

Simple Models for the Performance Evaluation of a Class of Two-Hop Relay Protocols *

Ahmad Al Hanbali^{§†}, Arzad A. Kherani[‡], and Philippe Nain[§]

[§] INRIA, Sophia-Antipolis, Cedex 06902, France

[‡] Dept. of Computer Science and Engineering, IIT Delhi, New Delhi, India

[†] Dept. of Computer Science, Université de Nice-Sophia Antipolis, France

Abstract. We evaluate the performance of a class of two-hop relay protocols for mobile ad hoc networks. The interest is on the multicopy two-hop relay (MTR) protocol, where the source may generate multiple copies of a packet and use relay nodes to deliver the packet (or a copy) to its destination, and on the two-hop relay protocol with erasure coding. Performance metrics of interest are the time to deliver a single packet to its destination, the number of copies of the packet at delivery instant, and the total number of copies that the source generates. The packet copies at relay nodes have limited lifetime (time-to-live TTL). Via a Markovian analysis, the three performance metrics of the MTR protocol are obtained in closed-form in the case where the number of the copies in the network is limited. Also, we develop an approximation analysis in the case where the inter-meeting times between nodes are arbitrarily distributed and the TTLs of the copies are constant and all equal. In particular, we show that exponential inter-meeting times yield stochastically smaller delivery delays than hyper-exponential inter-meeting times, and that exponential TTLs yield stochastically larger delivery delays than constant TTLs. Finally, we characterize the delivery delay and the number of transmissions in the two-hop relay protocol with erasure coding and compare this scheme with the multicopy scheme.

Keywords. Mobile ad hoc network; Two-hop relay protocol; Erasure coding; Mobility model; Analytical model; Markovian analysis; Performance evaluation.

1 Introduction

In mobile ad hoc networks (MANETs), since there is no fixed infrastructure and nodes are mobile, routes between nodes are set up and turn down dynamically. For this reason, MANETs often experience route failures and network disconnectivity, especially when the nodes are moving frequently and the network is sparse. Grossglauser and Tse [9] propose to make the mobile nodes serve as relays in order to increase the network throughput in MANETs. Their relay mechanism, called *two-hop relay* protocol, is simple: if there is no direct route between the source node and the destination node, the source node transmits its packets to the nearest neighbor node (called *relay* node) for delivery to the destination. Then, it was shown that with this protocol it is possible to

*The authors acknowledge the support of the European IST project BIONETS and of the Network of Excellence (NoE) EuroNGI.

schedule $\Theta(N)$ concurrent successful transmissions per time-slot, where N is the number of nodes [9]. After that, it was observed that the delay experienced by the packets under this protocol is large [7, 13, 17]. In order to reduce this delay it is proposed in [5, 13] to allow the source transmits the packet to all its neighboring nodes (not only to the nearest neighbor).

In this paper we evaluate the performance of two variants of the basic two-hop relay protocol: the *multicopy two-hop relay (MTR) protocol* and the two-hop relay protocol with *erasure coding* when N is finite. Before describing these two variants that aim at improving the delay, let first introduce the nodes mobility model. All nodes move independently of each other according to the same random mobility model inside a two-dimensional bounded region. Two nodes may only communicate at certain points in time, called *meeting times*. The time that elapses between two consecutive meeting times of a given pair of nodes is called the *inter-meeting time*.

The following will be assumed throughout: **(A1)** Transmission times are instantaneous. **(A2)** All inter-meeting times are independent and identically distributed (iid) random variables (rvs) with a common cumulative distribution function (CDF) $G(\cdot)$.

Assumption **(A1)** will be justified in delay tolerant networks, where the incurred delay to send a packet may be very large with respect to the transmission times [1].

For Assumption **(A2)**, in [15] it has been shown that when nodes move independently on a sphere and they have uniform (stationary) spatial distribution, the inter-meeting time distribution is approximately the exponential distribution with mean $1/\lambda$ for $\lambda t \ll 1$. Furthermore, in [8] the exponential approximation of inter-meeting times distribution was validated for the random mobility models inside a square of *non-uniform* spatial distribution such as the random waypoint model [3] and of uniform spatial distribution such as the random direction model [12]. Moreover, it was observed that assumption **(A2)** is “reasonable” as long the node transmission range is not “too large” with respect to the area where nodes move. On the other hand, if nodes are humans moving in a conference space it has been found that the inter-meeting times distribution shows a heavier-than-exponential tail [4]. In this paper, we will show the impact of considering arbitrary and exponential inter-meeting times distribution on the performance of the MTR protocol. See Section 3 for more details.

We consider the scenario consisting of $N + 1$ mobile nodes: one source node, one destination node, and $N - 1$ relay nodes. The source has a single packet to transmit to the destination. We now introduce the two variants.

MTR protocol. In the MTR protocol the source node may either transmit the packet directly to the destination node when both nodes come within transmission of one another, or use the relay nodes. In the latter case, if the source meets a relay node before meeting the destination, then it sends a copy of the packet to this relay node; this relay node will only transmit the packet to the destination when it comes close to it (as opposed to the *epidemic routing protocol* [20], also called the *unrestricted relay protocol* [8], where a relay node is allowed to send a copy (of its copy) to another relay node). Section 2 evaluates the performance of the MTR protocol in the case where the number of copies in the network is limited.

We define, T_d , the (packet) *delivery delay* as the first time when the destination receives the original packet or a copy, whichever arrives first to the destination. We

assume that the packet at the source cannot be dropped before the transmission has taken place, i.e., before time T_d if the packet joins the source at time $t = 0$. On the other hand, each copy has a time-to-live (TTL) associated with it: when a TTL expires, the relay node that holds the copy drops it. This relay node then becomes eligible to receive another copy. We assume that the source cannot transmit a copy to a relay node that already holds a copy.

Two-hop relay protocol with erasure coding. The relay protocol is the basic two-hop relay protocol. A node may only transmit a piece of information either directly to the destination or to at most one relay node.

In the two-hop relay protocol with erasure coding, the source introduces some redundancy in its transmissions and sends more data than the actual information. Upon receiving a piece of data (packet), the source produces n blocks of data. The transmission of the packet is completed when the destination receives the k th block, regardless of the identity of the $k \leq n$ blocks it has received [15, 19]. More details are provided in Section 4, where the Laplace-Stieljes transform (LST) of the delivery delay and the z -transform of the number of transmissions are derived in closed-form.

Note that under the above assumptions, the delivery delay obtained in our setting gives a lower-bound, as a consequence of the instantaneous transmission time. Second, the total number of copies per-packet generated gives an upper-bound. This is so because in the realistic context the source will not systematically transmit a packet to a relay node that it encounters.

The rest of the paper is organized as follows: Section 2 evaluates the performance of the MTR protocol with the assumptions that the number of packet copies in the network is limited, and that inter-meeting times and TTLs are distributed exponentially. In Section 3, we show the impact of considering arbitrary distribution of inter-meeting times and constant TTLs on the delivery delay of MTR. Section 4 finds the delivery delay and the number of transmissions for the two-hop relay protocol with erasure coding, and compares with erasure coding scheme.

2 Performance of MTR protocol with limited number of copies

In this section we consider the MTR protocol, with exponentially distributed node inter-meeting times with parameter λ (i.e., $G(t) = 1 - e^{-\lambda t}$), exponentially distributed TTLs with parameter μ , and where the number of copies of a packet in the network may not exceed K (including the packet at the source), where K is an arbitrary integer less than or equal to N (in [2, Sec. 3] $K = N$). We recall that we only focus on the transmission of a single packet between a given source and a given destination, and that the packet at the source has no TTL (only copies have a TTL).

The performance metrics of interest are: T_d the time needed to deliver the packet to the destination, C_d the number of copies of the packet before the delivery to the destination, and G_d the total number of transmissions before the delivery to the destination. Note that the latter metric is related to the energy needed to deliver a packet to the destination. In this section we derive closed-form expressions for the $E[T_d^n]$ for all $n \geq 1$, $P(C_d = j)$, and $E[G_d]$. We conclude this section by showing how these results can be used to find the value of K that minimizes $E[G_d]$, subject to a constraint on $E[T_d]$.

Under the above assumptions it is easily seen that the system can be modeled as a finite-state absorbing Markov chain $\mathbf{I} = \{I(t), t \geq 0\}$, where $I(t) \in \{1, 2, \dots, K\}$ gives the number of copies of the packet at time $t < T_d$, and $I(t) = a$ if $t \geq T_d$. The states $1, 2, \dots, K$ are the transient states and the state a is the absorbing state of \mathbf{I} . Let $q(i, j)$ denote the (i, j) -entry of \mathbf{Q} , the infinitesimal generator of \mathbf{I} . It is easily found that

$$\begin{aligned} q(i, i+1) &= (N-i)\lambda, \quad i = 1, \dots, K-1, \\ q(i, i-1) &= (i-1)\mu, \quad i = 2, \dots, K, \\ q(i, i) &= -[N\lambda + (i-1)\mu], \quad i = 1, \dots, K-1, \\ q(K, K) &= -[K\lambda + (K-1)\mu], \\ q(i, a) &= i\lambda, \quad i = 1, \dots, K, \\ q(i, j) &= 0, \text{ otherwise.} \end{aligned}$$

The matrix \mathbf{Q} can be written as

$$\mathbf{Q} = \left(\begin{array}{c|c} \mathbf{Q}_K & \mathbf{R} \\ \hline \mathbf{0} & 0 \end{array} \right), \quad (1)$$

where $\mathbf{Q}_K = [q(i, j)]_{1 \leq i, j \leq K}$, $\mathbf{R} = (q(1, a), \dots, q(K, a))^T$, and $\mathbf{0}$ is a K -dimensional row vector with all entries equal to 0.

We will show below that for any initial state $I(0)$, $E[T_d^n]$, $P(C_d = j)$, and $E[G_d]$ can be derived in closed-form if one has a closed-form expression for \mathbf{Q}_K^{-1} , the inverse of \mathbf{Q}_K .

We now derive \mathbf{Q}_K^{-1} in closed-form. We note that \mathbf{Q}_K can be decomposed as $\mathbf{Q}_K = \hat{\mathbf{Q}}_K + buu^T$, where $\hat{\mathbf{Q}}_K$ is the K -by- K sub-matrix composed of the first K rows and columns of the matrix \mathbf{Q}_N , $u = (0, \dots, 0, 1)^T$, and $b = \lambda(N - K)$. By applying the Sherman-Morrison formula [16, P. 76] we find that

$$\mathbf{Q}_K^{-1} = \hat{\mathbf{Q}}_K^{-1} - \frac{b}{1 + bu^T \hat{\mathbf{Q}}_K^{-1} u} \hat{\mathbf{Q}}_K^{-1} uu^T \hat{\mathbf{Q}}_K^{-1}. \quad (2)$$

Let $q_K(i, j)$ be the (i, j) -entry of \mathbf{Q}_K^{-1} and $\hat{q}_K(i, j)$ be the (i, j) -entry of $\hat{\mathbf{Q}}_K^{-1}$. Equivalently, (2) rewrites

$$q_K(i, j) = \hat{q}_K(i, j) - \frac{\lambda(N - K)\hat{q}_K(i, K)\hat{q}_K(K, j)}{1 + \lambda(N - K)\hat{q}_K(K, K)}. \quad (3)$$

It remains to find the entries $\hat{q}_K(i, j)$ of $\hat{\mathbf{Q}}_K^{-1}$.

The matrix \mathbf{Q}_N^{-1} was obtained in closed-form in [2, Appendix I]. On the other hand, simple algebra shows that the (i, j) -entry of $\hat{\mathbf{Q}}_K^{-1}$ is related to \mathbf{Q}_N^{-1} through the relation

$$\hat{q}_K(i, j) = q_N(i, j) + \frac{\lambda(N - K)q_N(i, K)q_N(K + 1, j)}{1 - \lambda(N - K)q_N(K + 1, K)} \quad (4)$$

for $1 \leq i, j \leq K$. A closed-form expression for $q_N(i, j)$ was obtained in [2, Eq. 25] for any i and j .

2.1 The performance metrics

Given that $I(0) = i$, the n th order moment of T_d , which is by definition equal to the time to absorption of \mathbf{I} , is given by [14, Chap. 2, Eq. 2.2.7]

$$E_i[T_d^n] = (-1)^n n! (\alpha_i \mathbf{Q}_K^{-n} \mathbf{e}) = (-1)^n n! \sum_{j=1}^K q_K^{(n)}(i, j) \quad (5)$$

for $i = 1, \dots, K$, where \mathbf{e} is a K -dimensional column vector with all entries equal to 1, α_i is a K -dimensional row vector with all entries equal to 0 except the i th one that is equal to 1, and $q_K^{(n)}(i, j)$ is the (i, j) -entry of \mathbf{Q}_K^{-n} that can be expressed in closed-form in terms of $q_K(i, j)$ which are given in (3)-(4).

Given that $I(0) = i$, the probability distribution of the number of copies just at delivery time is [2, Sec. 3.2]

$$P_i(C_d = j) = -j\lambda q_K(i, j), \quad i = 1, \dots, K. \quad (6)$$

Given that $I(0) = i$, the expected total number of transmissions before delivery is given by [2, Sec. 3.3]

$$E_i[G_d] = \frac{1}{2} \left[\alpha E_i[T_d] + E_i[C_d] + bq_K(i, K) + \delta \right], \quad (7)$$

where $\alpha := (\lambda N - \mu)$, $b = \lambda(N - K)$, and $\delta := \frac{1}{\rho} - i - 1$.

2.2 Minimizing the consumed energy

The total number of copies that are transmitted is directly related to the energy consumed to deliver the packet to the destination. Our objective now is to use the above results to find, K_{opt} , the optimal value of K which minimizes the expected total number of copies that are transmitted before the delivery of the packet to the destination, subject to a constraint on the expected delivery delay, that is,

$$\min_{\{K: E^{(K)}[T_d] \leq C\}} E^{(K)}[G_d],$$

where $E^{(K)}[T_d] := E_1[T_d]$ and $E^{(K)}[G_d] := E_1[G_d]$. The superscript (K) emphasizes the dependency in the variable K . Since the integer mappings $K \rightarrow E^{(K)}[T_d]$ and $K \rightarrow E^{(K)}[G_d]$ are strictly decreasing and strictly increasing, respectively, the solution to this constrained optimization problem is obtained for the smallest integer K in $[1, N]$ such that $E^{(K)}[T_d] \leq C$ if $E^{(N)}[T_d] \leq C$, and has no solution if $E^{(N)}[T_d] > C$. Figure 1 reports K_{opt} as a function of C for $\rho = 1$ and for three values of λ . We observe that K_{opt} has a sharp decay as C increases.

3 Impact of arbitrary inter-meeting and constant TTLs on MTR

In [4], it has been observed that the inter-meeting times distribution has heavier-than-exponential tail. This finding was the motivation to investigate the impact of arbitrary inter-meeting times distribution on the delivery delay of the MTR protocol.

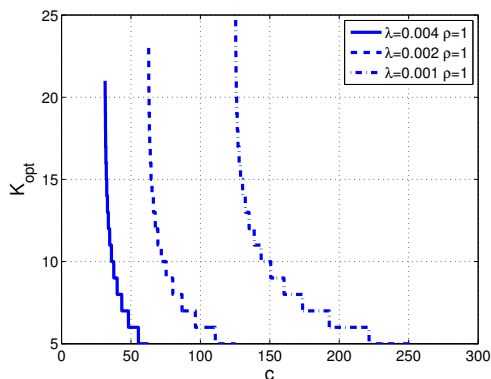


Fig. 1. The optimal maximum number of copies (K_{opt}) as a function of constraint (C) on $E[T_d]$ for $N = 100$ and $C \in [E^{(100)}[T_d], 2E^{(100)}[T_d]]$.

Throughout this section we assume that for any pair of nodes their inter-meeting times are iid with distribution $G(t)$, and all inter-meetings are mutually independent. Let X be a generic rv with distribution G . Also define $G^*(s) = E[e^{-sX}]$ the LST of X . We assume that TTLs are constant and all equal to T . As a result, the stochastic process \mathbf{I} is no longer a Markov process and a different approach has to be used in order to evaluate the delivery delay of the MTR protocol. For sake of simplicity we consider the case where $K = N$, i.e., there is no restriction on the number of packet copies in the network.

For convenience we label the nodes so that node 0 is the source, node N is the destination, and nodes $1, 2, \dots, N-1$ are the relay nodes. Since $K = N$, we have

$$T_d \stackrel{st}{=} \min(X_{sd}, D_1, \dots, D_{N-1}), \quad (8)$$

where $X_{sd} \stackrel{st}{=} X$ represents the inter-meeting time between the source and the destination, and D_i is the time needed for relay node $i = 1, 2, \dots, N-1$ to deliver a copy of the packet to the destination. Moreover, the rvs $X_{sd}, D_1, \dots, D_{N-1}$ are mutually independent and the rvs D_1, \dots, D_{N-1} are identically distributed. Hence,

$$P(T_d < t) = 1 - (1 - G(t))P(D_i > t)^{N-1}. \quad (9)$$

We need to determine $P(D_i > t)$. We shall actually find an approximation formula for $P(D_i > t)$ since finding an exact expression is a very difficult task, unless $G(t)$ is the exponential distribution that is considered in the end of this section.

From now on i is fixed in $\{1, \dots, N-1\}$. We assume that the source, destination and relay node i are in steady-state at time $t = 0$, and that the relay node i does not hold a copy of the packet at $t = 0$ (only the source holds the original packet at $t = 0$).

Let R record the number of times the relay node i has dropped a copy of the packet before it transmits it to the destination. On the event $R = m+1$, let $a_k > 0$ be the arrival time of the k th copy to the relay node i for $k = 1, \dots, m+1$, let $d_k > a_k$ be the time where the k th copy is dropped by relay node i for $k = 1, \dots, m$, and let e_{m+1} be

the time where copy $m + 1$ reaches the destination. Define $\hat{X} = a_1$, $Z_k = a_{k+1} - d_k$ for $k = 1, \dots, m$, and $\hat{Z} = e_{m+1} - a_{m+1}$. Clearly,

$$D_i = \hat{X} + Z_1 + \dots + Z_m + mT + \hat{Z} \quad (10)$$

on the event $R = m + 1$. Given that $R = m + 1$, the rvs $\hat{X}, Z_1, \dots, Z_m, \hat{Z}$ are mutually independent; moreover the rvs Z_1, \dots, Z_m are iid.

Let $D^*(s) := E[e^{-sD_i}]$ be the LST of D_i . We have

$$D^*(s) = E[e^{-s\hat{X}}]E[e^{-s\hat{Z}}] \sum_{m \geq 0} (e^{-sT} E[e^{-sZ_k}])^m P(R = m + 1). \quad (11)$$

1. *Evaluation of $Z_T^*(s) := E[e^{-sZ_k}]$* : Recall that X denotes a generic inter-meeting time and that its density probability is $g(\cdot)$.

Let $h_T(t) := dP(Z_k < t)/dt$ be the probability density of Z_k . The reason why we indicate the dependency on the parameter T in $h_T(t)$ will soon become apparent. If the source does not meet the relay node i in $(a_k, a_k + T)$ then $Z_k = a_{k+1} - a_k - T \stackrel{st}{=} X - T$, otherwise $Z_k = a_{k+1} - a_k - (T - (a'_k - a_k))$ where a'_k is first time the source meets the relay node i in $(a_k, a_k + T)$. The latter rewrites $Z_k \stackrel{st}{=} X_1 + X_2 - T$ with $X_j \stackrel{st}{=} X$ for $j = 1, 2$. From this we deduce that $Z_T^*(s) = \int_0^\infty h_T(t) e^{-st} dt$ satisfies the following renewal equation

$$Z_T^*(s) = \int_0^\infty e^{-st} g(T + t) dt + \int_0^T g(u) Z_{T-u}^*(s) du. \quad (12)$$

We have shown that $Z_T^*(s)$ satisfies an integral equation (of Fredholm type) from which $Z_T^*(s)$ can be obtained numerically using standard techniques [10].

2. *Probability distribution of R* : Finding the probability distribution of R is difficult. We will first assume that R is a geometric rv with parameter $\pi = 1 - P(R = 1)$. It is possible to find an integral equation for π . However, for sake of simplicity, we will assume that the destination node is at equilibrium at time a_1 , so that $\pi = 1 - G_e(T)$, with $G_e(t)$ the excess probability distribution of $G(t)$, that is, $G_e(t) = (1/E[X]) \int_0^t (1 - G(u)) du$. In summary, $P(R = m + 1) \approx (1 - \pi)\pi^m$, for $m \geq 0$. Plugging the approximation of $P(R = m + 1)$ in (11) gives that

$$D^*(s) \approx \frac{1 - \pi}{sE[X]} \frac{(1 - G^*(s)) E[e^{-s\hat{Z}}]}{1 - \pi e^{-sT} Z_T^*(s)}, \quad (13)$$

where $Z_T^*(s)$ is the solution of the integral equation (12), and $\pi = 1 - G_e(T)$ (Hint: $E[e^{-s\hat{X}}] = (1 - G^*(s))/sE[X]$). It remains to evaluate $E[e^{-s\hat{Z}}]$. Again, this is not an easy task. Clearly, $e^{-sT} \leq E[e^{-s\hat{Z}}] \leq 1$. For sake of simplicity, we will replace $E[e^{-s\hat{Z}}]$ by 1. This gives the final approximation

$$D^*(s) \approx \frac{1 - \pi}{sE[X]} \frac{(1 - G^*(s))}{1 - \pi e^{-sT} Z_T^*(s)}. \quad (14)$$

Finally, $P(D_i > t)$ is obtained by inverting $(1 - D^*(s))/s$ (since $(1 - D^*(s))/s = \int_0^\infty e^{-st} P(D_i > t) dt$) with the help of the complex inversion formula [18, Chap. 7], which yields

$$P(D_i > t) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} e^{ts} \frac{1 - D^*(s)}{s} ds, \quad t > 0, \quad (15)$$

where the integration has to be performed along a line $s = \gamma$ in the complex plane (in (15) i denotes the imaginary complex number). The real number γ must be chosen so that $s = \gamma$ lies to the right of all singularities. Note that since $P(D_i > t)$ is a bounded function it is sufficient to take $\gamma > 0$.

The approximation for $F_{T_d}(t) := P(T_d > t)$ has been computed when $G(t)$ is an hyper-exponential distribution, namely, $G(t) = 1 - \sum_{l=1}^H p_l e^{-\nu_l t}$, and compared to simulation results. The evaluation of the integral in (15) has been performed by using the procedure described in [6].

Numerical results are reported in Figure 2. Two hyper-exponential distributions, represented by the t-tuple $(H, \nu_1, \dots, \nu_H, p_1, \dots, p_N)$, have been considered: (H1) $(3, 0.09, 0.08, 0.07, 0.6, 0.3, 0.1)$ with mean 22.83sec., and (H2) $(3, 0.05, 0.04, 0.03, 0.6, 0.3, 0.1)$ with mean 11.84sec.. The numerical results of the mathematical model were done using a C program for a network composed of one source, one destination, and $N - 1$ relay nodes.

Let $T_d(app)$ and $T_d(sim)$ be the approximate and simulated delivery delays, respectively. Let $F_{T_d}(app)(\cdot)$ (resp. $F_{T_d}(sim)(\cdot)$) denotes the complementary cumulative distribution function (CCDF) of $T_d(app)$ (resp. $T_d(sim)$).

Figure 2.(a) displays the mappings $t \rightarrow F_{T_d}(app)(t)$ and $t \rightarrow F_{T_d}(sim)(t)$ for the hyper-exponential distributions H1 and H2. We observe that the approximation is accurate for moderate value of N .

Figure 2.(b) compares $F_{T_d}(sim)(t)$ with the CCDF of T_d in the case where the inter-meeting times are exponentially distributed, where the latter distribution has been obtained in closed-form. See (16) in the following for details. We conclude from these results that T_d under exponential inter-meeting times is stochastically smaller than T_d under hyper-exponential inter-meeting times. This is related to the fact that the hyper-exponential distribution has a fatter tail than the exponential distribution.

We conclude this section by briefly addressing the simple case where the inter-meeting times are distributed exponentially with rate λ , and the TTLs are constant and all equal to T . In this case we derive the closed-form expression of the delivery delay of the MTR protocol as follows

$$P(D_i > t) = e^{-\lambda t} \left[1 + \sum_{m=0}^{\lfloor \frac{t}{T} \rfloor} \sum_{k=0}^m \frac{(A_m)^{k+1} - (B_m)^{k+1}}{(k+1)!} \right], \quad (16)$$

where $A_m := \lambda(t - mT)$, $B_m := \lambda[t - (m+1)T]^+$, $\lfloor x \rfloor$ designates the largest integer $\leq x$, and $[x]^+ := \max(0, x)$. Plugging (16) in (9) gives the CDF of T_d . By comparing the CCDF of T_d with constant TTL = T and with TTLs distributed exponentially with mean T , we deduced that T_d with constant TTL is stochastically smaller than T_d with TTL exponential.

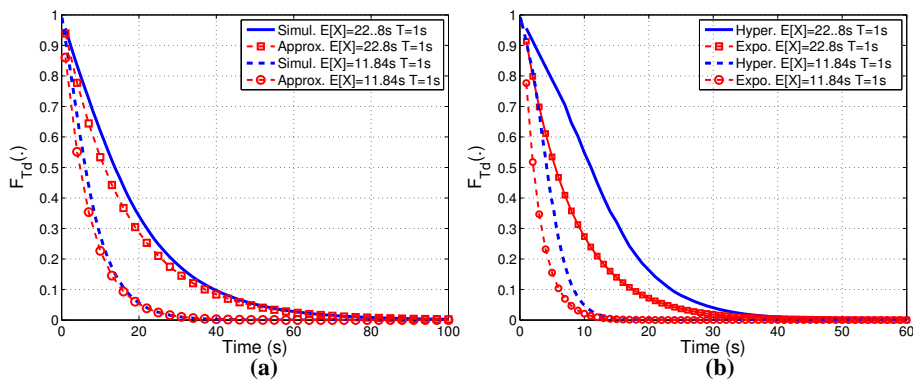


Fig. 2. (a) Mappings $t \rightarrow F_{T_d(app)}(t)$ and $t \rightarrow F_{T_d(sim)}(t)$ for two different hyper-exponential distributions ($N = 10$). (b) Comparison of CCDF of T_d in the case of hyper-exponential (simulated CCDF) and exponential (CCDF using (16)) inter-meeting time distributions ($N = 50$).

4 Two-hop Relay with Erasure Coding

We now consider a system where the source introduces some redundancy in its transmissions and sends more data than the actual information. The advantage of this mechanism is that it can considerably reduce the variance of the delivery delay at the expense of an increase of its expectation.

One of these techniques is known as erasure coding [11]. Erasure coding with replication factor r works as follow. Upon receiving a packet of size M , the source produces $n = r \cdot M/b$ equal sized code blocks of size b , such that any of the $k = (1 + \epsilon) \cdot M/b$ code blocks can be used to reconstruct the packet. Here ϵ is a small constant, close to zero [11]. Thus, the destination is able to decode the packet if it receives $k \leq n$ blocks. On the other hand, when $k = 1$, the size of a block becomes almost equal to M , the packet size, and in this case the destination needs to receive a block in order to decode the packet. Thus for $k = 1$, the erasure coding scheme is the same as a simple multicopy scheme in which the source sends exactly one copy of a packet to n different relay nodes [19]. We will exploit this observation to compare erasure coding with the multicopy scheme in the following.

Throughout this section the stochastic model is the following. There are N relay nodes, one source node, and one destination node. We assume that the source cannot send directly a packet to the destination. Inter-meeting times between any pair of nodes are exponentially distributed with rate λ , except for the pair source-destination. Under this setting, the only way to forward the data from the source to the destination is through the relay nodes. The source has only one packet to send to the destination, and the source implements the erasure coding algorithm with replication factor r and parameter k . Hence, the destination needs to receive $k \leq n$ of the blocks in order to decode the original packet. The forwarding mechanism used to deliver the blocks to the destination is the standard two-hop relay protocol. We assume that there is only one copy of a block in the network.

A relay node can only relay one block at a time, and it is possible for a relay node that already delivered a block to the destination receives a new block when it again encounters the source. There is no TTL associated with the blocks.

Let T_d and G_d be the delivery delay and the total number of source-relay transmissions at the time when the k th block reaches the destination, respectively. Introduce the joint transform $H(s, z) := E[e^{-sT_d} z^{G_d}]$, $s \geq 0$, $|z| \leq 1$.

We now evaluate $H(s, z)$. Let $A(t) = (B(t), R(t))$ denote a two-dimensional process such that $A(t) = (m, l)$, $1 \leq m \leq n$, $1 \leq l \leq k-1$, $m+l \leq n$, if there are m relay nodes that hold m blocks (one block for each relay node) and the destination has received l blocks at time $t < T_d$, and $A(t) = a$ when $t \geq T_d$ (a is an absorbing state). Under the above assumptions, $\{A(t), t \geq 0\}$ is a finite-state absorbing Markov process. Figure 3 displays the transition diagram of this Markov chain, where the y-axis represents $R(t)$ and the x-axis represents the sum $B(t) + R(t)$. More precisely, a point (i, l) , $i \geq l$, in the transition diagram means that the destination has received l blocks and that there are $i-l$ relay nodes that hold $i-l$ blocks (one block for each).

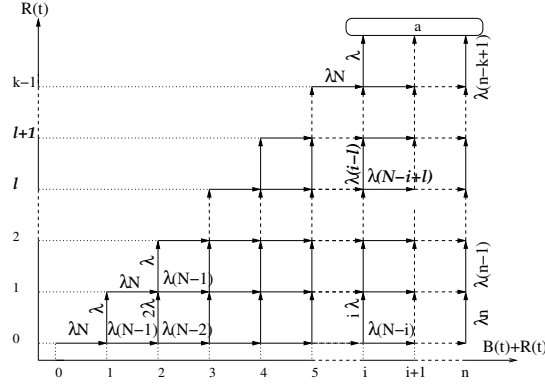


Fig. 3. Transition diagram of the Markov chain $\{A(t), t \geq 0\}$.

Let $j_i \geq 0$ denote the number of jumps (transitions) along the horizontal line of index $i \in \{0, \dots, k-1\}$. Let S_i denote the total number of jumps along the lines of index less than or equal to i . Given S_{i-1} , the probability of making j_i jumps along the horizontal line i is

$$\begin{cases} P_1(j_i) = \frac{(N-S_{i-1}+i)!}{(N-S_i+i)!} \frac{S_i-i}{N^{j_i+1}}, & S_i < n \\ P_2(j_i) = \frac{(N-S_{i-1}+i)!}{(N-n+i)!} \frac{1}{N^{j_i}}, & S_i = n, \end{cases} \quad (17)$$

for $0 \leq i \leq k-1$. Let m^* denote the index of the horizontal line such that $S_{m^*-1} < n$ and $S_{m^*} = n$. Conditioned on all the possible paths before absorption at $\{a\}$ and given that $A(0) = (0, 0)$, $H(s, z)$ can be written as

$$\begin{aligned} H(s, z) &= \sum_{j_0=1}^{n-1} \cdots \sum_{j_{k-1}=0}^{n-1} \left(\frac{zN\lambda}{s+N\lambda} \right)^{S_{k-1}+k} \prod_{l=0}^{k-1} P_1(j_l) + \sum_{j_0=1}^n \cdots \sum_{j_{m^*-1}=0}^{n-S_{m^*-2}} P_2(n-S_{m^*-1}) \\ &\quad \times \prod_{l=0}^{m^*-1} P_1(j_l) \left(\frac{zN\lambda}{s+N\lambda} \right)^{n+m^*} \prod_{l=m^*}^{k-1} \frac{\lambda(n-l)}{s+\lambda(n-l)}, \end{aligned} \quad (18)$$

We now evaluate the expectation and the variance of T_d for different values of k , n , and N . Let $\sigma_{T_d} := \sqrt{\text{var}(T_d)}/E[T_d]$ denote the normalized standard deviation of T_d . As noted earlier, erasure coding when $k = 1$ is similar to the multicopy scheme. Table 1 shows that $E[T_d]$ increases with k and that σ_{T_d} decreases with k . For instance, for $N = 20$ and $n = 10$ when k increases from 1 to 5 $E[T_d]$ increases by a factor of 3 while σ_{T_d} decreases by a factor of 7.5, thereby showing that erasure coding has a much lower variability than the multicopy scheme. A similar result was found in [19] under the assumption that the source instantaneously transmits all its n blocks to n different relay nodes.

Table 1. Erasure coding ($k > 1$) vs multicopy scheme ($k = 1$) for different value of n and N .

(N, n)	(20,10)			(40,10)		
k	1	2	5	1	2	5
$E[T_d]$ (sec.)	31.51	42.04	96.69	22	33.7	79.76
σ_{T_d}	8.7	4.9	1.17	2.01	1.38	0.63

We conclude this section by investigating the behavior of $T_N^*(s) := E[e^{-sT_d}]$, the LST of T_d , as N is large. We observed that as N becomes large the most probable path (MPP) is where all n blocks are first transmitted to n relay nodes, and then these relay nodes start to deliver these blocks to the destination. Further, it is easy to see that as N is large the system has a deterministic path MPP with probability one. Therefore,

$$T_N^*(s) \approx \left(\frac{N\lambda}{s + N\lambda} \right)^n \prod_{l=0}^{k-1} \frac{\lambda(n-l)}{s + \lambda(n-l)} \quad (19)$$

5 Concluding Remarks

In this paper, we have studied a class of two-hop relay protocols. The interest was on the multicopy two-hop relay (MTR) protocol and on the two-hop relay protocol with erasure coding. Closed-form expressions have been derived for n th order moment of the time to deliver a packet to its destination, the distribution of the number of copies at delivery, and the expected total number of copies generated before delivery in the case where copies have limited lifetime (TTL) and where the number of copies in the network is limited. Also, We investigated the impact of arbitrary inter-meeting times distribution and constant TTLs on the delivery delay of the MTR protocol. In particular, we show that exponential inter-meeting times yield stochastically smaller delivery delay than hyper-exponential inter-meeting times. Finally, for the two-hop relay protocol with erasure coding the joint generating function of the delivery delay and of the number of transmissions was derived in closed-form. By analyzing these results, we found that the delivery delay in the case of erasure coding has much lower dispersion than the delivery delay of the multicopy scheme.

As future work, we will study the delay when there are multiple sources with multiple packets to transmit to a set of destinations and where the relay nodes may have different mobility.

References

1. Delay Tolerant Research Group, <http://www.dtnrg.org>.
2. A. Al Hanbali, P. Nain, and E. Altman, Performance of Two-hop Relay Routing Protocol With Limited Packet Lifetime, *Proc. of Valuetools 2006*, Pisa, Italy, Oct. 2006.
3. C. Bettstetter, H. Hartenstein, and X. Prez-Costa. Stochastic Properties of the Random Waypoint Mobility Model. *ACM/Kluwer Wireless Networks, Special Issue on Modeling and Analysis of Mobile Networks*, vol. 10, no. 5, pp. 555-567, Sept. 2004.
4. A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, Impact of Human Mobility on the Design of Opportunistic Forwarding Algorithm, *Proc. of INFOCOM 2006*, Barcelona, Spain, Apr. 2006.
5. R. De Moraes, H. Sadjadpour, and J. Garcia-Luna-Aceves, Throughput-Delay Analysis of Mobile Ad-hoc Networks with a Multi-Copy Relaying Strategy, *Proc. of IEEE SECON*, Santa Clara, CA, Oct. 2004.
6. H. Dubner and J. Abate, Numerical inversion of Laplace transforms by relating them to the finite Fourier cosine transform, *Journal of the ACM*, Vol. 15, No. 1, pp. 115-123, Jan. 1968.
7. A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-Delay Trade-off in Wireless Networks *Proc. of INFOCOM 2004*, Hong Kong, Apr. 2004.
8. R. Groenevelt, P. Nain, and G. Koole, The Message Delay in Mobile Ad Hoc Networks, *Proc. of Performance 2005*, Juan-les-Pins, France, October 2005. Published in *Performance Evaluation*, Vol. 62, Issues 1-4, Oct. 2005, pp. 210-228.
9. M. Grossglauser and D. Tse, Mobility Increases the Capacity of Ad hoc Wireless Networks, *IEEE/ACM Transactions on Networking*, Vol. 10, No. 4, Aug. 2002, pp. 477-486.
10. S. G. Mikhlin, *Integral Equations*. Pergamon Press, Oxford, 1964.
11. M. Mitzenmacher, Digital Fountains: A Survey and Look Forward, *Proc. of IEEE Information Theory Workshop*, TX, USA, Oct. 2004.
12. P. Nain, D. Towsley, B. Liu, and Z. Liu, Properties of Random Direction Models *Proc. of INFOCOM 2005*, Miami, FL, USA, Mar. 2005.
13. M. J. Neely and E. Modiano, Capacity and Delay Tradeoffs for Ad-Hoc Mobile Networks, *IEEE Transactions on Information Theory*, Vol. 51, No. 6, June 2005.
14. M. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Johns Hopkins University Press, 1981.
15. E. Perevalov and R. Blum, Delay Limited Capacity of Ad hoc Networks: Asymptotically Optimal Transmission and Relaying Strategy, *Proc. of INFOCOM 2003*, San Francisco, USA, Apr. 2003.
16. W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1988.
17. G. Sharma, R. Mazumdar, and N. Shroff Delay and Capacity Trade-offs in Mobile Ad Hoc Networks: A Global Perspective *Proc. of INFOCOM 2006*, Barcelona, Spain, Apr. 2006.
18. M. R. Spiegel, *Schaum's Outline of Theory and Problems of Laplace Transforms*. McGraw-Hill, New York, 1965.
19. Y. Wang, S. Jain, M. Martonosi, and K. Fall, Erasure-coding based routing for opportunistic networks *Proc. of SIGCOMM Workshop on DTN*, Philadelphia, PA, USA, Aug. 2005.
20. X. Zhang, J. Neglia, J. Kurose, and D. Towsley, Performance Modeling of Epidemic Routing *Proc. of NETWORKING 2006*, Coimbra, Portugal, May 2006.