Delay- and diversity-aware buffer-aided relay selection policies in cooperative networks

Dimitrios Poulimeneas*, Themistoklis Charalambous†, Nikos Nomikos‡, Ioannis Krikidis§, Demosthenes Vouyioukas‡, Mikael Johansson*

*Automatic Control Lab, Royal Institute of Technology (KTH), Stockholm, Sweden

†Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden

‡Department of Information and Communication Systems Engineering, University of the Aegean, Samos, Greece

§Electrical and Computer Engineering Department, University of Cyprus, Nicosia, Cyprus

***Abstract*—In this paper, we propose novel relay selection policies that aim at reducing the average delay by incorporating the buffer size of the relay nodes into the relay selection process. More specifically, we propose two delay-aware protocols that are based on the $\max-\text{link}$ relay selection protocol. First, a delay-aware only approach while it reduces the delays considerably it starves the buffers and increases the outage probability of the system. Towards this end, we propose a delay- and diversity-aware buffer-aided relay selection policy that aims at reducing the average delay considerably and at the same time maintaining good diversity. The protocols are analyzed by means of Markov Chains and expressions for the outage, throughput and delay are derived. The performance and use of our proposed algorithms is demonstrated via extensive simulations and comparisons.**

***Index Terms*—Cooperative relaying, relay selection, buffer-aided relaying, delays, diversity, Markov chains.**

I. Introduction

Relay selection has been introduced as a spectrally efficient solution that achieves full diversity by requiring only one additional orthogonal channel. In addition, it was shown in [1] to reduce the complexity of the network since multi-relay schemes rely on distributed space-time codes and the coordination needed among the relays. In earlier works, in which relays were assumed to lack data buffers, relay selection was mainly based on the $\max-\min$ criterion and its variations (see, for example, [1]–[3] and references therein). As a result, the relay that received the source signal in the first slot is the same that is subsequently forwarding the signal towards the destination in the second slot. As a result, the performance of the $\max-\min$ relay selection policy is limited by the constraint that the links for each data transmission are determined concurrently and they are associated with the same relay.

Buffer-aided relaying breaks this coupling between the two slots, since relay nodes are now equipped with buffers that store packets. Therefore different relays can be selected for transmission and reception, thus allowing increased degrees of freedom (see [4]–[7]). In [6], for example, the $\max-\max$ relay selection policy is proposed where the links with the strongest source-relay and relay-destination channels are selected for reception and transmission, respectively, without necessarily using the same relay. $\max-\max$ offers significant coding gain over $\max-\min$, but the diversity gain remains unchanged. Going one step further, in [8] and subsequent works (see, for example, [9], [10]) the two-slot convention is relaxed and at each slot either the source transmits or a relay, thus allowing for increased diversity of the system. Buffering capabilities at the relay nodes enhances the performance of cooperative networks.

In the majority of the aforementioned works the main target is outage probability reduction or throughput improvement. An issue that arises from the use of buffer-aided relays in cooperative systems is that except for increasing diversity and robustness, buffers have increased average transmission delays (which can be seen as the amount of time required on average for packets to reach the receiver). It is clear that by reducing the average delay and establishing some guarantees, buffer-aided schemes can be of use in delay intolerant applications as well (e.g., video streaming, web browsing, file sharing).

The main question is whether there is a trade-off between the benefits offered by buffer-aided relaying and delays or whether there is a way to reduce delays considerably without compromising any of the aforementioned benefits. In our work, we aim at providing an answer to these questions and focus on reducing the average transmission delays of the system. We propose a different approach to relay selection by relaxing the requirement of link selection based on the link quality, since we consider transmissions with fixed rates and power levels. Instead the selected link is chosen based on its buffer size only, provided that transmission on the chosen link is feasible (i.e., the link is not in outage). More specifically, the contributions of the paper are as follows:

1) We investigate the performance of a corresponding delay-aware algorithm based on the $\max-\text{link}$ protocol presented in [8], where in each slot either a source-relay or a relay-destination link is chosen, and the aim is to minimize the delay of packets in the system. We show via simulations that while the algorithm reduces the average delay considerably, it does so at the expense of reduced diversity that results in an increase of the outage probability of the system. The poor performance of this protocol, leads us towards the investigation of a delay-aware algorithm that takes the diversity issue into account, which is our next contribution.

2) Then, a diversity- and delay-aware algorithm is proposed based on the $\max-\text{link}$ protocol, where it aims at guaranteeing that buffer sizes are not empty. Simulations show that the outage probability is reduced (even more than $\max-\text{link}$ itself) while reducing also the average delay with respect to that of $\max-\text{link}$. Analytical expression shows that the average delay in the high Signal-to-Noise Ratio (SNR) regime depends only on the number of buffers.

3) Finally, the performance as well as theoretical results of our proposed algorithms are demonstrated via extensive simulations and comparisons.

The remainder of this paper is organized as follows. In Section II, we introduce the system model and preliminary discussion. In Sections III and IV, we present in detail the relay selection schemes proposed herein. Simulations for each scheme are provided after its description. Finally, conclusions are discussed in Section V.

II. System Model and Preliminaries

A. System Model

We assume a relay-assisted network consisting of one source S, one destination D and a cluster $\mathcal{C}$ of K Half-Duplex (HD) Decode-and-Forward (DF) relays $R_k \in \mathcal{C}$ $(1 \leq k \leq K)$. For simplicity, it is assumed that a direct link between the source and the destination does not exist and communication can be established only via relays. Each relay R_k holds a buffer (data queue) Q_k of capacity L (maximum number of data elements), where it can store source data that has been decoded at the relay and can be forwarded to the destination. We denote by $Q \triangleq (Q_1, Q_2, \ldots, Q_K)$ the vector containing the buffer sizes of all relays. The system model is depicted in Figure 1.

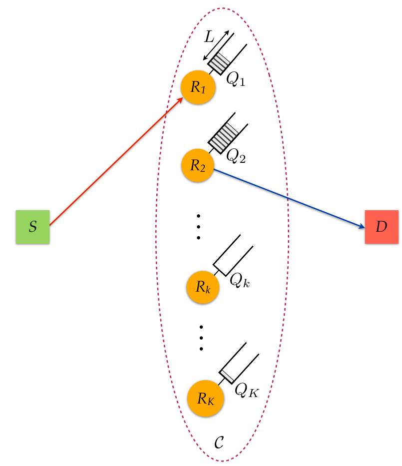

Fig. 1. The system model: Source S communicates with Destination D via a cluster of relays $R_k \in \mathcal{C}$, $k \in \{1, 2, \ldots, K\}$.

The quality of the wireless channels is degraded by Additive White Gaussian Noise (AWGN) and frequency non-selective Rayleigh block fading according to a complex Gaussian distribution with zero mean and variance σ_{ij}^2 for the $\{i \rightarrow j\}$ link. For simplicity, the variance of the AWGN is assumed to be normalized with zero mean and unit variance. The channel gains are $g_{ij} \triangleq |h_{ij}|^2$ and exponentially distributed. It is assumed that global Channel State Information (CSI) is available, unless otherwise stated. The power level chosen by the transmitter i is denoted by $P_i \in [0, P_{i,\max}]$, where $P_{i,\max}$ is the maximum power of a transmitter i. The variance of thermal noise is denoted by η, and it is assumed to be AWGN and the same on all nodes, for simplicity. Time is considered to be slotted and at each time-slot the source S or one of the relays R_k attempts to transmit a packet. The source node is assumed to be saturated (it has always data to transmit) and the information rate, when the transmission is successful, is fixed and equal to r_0. Equivalently, a transmission from a transmitter to its corresponding receiver is successful if the SNR of the receiver is greater or equal to a threshold γ_0, called the *capture ratio*. The value of γ_0 depends on the modulation and coding characteristics of the radio such as the required data rate of the application which is supported by the network and the error-correction coding technique. Hence, independently of nodal distribution and traffic pattern, a transmission from a transmitter i to its corresponding receiver j is successful (error-free) if the SNR of the receiver j, denoted by γ_j, is greater or equal to the *capture ratio* γ_0. Therefore, we require that

$$\gamma_j(P_i) \triangleq \frac{g_{ij} P_i}{\eta} \geq \gamma_0. \tag{1}$$

Link $\{i \rightarrow j\}$ is in outage if $\gamma_j(P_{i,\max}) < \gamma_0$, i.e., $\frac{g_{ij} P_{i,\max}}{\eta} < \gamma_0$, and the probability of outage is given by

$$p_{\text{out}} = \mathbb{P}\left[g_{ij} < \frac{\gamma_0 \eta}{P_{i,\max}}\right].$$

This framework has been widely used and is equivalent to the *capture model* introduced in [11]. Hence, the instantaneous SNR from S to R_i and the instantaneous SNR from R_j to D, when relay R_i is selected for reception and relay R_j is selected for transmission, are expressed as $\gamma_{R_i}(P_S) = \frac{g_{SR_i} P_S}{\eta} \geq \gamma_0$, and $\gamma_D(P_{R_j}) = \frac{g_{R_j D} P_{R_j}}{\eta} \geq \gamma_0$, respectively.

The retransmission process is based on an Acknowledgement/Negative-Acknowledgement (ACK/NACK) mechanism, in which short-length error-free packets are broadcasted by the receivers over a separate narrow-band channel.

The framework introduced herein assumes there exists knowledge of the CSI of the links and instead of maximizing the throughput, we minimize the power expenditure which can be seen as its dual problem. If the CSI of the links is not known and instead only the channel connectivity state is known (i.e., whether or not the links are in outage, provided they transmit with a fixed power) at the source and relays, the procedure, results and algorithms proposed herein remain the same. The channel connectivity state can be obtained, for instance, by using a one-bit feedback message from the receiver utilizing the ACK/NACK mechanism.

Let $b_{SR} \triangleq (b_{SR_1}, b_{SR_2}, \ldots, b_{SR_K})$ and $b_{RD} \triangleq (b_{R_1 D}, b_{R_2 D}, \ldots, b_{R_K D})$ be the binary representation of the feasible links due to the fulfillment of the channel conditions (i.e., if transmission on link $R_i D$ is possible, then $b_{R_i D} = 1$). Similarly, let $q_{SR} \triangleq (q_{SR_1}, q_{SR_2}, \ldots, q_{SR_K})$ and $q_{RD} \triangleq (q_{R_1 D}, q_{R_2 D}, \ldots, q_{R_K D})$ be the binary representation of the feasible links due to the fulfillment of the queue conditions (i.e., for a $\{S\to R\}$ link the buffer is not full and for a $\{R\to D\}$ link the buffer is not empty). By $\mathcal{F}_{SR}$ and $\mathcal{F}_{RD}$, we denote the $\{S\to R\}$ and $\{R\to D\}$ links that are feasible. In practice, no attempt for transmission is made on link $\{i\to j\}$ if $b_{ij}=0$ or $q_{ij}=0$ and we say that the system is in *outage*.

B. Preliminaries

A mechanism that considers not only the channel state but also the queue state could have beneficial effect on balancing the queues of the system, a fact that in the systems of interest could reduce delay considerably. We aim at reducing the average packet transmission delay without at the same time, sacrificing the advantages of buffer-aided relaying. Thus, we build upon two existing relay selection schemes, namely the HRS [6] and the $\max-\text{link}$ [8], by providing a novel way to direct the packets towards the destination. In the case of the HRS, each time-frame consists of two time-slots and, thus, two different links need to be selected per frame, whereas in the case of $\max-\text{link}$ one link is selected in every time-slot, since time-frame and time-slot durations are identical.

In [8] a framework based on Discrete Time Markov Chains (DTMC) is proposed to analyze the $\max-\text{link}$ algorithm. This framework has been used in many subsequent works in the field to analyze other buffer-aided relay selection protocols whose buffer is finite. Below, the general picture is provided.

States of the DTMC. The states of the DTMC represent all the possible states of the buffers. The transitions between the states are given by the probabilities of successful transmissions of packets either to or from a relay. The state of the DTMC can be represented by

$$S_r \triangleq \left(Q_1^{(r)}Q_2^{(r)}\dots Q_K^{(r)}\right),\quad r\in\mathbb{N}_+, 1\le r\le (L+1)^K.$$

The states are predefined in a random way as all the possible $(L+1)^K$ combinations of the buffer sizes combined with the destination state, and are considered as a data input for the proposed selection policy.

Construction of the state transition matrix of the DTMC. Let $\mathbf{A}\in\mathbb{R}^{(L+1)^K\times(L+1)^K}$ denote the state transition matrix of the DTMC, in which the entry

$$\mathbf{A}_{i,j}=\mathbb{P}\big(S_j\to S_i\big)=\mathbb{P}\big(X_{t+1}=S_i|X_t=S_j\big)$$

is the transition probability to move from state S_j at time t to state S_i at time $(t+1)$. In order to construct the state transition matrix $\mathbf{A}$, we need to identify the connectivity between the different states of the buffers. For each time slot, the buffer status can be modified as follows: (a) the number of elements of one relay buffer can be decreased by one, if a relay node is selected for transmission and the transmission is successful, (b) the number of elements of one buffer can be increased by one, if the source node is selected for transmission and the transmission to the relay is successful and the transmission to the destination is unsuccessful, (c) the buffer state (not the DTMC state) remains unchanged when there is an outage event (i.e., all the $\{S\to R\}$ and $\{R\to D\}$ links are in outage).

Properties of the DTMC. Due to the fact that the buffer of each relay is finite, the DTMC can be easily shown to be Stationary, Irreducible and Aperiodic (SIA) [8]. The existence of a steady state distribution by Little's law implies finite average packet delay. In what follows, we provide *general* analytical expressions for the outage probability, average throughput and average packet delay, that *can be used in any policy defined for finite-length buffer-aided relay selection*.

Derivation of the outage probability. An outage event occurs *only when there is no change in the buffer state*. Hence, the outage probability of the system is given by the sum of the product of the probabilities of being at a stage r and having an outage event, i.e.,

$$p_{\text{out}}=\sum_{r=1}^{(L+1)^K}\pi_r\overline{p}_r=\text{diag}(\mathbf{A})\pi. \tag{2}$$

Eq. (2) shows that the construction of the state matrix $\mathbf{A}$ and the computation of the related steady state π comprises a simple theoretical framework for the computation of the outage probability for a buffer-aided relay selection policy.

Derivation of the average throughput. If there is only one transmission per time-slot (e.g., [6], [8]), the average data rate ρ is $1/2$ since two hops are required to reach the destination; in schemes with successive transmissions though, ρ is approaching 1. The proportion of the packets that make it through is $(1-p_{\text{out}})$. Hence, the average throughput is given by $\mathbb{E}[T]=\rho(1-p_{\text{out}})$, where $\rho\in\{1/2,1\}$. Note that if the links are i.i.d., then the average throughput of relay R_j is given by

$$\mathbb{E}[T_j]=\frac{\rho(1-p_{\text{out}})}{K}. \tag{3}$$

Derivation of the average packet delay. The delay of a packet is the duration of time between the time it arrives at a relay until the time it reaches the destination (i.e., no delay is measured when the packet resides at the source). The average packet delay under this framework was recently presented in [12]. We summarize the results herein for completeness. For i.i.d. channels, the average delay is the same on all relays. Hence, it is enough to analyze the average delay on a single relay. By Little's law, the average packet delay at relay R_j, denoted by $\mathbb{E}[d_j]$ is given by

$$\mathbb{E}[d_j]=\frac{\mathbb{E}[L_j]}{\mathbb{E}[T_j]}, \tag{4}$$

where $\mathbb{E}[L_j]$ and $\mathbb{E}[T_j]$ are the average queue length and average throughput, respectively. The average queue length at relay R_j is given by

$$\mathbb{E}[L_j]=\sum_{r=1}^{(L+1)^K}\pi_r Q_j^{(r)}, \tag{5}$$

and the average throughput is given in (3). Substituting (2), (3) and (5) into (4) we have that the average delay is given by

$$\mathbb{E}[d_j]=\frac{K\sum_{r=1}^{(L+1)^K}\pi_r Q_j^{(r)}}{\rho\left(1-\sum_{r=1}^{(L+1)^K}\pi_r\overline{p}_r\right)}. \tag{6}$$

III. DELAY-AWARE $\max-\text{link}$

During the execution of $\max-\text{link}$, at each time-slot, both $\{S\rightarrow R\}$ and $\{R\rightarrow D\}$ links compete for being selected at the specific time slots, since the two-slots per frame convention is removed and only one slot is available. However, if we take delay into account, the intuitive would be that $\{R\rightarrow D\}$ links should be given priority, in order to enforce packets to leave the buffers sooner than later and hence, reduce the average length.

Based on this observation, we proceed with the description of the Delay-Aware $\max-\text{link}$ ($\text{DA}-\max-\text{link}$) Relay Selection algorithm. In every time slot, from all the relays that can transmit packet to the destination, the criterion of relay selection is the number of packets waiting on the buffer. The relay with the maximum number of packets in the buffer will be prioritized over all the other relays that could be selected to transmit. If this choice is not available (i.e., no nonempty relay can transmit due to bad channel conditions), the relay with the minimum buffer size available is granted priority to receive a packet from the source. If there does not exist a relay with a feasible $\{S\rightarrow R\}$ link, no packet transmission takes place during this slot and the system is in outage. When more than one relays have the same number of packets in the data queue, a relay is randomly selected.

The $\text{DA}-\max-\text{link}$ Relay Selection algorithm for a single time-slot is as follows:

Algorithm 1 Delay-Aware $\max-\text{link}$ Relay Selection

input Q, $\mathcal{F}_{SR}$, $\mathcal{F}_{RD}$
if $\mathcal{F}_{SR}=\emptyset$ **and** $\mathcal{F}_{RD}=\emptyset$ **then**
  No packet transmission takes place.
else
  if $\mathcal{F}_{RD}\neq\emptyset$ **then**
    $j=\arg\max_{i\in\mathcal{F}_{RD}} Q_i$ ($\{R\rightarrow D\}$ link)
  else
    $k=\arg\min_{i\in\mathcal{F}_{SR}} Q_i$ ($\{S\rightarrow R\}$ link)
  end if
end if
Output Link $\{R_j\rightarrow D\}$ or $\{S\rightarrow R_k\}$ for transmission.

A. Numerical evaluation

The effect of the different buffer sizes on the average delay and the outage probability for the $\text{DA}-\max-\text{link}$ algorithm with respect to the $\max-\text{link}$ algorithm is depicted in Figure 2. It can be seen that for the $\text{DA}-\max-\text{link}$ algorithm the buffer size has no effect on the average delay, but it affects the outage probability. While the gain on the average delay is remarkable, this comes at the expense of increased outage probability. When the channel is good, the difference in outage probability is of the order of magnitude bigger for the $\text{DA}-\max-\text{link}$ algorithm than that of the $\max-\text{link}$ algorithm. This is due to the fact that most of the relays have empty buffers throughout its operation. This, of course, results in diversity loss of the system and this is evident in the slopes for the two schemes, as shown in Figure 2.

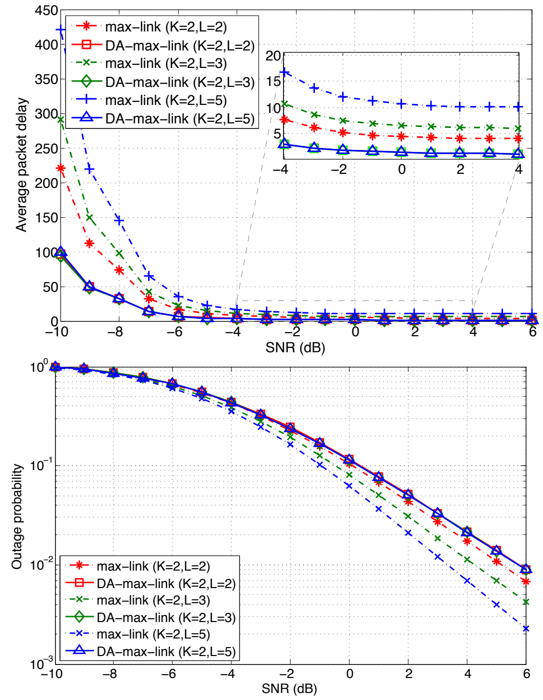

Fig. 2. Average delay (top) and outage probability (bottom) of $\text{DA}-\max-\text{link}$ algorithm for $K=2$ and $L=2,3,5$.

The effect of the different number of relays on the average delay and the outage probability for the $\text{DA}-\max-\text{link}$ algorithm with respect to the $\max-\text{link}$ algorithm is depicted in Figure 2. Again, since one relay is being used at each time-slot, for the $\text{DA}-\max-\text{link}$ algorithm, the number of relays do not change its performance in terms of average delay. However, the effect on the outage probability is more evident. For $K=2$ and $L=5$ the difference in outage probability is in the order of a magnitude.

IV. DELAY- AND DIVERSITY-AWARE RELAY SELECTION

Prioritizing the packets that have already entered the network over the packets that await in the source was identified to have negative effect on the diversity of the system for the $\max-\text{link}$ algorithm, causing an increase of the outage probability. This is illustrated for bigger buffer sizes in Figure 4.

In this section, a protocol is proposed that overcomes this negative effect by sacrificing part of the average delay reduction achieved in favor of maintaining high diversity. Note that the effect of reduced diversity does not occur in the delay-aware HRS. This can be explained from the way the decision is made in the different slots consisting a frame: During the first slot, a packet is selected to be sent to the available relay that has the minimum number of packets in its queue, which means that we prioritize the empty queues over all the other available choices. During the second slot, a packet is selected to be sent from the available relay that has the maximum number of packets in its queue, which in this case means that the relays with one packet are the last choice to transmit towards the destination.

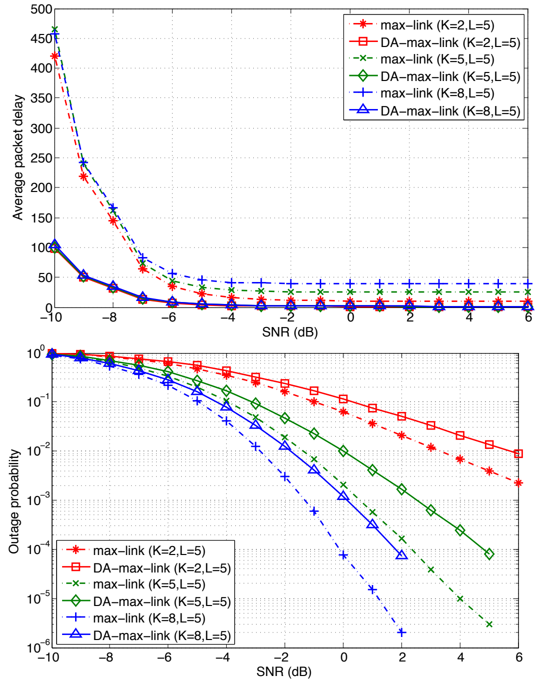

Fig. 3. Comparison of average delay (top) and outage probability (bottom) between $\mathrm{DA-max-link}$ and $\mathrm{max-link}$ algorithms for $K = 2, 3, 4$ and $L = 5$.

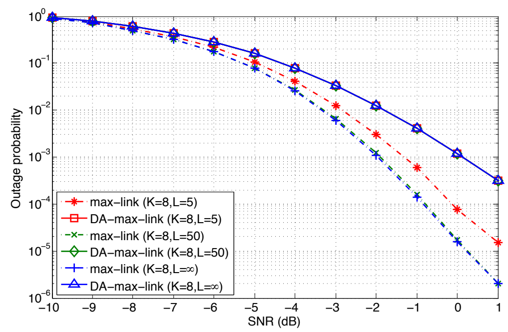

Fig. 4. Comparison between $\mathrm{max-link}$ and $\mathrm{DA-max-link}$ algorithms of outage probability for $K = 8$ and $L = 5, 50, \infty$.

In what follows, we describe in detail the diversity- and delay- aware $\mathrm{max-link}$ ($\mathrm{DDA-max-link}$) Relay Selection algorithm:

A. Numerical evaluation

In Figure 5, the performance of the $\mathrm{DDA-max-link}$ algorithm is compared to that of the $\mathrm{max-link}$ for $K = 2$ and for different values of the buffer size L. Regarding the delays, the $\mathrm{max-link}$ performs better for small buffer sizes ($L < 4$). For $L = 5$, the $\mathrm{max-link}$ is outperformed by the proposed algorithm. Note that the examples of the $\mathrm{DDA-max-link}$ algorithm (and the $\mathrm{max-link}$ algorithm) reach the theoretical values derived in the analysis.

In Figure 6, we show another example where the buffer size is fixed to $L = 5$ and the number of relays changes. Since $L \geq 4$, as mentioned in the analysis, the proposed algorithm outperforms the $\mathrm{max-link}$ algorithm in terms of delays. Note also, that $\mathrm{DDA-max-link}$ outperforms

Algorithm 2 Diversity- and Delay-Aware $\mathrm{max-link}$ Relay Selection

input Q, $\mathcal{F}_{SR}$, $\mathcal{F}_{RD}$
if $\mathcal{F}_{SR} = \emptyset$ **and** $\mathcal{F}_{RD} = \emptyset$ **then**
  No packet transmission takes place.
else
  if $\mathcal{F}_{SR} = \emptyset$ **then**
    $j = \arg\max_{i \in \mathcal{F}_{RD}} Q_i$ ($\{R \to D\}$ link)
  else
    $\widetilde{\mathcal{F}}_{SR} \triangleq \{i : i \in \mathcal{F}_{SR}, Q_i \leq 1\}$
    if $\widetilde{\mathcal{F}}_{SR} \neq \emptyset$ **then**
      $k = \arg\min_{i \in \widetilde{\mathcal{F}}_{SR}} Q_i$ ($\{S \to R\}$ link)
    else
      $\widetilde{\mathcal{F}}_{RD} \triangleq \{i : i \in \mathcal{F}_{RD}, Q_i \geq 2\}$
      if $\widetilde{\mathcal{F}}_{RD} \neq \emptyset$ **then**
        $j = \arg\max_{i \in \widetilde{\mathcal{F}}_{RD}} Q_i$ ($\{R \to D\}$ link)
      else
        $\hat{\mathcal{F}}_{SR} \triangleq \{i : i \in \mathcal{F}_{SR}, Q_i \geq 2\}$
        $k = \arg\min_{i \in \hat{\mathcal{F}}_{SR}} Q_i$ ($\{S \to R\}$ link)
      end if
    end if
  end if
end if
Output Link $\{R_j \to D\}$ or $\{S \to R_k\}$ for transmission.

$\mathrm{max-link}$ in terms of outage probability as well, compared to $\mathrm{DA-max-link}$ who performed very poorly.

B. Analysis

For $\mathrm{max-link}$ the derivation for the average delay at the high SNR regime is given in [12]. First the throughput of each relay is found: since the selection of a relay is equiprobable, the average throughput at any relay R_j is ρ/K, where ρ is the average data rate; since we have half-duplexity $\rho = 1/2$ and therefore $\mathbf{E}[T_j] = 1/2K$. Also, it can be easily shown that the average queue length at any relay is $\mathbf{E}[L_j] = \frac{L}{2}$. Hence, by Little's law, $\mathbf{E}[d_j] = \mathbf{E}[d] = KL$. That means that, as either the number of relays or the buffer size increases, the average delay of the $\mathrm{max-link}$ algorithm increases.

In what follows, we derive the average delay of the $\mathrm{DDA-max-link}$ at the high SNR regime and we show that it is independent of the buffer size L. In addition, we show that for $L \geq 4$, the average delay of $\mathrm{DDA-max-link}$ is smaller than that of $\mathrm{max-link}$, for any number of relays.

First, we construct the DTMC with the admissible states that our proposed algorithm enforces in the high SNR regime. Our algorithm as structured, aims to have 2 packets in each relay. Hence, at the high SNR regime, one packet from a relay is transmitted to the destination and in the next slot, the gap in the queue is covered by a packet from the source to that relay. As a result, the number of states is reduced considerably.

For a network of K relays it can be easily deduced by the construction of the DTMC that the steady state distribution is given by $\pi = \begin{bmatrix} 1/2 & 1/2K & 1/2K & \dots & 1/2K \end{bmatrix}$, and thus, the average queue length at relay R_j is $\mathbf{E}[L_j] =$

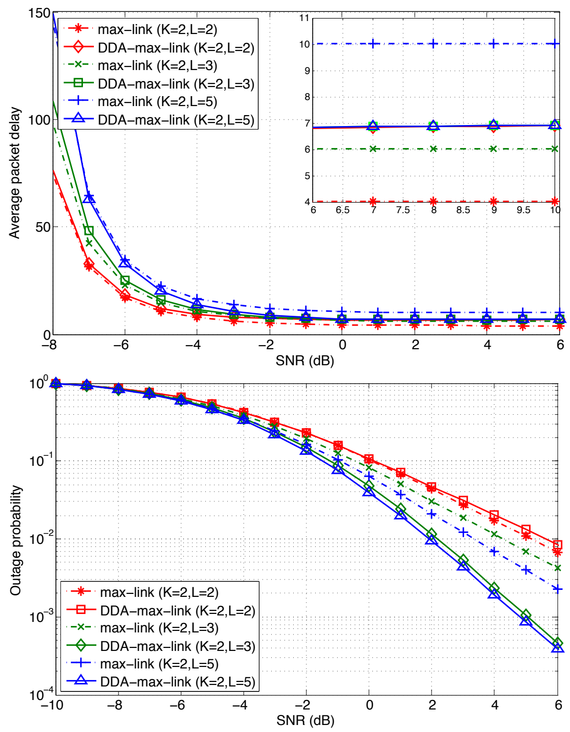

Fig. 5. Average delay (top) and outage probability (bottom) for the $\mathrm{DDA}-\max-\mathrm{link}$ algorithm for $K=2$ and $L=2,3,5$.

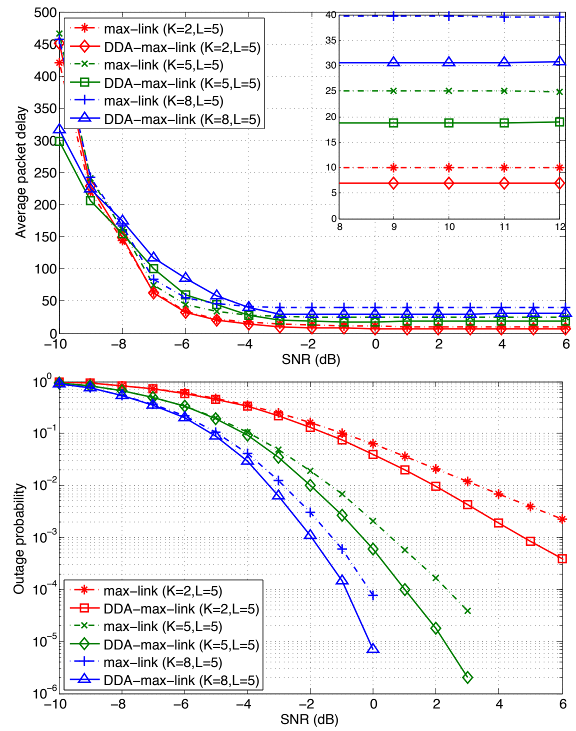

Fig. 6. Average delay (top) and outage probability (bottom) for the $\mathrm{DDA}-\max-\mathrm{link}$ algorithm for $K=2,5,8$ and $L=5$.

$(4K-1)/2K$. Hence, by Little's law, the average delay is given by $\mathbf{E}[d_j]=4K-1$, a quantity which is independent of the buffer size L. This is expected since the state transition matrix does not depend on the buffer size. The average delay of $\mathrm{DDA}-\max-\mathrm{link}$ is lower than that of $\max-\mathrm{link}$ when $4K-1\leq KL$; hence, for $L\geq 4$, the $\mathrm{DDA}-\max-\mathrm{link}$ algorithm performs better than the $\max-\mathrm{link}$ algorithm for any size of network. From a realistic point of view, the buffer size is much bigger than $L=4$, suggesting that this scheme should be preferred over the $\max-\mathrm{link}$. The theoretical results are justified via the simulations in Section IV-A.

V. Conclusions

In this paper, we propose relay selection policies based on $\max-\mathrm{link}$ that aim at reducing the delays inherent to buffer-aided relaying, so that delay-sensitive applications can be supported. When the overall performance of the protocols discussed is considered, the diversity- and delay-aware protocol is a better option, even though the reduction of average transmission delay might be smaller compared to that of the protocols that do not consider diversity. The performance of the proposed scheme demonstrate its effectiveness.

References

[1] A. Bletsas, A. Khisti, D. Reed and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Select. Areas Commun.*, vol. 24, pp. 659–672, March 2006.

[2] D. S. Michalopoulos and G. K. Karagiannidis, "Performance analysis of single relay selection in Rayleigh fading," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 3718–3724, Oct. 2008.

[3] A. Bletsas, H. Shin and M. Z. Win, "Cooperative communications with outage-optimal opportunistic relaying," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 3450–3460, Sept. 2007.

[4] N. B Mehta, V. Sharma and G. Bansal, "Performance analysis of a cooperative system with rateless codes and buffered relays," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1069–1081, April 2011.

[5] R. Wang; V. K. N Lau and H. Huang, "Opportunistic buffered decode-wait-and-forward (OBDWF) protocol for mobile wireless relay networks," *IEEE Trans. Wireless Commun.* , vol. 10, no. 4, pp. 1224–1231, April 2011.

[6] A. Ikhlef, D. S. Michalopoulos and R. Schober, "Max-max relay selection for relays with buffers," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 1124–1135, March 2012.

[7] A. Ikhlef, K. Junsu, and R. Schober, "Mimicking full-duplex relaying using half-duplex relays with buffers," *IEEE Trans. Vehicular Tech.*, vol. 61, pp. 3025–3037, Sept. 2012.

[8] I. Krikidis, T. Charalambous and J. S. Thompson, "Buffer-aided relay selection for cooperative diversity systems without delay constraints," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 1957–1967, May 2012.

[9] T. Charalambous, N. Nomikos, I. Krikidis, D. Vouyioukas and M. Johansson, "Modeling Buffer-Aided Relay Selection in Networks With Direct Transmission Capability", *IEEE Commun. Letters*, vol. 19, no. 4, pp. 649–652, April 2015.

[10] N. Nomikos, T. Charalambous, I. Krikidis, D. N. Skoutas, D. Vouyioukas and M. Johansson, "A buffer-aided successive opportunistic relaying selection scheme with power adaptation and inter-relay interference cancellation for cooperative diversity systems," *IEEE Trans. on Commun.*, vol. 63, no. 5, pp. 1623–1634, May 2015.

[11] B. Hajek, A. Krishna, and R. O. LaMaire, "On the capture probability for a large number of stations," *IEEE Trans. on Commun.*, vol. 45, no. 2, pp. 254–260, Feb. 1997.

[12] Z. Tian, G. Chen, Y. Gong, Z. Chen, and J. Chambers, "Buffer-Aided Max-Link Relay Selection in Amplify-and-Forward Cooperative Networks," *IEEE Trans. on Veh. Techn.*, vol. 64, no. 2, pp. 553–565, Feb. 2015.

[13] M. S. Neiman, "The Principle of Reciprocity in Antenna Theory," Proceedings of the IRE , vol.31, no.12, pp. 666–671, Dec. 1943.