On the Design of a Relay–Assisted Network

Ahmed El Shafie†, Tamer Khattab*, Ahmed Sultan*, H. Vincent Poor††
†Electrical Engineering Dept., University of Texas at Dallas, USA.
*Electrical Engineering, Qatar University, Doha, Qatar.
††King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia.

Abstract—A relay network composed of a source, a set of relays, and a destination is considered. A queuing framework is proposed, where each node maintains a buffer to store its incoming traffic. An optimization-based formulation is proposed to maximize the mean service rate of the source queue. The relay selection technique is based on the statistics of the system parameters. Three packet decoding strategies at the relays are investigated and their performance is compared. The proposed decoding strategies, which do not require the channel state information at the transmitters, can outperform the max-max relay selection strategy.

I. INTRODUCTION

Relays with buffers are considered in [1]–[3]. The max-max relay selection policy is proposed in [1]. In the max-max policy, the buffered relays enable the selection of the relay with the best source-relay channel for reception and the best relay-destination channel for transmission.

In this work, we consider buffered relays. Unlike the existing work in buffered relays and relay selection [1]–[3], we assume that the instantaneous channels are unknown at the transmitters, hence, our protocol does not involve relay selection on the basis of instantaneous channel gains, which are difficult to track. We also assume the presence of a direct link between the source and its destination.

We can summarize the contributions in this paper as follows. We consider a network composed of a source, N relays and a destination. The relays are used to aid the source transmission. We propose an ordered acceptance strategy, denoted by $S_{OD}$, in which the relays are ordered in terms of accepting undelivered packets of the source into their queues. We also propose a random assignment decoding strategy, denoted by $S_{RD}$, and a round robin decoding strategy, denoted by $S_{RR}$, in which each relay is individually assigned to the decoding role for a fraction of the time slots. We optimize the mean service rate of the source queue subject to the stability of all queues in the network, which guarantees the arrival of all packets to the destination as time progresses. We compare our proposed strategies with the max-max relay selection strategy and show that the performance of our proposed strategies can outperform it.

II. SYSTEM MODEL

The network is composed of a source ‘s’, a destination ‘d’, and a set of N relays labeled as 1, 2, 3, …, N. The relays are half-duplex terminals, which means that they either transmit or receive but cannot do both at the same time. Time is partitioned into slots. Each terminal maintains an infinite buffer for storing fixed-length packets. The source operates as a discrete-time Geo/Geo/1 queue [4].1 The arrivals at the source queue are independent and identically distributed (i.i.d.) Bernoulli random variables from slot to slot with average $\lambda_n \in [0, 1]$ packets per time slot. The queue of the source is denoted by $Q_s$, whereas the queue of relay $k \in \mathcal{R}$, $\mathcal{R} = \{1, 2, \ldots, N\}$, is denoted by $Q_k$. The relays help the source to deliver its packets during its periods of silence. If a terminal transmits during a time slot, it sends exactly one packet to the respective receiver.

We assume that a single-bit feedback is implemented by the receiving nodes to inform the decodability status of the packets. This single-bit feedback signal is assumed to be correctly decoded at all nodes in the network. If the feedback is positive, which means that the packet is decoded correctly, we refer to this feedback as an ACK. If the feedback is negative, which means that the packet is decoded erroneously, we refer to this feedback as a NACK. The time slot structure under each of our proposed strategies is depicted in Fig. 1.

1) Ordered Acceptance Decoding, $S_{OD}$: We start with a description of the ordered acceptance decoding strategy, $S_{OD}$. If a relay senses the source to be busy (i.e. sending a packet), it receives the source packet till the end of its transmission. If the destination declares the correct reception of the transmitted packet by sending an ACK signal, the relays discard what they have received from the source. If the destination declares its failure to decode the received packet properly by generating a NACK signal, the relays reveal their ability to decode the

1The notion of a discrete-time Geo/Geo/1 queue is used to describe a queuing system with a Bernoulli arrival process and geometrically distributed service times.

2Even though a relay may be scheduled for transmission in a particular time slot, it will only transmit in this time slot if the source is idle.
received packet through sending the feedback signals based on their packet acceptance orders. That is, if the first-ranked relay can decode the packet, an ACK is transmitted by the relay to inform the source to drop the packet from its queue and to notify the other relays that the packet has already been decoded correctly. If the first-ranked relay fails to decode the packet (due to channel outage), it remains idle and the second-ranked relay sends an ACK if it has decoded the packet properly. Generally, a relay, depending on its acceptance rank, accepts a packet only when its preceding relays could not decode it. Since the channel bandwidth is $1/W$, the total time needed to implement the feedback process is $(N+1)/W$, where $1/W$ is the time duration needed for a one-bit feedback transmission.

The order of source packet acceptance is the $N$-tuple $m_n = (m_1, m_2, \ldots, m_N)$, where $m_z \in \{1, 2, \ldots, N\}$ and $m_z \neq m_k, \forall z, k \in \mathcal{R} : z \neq k$. The $N$-tuple $(m_1, m_2, \ldots, m_N)$ means that relay 1 is assigned the $m_1$th acceptance rank, relay 2 is assigned the $m_2$th rank and so on. It is evident that $m_n$ is a permutation over the set $\mathcal{R}$ and there are $N!$ such permutations constituting the set $\mathcal{M}$. We define the probability $p_n$, $n \in \{1, 2, \ldots, N!\}$, as the probability of the decoding order $m_n = (m_1, m_2, \ldots, m_N)$. The sum of the probabilities of permutations which assign a rank to a certain relay denotes the fraction of time slots with the relay having the aforementioned probabilities, $p_n$, by $\rho$.

The MAC operation under $S_{OD}$ can be summarized as follows: At the beginning of a time slot, the source transmits the packet at the head of its queue to the destination. Due to the broadcast nature of the wireless channel, the relays can listen to the transmitted packet. The relays sense the channel over the first $\tau$ seconds of the time slot. If the source is active, the destination and relays attempt to decode the transmitted packet. If the packet is received correctly by the destination, an ACK message is fed back from the destination. The packet is then dropped from the source queue. The relays also discard what they have received. If the source packet is not received correctly at the destination, a NACK message is fed back from the destination. Based on the decodability status of the packet at the relays and their packet acceptance ranking, the relays send their feedback signals. If no relay could decode the packet, it is kept in the source's queue for retransmission. If the source is idle (i.e. has no packets to send), the relays start transmitting the packets at the heads of their queues.

2) Random Assignment Decoding, $S_{RD}$, and Round Robin Decoding, $S_{RR}$: The difference between random assignment decoding, $S_{RD}$, and $S_{OD}$ is that in $S_{RD}$ only one relay is scheduled to decode the undelivered source packet in any slot. The probability that relay $k$ is assigned the decoding role in a time slot is denoted as $\beta_k$. We define the vector $\beta = (\beta_1, \beta_2, \ldots, \beta_N)$ with the constraint $\sum_{k=1}^{N} \beta_k = 1$. The operation of the relays can be summarized as follows: At the beginning of each time slot, the index, $k$, of the randomly selected relay is generated according to $\beta$. If the source packet is not received correctly, a NACK message is fed back from the destination. The relay that is assigned for packet decoding attempts to decode the undelivered packet. If the packet is decoded correctly, the relay admits the packet to the relaying queue and sends an ACK message, thereby inducing the source to drop the packet. If the relay assigned for decoding fails to decode the packet, it is kept in the source queue for retransmission.

Round robin decoding, $S_{RR}$, is a simplification of $S_{RD}$ in which the decoding assignment is deterministic and happens in a cyclic manner such that on the average a relay is selected for decoding a packet with probability $1/N$ (i.e. equiprobable relay selection). That is, $\beta_k = 1/N$, where $k \in \{1, 2, \ldots, N\}$.

B. Physical (PHY) Layer

The outage probability of a channel is calculated as follows. Assuming that the number of bits in a packet is $b$, the time slot duration is $T$, and the feedback process duration is $T_f$, the transmission rate is

$$ r_i = \frac{b}{T (1 - \frac{T_f}{T})} \quad (1) $$

with $T_f = i\tau + T_f < T$. For the source, the index $i = 0$ as its transmission proceeds at the early beginning of the time slot, whereas for the scheduled relay for transmission, $i = 1$ as its transmission is preceded by a spectrum sensing period of $\tau$ seconds. Outage of a link occurs when the capacity is lower than the transmission rate. Hence, the outage probability of the link between node $j \in \{1, 2, \ldots, N, s\}$ and node $k \in \{1, 2, \ldots, N, d\}$, $\forall j \neq k$, is given by [4]

$$ P_{j,k} = \Pr \left\{ r_i > W \log_2 (1 + \theta_j \theta_{j,k} h_{j,k}) \right\}, \quad (2) $$

where $W$ is the bandwidth of the channel, $\theta_{j,k}$ is the received signal-to-noise ratio (SNR) when the channel gain is equal to unity, and $h_{j,k}$ is the channel gain (i.e. the squared value of the absolute of the complex channel coefficient), which is exponentially distributed in the case of Rayleigh fading. The channel gain, $h_{j,k}$, is assumed to be independent from slot to slot and link to link. Note that $T - T_f$ and $W$ are large enough so that the time slot can contain a codeword comprised of $n \to \infty$ Gaussian distributed symbols onto which the $b$ bits are mapped. Assuming that the mean value of $h_{j,k}$ is $\sigma_{j,k}$ and letting $P_{j,k} = 1 - P_{j,k}$ be the probability of correct reception, we have

$$ P_{j,k} = \exp \left( -2 \frac{T_f \frac{W}{\theta_{j,k} \sigma_{j,k}}}{b} \right). \quad (3) $$

The value of the feedback process, $T_f$, changes based on the adopted strategy and the node that uses the spectrum. Since

![Image](1089-7798_2015_IEEE_Commun_Lett_2195717.pdf)
the relays send packets when the source’s queue is empty, there is only one feedback signal from the destination and consequently \( T_f = \tau_f \), where \( \tau_f = 1/W \) is the duration of one feedback signal. For the source transmission, the impact of the adopted strategy on \( T_f \) is explained as follows. In the case of \( S_{OD} \), the relays are ordered in terms of sending the feedback signals. If each relay needs \( \tau_f = 1/W \) seconds, then the overall feedback duration is \( T_f = (N + 1)\tau_f \) given that the destination also needs \( 1/W \) seconds to acknowledge the reception of a data packet. On the other hand, \( S_{RR} \) need only \( T_f = 2\tau_f = 2/W \) to establish the feedback process. Note that increasing the feedback duration boosts the outage probability of the channel, as understood from Eqn. (3) where \( T \) is explained as follows. In the ordered acceptance case with the relevant stability [4].

### III. Problem Formulation and Optimization Problem

#### A. Average Arrival and Service Rates

1) Ordered Acceptance: A packet at the head of \( Q_s \) can be served in either one of the following events: the channel between the source and the destination is not in outage; or it is in outage but one of the relays decodes correctly the source packet. Note that for relay \( k \) to admit the packet, all the relays assigned a higher priority in accepting the packet will fail to receive (decode) the packet correctly due to channel outage. Hereinafter, we adopt the notation \( \lambda = 1 - \lambda' \). The average service rate at queue \( Q_s \) is given by

\[
\mu_s = 1 - P_{s,d} \prod_{k=1}^{N} P_{s,k}. \tag{4}
\]

The probability that \( Q_s \) is empty is given by

\[
\pi = 1 - \frac{\lambda_s}{\mu_s} \tag{5}
\]

with \( \lambda_s > \lambda_s \) as a necessary and sufficient condition for queue stability [4].

Let \( \lambda_k \) be the arrival rate at the queue \( Q_k \). Note that for an arrival event to occur at \( Q_k \), the source queue should be nonempty and all higher order queues should fail to decode. The expression for the average arrival rate follows directly from (4) and is given by

\[
\lambda_k = \pi P_{s,d} \sum_{m_n \in \mathcal{M}} P_{s,k} \prod_{m_n \in \mathcal{M}} P_{s,m_n}. \tag{6}
\]

For a relay to transmit, the source queue should be empty. The average service rate, \( \mu_k \), for \( Q_k \) is given by

\[
\mu_k = \omega_k \pi P_{k,d}. \tag{7}
\]

2) Random Assignment Decoding and Round Robin: In \( S_{RD} \), the \( k \)th relay is scheduled to decode the transmitted packet with probability \( \beta_k \). Hence, the average service rate of the source and the average arrival rate to the relaying queue \( Q_k \) are given by

\[
\mu_k = P_{s,d} + P_{s,d} \sum_{k=1}^{N} T_{s,k} \beta_k, \quad \lambda_k = P_{s,d} \sum_{k=1}^{N} T_{s,k} \beta_k. \tag{8}
\]

The average service rate of the relaying queue \( Q_k \) is the same as in the ordered acceptance case.

In \( S_{RR} \), the expressions are similar to \( S_{RD} \) with the substitution \( \beta_k = 1/N \).

#### B. Optimization Problem

We aim at maximizing the source average service rate given \( \lambda_s \) and \( N \) subject to the stability of all queues. Under the ordered acceptance strategy, \( S_{OD} \), the maximum source mean service rate can be obtained by solving the following problem:

\[
\max_{\omega,\rho} \mu_s \quad \text{s.t.} \quad \lambda_s < \mu_s, \quad \lambda_s < \mu_k, \forall k, \quad 0 \leq \omega, \rho, \quad \|\omega\|_1, \|\rho\|_1 = 1, \tag{9}
\]

where the constraints \( \mu_s > \lambda_s \) and \( \mu_k > \lambda_k \) represent the stability of the source queue and relay \( k \) queue, respectively.

The notation \( w \leq y = (y_1, y_2, \ldots, y_m, \ldots) \) is an element wise condition on the vector \( y \) implying that \( w \leq y_m, \forall m, \) and \( \|y\|_1 \) is the \( \ell_1 \)-norm of the vector \( y \) defined as \( \|y\|_1 = \sum_m |y_m| \). We note that under the ordered acceptance the total number of optimization parameters is \( N! + N \). The optimization problem is a linear program, which can be solved efficiently and reliably using any convex optimization toolbox (e.g., CVX [5]). Note that in case of \( S_{OD}, \mu_s \) is independent of \( \omega \) and \( \rho \); hence, the problem is a convex feasibility problem where we search for the solution that satisfies the linear constraints.

The optimization problem is solved at the source which then supplies the required information to the relays. The optimal parameters are functions of many long-term average system parameters such as the channels’ average outage probabilities, source mean arrival rate, and number of relays. Thus, the optimal parameters will need to be recalculated only when these long-term averages change, which happens over relatively long periods of time. Once the optimal parameters are obtained, the source generates a long sequence of decoding orders and a time slot accessing distribution to be supplied to the relay stations during the entire operational time of the system until the next long-term parameters recalculation is due.

For random decoding, the optimal mean service rate of the source can be obtained by solving a similar optimization problem as in the ordered decoding case with the relevant optimization variables, e.g., \( \beta \) and \( \omega \). The total number of optimization parameters in the case of random decoding is \( 2N \). The optimization problem is a linear program for a given \( \omega \).

Setting all elements of \( \beta \) to \( 1/N \), the maximum source mean service rate can be obtained for \( S_{RR} \). The optimization problem in this case is a linear (convex) feasibility problem.

**Proposition 1.** For a network with a source, \( N \) relays and a destination, and a negligible feedback duration per relay, i.e., \( \tau_f = 1/W \approx \kappa \ll T \) and \( \kappa \rightarrow 0 \), which is the case when \( W \) and \( T \) are large, the ordered acceptance strategy outperforms the random decoding and round robin decoding strategies.

**Proof.** Since the channel outage probabilities are the same for all strategies when \( \tau_f \) is negligible, we have the following:

\[
\mu_s^{(OD)} = P_{s,d} + P_{s,d} \sum_{k=1}^{N} T_{s,k} \beta_k \leq P_{s,d} + P_{s,d} P_{\text{max}} \leq 1 - P_{s,d} P_{\text{max}} \leq 1 - P_{s,d} P_{s,1} P_{s,2} \ldots P_{s,N} = \mu_s^{(OD)} \tag{10}
\]
where $P_{\text{max}} = \max \{ P_{S1}, P_{S2}, \ldots, P_{SN} \}$ and the superscript represents the used decoding strategy. The average arrival rates to the relaying queues under $S_{\text{RD}}$ are given by

$$\lambda_k^{(\text{RD})} = P_{s,d} \frac{\lambda_s}{\mu_k} \geq P_{s,d} \frac{\lambda_s}{\mu_k^{(\text{OD})}}. \quad (11)$$

We can rewrite the arrival rate to the $k$th queue under $S_{\text{OD}}$ as follows:

$$\lambda_k^{(\text{OD})} = \frac{\lambda_s}{\mu_k^{(\text{OD})}} P_{s,d} P_{a,k} \left( \sum_{m_n \in M_{1,k}} \rho_n + \sum_{m_n \in M, n \in M_{1,k}} \rho_n \prod_{v=1}^{N} P_{a,v} \right). \quad (12)$$

The term $\sum_{m_n \in M_{1,k}} \rho_n$ represents the probability of assigning relay $k$ to the first acceptance order where $M_{1,k}$ is the set of permutations in which relay $k$ is assigned to the first acceptance order. Letting $\sum_{m_n \in M_{1,k}} \rho_n = g_k$, we get

$$\lambda_k^{(\text{OD})} = \frac{\lambda_s}{\mu_k^{(\text{OD})}} P_{s,d} P_{a,k} \left( g_k + \sum_{m_n \in M} \rho_n \prod_{v=1}^{N} P_{a,v} \right). \quad (13)$$

Setting $g_k$ to $\beta_k$ and the other ordering assignment probabilities to zero, we get

$$\lambda_k^{(\text{OD})} = \frac{\lambda_s}{\mu_k^{(\text{OD})}} P_{s,d} P_{a,k} \beta_k. \quad (14)$$

From (11) and (14), we note that the lower bound on the arrival rate to the $k$th relaying queue under random decoding is achieved by ordered acceptance decoding via setting the optimization parameters to special values. From the equations for the mean service and arrival rates, the service rates of all queues under $S_{\text{OD}}$ are always greater than those of $S_{\text{RD}}$, whereas the mean arrival rates of the queues under $S_{\text{OD}}$ are always smaller than those of $S_{\text{RD}}$. Thus, the performance of $S_{\text{OD}}$ is always better than that of $S_{\text{RD}}$. Hence, the mean service rates of the source and the relays as well as the end-to-end queueing delay of source packets under $S_{\text{OD}}$ are better than under $S_{\text{RD}}$. Since $S_{\text{RR}}$ is a special case of $S_{\text{RD}}$, $S_{\text{RD}}$ is always better than $S_{\text{RR}}$ and hence $S_{\text{OD}}$.

It is worth noting that under the stability conditions, the system throughput is directly equal to the arrival rate to the source while the service rate represents the maximum achievable throughput.

IV. NUMERICAL RESULTS AND CONCLUSIONS

In this section, we provide some numerical results for the optimization problems considered in this paper. The parameters used to generate the results are depicted in Table I. Fig. 2 reveals the impact of increasing the number of relays on the optimal mean service rate of the source queue. It is noted that for low $\tau_f$, ordered acceptance decoding outperforms random decoding and round robin decoding for $N \in \{2, 3, 4, 5\}$, and the relation between $\mu_s$ and $N$ is monotonically nondecreasing for all strategies. The performance of the max-max relay selection strategy [1] is plotted for comparison purposes. To implement this selection strategy, each relay sends a single-bit feedback sequentially to the source to employ channel estimation. Thus, the total time needed to estimate all channels and select the maximum one is $(N+1)\tau_f = (N+1)/W$ which exactly equals the time needed for executing the feedback process for the ordered acceptance decoding strategy, but at the expense of channel estimation errors in the case of max-max selection. Moreover, max-max relay selection, the control unit needs to know all channel gains between the source and the relays and between the relays and the destination. In addition, the selected relay must be informed by the control unit. Hence, a large portion of the time slot is consumed for signaling relative to our proposed scheme. In this letter, we assume that the channel estimation for the max-max selection is perfect and that the control unit can know the channels of all nodes. Thus, we compare with an upper bound for the max-max relay selection strategy. As shown in Fig. 2, our proposed ordered acceptance decoding achieves this upper bound.

REFERENCES


