpha, x) \rightarrow 0} \left\{ \frac{1}{p(\alpha, x)} \ln(1 - p(\alpha, x)) \right\} \\ &\quad + 2 + \lim_{x \rightarrow +\infty} \{q \ln 2 \cdot x\} \\ &= +\infty. \end{aligned}$$

Based on the monotonically increasing property, $T(\alpha, x) = 0$ has only one positive root denoted as $x = x^*$. Thus, we have

$$\begin{cases} T(\alpha, x) < 0, & \text{if } 0 < x < x^* \\ T(\alpha, x) = 0, & \text{if } x = x^* \\ T(\alpha, x) > 0, & \text{if } x > x^*. \end{cases}$$

Since $qt(\ln 2)2^{qx} > 0$ and $(\partial f(\alpha, x)/\partial x) = qt(\ln 2)2^{qx}T(\alpha, x)$ [from (12)], we have

$$\begin{cases} \frac{\partial f(\alpha, x)}{\partial x} < 0, & \text{if } 0 < x < x^* \\ \frac{\partial f(\alpha, x)}{\partial x} = 0, & \text{if } x = x^* \\ \frac{\partial f(\alpha, x)}{\partial x} > 0, & \text{if } x > x^*. \end{cases}$$

This completes the proof. ■

APPENDIX C PROOF OF LEMMA 2

Proof: From the proof of Lemma 1, $(\partial Q(\alpha, x)/\partial x) = 0$ has a single root $x_0(\alpha)$, which satisfies

$$\begin{aligned} &\frac{1 - p(\alpha, x_0(\alpha))}{p(\alpha, x_0(\alpha))} \ln(1 - p(\alpha, x_0(\alpha))) \\ &\quad + t(\ln 2) \frac{L}{T} \frac{x_0(\alpha)}{\alpha} 2^{\frac{L}{T} \frac{x_0(\alpha)}{\alpha}} = 0 \end{aligned} \quad (13)$$

where $p(\alpha, x_0(\alpha)) = \exp(t(1 - 2^{(L/T)(x_0(\alpha)/\alpha)})$. Thus, finding $x_0(\alpha)$ is equivalent to finding the single root y_0 to the following equation:

$$\begin{aligned} &\frac{1 - \exp(t(1 - 2^{Ly/T}))}{\exp(t(1 - 2^{Ly/T}))} \ln \{1 - \exp(t(1 - 2^{Ly/T}))\} \\ &\quad + t(\ln 2) \frac{Ly}{T} 2^{Ly/T} = 0 \end{aligned} \quad (14)$$

which does not depend on α . It can be seen that $x_0(\alpha) = \alpha y_0$. Therefore, $x_0(\alpha)$ linearly scales with α . ■

APPENDIX D PROOF OF LEMMA 3

Proof: To prove Lemma 3, we first show that, for an infinitely small positive value ϵ , we have $m^*(\alpha_0 + \epsilon) = n + 1$.

First, note that, for $\alpha_0 \in (a, b)$, $m^*(\alpha_0) = n + 1$ means that $x_0(\alpha_0) \in (n, n + 1]$ and $Q(\alpha_0, n) > Q(\alpha_0, n + 1)$.⁵ From Lemma 1,

⁵When $Q(\alpha_0, n) = Q(\alpha_0, n + 1)$, i.e., n and $n + 1$ relays give the same system throughput, we choose $m^*(\alpha) = n$ to reduce the total energy consumed by relays.

$Q(\alpha_0, x)$ decreases when x increases from n to $x_0(\alpha_0)$ and increases when x increases from $x_0(\alpha_0)$ to $n+1$. Since $Q(\alpha_0, n) > Q(\alpha_0, n+1)$, there exists $z \in (n, x_0(\alpha_0))$ such that $Q(\alpha_0, z) = Q(\alpha_0, n+1)$.

Taking the first derivative of $Q(\alpha, x) = (1 - p(\alpha, x))^x$ with respect to α , we get

$$\begin{aligned} & \frac{\partial Q(\alpha, x)}{\partial \alpha} \\ &= x(1 - p(\alpha, x))^{x-1} \cdot (-1) \cdot \frac{\partial p(\alpha, x)}{\partial \alpha} \\ &= \ln 2 \cdot t \cdot 2^{\frac{Lx}{\alpha T}} x^2 (1 - p(\alpha, x))^{x-1} p(\alpha, x) \frac{L}{T} \left(-\frac{1}{\alpha^2} \right) \\ &= \ln 2 \cdot t \cdot 2^{\frac{Lx}{\alpha T}} x^2 \frac{Q(\alpha, x) p(\alpha, x)}{1 - p(\alpha, x)} \frac{L}{T} \left(-\frac{1}{\alpha^2} \right) < 0 \end{aligned} \quad (15)$$

which means that $Q(\alpha, x)$ decreases as α increases. In addition, as proven in Appendix E, for any given $\alpha \in [0, 1)$, if $Q(\alpha, z) = Q(\alpha, n+1)$ and $0 < z < x_0(\alpha) < n+1$, we then have

$$\left| \frac{\partial Q(\alpha, x)}{\partial \alpha} \Big|_{x=z} \right| < \left| \frac{\partial Q(\alpha, x)}{\partial \alpha} \Big|_{x=n+1} \right|. \quad (16)$$

Since $Q(\alpha_0, z) = Q(\alpha_0, n+1)$ and ϵ is an infinitely small positive value, we have

$$\begin{aligned} & Q(\alpha_0 + \epsilon, z) - Q(\alpha_0 + \epsilon, n+1) \\ &= \epsilon \left(\frac{\partial Q(\alpha, x)}{\partial \alpha} \Big|_{x=z, \alpha=\alpha_0} - \frac{\partial Q(\alpha, x)}{\partial \alpha} \Big|_{x=n+1, \alpha=\alpha_0} \right) \\ &+ O(\epsilon^2) > 0. \end{aligned} \quad (17)$$

Furthermore, since $z \in (n, x_0(\alpha))$ and $x_0(\alpha_0 + \epsilon) = x_0(\alpha_0)(\alpha_0 + \epsilon)/\alpha_0 > x_0(\alpha_0)$, we have $z \in (n, x_0(\alpha_0 + \epsilon))$. Note that, from Lemma 1, $Q(\alpha_0 + \epsilon, x)$ decreases when x changes from n to $x_0(\alpha_0 + \epsilon)$. Consequently, we have

$$Q(\alpha_0 + \epsilon, n) > Q(\alpha_0 + \epsilon, z) \stackrel{\text{from (17)}}{>} Q(\alpha_0 + \epsilon, n+1) \quad (18)$$

and therefore, $m^*(\alpha_0 + \epsilon) = n+1$. From $m^*(\alpha_0 + \epsilon) = n+1$, we have $m^*(\alpha_0 + 2\epsilon) = m^*((\alpha_0 + \epsilon) + \epsilon) = n+1$, using the same proof as previously mentioned. Keeping adding ϵ , we can eventually have $m^*(\alpha) = n+1$ for any $\alpha \in [\alpha_0, b]$. This completes the proof. ■

APPENDIX E PROOF OF (16)

Proof: For any given $\alpha \in [0, 1)$, from (15) and the fact that $Q(\alpha, z) = Q(\alpha, n+1)$, we have

$$\frac{\left| \frac{\partial Q(\alpha, x)}{\partial \alpha} \Big|_{x=z} \right|}{\left| \frac{\partial Q(\alpha, x)}{\partial \alpha} \Big|_{x=n+1} \right|} = \frac{\frac{1-p(\alpha, x)}{qt(\ln 2)x^2 2^{qx} p(\alpha, x)} \Big|_{x=n+1}}{\frac{1-p(\alpha, x)}{qt(\ln 2)x^2 2^{qx} p(\alpha, x)} \Big|_{x=z}} \quad (19)$$

where $q = L/(\alpha T)$. Since $Q(\alpha, z) = Q(\alpha, n+1)$, we have $(1 - p(\alpha, z))^z = (1 - p(\alpha, n+1))^{n+1}$, which leads to $(z/n+1) = \ln(1 - p(\alpha, n+1))/\ln(1 - p(\alpha, z))$. Thus, (19) can be rewritten as

$$\begin{aligned} & \frac{\left| \frac{\partial Q(\alpha, x)}{\partial \alpha} \Big|_{x=z} \right|}{\left| \frac{\partial Q(\alpha, x)}{\partial \alpha} \Big|_{x=n+1} \right|} = \frac{\frac{(1-p(\alpha, x)) \ln(1-p(\alpha, x))}{qt(\ln 2)x^2 2^{qx} p(\alpha, x)} \Big|_{x=n+1}}{\frac{(1-p(\alpha, x)) \ln(1-p(\alpha, x))}{qt(\ln 2)x^2 2^{qx} p(\alpha, x)} \Big|_{x=z}} \\ &= \frac{|A(\alpha, n+1)/B(\alpha, n+1)|}{|A(\alpha, z)/B(\alpha, z)|} \end{aligned} \quad (20)$$

where $A(\alpha, x) \triangleq (1 - p(\alpha, x)) \ln(1 - p(\alpha, x))/p(\alpha, x)$, and $B(\alpha, x) \triangleq qt(\ln 2)x^2 2^{qx}$.

It can be seen that the $f(\alpha, x)$ in (7) can be expressed as $f(\alpha, x) = A(\alpha, x) + B(\alpha, x)$. In the proof of Lemma 1, it has been shown that $f(\alpha, x)$ has the following property:

$$\begin{cases} f(\alpha, x) < 0, & \text{if } 0 < x < x_0(\alpha) \\ f(\alpha, x) = 0, & \text{if } x = x_0(\alpha) \\ f(\alpha, x) > 0, & \text{if } x > x_0(\alpha). \end{cases}$$

For $x \in (0, +\infty)$, we have $p(\alpha, x) \in (0, 1)$, $A(\alpha, x) < 0$, and $B(\alpha, x) > 0$. Thus, from the preceding property of $f(\alpha, x)$, we have

$$\begin{cases} |A(\alpha, x)| > |B(\alpha, x)|, & \text{if } 0 < x < x_0(\alpha) \\ |A(\alpha, x)| = |B(\alpha, x)|, & \text{if } x = x_0(\alpha) \\ |A(\alpha, x)| < |B(\alpha, x)|, & \text{if } x > x_0(\alpha). \end{cases}$$

For $0 < z < x_0(\alpha) < n+1$, we have

$$\left| \frac{A(\alpha, z)}{B(\alpha, z)} \right| > 1 > \left| \frac{A(\alpha, n+1)}{B(\alpha, n+1)} \right|.$$

Together with (20), it can be concluded that

$$\left| \frac{\partial Q(\alpha, x)}{\partial \alpha} \Big|_{x=z} \right| < \left| \frac{\partial Q(\alpha, x)}{\partial \alpha} \Big|_{x=n+1} \right|.$$

This completes the proof. ■

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor Dr. O. B. Akan and the anonymous reviewers for their constructive comments and suggestions, which helped to improve the quality of the paper.

REFERENCES

- [1] H. Jiang, W. Zhuang, X. Shen, and Q. Bi, "Quality-of-service provisioning and efficient resource utilization in CDMA cellular communications," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 1, pp. 4–15, Jan. 2006.
- [2] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.
- [3] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: A survey," *Comput. Netw.*, vol. 47, no. 4, pp. 445–487, Mar. 2005.
- [4] R. Fan, H. Jiang, S. Wu, and N. Zhang, "Ranging error-tolerable localization in wireless sensor networks with inaccurately positioned anchor nodes," *Wireless Commun. Mobile Comput.*, vol. 9, no. 5, pp. 705–717, May 2009.
- [5] K. T. Phan, R. Fan, H. Jiang, S. A. Vorobyov, and C. Tellambura, "Network lifetime maximization with node admission in wireless multimedia sensor networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3640–3646, Sep. 2009.
- [6] B. Rong, Y. Qian, K. Lu, H.-H. Chen, and M. Guizani, "Call admission control optimization in WiMAX networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2509–2522, Jul. 2008.
- [7] R. Fan and H. Jiang, "Channel sensing-order setting in cognitive radio networks: A two-user case," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 4997–5008, Nov. 2009.
- [8] J. She, P. Ho, and L. Xie, "IPTV over WiMax: Key success factors, challenges, and solutions," *IEEE Commun. Mag.*, vol. 45, no. 8, pp. 87–93, Aug. 2007.
- [9] S. Kota, Y. Qian, E. Hossain, and R. Ganesh, "Advances in mobile multimedia networking and QoS," *IEEE Commun. Mag.*, vol. 45, no. 8, pp. 52–53, Aug. 2007.
- [10] K. Kang and T. Kim, "Improved error control for real-time video broadcasting over CDMA2000 networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 188–197, Jan. 2009.
- [11] C.-W. Bao and W. Liao, "Performance analysis of reliable MAC-layer multicast for IEEE 802.11 wireless LANs," in *Proc. ICC*, May 2005, pp. 1378–1382.

- [12] J. Metzner, "An improved broadcast retransmission protocol," *IEEE Trans. Commun.*, vol. COM-32, no. 6, pp. 679–683, Jun. 1984.
- [13] A. Majumdar, D. Sachs, I. Kozintsev, K. Ramchandran, and M. Yeung, "Multicast and unicast real-time video streaming over wireless LANs," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 524–534, Jun. 2002.
- [14] J. Nonnenmacher, E. Biersack, and D. Towsley, "Parity-based loss recovery for reliable multicast transmission," *IEEE/ACM Trans. Netw.*, vol. 6, no. 4, pp. 349–361, Aug. 1998.
- [15] D. Nguyen, T. Nguyen, and B. Bose, "Wireless broadcast using network coding," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 914–925, Feb. 2009.
- [16] X. Zhang and Q. Du, "Adaptive low-complexity erasure-correcting code-based protocols for QoS-driven mobile multicast services over wireless networks," *IEEE Trans. Veh. Technol.*, vol. 55, no. 5, pp. 1633–1647, Sep. 2006.
- [17] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.
- [18] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.
- [19] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [20] P. Liu, Z. Tao, Z. Lin, E. Erkip, and S. Panwar, "Cooperative wireless communications: A cross-layer approach," *IEEE Wireless Commun.*, vol. 13, no. 4, pp. 84–92, Aug. 2006.
- [21] I. Güvenc, U. C. Kozat, M.-R. Jeong, F. Watanabe, and C.-C. Chong, "Reliable multicast and broadcast services in relay-based emergency communications," *IEEE Wireless Commun.*, vol. 15, no. 3, pp. 40–47, Jun. 2008.
- [22] S. Raza, D. Li, C.-N. Chuah, and G. Cheung, "Cooperative peer-to-peer repair for wireless multimedia broadcast," in *Proc. IEEE ICME*, Jul. 2007, pp. 1075–1078.
- [23] X. Liu, S. Raza, C.-N. Chuah, and G. Cheung, "Network coding based cooperative peer-to-peer repair in wireless ad-hoc networks," in *Proc. IEEE ICC*, May 2008, pp. 2153–2158.
- [24] J. Jin and B. Li, "Cooperative multicast scheduling with random network coding in WiMAX," in *Proc. 17th IWQoS*, Jul. 2009.
- [25] P. Fan, C. Zhi, C. Wei, and K. B. Letaief, "Reliable relay assisted wireless multicast using network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 749–762, Jun. 2009.
- [26] A. del Coso, O. Simeone, Y. Bar-ness, and C. Ibars, "Outage capacity of two-phase space-time coded cooperative multicasting," in *Proc. 40th ACSAC*, Oct./Nov. 2006, pp. 666–670.
- [27] G. Jakllari, S. V. Krishnamurthy, M. Faloutsos, and P. V. Krishnamurthy, "On broadcasting with cooperative diversity in multi-hop wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 484–496, Feb. 2007.
- [28] B. Sirkeci-Mergen and A. Scaglione, "On the power efficiency of cooperative broadcast in dense wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 497–507, Feb. 2007.
- [29] H. Kubo, A. Kawamura, R. Shinkuma, and T. Takahashi, "Loss compensation using route diversity in wireless multicast," in *Proc. 14th Asia-Pacific Conf. Commun.*, Oct. 2008.
- [30] I. Maric and R. D. Yates, "Cooperative multi-hop broadcast for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1080–1088, Aug. 2004.
- [31] I. Maric and R. D. Yates, "Cooperative multicast for maximum network lifetime," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 127–135, Jan. 2005.
- [32] O. Alay, T. Korakis, Y. Wang, E. Erkip, and S. Panwar, "Layered wireless video multicast using omni-directional relays," in *Proc. IEEE ICASSP*, Mar./Apr. 2008, pp. 2149–2152.
- [33] W. Ge, J. Zhang, and X. Shen, "A cross-layer design approach to multicast in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 1063–1071, Mar. 2007.

Performance of Differentiated Rate Scheduling Using Contention-Based CSI Feedback

Seung Young Park, *Senior Member, IEEE*

Abstract—In this paper, we address the performance of downlink packet delivery under differentiated rate scheduling, where the base station supports different arrival rates to different subsets of users. To reduce the uplink feedback load, the user whose downlink channel gain is larger than a threshold is allowed to send its channel-state information with an access probability through a spread-spectrum-based contention channel. Under this framework, we show that the differentiated rates can be supported among the subsets by adjusting their access probabilities according to their arrival rates. Using large deviation techniques (LDTs), we show that the throughput can be maintained if the performance degradations in the delay and the queue length are allowed for the low-arrival-rate user subset. However, the queue-length performance degradation of the low-rate subset is less severe compared with that of the delay performance. In addition, our simulation results show that the delay performance estimation from LDT analysis is valid even for a delay that is not too large, although it is developed under the assumption of large delay asymptote.

Index Terms—Contention channel, differentiated rate scheduling, large deviation techniques (LDTs), multiuser diversity (MUD).

I. INTRODUCTION

WIRELESS packet scheduling has recently been considered because it is possible to obtain an improved system throughput through improved scheduling. These gains are obtained due to the fact that multiuser diversity (MUD) can be exploited. To fully exploit MUD, a greedy scheduling policy that allocates radio resources to the user whose channel is near its peak has been employed [1]. In addition, this greedy scheduling policy is known to maximize the system throughput only when there is no constraint on scheduling [1]. If every user has a certain probability of being selected, then there must be some sort of associated delay distribution and, thus, a certain probability of the delay exceeding a given threshold. More accurately, the greedy policy does not guarantee a specified maximum delay, e.g., for a certain class of service. In practical situations, the average signal-to-noise power ratio (SNR) may be different because the users are distributed at different distances from the base station and have different levels of shadowing. In this case, the maximum delay performance of the greedy policy becomes more severely degraded. The reason is that the user with a high average SNR is more frequently selected. On the other hand, for a given delay constraint, the total system throughput of the greedy policy in a Rayleigh fading channel eventually decreases to zero as the number of users becomes sufficiently large [2]. Thus, it is important to investigate how a large delay is required to maintain most of the MUD gain with the number of users. A queue-length-based (QLB) scheduling policy, in which the user with the best queue-length-weighted channel state is served, has been known to mitigate this problem [3], [4]. Particularly, it has been shown that MUD can be exploited through QLB scheduling when the delay constraint is suitably adjusted with the number of users [5].

Manuscript received June 5, 2009; revised November 16, 2009 and January 25, 2010; accepted April 6, 2010. Date of publication May 3, 2010; date of current version July 16, 2010. This work was supported by Kangwon National University, Korea, under the 2009 Research Grant Program. The review of this paper was coordinated by Prof. W. A. Krzymień.

The author is with the School of Information Technology, Kangwon National University, Chuncheon 200-701, Korea (e-mail: s.young.park@kangwon.ac.kr).

Digital Object Identifier 10.1109/TVT.2010.2049593