On the Outage Probability and Effective Capacity of Multiple Decode-and-Forward Relay System

Yulin Hu, James Gross
UMIC Research Center
RWTH Aachen University, Germany
Email: {hu|gross}@umic.rwth-aachen.de

***Abstract*—This paper[1] investigates the quality-of-service performance of multiple decode-and-forward relay system. In particular, we study the outage probability and effective capacity performance of the system based on average channel gain. First, a closed-form expression of outage probability is derived which is more concise than the previous work. Second, we introduce four retransmission strategies and derive the effective capacities of the multi-relay system under these strategies. By means of simulation, we provide the numerical verification of our formula derivation. The simulation results illustrate perfect agreement for our theoretical analysis. In addition, we evaluate the system queue behavior under the strategies. The simulation shows that relay can significantly improve the system queue behavior and that the combined signal processing only pays off under relay-driven retransmission strategies.**

***Index Terms*—decode-and-forward, multiple relays, outage probability, effective capacity, QoS**

I. Introduction

Future wireless communication networks are expected to provide reliable service to high traffic and real-time applications, such as video conference or VoIP. In order to provide quality of service (QoS) guarantees the system is typically required to meet stringent deadlines while requiring high reliability. Therefor, delay and outage probability (OP) are two key metrics in a QoS-constrained system.

In [1], Wu and Negi propose the effective service capacity model to analyze systems performance under statistical QoS constraint pair, which includes delay and delay-based OP. The paper also provides the concept of maximum sustainable rate that can be supported by a given transmission process while satisfying the QoS constraints. Furthermore the effective service capacity formulations are derived for different channel models in [2], [3]. Effective service capacity and relative maximum sustainable rate are actually widely-accepted measures for evaluating the QoS-constrained performance.

On the other hand, previous works have shown that the relay association is a promising way to improve the performance of the system [4], [5]. Specifically, deploying multiple relays can significantly improve the transmission capacity and QoS [6], [7]. Based on the relay system [8] derives the expression of maximum sustainable rate for single decode-and-forward (DF) relay system with perfect channel state information (CSI). [9] maximizes the effective service capacity for multi-relay system by time-slot allocation based on one approximate expression of effective service capacity. With perfect CSI feedback, [10] obtains the expression of effective service capacity for single and multiple relays transmissions. However, the theoretical expression for the effective service capacity or maximum sustainable rate for the opportunistic multi-DF-relay transmission (which only has the average CSI) is still an open problem. Without perfect CSI the source does not know how many and which relays will decode the transmitted data correctly. It is a challenge to obtain a simple closed-form expression on the system performance when the number of forwarding relays and the received signal-to-noise ratio (SNR) at the destination are unpredictable.

This paper is motivated by studying the case that without perfect CSI how the powerful but unpredictable forwarding relays affect the QoS performance of the system. This paper takes both OP and maximum sustainable rate into account and theoretically derives their corresponding closed-form expression for a multi-DF-relay system. Our contributions are:

1) We derive the close-form expression for the OP of multi-DF-relay two-hop transmission. Different from [11], we prove that the value of the received SNR at the destination obeys a gamma distribution, and this directly makes the expression of OP much more concise than [11].
2) According to different protocol complexities and memory requirements, we introduce four strategies and analyze their maximum sustainable rate.
3) By means of simulation, we show the numerical results perfectly match our theoretical analysis. In addition, we show that relays can significantly improve the system queue behaviors. Furthermore the combined signal processing at the destination is only effective under relay-driven retransmission strategies.

The rest of the paper is organized as follows. Section II introduces the overall system model, discusses the problem we are interested in, and briefly summarizes the related work. Section III is the core part of this paper. In this section we first derive the closed-form OP for the two-hop multi-DF-relay transmission, then we theoretically analyze the maximum sustainable rate for the four different retransmission strategies. In section IV, this paper provides numerical simulation results. Finally, we conclude the paper in Section V.

[1]This work was supported by China Scholarship Council and the DFG Cluster of Excellence on Ultra High-Speed Mobile Information and Communication (UMIC), German Research Foundation grant DFG EXC 89.

978-1-4673-4404-3/12/$31.00 ©2012 IEEE

II. PRELIMINARIES

This section first introduces the overall system model and then presents the problem. Contributing to this problem requires us to utilize the framework of the effective service capacity, which we summarize briefly afterwards.

A. System Model

We consider a simple relaying scenario with multiple transceivers as schematically shown in Figure 1. Among the transceivers there is one which is the source of a data flow. The flow needs to be conveyed to another transceiver which is the destination. The remaining J transceivers operate as relays. Relaying in this paper refers exclusively to DF type of relaying. The entire system operates in a slotted fashion where time is divided into frames of length T_{f}. At the source a constant data flow originates and needs to be transmitted to the destination. The flow arrival rate equals r bits per frame duration T_{f}. The transmission of this data is subject to QoS requirements $\{d, \mathcal{P}_d\}$ as observed by the destination where d stands for a maximum tolerable delay and $\mathcal{P}_d$ denotes the delay violation probability, i.e. the probability that the delay target is not met. Data that can not be transmitted immediately at the source is put into a first in first out queue (of infinite size). Finally, we denote the cumulative arrival process to the link layer at the transmitter up to frame i by $A_i = \sum_{1}^{i} a_i = i \cdot r$. During each frame any transceiver conveys N symbols. Also,

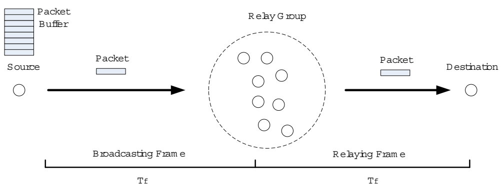

Fig. 1. Example of the considered multiple relay system scenario.

all transceivers apply a (fixed) transmit power of P_{tx} [Watt] for data transmission. We do not assume the presence of any external source of interference. The relaying system operates overall in the following way. Initially, the source takes a fixed data amount out of the queue and forwards this data to the relays. We refer to this initial phase (of duration T_{f}) as broadcasting phase. Next, all relays that decoded the initial transmission successfully forward the packet simultaneously to the destination. We refer to this phase as relaying phase. Note that its duration is again T_{f}. After that the destination informs the source and/or the relays about the decoding success. If the packet was not received successfully, a retransmission strategy is applied until the packet finally reaches the destination. The main focus of our paper is the investigation of the different retransmission strategies on the queuing performance of the system as discussed later in this section.

The successful reception of the packet at either the destination or the relays depends on the instantaneous channel quality between the respective transmitter and receiver. Denote the instantaneous channel gain during the broadcasting phase of frame i from the source to relay j by $h_{i,\mathrm{s},j}^2$. Then the channel quality is given by the signal-to-noise-ratio (SNR) defined as

$$\gamma_{i,\mathrm{s},j} = P_{\mathrm{tx}} \cdot h_{i,\mathrm{s},j}^2 / \sigma_0^2 \tag{1}$$

where σ_0^2 denotes the noise power. Correspondingly, the channel gain and SNR from relay j to the destination during phase i are given by $h_{i,j,\mathrm{d}}^2$ and $\gamma_{i,j,\mathrm{d}} = P_{\mathrm{tx}} \cdot h_{i,j,\mathrm{d}}^2 / \sigma_0^2$. Given an instantaneous SNR γ during one frame (with N symbols) at most $N \cdot \log_2(1+\gamma)$ bits can be conveyed correctly. We assume the channel states to vary randomly due to fading while the transmitters (either source or relays) do not have instantaneous information about the current channel states. Hence, a currently transmitted packet of size ρ is successfully received if the SNR of the link is above the threshold $\gamma^* = 2^{\rho/N} - 1$. For the fading process we assume a block-fading Rayleigh-distributed type of model. Hence, the channel gains are exponentially distributed while the fading states are independent from frame to frame. Due to spatial diversity they are also independent among different transmitters.

Due to the fading, a varying amount of relays will decode the packet successfully during the broadcasting phase. We call this set the forwarding relay set (FRS) and denote the relays in this set during frame i by Ω_i. As the successive relaying phase enables all relays in set Ω_i to forward the packet to the destination, the SNR at the destination is simply the sum of the SNR of the individual links. Therefore, we have $\gamma_{i,\mathrm{d}} = \sum_{\forall j \in \Omega_i} \gamma_{i,j,\mathrm{d}}$. Hence, the packet reception at the destination during the relaying phase depends on this joint SNR of all forwarding links.

From the above system operation we yield a service process in the following manner. Assume that at frame $i-1$ the system successfully conveyed a packet. If the backlog is bigger than ρ the source will form the next packet and attempt to transmit it to the destination via the relays. The transmission of this packet will take a random amount of frames. Denote this random number by τ. Hence, the service – i.e. the amount of bits that effectively leave the queue – at time i can be modeled by the ratio ρ/τ. In fact, the service which the queue receives will be ρ/τ for the next τ time units. Given this service process s_i, we finally denote the cumulative service process by $S_i = \sum_{n=0}^{i} s_n$.

B. Problem statement

In this work we are interested in the queuing performance impact that different retransmission schemes have in combination with different signal processing approaches. Let us first discuss the different schemes considered in detail. In all cases, initially a packet is taken out of the queue and a transmission attempt is made via the relays as discussed above. If the packet is not received successfully, at least four different retransmission variants can be implemented with very different trade-offs regarding protocol complexity, memory requirement and feedback overhead. To ensure reliable reception of the packet at the destination, perfect automatic repeat request feedback is assumed in our system, which means the feedback can be received reliably and immediately.

- **Source-driven retransmission with independent signal processing** The first scheme considered assumes that a negative acknowledgement is transmitted back to the source and the source then initiates a retransmission during a newly upcoming frame. In addition, in this scheme we assume that the destination does not exploit combined signal processing of the initial transmission attempt and the successive ones.
- **Source-driven retransmission with combined signal processing** In the second variant considered, the retransmission is again initiated by the source (upon receiving a negative acknowledgement from the destination). However, during the reception of the retransmitted packet the destination applies combined signal processing. This means that the originally received set of digital samples will be used for decoding the packet in combination with the set of digital samples of the retransmission. In order to model this, denote the frame times of the first and second relaying phase by i and $i + \Delta$. Correspondingly, the FRSs of the two phases are Ω_i and $\Omega_{i+\Delta}$. Then, our model of combined signal processing simply increases the SNR during the retransmission phase by assuming $\gamma_{i+\Delta,\mathrm{d}} = \sum_{\forall j \in \Omega_i} \gamma_{i,j,\mathrm{d}} + \sum_{\forall j \in \Omega_{i+\Delta}} \gamma_{i+\Delta,j,\mathrm{d}}$. Hence, combined signal processing always leads to an equal or better reception probability at the destination.
- **Relay-driven retransmission with combined signal processing and fixed forwarding set** The third considered variant assumes that the destination only informs the relays of the packet loss and the relays immediately attempt a retransmission. This only applies to the relays in the forwarding set of the original relaying phase. All these relays will retransmit the packet until it is successfully received. The destination applies again combined signal processing as discussed above.
- **Relay-driven retransmission with combined signal processing and dynamic forwarding set** The last variant assumes that during the initial relaying phase those relays that are not contained in the FRS Ω can nevertheless try to decode the packet again. Hence, if the destination does not decode the packet successfully, the joint retransmission of the relays can be performed now by a bigger set of relays. In fact, we assume that for all retransmission attempts in this case all relays will participate as any relay that did not receive the packet successfully during the broadcasting phase will receive it successfully during the first relaying phase. Again, the destination applies combined signal processing.

The above variants differ regarding to their performance, protocol complexity and memory requirements. Source-driven retransmissions do not require the relays to store the relayed packet which lowers the amount of state (and memory) that has to be maintained at the relays. However, they are not flexible that one single retransmission cost is fixed as two frames. Combined signal processing obviously puts a larger burden on the destination as more memory and more processing power is required. Finally, relay-driven retransmissions can be performed more flexible at the price of storing the packets at the relays upon a positive acknowledgement from the destination which requires more memory but also a more complex protocol implementation to realize it.

Our main focus in this work is to investigate the QoS-constrained performance of these systems. In fact, we will focus in the following on the implications of the different retransmission variants on the queuing process at the source. How much more delay is caused for a fixed rate arrival flow if we switch from relay-driven retransmissions to source-driven retransmissions? How much more rate can we support if we implemented in the system combined signal processing instead of independent signal processing? Our approach is to estimate the queuing behavior by the framework of the effective service capacity, as summarized in the next section. We will thus derive the effective service capacity of every variant and make numerical statements afterwards.

C. *Effective Service Capacity*

The mathematical framework of the effective service capacity allows to approximate the distribution of the steady-state queue length of a stable queuing system. It is therefore a tool for analysis of arbitrary service processes in a queuing system. The framework was originally applied to characterize the queue length for arbitrary arrival processes (source flows) which are served by a constant rate queuing system. In this context, deriving the so called *effective bandwidth* of the arrival process allows to bound the queue length distribution [12]. Interestingly, this analysis technique can also be turned 'up side down' such that the effective service capacity of a random service process has to be derived in order to bound the queue length distribution assuming constant arrivals. In the following we give a brief introduction to this analytical framework. The starting point for the analysis is Reich's equation which states that for the considered queuing system the queue length at time i is given by:

$$Q_i = \max_{0 \leq k \leq i} \left((i-k) \cdot r - (S_i - S_k) \right) . \qquad (2)$$

Let us consider that the arrival and service process are stationary. Furthermore, assume that the queue is stable as the average service rate is larger than the average arrival rate. Hence, the random queue length Q_i at time i converges to the steady-state random queue length Q. We are interested in characterizing the long-term statistics $\Pr.\{Q\}$ of the queue length. The framework of effective service capacity gives us the following upper bound:

$$\Pr.\{Q > x\} \leq K \cdot e^{-\theta^* \cdot x} , \qquad (3)$$

where K is the probability that the queue is non-empty and θ^* is the so called quality-of-service exponent. Due to several mathematical derivation steps [13], for a constant bit rate source with r bits per time unit arrival rate, the exponent θ^* has to fulfill the following constraint:

$$r < \Lambda\left(-\theta^*\right)/\theta^* . \qquad (4)$$

$\Lambda(\theta)$ is called the log-moment generating function of the increments of the cumulative service process S_i defined as (assuming the increments to be stationary as well):

$$\Lambda(\theta) = \lim_{i \to \infty} \frac{1}{i} \log \mathrm{E}\left[e^{\theta \cdot (S_i - S_0)}\right] . \quad (5)$$

Finally, the ratio $\Lambda(-\theta)/\theta$ is called the effective service capacity, as the exponential decay of the distribution in Equation (3) is only witnessed if the ratio $\Lambda(-\theta)/\theta$ is bigger than the constant arrival rate r of the source for some θ^*.

So far we have only considered the random queue length. Denote by D_i the random queuing delay of the head-of-line bit during frame i. This random variable converges in the long-run to the random steady state queuing delay D of the head-of-line bit. As the arrival process has a fixed rate, the steady-state queue length statistics are related to the steady-state delay statistics of the head-of-line bit. Hence, a queue length of $Q = q$ is associated with a current delay of the head-of-line bit of $D = q/r$. This yields the following approximation for the steady-state delay distribution which is based on Equation 3:

$$\Pr.\{D > d\} \leq K \cdot e^{-\theta^* \cdot r \cdot d} . \quad (6)$$

A considerable challenge in determining the effective service capacity is the characterization of the log-moment generating function. If the service process s_i can be assumed to be i.i.d., a convenient simplification is to obtain the log-moment generating function via the law of the large numbers [14]. Hence, the effective service capacity can be obtained by:

$$\frac{\Lambda(-\theta)}{\theta} = \lim_{i \to \infty} \frac{1}{i \cdot \theta} \log \mathrm{E}\left[e^{-\theta \cdot s_i}\right] = \mathrm{E}[s_i] - \frac{\theta}{2} \mathrm{Var}[s_i] . \quad (7)$$

It is therefore sufficient to determine the average and the variance of the instantaneous service process s_i.

The above analysis allows to determine a bound on the maximum OP if the (constant) arrival rate is given. In contrast, we can also fix the delay and outage target and derive the maximum sustainable rate r^* that can be supported by the random service process. From Equation (6) we obtain the following (upper bounding K by 1):

$$\frac{-\ln(\mathcal{P}_d) + \ln(K)}{d} \geq r \cdot \theta \quad \Leftrightarrow \quad r^* \cdot \theta \approx -\frac{\ln(\mathcal{P}_d)}{d} , \quad (8)$$

where the approximation results from the fact that upper bounding K by one can underestimate the maximum sustainable rate if especially the delay target is quite low. Next, from Equation (4) and (7) we obtain in general for θ:

$$\theta = 2\left(\mathrm{E}[s_i] - r\right)/\mathrm{Var}[s_i] . \quad (9)$$

We use this expression and substitute it in Equation (8). We then obtain the following relationship for the maximum sustainable rate r^* which has been first proposed by [14]:

$$r^* \approx \frac{\mathrm{E}[s_i]}{2} + \frac{1}{2}\sqrt{\left(\mathrm{E}[s_i]\right)^2 + \frac{2 \cdot \ln(\mathcal{P}_d)}{d} \cdot \mathrm{Var}[s_i]} . \quad (10)$$

In the following, we use this equation to determine the maximum sustainable rate. So it becomes the major difficulty that obtaining the mean and variance of the instantaneous service process s_i.

III. Derivation of the Effective Service Capacities

In this section, we will derive the corresponding effective service capacities. As discussed in Section II, it is sufficient to obtain analytical expressions for the mean and variance of the service process s_i. Let us initially discuss the service process and ways of modeling it. For this, we assume that a packet just has been successfully transmitted and that during the upcoming time frame i we take the next packet of size ρ out of the queue for transmission. As we are considering more complex systems with different retransmission schemes, the service of this packet depends crucially on the (random) number of time frames τ that it takes to successfully transmit this packet. For the $\tau - 1$ frames prior to the successful transmission, in the real system the amount of service that the queue receives is actually 0 while for frame τ then the service is ρ. In order to make the system analytically tractable, we instead model the service process of the above example such that during all τ frames the service equals ρ/τ. This causes two discrepancies from the real queuing behavior. For one, bits of the transmitted packet arrive at the destination earlier. However, we can adjust for this discrepancy by introducing a palletization effect to the framework of Section II-C modifying Equation 3 by substituting x with $x + \rho$ and considering this in all subsequent derivations. However, a second discrepancy is that during the τ time frame in our model the queue size has increments r as well as decrements ρ/τ whereas in the real system during the $\tau - 1$ frames the system only has increments r. That leads to maximum sustainable rate predicted analytically (slightly) bigger than the real system. We show in the numerical section that this discrepancy is actually low and so we can apply our simplifying model.

The advantage of our proposed model is that for the mean $\mathrm{E}[s_i]$ and the variance $\mathrm{Var}[s_i]$ it is sufficient to clarify the distribution of τ under the different system operations as presented above. Once that we have the distributions, we can easily obtain the mean and the variance. This is what we present in the following. We first analyze the outage probability of the broadcasting frame and the relaying frame. We then obtain the probability of retransmission times based on these outage probabilities, and further analyze the mean and variance of service process according to the model above.

A. Outage Probability of Initial Transmission

In this subsection, we study the OP of two-hop transmission (without retransmission) for a new packet from frame i to frame $i + 1$. As introduced in Section II, there is an outage when the received SNR is smaller than threshold γ^*.

Based on the Rayleigh block-fading channel, the OP of every single broadcasting link in frame i can be given by:

$$P_{1,i,j} = 1 - \exp\left(-\gamma^* \sigma_0^2 / 2\bar{h}_{i,\mathrm{s,j}}^2 P_{\mathrm{tx}}\right) , \quad (11)$$

where $\bar{h}_{i,\mathrm{s,j}}^2$ is the average channel gain of the link from source to relay j. Thus, for frame i the probability of the FRS Ω_{i+1} is given by:

$$\Pr(\Omega_i) = \prod_{j \notin \Omega_{i+1}} P_{1,i,j} \prod_{j \in \Omega_{i+1}} \left(1 - P_{1,i,j}\right) . \quad (12)$$

From this we immediately obtain the OP of the relaying frame:

$$P_{2,i+1} = 1 - \exp\left(-\gamma^* \sigma_0^2 \Big/ 2 \sum_{j\in\Omega_{i+1}} \bar{h}_{i,\mathrm{s,j}}^2 P_{\mathrm{tx}}\right) . \quad (13)$$

However, this fairly general case is hard to analyze further since the Ω_{i+1} is unpredictable.

We therefore consider in the following a simplification in the topology where we assume that the distance among relays is significantly small compared to the distance to the source. Likewise, we also assume this to be the case for the distance to the destination from the relays. This simplification makes all the relays have same value of OP in broadcasting frame which means $P_{1,i,j} = P_{1,i}, \forall j$. This is very good, as it turns the probability distribution of the size of the forwarding relay set Ω into a binomial distribution. Furthermore, the sum signal transmitted from all the relays to the destination becomes Gamma distributed. Notice that slow power control at the relays can generate an identical situation for significantly different channel gains from the relays to the destination. Hence, we denote the average channel gains of broadcasting and relaying links by $\bar{h}_1^2$ and $\bar{h}_2^2$. Therefore, denoted by $J_{\mathrm{F,i+1}}$ the FRN in frame $i+1$ becomes binomial with base probability $P_{1,i}$. Denote the probability that $J_{\mathrm{F,i+1}} = n$ by $\mathrm{P}_B\left(n; J_{\mathrm{F,i+1}}, P_{1,i}\right) = \Pr\left(J_{\mathrm{F},i+1} = n\right)$.

We move on to the relaying phase during frame $i+1$. Due to either similar distances between the relays and destination (or due to power control), the sum of the fading signals from all the relays in the forward set becomes a Gamma distributed random variable [15]:

$$\gamma_{i+1,\mathrm{d}} = \frac{P_{\mathrm{tx}}}{\sigma_0^2} \sum_{i=1}^{J_{\mathrm{F,i+1}}} \left|h_{i+1,j,\mathrm{d}}\right|^2 \sim \Gamma(J_{\mathrm{F},i+1}, \frac{2P_{\mathrm{tx}}\bar{h}_2}{\sigma_0^2}) . \quad (14)$$

Based on this characteristic of joint SNR at the destination, the OP of relaying phase during frame $i+1$ is obtained based on the Gamma cumulative distribution function as:

$$P_{2,\mathrm{i}+1} = \Pr\left\{\gamma_{i+1,\mathrm{d}} \le \gamma^*\right\} = \mathrm{F}\left(\gamma^*; J_{\mathrm{F},i+1}, \beta\right) , \quad (15)$$

where we denote the average of the signal strength $\beta = 2P_{\mathrm{tx}}\bar{h}_2/\sigma_0^2$. If $J_{\mathrm{F,i+1}}$ equals 0, it means all relays fail in decoding and the OP equals 1. Otherwise, $J_{\mathrm{F,i+1}}$ is a positive integer. According to [15], $P_{2,\mathrm{i}+1}$ can be expressed as:

$$P_{2,\mathrm{i}+1} = \begin{cases} 1 - \sum_{j=0}^{J_{\mathrm{F},i+1}-1} \frac{1}{j!}\left(\frac{\gamma^*}{\beta}\right)^j e^{-\frac{\gamma^*}{\beta}}; J_{\mathrm{F},i+1} > 0 \\ 1; \quad J_{\mathrm{F,i+1}} = 0 \end{cases} . \quad (16)$$

Therefore based on the law of total probability, the mathematical expectation of OP of total two-frame transmission (i and i+1) can be obtained by:

$$\begin{aligned} \Pr_i^{\mathrm{out}} &= \sum_{n=0}^{J} F\left(\gamma^*; n, \beta\right) \mathrm{P}_B\left(n; J, P_{1,i}\right) \\ &= {P_{1,i}}^J + \sum_{j=0}^{J_{\mathrm{F},i+1}-1} \frac{1}{j!}\left(\frac{\gamma^*}{\beta}\right)^j e^{-\frac{\gamma^*}{\beta}} \mathrm{P}_B\left(n; J, P_{1,i}\right). \end{aligned} \quad (17)$$

Hence we have derived the closed-form expression of OP for this multi-DF-relay enhanced two-hop transmission.

We now turn to retransmission schemes. As the analysis of the third strategy (which relays only participate in the retransmission) needs taking the direct link into account, therefore we provide the OP here which is similar to Equation (11) as:

$$P_{0,i} = 1 - \exp\left(-\gamma^* \sigma_0^2 / 2\bar{h}_0^2 P_{\mathrm{tx}}\right) . \quad (18)$$

where $\bar{h}_0^2$ denotes the average gain of the channel from the source to the destination.

B. Distribution of Service

As we have briefly introduced in Section II, one single retransmission of source-driven retransmission strategies costs two frames. For the relay-driven retransmission, one single retransmission (no matter re-broadcasting or re-relaying) only costs one frame. For the system serving a new packet from frame i, respectively we denote number of retransmission times of source-driven retransmission as K_i^{R} and re-broadcasting times, re-relaying times of relay-driven retransmission as K_i^{s}, K_i^{r}. For example, $K_i^s = 0$ means successful transmission without retransmitting of broadcasting phase under relay-driven retransmission strategy and $K_i^{\mathrm{R}} = k$ means former $k-1$ times retransmission fail and only the last success under source-driven retransmission strategy. $k = 0, 1, 2..., k_{\max}$, where $k_{\max}$ is maximum number of retransmitting supported by the system which is closed to real systems. In other words, the system will stop retransmitting the packet when retransmitting times is out of the limit. Note that we are investigating the queuing behavior, so unlike dropping it in real system, the packet out of $k_{\max}$'s limit will still stay in the queue for next round transmission and retransmission.

Now we are going to further study the distribution of K_i^{R} or K_i^{s} and K_i^{r} under the strategies.

1) Independent Signal Processing (ISP): ISP is a source-driven retransmission case that retransmitting should be occurred after the destination failed in decoding or received nothing (no relay forwarding in the second phase). And the destination doesn't combine the signals of source or the previous fail-decoded-transmissions with the new arriving retransmitted signal. From frame i to $i+1+2k$, the total transmission frames number of ISP is $\tau_{\mathrm{ISP}} = 2(k+1)$ where the retransmission time is k. Based on the OP, the probability of $K_i^{\mathrm{R}} = k$ can be given as:

$$\Pr_{ISP}(K_i^{\mathrm{R}} = k) = \left(\Pr_i^{\mathrm{out}}\right)^k \cdot \left(1 - \Pr_i^{\mathrm{out}}\right) . \quad (19)$$

2) Combined Signal Processing (CSP): CSP is also a source-driven retransmission. From i to $i+1+2k$ frame, the FRN from initial transmitting to kth retransmitting are $J_{\mathrm{F},i+1}, J_{\mathrm{F},i+3}, ..., J_{\mathrm{F},i+1+2k}$. Denote $J_{\mathrm{v}}(k) = J_{\mathrm{F},i+1} + J_{\mathrm{F},i+3} + ... + J_{\mathrm{F},i+1+2k}$, after received the kth retransmitted signal, the combined signal for decoding has joint channel gain from $J_{\mathrm{v}}(k)$ relays. We further treat the combined signal as a joint signal form $J_{\mathrm{v}}(k)$ relays in once transmission, so the $J_{\mathrm{v}}(k) \sim B((k+1)J, P_{1,i})$. After k retransmission

$$\Pr_{CSP}(K_i^{\mathrm{R}}=k)=\sum_{J_v(k)=0}^{kJ}\left\{\left[F\left(\gamma^*;J_v(k),\beta\right)-\sum_{n'=0}^{J}\left[F\left(\gamma^*;J_v(k)+n',\beta\right)\cdot \mathrm{P}_B\left(n';J,P_{1,i}\right)\right]\right]\cdot \mathrm{P}_B\left(J_v(k);kJ,P_{1,i}\right)\right\}. \quad (20)$$

$$\Pr_{RTR}\{K_i^{\mathrm{s}}=k_1,K_i^{\mathrm{r}}=k_2\}=\begin{cases}1-P_{0,i}\,;k_1=0;\\ P_{0,i}\cdot\sum_{n=1}^{J}\left\{{P_{1,i}}^{J\cdot k_1}\cdot P_B(n,J,P_{1,i})\cdot[\mathrm{F}(\gamma^*;k_2n,\beta)-\mathrm{F}(\gamma^*;(k_2+1)n,\beta)]\right\};k_1>0;\end{cases}. \quad (21)$$

$$\Pr_{ARR}\{K_i^{\mathrm{s}}=k_1,K_i^{\mathrm{r}}=k_2\}=\sum_{n=1}^{J}\left\{{P_{1,i}}^{J\cdot k_1}\cdot P_B(n,J,P_{1,i})\cdot[\mathrm{F}(\gamma^*;Max\{(k_2-1)J+n,0\},\beta)-\mathrm{F}(\gamma^*;k_2j+n,\beta)]\right\}. \quad (22)$$

attempts, the total transmission frames number can be given by $\tau_{\text{CSP}}=2(k+1)$. And the probability $K_i^{\mathrm{R}}=k$ is given by Equation (20).

3) Retransmitting Relay (RTR): RTR is a relay-driven retransmission strategy. It requires all the relays to try to decode the packet from the source. However, the relays which have decoded the packet not participate the first transmitting but only retransmit the packet to the destination if the destination fail to decode the packet by the the direct link. In order to simplify the combined signal expression, we only consider combining the signals from relays after direct link transmission fail. Rebroadcasting happens only if all the relay fail to decode the packet, otherwise the FRS will retransmit till the destination successfully decode it. The prime transmission makes $J_{\mathrm{F},i+1}$ relays for forwarding in frame $i+1$. If $J_{\mathrm{F},i+1}>0$, after k times retransmission, the combined signal at the destination has joint channel gain from $J_{\mathrm{F},i+1}\cdot k$ relays. However, if $J_{F,(i+1):(i+k_1)}=0$ and $J_{F,i+1+k_1}>0$, it means there is no relays or destination decoded out the packet till k_1th rebroadcasting. So, after k_2 retransmission attempts of relays, the combined signal at the destination has joint channel gain from $J_{\mathrm{F},i+1+k_1}\cdot k$ relays with $\tau_{\text{RTR}}=k_1+k_2+1$ when $K_i^{\mathrm{s}}=k_1$ and $K_i^{\mathrm{r}}=k_2$.

In our system, relays are much close to the source than destination, so the probability that all the relays fail while destination successfully decode the packet is negligible. In fact, in order to be fair, all of the four strategies do not take direct link's signal into account for signal combining at the destination. Hence the probability that $J_{F,(i+1):(i+k_1)}=0$ while $J_{F,i+1+k_1}=n>0$ can be given as:

$$Pr\{K_i^{\mathrm{s}}=k_1,J_{F,i+1+k_1}=n\}={P_{1,i}}^{J\cdot(k_1-1)}\cdot P_B(n,J,P_{1,i})\,, \quad (23)$$

And the conditional probability of $K_i^{\mathrm{r}}=k_2$ when $K_i^{\mathrm{s}}=k_1$ and $J_{F,i+1+k_1}=n>0$:

$$\Pr\{K_i^{\mathrm{r}}=k_2\,|N_F=n\}=\mathrm{F}\left(\gamma^*;n,\beta\right)^{k_2-1}\left(1-\mathrm{F}(\gamma^*;n,\beta)\right). \quad (24)$$

So the joint probability of $K_i^{\mathrm{s}}=k_1,K_i^{\mathrm{r}}=k_2$ can be given as Equation (21).

4) All Relays Retransmitting (ARR): All relays fail to decode the packet in broadcasting phase will receive the packet again during the relaying phase from FRS. Since the relays are very close to each other, all the relays will have the packet after the first relaying frame. So for the re-relaying frame, the FRS Ω contains all the relays. Also, ARR is a relay-driven retransmission strategy, and re-broadcasting happens only if all relays fail to decode the packet at the broadcasting phase. Also denote the first transmission makes $J_{\mathrm{F},i+1}$ forwarding relays for the relaying in frame $i+1$. If $J_{\mathrm{F},i+1}>0$ till k_2th re-relaying success, the combined signal for decoding has joint channel gain from $J_{\mathrm{F},i+1}+k_2\cdot J$ relays, and the system has spent $\tau_{\text{ARR}}=2+k_2$ for this packet. For the case $J_{\mathrm{F},(i+1):(i+k_1)}=0$ only when k_1th rebroadcasting makes the $J_{\mathrm{F},i+k_1+1}=n>0$, after the relays' k_2 times re-relaying, the destination has joint channel gain from $J_{\mathrm{F},i+1+k_1}+k_2\cdot J$ relays. So the total transmission frames number $\tau_{\text{ARR}}=k_1+k_2+2$. And similar to RTR, the joint probability the case $K_i^{\mathrm{s}}=k_1,K_i^{\mathrm{r}}=k_2$ under ARR can be given as Equation (22).

So far, we have studied the distribution of total transmission frames number τ_v with relative probability under the four strategies, where v is the index of ISP, CSP, RTR, ARR. And obviously the data rate per frame of ISP and CSP are satisfied the single-variable discrete distribution. So with the packet size ρ, the corresponding mean and variance can be obtained by:

$$\mathrm{E}\left[s_i\right]=\sum_{l=0}^{+\infty}\left\{\sum_{k=0}^{k_{\max}}\frac{\rho}{\tau_v}\Pr(K_i^{\mathrm{R}}=k){\mathrm{P}_*}^l\right\}, \quad (25)$$

$$\mathrm{Var}\left[s_i\right]=\sum_{l=0}^{+\infty}\left\{\sum_{k=0}^{k_{\max}}(\frac{\rho}{\tau_v}-E_v)^2\Pr(K_i^{\mathrm{R}}=k){\mathrm{P}_*}^l\right\}. \quad (26)$$

where $\mathrm{P}_*=1-\sum_{k=0}^{k_{\max}}\Pr(K_i^{\mathrm{R}}=k)$ means the probability that the packet still fail to transmit after a round of retransmission under limit time of $k_{\max}$.

Similarly, for the RTR and ARR, the corresponding mean an variance of data rate can be given as (27) and (28).

$$\mathrm{E}\left[s_i\right]=\sum_{l=0}^{+\infty}\left\{\sum_{k_1=0}^{k_{\max}}\sum_{k_2=0}^{k_{\max}-k_1}\frac{\rho}{\tau_v}\Pr_v\left\{\begin{matrix}K_i^{\mathrm{s}}=k_1,\\K_i^{\mathrm{r}}=k_2\end{matrix}\right\}{\mathrm{P}_*}^l\right\}, \quad (27)$$

$$\mathrm{Var}\left[s_i\right]=\sum_{l=0}^{+\infty}\left\{\sum_{k_1=0}^{k_{\max}}\sum_{k_2=0}^{k_{\max}-k_1}\left(\frac{\rho}{\tau_v}-E_v\right)^2\Pr_v\left\{\begin{matrix}K_i^{\mathrm{s}}=k_1,\\K_i^{\mathrm{r}}=k_2\end{matrix}\right\}{\mathrm{P}_*}^l\right\}. \quad (28)$$

where $\mathrm{P}_*=1-\sum_{k_1=0}^{k_{\max}}\sum_{k_2=0}^{k_{\max}-k_1}\Pr_v\left\{\begin{matrix}K_i^{\mathrm{s}}=k_1,\\K_i^{\mathrm{r}}=k_2\end{matrix}\right\}$.

Finally, we have obtained the expression on the mean and variance of the service process s_i. Therefore the corresponding maximum sustainable rate can be calculated by (10).

IV. NUMERICAL RESULTS AND DISCUSSION

In this section we evaluate our analytical expressions and system queue behavior numerically. In our simulation, we randomly deploy 9 relays over a circle with radius R, and the distances of broadcasting and relaying links are both set as 200m. The system has 2048 subcarriers, the center frequency is 2GHz, and the frame length $T_{\mathrm{f}} = 10ms$. The transmitting power of source and relay and the power of background noise are 25dBm, 20dBm and -95dBm, respectively. At last, we utilize the COST231 model for calculating path loss.

We obtain two groups of numerical results from simulation. One group's results are the theoretical values which calculated by varying parameters in the formulas of our theoretical analysis. The other group's results we call them simulation values. Simulation values are obtained statistically from the repetitious simulations on the transmissions and retransmissions under real decode-and-forward protocols. For the simulation values of OP, we generate channels randomly and determine whether outage will happen in a frame based on the result of comparison between packet size and the Shannon capacity of the combined channel. At last, we obtain the simulation values of maximum sustainable rate by observing the queue performance (at source) over the numerous frames.

A. *Verification of the expression of outage probability*

As introduced in Section III this paper assumes that the distances among relays are negligible if they are much smaller than the transmission distance of broadcasting and relaying frames. Based on this assumption, this paper derives a clear closed-form expression of system OP. In this subsection, we are going to verify the theoretical expression by means of simulation considering distances difference.

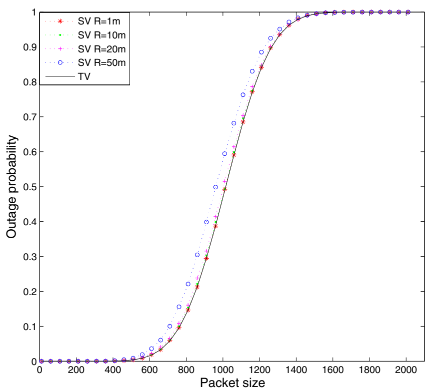

Fig. 2. The comparison on OP between theoretical values and simulation values with different radius of relay-group area.

Figure 2 shows the performance gap between theoretical value and other simulation values with different radius of relay-group area. Firstly, we can see that the bigger the relay-group area is the larger the performance mismatch is. Secondly, the results also proves that the performance mismatch is so tiny even when R = 20m which means the maximum distance among relays could be 1/10 of the distance from source to destination. So, as an important contribution of this paper, our closed-form expression of OP has high credibility when the relay region is small.

B. *Verification of the mean and variance of data rate*

As another important work of this paper, we derive the mean and variance of data rate for four retransmission strategies.

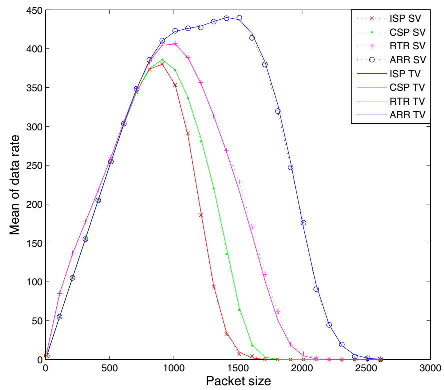

Fig. 3. The comparison between theoretical values and simulation values on mean of data rate while R = 10m and $k_{max} = 4$.

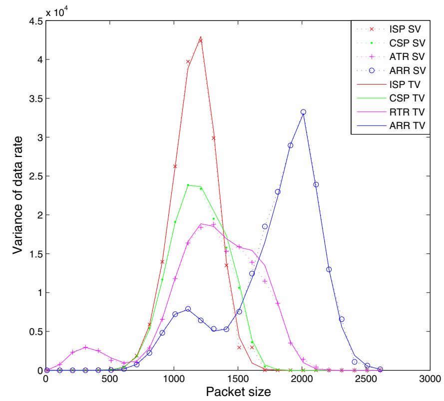

Fig. 4. The comparison between theoretical values and simulation values on variance of data rate while R = 10m and $k_{max} = 4$.

Figure 3 and 4 show the comparison between theoretical values and simulation values on the mean and variance of data rate respectively. It is clear that the simulation results in these figures illustrate perfect agreement for the analytical results. The results shown here is obtained by the simulation from the system with 9 relays. We have also done the simulation for the scenarios with 5 and 3 relays, the corresponding simulation results also match our theoretical analysis.

C. *Queue behavior evaluation*

The numerical curves of the maximum sustainable rate (in bit/T_f) are provided in Figure 5.

First, the simulation results prove our analysis at the beginning of Section III that the maximum sustainable rate predicted analytically is slightly bigger than the real system.

Second, for larger packets the performance sequence of the four strategies in decreasing order is ARR $>$ RTR $>$ CSP $>$ ISP. The CSP is superior to ISP by having more diversity from time domain. And ARR and RTR are better than CSP and ISP by more flexible retransmission mechanism which only need one frame for once retransmitting. ARR outoerforms RTR since it has more relays and higher received SNR for retransmission. However, with small packet size, which means lower OP and less probability of retransmitting, the RTR is the best strategy due to its minimum transmission can be one frame. As introduced in Section III, the initial transmission of RTR is direct-link transmission. So we can see the in Figure 5, RTR has the same performance with direct link transmission at the very beginning (where has lower probability of retransmission). Therefore RTR is a simple tradeoff between direct-link transmission and relay enhanced transmission. As a result, from both theoretical and simulation results, it is always the best choice over the others when the packet size is relative small (smaller than 600).

Third, it is also observed that while increasing the packet size all the strategies' maximum sustainable rate have convex curves. The reasons are as follows. Small packet size cannot support high maximum sustainable rate, so the rate be increased by extending the packet size at the beginning of the curves. However the channel capacity has the limit, as shown in Figure 2 the OP is getting worse by increasing the packet size. So with a very big packet size, the retransmission times becomes considerable, hence directly aggravates the delay. Consider the effective service capacity is defined based on the certain QoS requirement of delay, thus maximum sustainable rate tends to decreasing at the end of the curves.

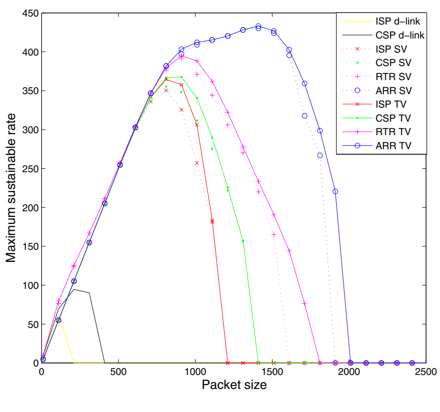

Fig. 5. The comparison between theoretical values and simulation values on maximum sustainable rate with QoS requirement$\{10T_f, 10^{-2}\}$ and R = 10m.

More importantly, Figure 5 shows that relays can significantly improve the system queue behavior in two ways. For one the relay enhanced transmission achieves much higher maximum sustainable rate than direct link transmission. For the other one, with more relays joining in re-relaying, the ARR's performance is dramatically superior over the other three retransmission strategies. In addition, it also can be seen from the curves that the gain from combining signal is not so significant when comparing the CSP with ISP. However, when comparing the RTR and ARR with ISP, this gain becomes considerable since more relays' signal joining in the combined signal set. In other words, both theoretical and simulation curves show that the combined signal processing is only effective under relay-driven retransmission strategies.

V. Conclusion

This paper derives the closed-form expression of outage probability for two-hop multi-DF-relay transmission. In addition, we introduce four retransmission strategies and analyze their maximum sustainable rate. By means of simulation, we show the numerical results perfectly match our theoretical analysis. We show that relays can significantly improve the system queue behaviors and that the combined signal processing at the destination is only effective under relay-driven retransmission strategies.

References

[1] D. Wu and R. Negi "Effective capacity: a wireless link model for support of quality of service, IEEE Trans. Wireless Commun., vol.2,no. 4, pp.630-643. Jul. 2003.

[2] D. Wu and R. Negi, "Effective Capacity Channel Model for Frequency-selective Fading Channels," ACM Wireless Networks, vol. 13, no. 3, pp. 299–310, Jun. 2007.

[3] Q. Wang, D. Wu, P. Fan, "Effective Capacity of a Correlated Rayleigh Fading Channel," Wireless Communications & Mobile Computing, vol. 11, no. 11, pp. 1485–1494, Nov. 2011.

[4] Y. Kim and H. Liu, "Infrastructure Relay Transmission With Cooperative MIMO", IEEE Trans. Vehicular Technology, vol. 57, no. 4, pp. 2180–2188, Jul. 2008.

[5] A. Adinoyi and H. Yanikomeroglu, "Cooperative Relaying in Multi-Antenna Fixed Relay Networks," IEEE Trans. Wireless Commun., vol. 6, no. 2, pp. 533–544, Feb. 2007.

[6] A. Bletsas, H. Shin, and M. Z. Win, "Cooperative communications with outage-optimal opportunistic relaying," IEEE Trans. Wireless Commun., vol. 6, No. 9, pp. 3450–3460, Sep. 2007.

[7] Y. Hu and L. Qiu, "A Novel Multiple Relay Selection Strategy for LTE-Advanced Relay Systems," in Proc. of the IEEE Vehicular Technology Conference (VTC), Budapest, Hungary, May 2011.

[8] D. Qiao, M. C. Gursoy, and S. Velipasalar, "On the Effective Capacity of Two-Hop Communication Systems," in Proc. of the IEEE International Conference on Communications (ICC), Kyoto, Japan, Jun. 2011.

[9] S. Ren. and K. B. Letaief, "Maximizing the Effective Capacity for Wireless Cooperative Relay Networks with QoS Guarantees," IEEE Trans. Commun., vol. 57, no. 7, pp. 2148–2159 Jul. 2009.

[10] L. Musavian, S. Aissa, "Effective Capacity for Interference and Delay Constrained Cognitive Radio Relay Channels," IEEE Trans. Wireless Commun., vol. 9, no. 5, May 2010.

[11] N.C. Beaulieu and J. Hu, "A Closed-Form Expression for the Outage Probability of Decode-and-Forward Relaying in Dissimilar Rayleigh Fading Channels," IEEE Commun. Letters, vol. 10, no. 12, pp. 813–815 Dec. 2006.

[12] C. Chang, "Stability, queue length and delay of deterministic and stochastic queueing networks, " IEEE Trans. Automat. Contr., vol. 39, no. 5, pp. 913–931, 1994.

[13] A. Kumar, D. Manjunath, and J. Kuri, Communication Networking: An Analytical Approach. Morgan Kaufmann, 2004.

[14] B. Soret, M. C. Aguayo-Torres, and J. T. Entrambasaguas, "Capacity with Explicit Delay Guarantees for Generic Sources over Correlated Rayleigh Channel, " IEEE Trans. Wireless Commun., vol. 9, no. 6, pp. 1901–1911, 2010.

[15] A. Papoulis and S.U. Pillai Probability, RandomVariables, and StochasticProcesses (FourthEdition). McGraw-Hill, Boston, 2002.