



HAL
open science

Performance Evaluation of Packet Relaying in Ad Hoc Networks

Ahmad Al Hanbali, Philippe Nain, Eitan Altman

► **To cite this version:**

Ahmad Al Hanbali, Philippe Nain, Eitan Altman. Performance Evaluation of Packet Relaying in Ad Hoc Networks. [Research Report] RR-5860, INRIA. 2006, pp.24. inria-00070166

HAL Id: inria-00070166

<https://inria.hal.science/inria-00070166>

Submitted on 19 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Performance Evaluation of Packet Relaying in Ad Hoc Networks

Ahmad Al Hanbali, Philippe Nain, and Eitan Altman

N° 5860

Mars 2006

Thème COM



*Rapport
de recherche*

Performance Evaluation of Packet Relaying in Ad Hoc Networks

Ahmad Al Hanbali, Philippe Nain, and Eitan Altman

Thème COM — Systèmes communicants
Projets Maestro

Rapport de recherche n° 5860 — Mars 2006 — 24 pages

Abstract: Considered is a mobile ad hoc network consisting of three types of nodes (source, destination and relay nodes) and using the two-hop relay routing protocol. Packets at relay nodes are assumed to have a limited lifetime in the network. All nodes are moving inside a bounded region according to some random mobility model. Both closed-form expressions, and asymptotic results when the number of nodes is large, are provided for a number of QoS metrics related to the packet delivery delay and the energy needed to transmit a packet from the source to its destination. Markov models and mean-field approximations are used. We also introduce and evaluate a variant of the two-hop relay protocol that limits the number of copies of a packet in the network. Our model is validated through simulations for two mobility models (random waypoint and random direction mobility models), numerical results for the two-hop relay protocols are reported, and the performance of the two-hop routing and of the epidemic routing protocols are compared.

Key-words: Ad hoc network, Two-hop relay protocol, Limited packet lifetime, Mobility model, Markovian analysis, Mean-field approximation, QoS metric.

Évaluation de Performance des Protocoles à base de Relais dans les Réseaux Ad Hoc

Résumé : Nous considérons un réseau ad hoc mobile constitué de trois types de nœuds: source, destination et nœud relais. Le rôle du nœud relais est de transmettre les paquets entre la source et la destination en l'absence d'une route directe entre ces derniers. Nous supposons que la durée de vie des paquets transmis par les noeuds relais est limitée. Tous les nœuds se déplacent indépendamment les uns des autres, selon le même modèle de mobilité. Nous établissons les formules de la distribution de probabilité des temps de délai pour l'envoi des paquets à la destination, ainsi que la quantité d'énergie consommée par le réseau pour l'envoi des paquets. Ces formules ont été trouvées en utilisant un modèle markovien et un modèle fluide. En plus, nous étudions l'effet sur les temps de délai de limiter la quantité d'énergie à consommer par le réseau . En comparant les résultats théoriques trouvés et les résultats des simulations, nous avons trouvé que notre modèle est précis surtout pour les petites valeurs de portée de transmission des noeuds et lorsque le nombre des noeuds est fini.

Mots-clés : réseaux ad hoc, relais des paquets, paquet de durée de vie limitée, modèles de mobilité, analyse Markovienne, évaluation de performance.

1 Introduction

Ad hoc Networks are complex distributed systems, that are composed of wireless mobile or static nodes that can freely and dynamically self-organize. In this way they form arbitrary, and temporary “ad hoc” network topologies, allowing devices to seamlessly interconnect in areas with no pre-existing infrastructure.

In a Mobile Ad Hoc Network (MANET), since there is no fixed infrastructure and nodes are mobile, links between nodes are set up and turn down dynamically. A link between two nodes is up when these nodes are inside one another communication range, and a link is down otherwise. The establishment of a route from a source node to a destination node requires the simultaneous availability of a number of links that are all up, one originating at the source node and another one ending at the destination nodes. Indeed, MANETs often experience route failures and network disconnectivity, especially when the nodes are moving frequently and the network is sparse. Grossglauser and Tse [9] have observed that mobility in MANETs can be used to increase the average network throughput. Their idea was to look at the diversity gain achieved by using the mobile nodes as relays. Their relay mechanism, called *two-hop relay* protocol, is simple: if there is no route between the source node and the destination node, the source node transmits its packets to all neighboring nodes (called *relay* nodes) that it meets for delivery to the destination. A relay node is only allowed to send a packet to its destination node, and it is not allowed to send the packet to another relay node, thereby justifying the name of this protocol. It was then shown in [6] that a bounded delay can be guaranteed under this relaying mechanism. The aim of these studies (see also [10]) is the scaling property of the throughput or delay as the number of nodes in the network becomes large. Our interest in the present work is in the performance of the above mentioned relaying mechanism in a network consisting of a fixed finite number of nodes.

It is important to mention that most of the studies of scaling laws of delay and throughput in wireless MANETs assume a uniform spatial distribution of nodes, which is the case, for example, when the nodes perform a symmetric Random Walk over the region of interest [6, 9], or when nodes move according to the Random Direction model [13]. In the present work, we replace this assumption by assuming that the *inter-meeting* time between two nodes, defined as the time duration time between two consecutive points in time where these nodes meet (i.e. come within transmission range of one another), is exponentially distributed. The validity of this assumption has been discussed in [8], and its accuracy has been shown for a number of mobility models (Random Walker, Random Direction, Random Waypoint) in the case when the node transmission range is small with respect to the area where the nodes evolve. It is worth pointing out that for some of these mobility models (non-symmetric Random Walk and Random Waypoint) nodes are not uniformly distributed over the area of interest.

The objective of this paper is to study a number of quality of service (QoS) metrics bearing on the packet delivery delay and the overhead induced by the two-hop relay protocol (see Section 2). This will be done under the assumption that, unlike in [8, 15], packets at relay nodes have a limited lifetime in the network.

Another relay protocol closely related to the two-hop relay protocol is the so-called *epidemic routing* protocol [14, 16]. This protocol is identical to the two-hop relay protocol, except that in the

epidemic routing protocol a relay node is allowed to transmit a packet to any node that it meets, including another relay node. Epidemic routing decreases the delivery delay of packets at the cost of increasing the energy consumption by the network. The performance of both the two-hop relay protocol and the epidemic routing protocol will also be compared in this paper.

The rest of the paper is organized as follows: Section 2 gives a careful description of the two-hop relay protocol, sets the modeling assumptions, and defines the QoS metrics of interest (delivery delay, overhead in terms of the number of copies of a packet). In Section 3 we develop a Markovian analysis that yields closed-form expressions for these QoS metrics. In Section 4, we propose and evaluate a modification of the two-hop relay protocol, called K -limited two-hop relay protocol, that aims at limiting the overall energy consumption. Section 5 presents an asymptotic analysis of the QoS metrics as the number of nodes is large; this analysis uses a mean-field approximation. Validation of our model, and comparison of the performance of the two-hop relay protocol and the epidemic routing protocol are given in Section 6. Section 7 concludes the paper and suggests some research directions.

2 The MANET model

We consider the MANET model introduced in [8]. In this model the characteristics of a MANET are captured through a single parameter, $1/\lambda$, representing the expected inter-meeting time between any pair of nodes. More precisely, there are $N + 1$ nodes, one source node, one destination node and $N - 1$ relay nodes. 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. In [8] it is assumed that inter-meeting times are mutually independent and identically distributed (iid) random variables (rvs), with an exponential distribution with intensity $\lambda > 0$.

Throughout this paper we address the unique scenario where the source wants to send a single packet to the destination node. To this end the source node may use the relay nodes, as explained below. In this paper, we will focus on the *two-hop relay* protocol [9].

The two-hop relay protocol works as follows. The source node keeps sending a copy of the packet to all nodes that it meets and that do not have a copy, including the destination node, until the destination node has received a copy of the packet. Transmissions are assumed to be instantaneous. The way the source node is notified that the destination node has received the packet, either directly from it or from a relay node, is irrelevant for the metrics that we will consider (see below).

A relay node that possesses a copy of the packet is only allowed to send it to the destination node, thereby justifying the name of this protocol (two-hop relay protocol).

In addition to the model in [8] we assume throughout this paper that each *copy* of the packet has a *Time-To-Live* (TTL). When the TTL of a copy expires then the copy is destroyed. TTLs are assumed to be iid rvs with an exponential distribution with rate $\mu > 0$. The packet to be sent by the source has no TTL associated with it, so that the source is always able to send a copy to another node (if the packet at the source has a TTL then there is a non-zero probability that the destination node will never receive the packet. This scenario is not considered in this paper).

We assume that the source is ready to transmit the packet to the destination at time $t = 0$. The (packet) delivery time (or delivery delay), T_d , is the first time after $t = 0$ when the destination node receives the packet (or a copy of the packet).

In the following we will investigate the delivery delay, the number of copies in the system at the delivery time, and the total number of copies generated by the source before the delivery time (Section 3). The latter is related to the overhead induced by the two-hop relay protocol and, in particular, to the total energy needed to transmit the packet to the destination (Section 4).

A word on the notation: throughout $\mathbf{1}_A$ will designate the indicator function of any event A ($\mathbf{1}_A = 1$ if A is true and 0 otherwise) and $\text{diag}(a_1, \dots, a_n)$ will define a n -by- n diagonal matrix with (i, i) -entry a_i .

3 Markovian Analysis

The state of the system is represented by the random variable $I(t) \in \{1, 2, \dots, N, a\}$, where $I(t) \in \{1, 2, \dots, N\}$ gives the number of copies when the packet has not been delivered to the destination (i.e. for $0 \leq t < T_d$) and $I(t) = a$ for $t \geq T_d$. Under the assumptions made in Section 2, $\{I(t), t \geq 0\}$ is an absorbing, finite-state, continuous-time Markov chain, with transient states $\{1, 2, \dots, N\}$ and absorbing state a . Let $\mathbf{P} = [p(i, j)]$ be the one-step transition matrix of the absorbing, finite-state, discrete-time Markov chain (referred to as **MC** from now on) *embedded* just before the jump times of the Markov chain $\{I(t), t \geq 0\}$. From the transition rate diagram of Markov chain $\{I(t), t \geq 0\}$ in Figure 1 we find

$$\begin{aligned} p(i, i+1) &= \frac{(N-i)\lambda}{N\lambda + (i-1)\mu} = \frac{(N-i)\rho}{N\rho + i - 1}, & i = 1, \dots, N-1, \\ p(i, i-1) &= \frac{(i-1)\mu}{N\lambda + (i-1)\mu} = \frac{i-1}{N\rho + i - 1}, & i = 2, \dots, N, \\ p(i, a) &= \frac{i\lambda}{N\lambda + (i-1)\mu} = \frac{i\rho}{N\rho + i - 1}, & i = 1, \dots, N, \\ p(i, j) &= 0, & \text{otherwise,} \end{aligned}$$

with $\rho := \lambda/\mu$.

The transition matrix \mathbf{P} of the Markov chain **MC** can be written as

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

where $\mathbf{Q} = [p(i, j)]_{1 \leq i, j \leq N}$, $\mathbf{R} = (p(1, a), \dots, p(N, a))^T$, and $\mathbf{0}$ is the row vector of dimension N whose all components are equal to 0.

Define $\mathbf{M} = (\mathbf{I} - \mathbf{Q})^{-1}$, the *fundamental matrix* of the absorbing Markov chain **MC**. The (i, j) -entry of \mathbf{M} , denoted by $m(i, j)$, gives the expected number of visits to state j given that $I(0) = i$ [7, Chap. 11, Theorem 11.4].

The matrix \mathbf{M} is computed in explicit form in Lemma 2 in the Appendix I, and $m(i, j)$ is given in (17).

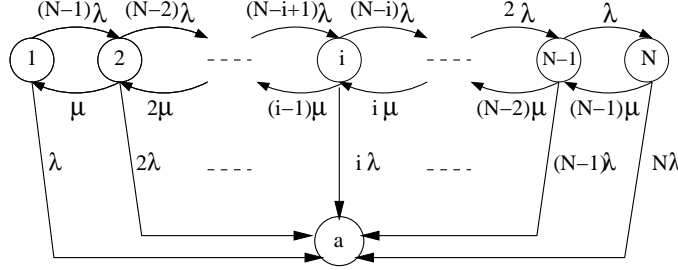


Figure 1: Transition rate diagram of the Markov chain $\{I(t), t \geq 0\}$.

We are now in position to compute the expected delivery delay, the distribution of the number of copies at the delivery instant, and the expected number of copies generated by the source.

3.1 Delivery delay

In this section we first determine $E_i[T_d]$, the expected delivery delay given that $I(0) = i \in \{1, 2, \dots, N\}$, from which the expected delivery delay $\bar{T}_d = E_1[T_d]$ will follow.

$E_i[T_d]$ is the expected time before absorption starting from the transient state i . Let n_{ij} be the number of visits to state j before absorption given that the chain starts in state i , and let T_{jl} be the sojourn time in state j at the l th visit to that state. Observe that $E[n_{ij}] = m(i, j)$, where $m(i, j)$ is given in Lemma 2, and that $E[T_{jl}] = 1/(N\lambda + \mu(j-1))$ for $j = 1, 2, \dots, N$ (see Figure 1). Hence,

$$E_i[T_d] = E \left[\sum_{j=1}^N \sum_{l=1}^{n_{ij}} T_{jl} \right] = \sum_{j=1}^N E \left[\sum_{l=1}^{n_{ij}} T_{jl} \right] = \sum_{j=1}^N m(i, j) E[T_{jl}],$$

where the last equality follows from Wald's identity, since $n_{i,j}$ is independent of the rvs $\{T_{jl}\}_{j,l}$. Plugging the value found in (17) for $m(i, j)$ in the latter equation gives

$$E_i[T_d] = -\frac{1}{\mu} \left(\binom{N-1}{i-1} \rho^{i-1} \right)^{-1} \sum_{k=1}^N \frac{\Psi_i^k}{z_k \Psi^k \tau^2 (\Psi^k)^T} \Psi^k \mathbf{1}^T, \quad (1)$$

with $\mathbf{1}^T$ the N -dimensional column vector whose all components are equal to 1. Quantities Ψ_i^k , τ and z_k are defined in Lemma 2 in the Appendix I.

More generally, the probability distribution of T_d , starting from $I(0) = i$ is given by

$$P_i(T_d < t) = 1 - \left(\binom{N-1}{i-1} \rho^{i-1} \right)^{-1} \sum_{k=1}^N \frac{\Psi_i^k}{\Psi^k \tau^2 (\Psi^k)^T} \Psi^k \mathbf{1}^T e^{z_k \mu t}.$$

The proof of this result is reported in Appendix II.

3.2 Number of copies in the system at delivery time

Let $P_i[C_d = j]$ be the probability that the number of copies in the network at the delivery time is j , given there are i copies in the network at time $t = 0$. We assume without loss of generality that the Markov chain \mathbf{MC} is left-continuous so that $P_i[C_d = j] = P[I(T_d-) = j]$ (by convention $I(t-)$ is the state of the process \mathbf{MC} just before time t). In words, $P_i[C_d = j]$ is the probability that the last visited state before absorption is j , given that the initial state is i .

If we split the absorbing state a into N absorbing states a_1, \dots, a_N , as shown in Figure 2, we will not affect the dynamics of the original Markov chain before absorption. This means that the fundamental matrix of the modified absorbed Markov chain is the same as the fundamental matrix of the original absorbed Markov chain. Clearly, $P_i[C_d = j]$ is now equal to the probability that the modified chain is absorbed in state a_j . Let b_{i,a_j} denote this probability. From the theory of absorbing Markov chains, we see that [7, Chap. 11, Theorem 11.6]

$$b_{i,a_j} = \sum_{k=1}^N m(i, k) r(k, a_j),$$

where $r(k, a_j)$ is the one-step transition probability from state k to the absorbing state a_j in the modified Markov chain. Clearly (see Figure 2) $r(k, a_j) = j\lambda / (N\lambda + (j-1)\mu) = j\rho / (N\rho + j-1)$ if $k = j$ and $r(k, a_j) = 0$ if $k \neq j$. Therefore,

$$P_i[C_d = j] = m(i, j) r(j, a_j) = -j \left(\binom{N-1}{i-1} \rho^{i-2} \right)^{-1} \sum_{k=1}^N \frac{\Psi_i^k \Psi_j^k}{z_k \Psi^k \tau^2 (\Psi^k)^T}.$$

The n th-order moment of C_d is equal to (with $\mathbf{J}_n := (1, \dots, i^n, \dots, N^n)$)

$$E_i[C_d^n] = - \left(\binom{N-1}{i-1} \rho^{i-2} \right)^{-1} \sum_{k=1}^N \frac{\Psi_i^k}{z_k \Psi^k \tau^2 (\Psi^k)^T} \Psi^k \mathbf{J}_{n+1}^T. \quad (2)$$

Coming back to the original problem, the probability distribution of the number of copies at delivery time is given by $P_1[C_d = j]$, and the n -th order moment is given by $E_1[C_d^n]$.

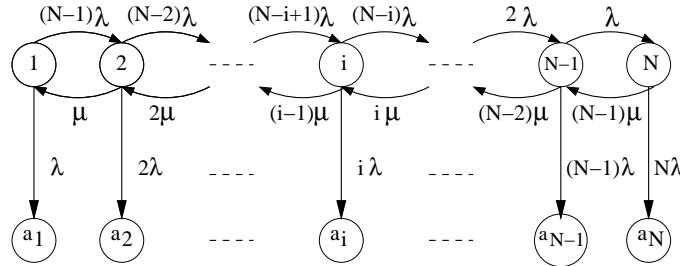


Figure 2: The modified absorbing Markov chain with N absorbing states.

3.3 Total number of copies generated by the source before delivery time

The objective is to find, $\overline{G_d}$, the expected number of copies generated by the source before the delivery time (or equivalently, before absorption). Let $G_d^{i,j}$ be the number of copies generated by the source before absorption given that the chain starts in state i and that state j is the last state visited before absorption (i.e. $I(T_d-) = j$). Introduce $J^{i,j}(k, k+1)$ (resp. $J^{i,j}(k+1, k)$) the number of transitions from state k (resp. state $k+1$) to state $k+1$ (resp. state k) given that $I(0) = i$ and $I(T_d-) = j$. It is easy to see that

$$J^{i,j}(k, k+1) = J^{i,j}(k+1, k) + \mathbf{1}_{\{i \leq k < j\}} - \mathbf{1}_{\{j \leq k < i\}}, \quad k = 1, 2, \dots, N-1. \quad (3)$$

A copy of the packet is generated by the source each time there is a transition from state k to state $k+1$ for all states $k = 1, 2, \dots, N-1$. Hence,

$$G_d^{i,j} = \sum_{k=1}^{N-1} J^{i,j}(k, k+1).$$

On the other hand, $n_{i,k}^j$, the total number of visits to state k given that $I(0) = i$ and $I(T_d-) = j$, satisfies the relation

$$\sum_{k=1}^N n_{i,k}^j = \sum_{k=1}^{N-1} J^{i,j}(k, k+1) + \sum_{k=1}^{N-1} J^{i,j}(k+1, k) + 1.$$

From (3) we find that

$$\sum_{k=1}^{N-1} J^{i,j}(k+1, k) = \sum_{k=1}^{N-1} J^{i,j}(k, k+1) + i - j.$$

Combining the three last identities gives

$$G_d^{i,j} = \frac{1}{2} \left[\sum_{k=1}^N n_{i,k}^j + j - i - 1 \right].$$

The expected number of copies given that $I(0) = i$, denoted by $\overline{G_d^i}$, is given by (Hint: remove the conditioning on $C_d = j$)

$$\overline{G_d^i} = \sum_{j=1}^N G_d^{i,j} P(C_d = j) = \frac{1}{2} \left[\sum_{k=1}^N m(i, k) + E_i[C_d] - i - 1 \right]$$

where $m(i, k)$ is given in Lemma 2 and $E_i[C_d]$ is given in (2) (with $n = 1$).

Finally, $\overline{G_d} = \overline{G_d^1}$. We will see in the next section how $\overline{G_d}$ can be used to compute the overall energy needed to transmit a packet to the destination.

4 Limiting the energy consumption

We will only consider the energy consumption due to packet transmission and decoding. Let p_t be the energy needed at the sender to transmit a packet to another node and let p_r be the energy needed at the receiver to decode a packet. The energy consumed by the source before the packet is delivered to the destination $P_s = p_t \overline{G_d}$, since the source needs to generate on the average $\overline{G_d}$ copies of the packet before one copy reaches the destination. The energy consumed by all nodes before the delivery time is given by $P_d = (p_t + p_r) \overline{G_d}$.

In this section we introduce and evaluate a new two-hop relay scheme that limits the energy consumption by limiting the number of copies that the source can generate before the packet reaches the destination. A similar scheme was introduced in [15] to limit the energy consumption of epidemic routing. We now assume that the source can generate at most K copies of the packet. In the following this scheme will be referred to as the K -limited two-hop relay protocol. Alike in the original protocol in Section 3 (corresponding to $K = \infty$), we will compute the expected delivery delay and the expected number of copies generated before the delivery time for the K -limited two-hop relay protocol. The behavior of the K -limited two-hop relay protocol can be modeled as a two-dimensional, finite-state, absorbing and continuous-time Markov chain (referred to as $\overline{\text{MC}}_K$) with state (i, c) , where $i \in \{1, 2, \dots, N\}$ gives the number of copies in the network, and $c \in \{0, 1, \dots, K\}$ records the total number of copies generated by the source. It is easy to see that the one-step probability transition matrix $\mathbf{P}_K = [p_K((i, c), \cdot)]$ of the absorbing, finite-state, discrete-time Markov chain (referred to as MC_K) embedded just before the jump times of $\overline{\text{MC}}_K$ is given by

$$\begin{aligned}
p_K((i, c), (i+1, c+1)) &= \frac{(N-i)\rho}{N\rho + (i-1)}, \quad i = 1, \dots, K_m, c = i-1, \dots, K-1, \\
p_K((i, c), (i-1, c)) &= \frac{(i-1)}{N\rho + (i-1)}, \quad i = 2, \dots, K_m, c = i-1, \dots, K-1, \\
p_K((i, c), a) &= \frac{i\rho}{N\rho + (i-1)}, \quad i = 1, \dots, K_m, c = i-1, \dots, K-1, \\
p_K((N, c), (N-1, c)) &= \frac{(N-1)}{N\rho + (N-1)} \mathbf{1}_{\{K \geq N\}}, \quad c = N-1, \dots, K-1, \\
p_K((N, c), a) &= \frac{N\rho}{N\rho + (N-1)} \mathbf{1}_{\{K \geq N\}}, \quad c = N-1, \dots, K-1, \\
p_K((i, K), (i-1, K)) &= \frac{(i-1)}{i\rho + (i-1)}, \quad i = 2, \dots, K_m + 1, \\
p_K((i, K), a) &= \frac{i\rho}{i\rho + (i-1)}, \quad i = 1, \dots, K_m + 1, \\
p_K((i, c), (j, d)) &= 0, \quad \text{otherwise,}
\end{aligned}$$

with $K_m := \min(K, N-1)$, and where a is the absorbing state. If $K \leq N-1$ then the total number of transient states is $L_1 := (K+1)(K+2)/2$ whereas if $K > N$ this number is equal to $L_2 := N(2K - N + 3)/2$.

If we label the transient states $(1, 0)$ as 1, $(2, 1)$ as 2, \dots , (i, c) as $\frac{(c+1)(c+2)}{2} - i + 1$ for $c \leq K_m$ and $i \leq c + 1, \dots, (i, c)$ as $\frac{N(2c-N+1)}{2} + N - i + 1$ for $c > K_m$ and $i \leq N$, then we can write the matrix \mathbf{P}_K as

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

where \mathbf{Q}_K is an L -by- L matrix giving the one-step transition probability between two transient states, \mathbf{R}_K is an L -by-1 matrix giving the one-step transition probability from a transient state to the absorbing state a , and $\mathbf{0}$ is the 1-by- L zero matrix.

The fundamental matrix associated with the absorbing Markov chain \mathbf{MC}_K is $\mathbf{M}_K = (\mathbf{I} - \mathbf{Q}_K)^{-1}$. Let $m_K(i, j)$ be the (i, j) -entry of \mathbf{M}_K . The matrix \mathbf{M}_K is an upper triangular matrix and, unlike in the original protocol in Section 3, it can no longer be obtained in explicit form.

Once the matrix \mathbf{M}_K has been (numerically) computed the main QoS metrics can easily be deduced, as shown below.

4.1 Expected delivery delay

We distinguish the cases $K \leq N - 1$ and $K \geq N$. In the former case, the expected delivery delay given that the chain starts in state $(1, 0)$ reads

$$\overline{T_d^K} = \sum_{j=1}^{L_1} m_K(1, j) E[T_j] = \frac{1}{\mu} \left(\sum_{i=1}^{K+1} \frac{m_K(1, L_1 - i + 1)}{i\rho + i - 1} + \sum_{i=1}^K \frac{\sum_{j=i-1}^{K-1} m_K(1, a(i, j))}{N\rho + i - 1} \right),$$

where $a(i, j) := 1 - i + \frac{(j+1)(j+2)}{2}$, and T_j is the sojourn time in the transient state with label j .

If $K \geq N$ we find

$$\overline{T_d^K} = \frac{1}{\mu} \left(\sum_{i=1}^N \frac{m_K(1, L_2 - i + 1)}{i\rho + i - 1} + \sum_{i=1}^N \frac{\sum_{j=i-1}^{N-1} m_K(1, a(i, j)) + \sum_{j=N}^{K-1} m_K(1, b(i, j))}{N\rho + i - 1} \right),$$

where $b(i, j) := N - i + 1 + \frac{N(2j-N+1)}{2}$.

4.2 Expected number of copies

The expected number of copies generated by the source before the delivery time, given that the chain starts in state $(1, 0)$ is given by

$$\overline{G_d^K} = \sum_{c=1}^K c \sum_{i=1}^{K_m} P(\text{absorption occurs in state } (i, c)).$$

If $K \leq N - 1$ then

$$\overline{G_d^K} = \rho \left[\sum_{c=1}^{K-1} \sum_{i=1}^{c+1} i c \frac{m_K(1, a(i, c))}{N\rho + i - 1} + K \sum_{i=1}^{K+1} i \frac{m_K(1, a(i, K))}{i\rho + i - 1} \right],$$

while if $K \geq N$

$$\overline{G}_d^K = \rho \left[\sum_{c=1}^{N-1} \sum_{i=1}^{c+1} ic \frac{m_K(1, a(i, c))}{N\rho + i - 1} + \sum_{c=N}^{K-1} \sum_{i=1}^N ic \frac{m_K(1, b(i, c))}{N\rho + i - 1} + K \sum_{i=1}^N i \frac{m_K(1, b(i, K))}{i\rho + i - 1} \right],$$

where $a(i, c)$ and $b(i, c)$ are defined in Section 4.1.

The energy consumed by the source before the packet is delivered to the destination is given by $p_t \overline{G}_d^K$ while the energy consumed by all nodes during this period is $(p_t + p_d) \overline{G}_d^K$.

5 Asymptotic analysis

In this section we derive asymptotic results for the expected delivery delay and the expected number of copies at delivery instant in the two-hop relay protocol when the number of nodes N is large. Deriving these asymptotic results from the explicit formulas in (1) and (2), respectively, is not easy (in the more simpler case when there are no timeouts getting asymptotics from the explicit results were already quite involved [8, Appendix A]).

We shall instead follow a mean field approach to find approximations of these asymptotics. The same approach was used in [14] and in [16] to derive asymptotic results for epidemic models.

The mean field approximation says that $X(t)$ (rep. $G(t)$), the *expected* number of copies (resp. of copies generated by the source) in the network at time t , before absorption, when N is large, can be approximated by the solution of the following 1st-order differential equations (see [12] for the general theory)

$$\dot{X}(t) = \lambda(N - X(t)) - \mu(X(t) - 1), \quad t > 0. \quad (4)$$

$$\dot{G}(t) = \lambda(N - X(t)), \quad t > 0. \quad (5)$$

The first equation simply reflects the fact that at time t $X(t)$ increases with the rate $\lambda(N - X(t))$ and decreases with the rate $\mu(X(t) - 1)$. However the second equation reflects the fact that at time t $G(t)$ is a increasing function with rate $\lambda(N - X(t))$. We need to complement these equations with another equation whose the solution approximates $D(t) := P(T_d < t)$, the probability distribution of the delivery delay. It was found in [14] that

$$\dot{D}(t) = \lambda X(t)(1 - D(t)), \quad t > 0. \quad (6)$$

Solving (4)-(6) with the initial conditions $X(0) = x_0$ ($x_0 = 1$ in our model), $G(0) = 0$, and $D(0) = 0$ yields

$$X(t) = \frac{N\lambda + \mu}{\lambda + \mu} + \left(x_0 - \frac{N\lambda + \mu}{\lambda + \mu} \right) e^{-(\lambda + \mu)t} \quad (7)$$

$$G(t) = \lambda N t - f_N(t) \quad D(t) = 1 - e^{-f_N(t)} \quad (8)$$

$$f_N(t) = \frac{\lambda}{\lambda + \mu} \left[(N\lambda + \mu)t + \left(x_0 - \frac{N\lambda + \mu}{\lambda + \mu} \right) (1 - e^{-(\lambda + \mu)t}) \right] \quad (9)$$

It can be checked that $D(0) = 0$, $\lim_{t \rightarrow \infty} D(t) = 1$ and $t \rightarrow D(t)$ is nondecreasing, so that $D(t)$ is indeed a probability distribution of a proper rv. As expected from the very definition of $X(t)$, we note that $X(\infty) = (N\lambda + \mu)/(\lambda + \mu)$ is the expected stationary number of customers in a finite-state birth and death process, with birth rate (resp. death rate) $\lambda(N - i)$ (resp. $\mu(i - 1)$) in state $i \in \{1, 2, \dots, N\}$.

5.1 Delivery delay

By definition, $E[T_d] = \int_0^{+\infty} P(T_d > t) dt$, so that from (7) $E[T_d]$ can be approximated by

$$E[T_d] \approx \int_0^{+\infty} e^{-f_N(t)} dt$$

when N is large. When N is large it is easily seen that the dominant contribution of $e^{-f_N(t)}$ to the above integral comes from small values of t since $f_N(t)$ is a nondecreasing function of N . We distinguish between the following two cases of x_0 : (1) x_0 is constant, (2) $x_0 = \alpha N$ with $\alpha \leq 1$.

x_0 constant: In this case, $e^{-f_N(t)}$ can be approximated by $e^{-f_N''(0)t^2/2}$ since $f_N(0) = 0$ and since $f_N'(0) = \lambda x_0$ does not depend on N , with $f_N''(0) = \lambda(N\lambda + \mu - (\lambda + \mu)x_0)$. For $0 \leq x_0 < X(\infty)$ this gives the 1st-order asymptotics

$$E[T_d] \approx \sqrt{\frac{\pi}{2\lambda(N\lambda + \mu - (\lambda + \mu)x_0)}} \approx \frac{1}{\lambda} \sqrt{\frac{\pi}{2N}} \quad (N \rightarrow \infty). \quad (10)$$

The 2nd-order asymptotics of $E[T_d]$ can be obtained by expanding $f_N(t)$ in Taylor series at the order three in the vicinity of $t = 0$. We find

$$\begin{aligned} E[T_d] &\approx \int_0^{+\infty} e^{-\frac{f_N''(0)}{2!}t^2} \left(1 - \frac{f_N^{(3)}(0)}{3!}t^3\right) dt \\ &= \sqrt{\frac{\pi}{2\lambda(N\lambda + \mu - (\lambda + \mu)x_0)}} + \frac{(\lambda + \mu)(N\lambda + \mu - (\lambda + \mu)x_0)}{3\lambda^3(N - 1)^2} \quad (N \rightarrow \infty). \end{aligned} \quad (11)$$

Figure 3 displays the 1st-order and 2nd-order asymptotics of $E[T_d]$, given in (10) and in (11), respectively, as a function of N , and compare them with the exact value obtained in (1). We observe that, as N increases, both asymptotics converge to the exact result.

x_0 linear with N : In this case $x_0 = \alpha N$ with $\alpha \leq 1$ is constant, $f_N(t)$ is approximated by $f_N(t) \approx f_N'(0)t$, since $f_N'(0)$ depend on N . So, this gives a 1st-order asymptotic

$$E[T_d] \approx \frac{1}{\lambda\alpha N} \quad (N \rightarrow \infty). \quad (12)$$

The 2nd-order asymptotics of $E[T_d]$ can be obtained by expanding $f_N(t)$ in Taylor series at the order three in the vicinity of $t = 0$. We find

$$E[T_d] \approx \frac{1}{\alpha\lambda N} - \frac{\lambda - (\lambda + \mu)\alpha}{\alpha^3\lambda^2 N^2} + \frac{(\lambda + \mu)(\lambda - (\lambda + \mu)\alpha)}{\alpha^4\lambda^3 N^3} \quad (N \rightarrow \infty) \quad (13)$$

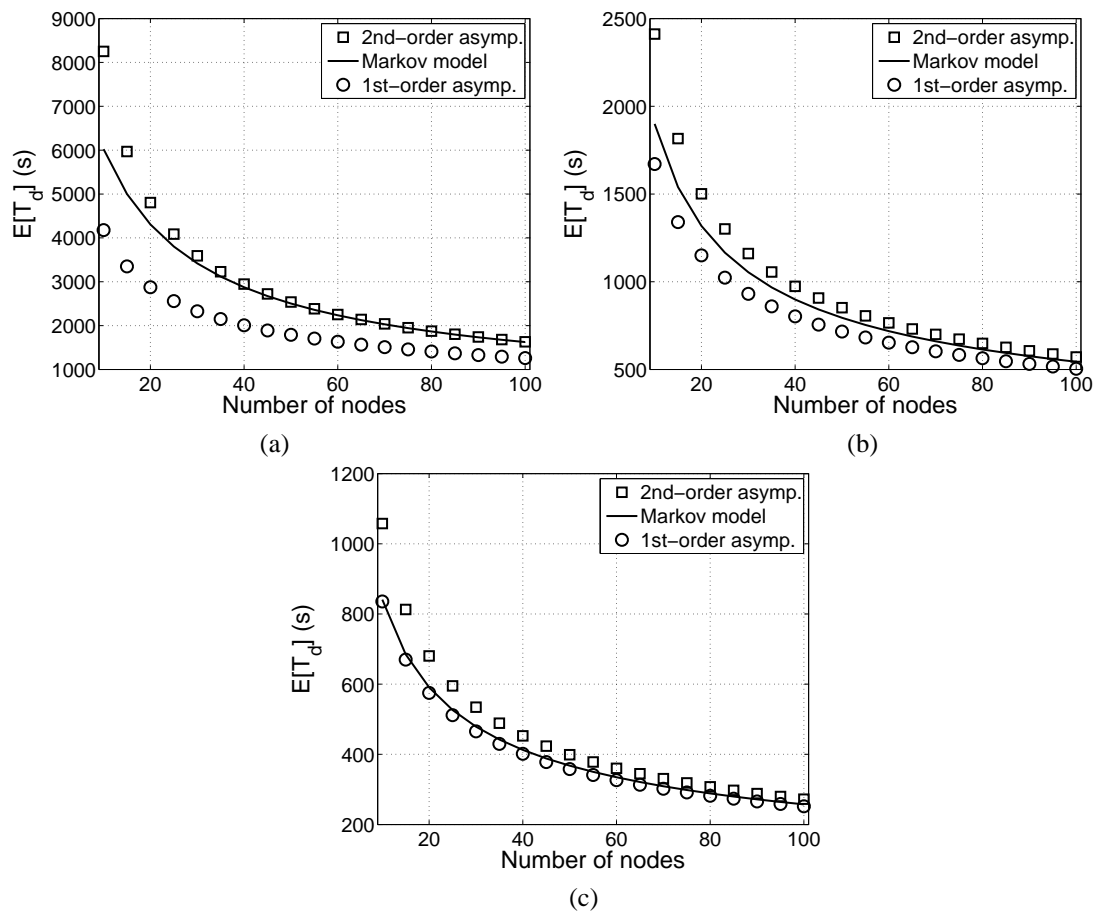


Figure 3: Comparing asymptotics for the expected delivery delay to the exact result ($\mu=0.001$: (a) $\lambda=0.0001$, (b) $\lambda=0.00025$, (c) $\lambda=0.0005$.)

Figures 4, and 5 displays the 1st-order and 2nd-order asymptotics of $E[T_d]$, given in (12) and in (13), respectively, as a function of α , and compare them with the exact value obtained in (1). We observe that, as N increases, both asymptotics converge to the exact result and that 2nd-order asymptotic is closer to the exact value than the 1st-order asymptotic.

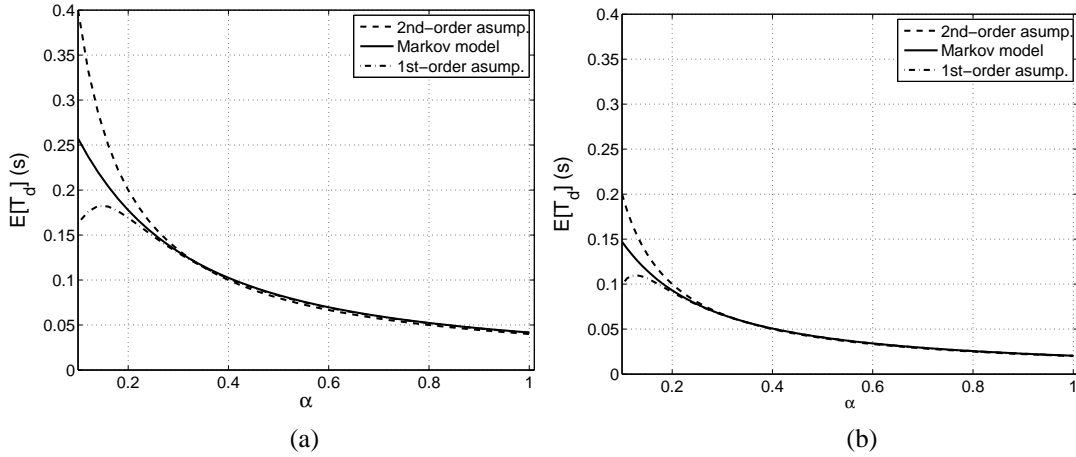


Figure 4: Comparing asymptotics for the expected delivery delay to the exact result as a function of α ($\mu=1$ and $\lambda = 0.5$): (a) $N=50$, (b) $N=100$)

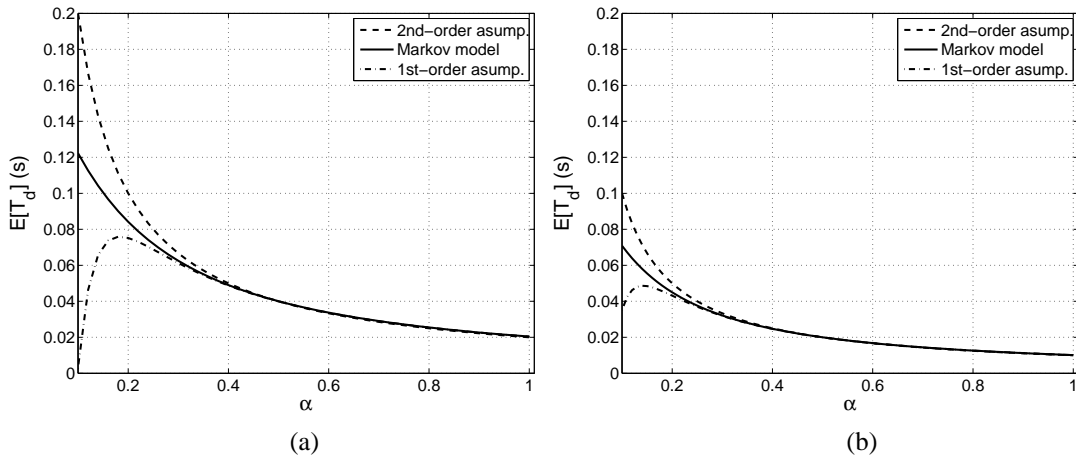


Figure 5: Comparing asymptotics for the expected delivery delay to the exact result as a function of α ($\mu=1$ and $\lambda = 1$): (a) $N=50$, (b) $N=100$)

Asymptotics, as N is large, for any order moment of T_d can be derived using a similar approach as follows.

Lemma 1 *When N is large, the n th-order moment of T_d is equal to*

$$\frac{E[T_d^n]}{E[T_d]} \approx \frac{n!}{(\lambda x_0)^{n-1}}. \quad (14)$$

◇

Proof: the n th-order moment of T_d is equal to

$$\begin{aligned} E[T_d^n] &= n \int_0^{+\infty} t^{n-1} (1 - P(t)) dt \\ &= \lambda \int_0^{+\infty} t^n f'_N(t) e^{-f_N(t)} dt \\ &= \lambda X(\infty) \int_0^{+\infty} t^n e^{-f_N(t)} dt + \lambda(x_0 - X(\infty)) \int_0^{+\infty} t^n e^{-f_N(t) - (\lambda + \mu)t} dt, \end{aligned} \quad (15)$$

where we integrate by parts and we used (9). When N is large the integral

$$\int_0^{+\infty} t^n e^{-f_N(t) - (\lambda + \mu)t} dt \approx \int_0^{+\infty} t^n e^{-f_N(t)} dt.$$

So the n th-order moment of T_d becomes equal to

$$E[T_d^n] \approx \lambda x_0 \int_0^{+\infty} t^n e^{-f_N(t)} dt = \frac{\lambda x_0}{n+1} E[T_d^{n+1}], \quad (16)$$

which by recurrence gives the Lemma. □

5.2 Expected number of copies at delivery instant

When N is large, $E[C_d]$, the mean number of copies at the delivery time T_d , is approximated by $\int_0^{+\infty} X(t) dD(t)$. With the use of (7) an integration by part gives

$$E[C_d] \approx x_0 + (N\lambda + \mu - (\lambda + \mu)x_0) \int_0^{\infty} e^{-f_N(t) - (\lambda + \mu)t} dt \quad (N \rightarrow \infty).$$

By using again the property that the dominant contribution of $e^{-f_N(t) - (\lambda + \mu)t}$ to the above integral comes from small values of t . we may approximate $e^{-f_N(t) - (\lambda + \mu)t}$ by $e^{-f_N''(0)t^2/2}$ (Hint x_0 fixed). Hence,

$$E[C_d] \approx x_0 + \sqrt{\frac{\pi}{2\lambda}} \sqrt{N\lambda + \mu - (\lambda + \mu)x_0} \approx \sqrt{\frac{\pi N}{2}} \quad (N \rightarrow \infty).$$

5.3 Expected number of copies

When N is large, $E[G_d]$, the mean number of copies generated by the source before delivery time T_d , is approximated by $\int_0^{+\infty} G(t)dD(t)$. With the use of (8) an integration by part gives

$$E[G_d] \approx \lambda N \int_0^{+\infty} e^{-f_N(t)} dt - 1.$$

Hence,

$$E[G_d] \approx \sqrt{\frac{\pi N}{2}} \quad (N \rightarrow \infty),$$

for x_0 fixed.

6 Validation and Numerical Results

In this section we first validate the Markov model introduced in Section 3 by comparing its performance (expected delivery delay) to that obtained by simulations, for two different mobility models (Random Waypoint (RWP) and Random Direction (RD) models). We then compare the expected delivery delay and the energy consumption induced by the two-hop relay protocol and the epidemic protocol. We conclude by investigating the performance of the K -limited two-hop relay protocol.

6.1 Model validation

We have simulated the two-hop relay protocol with exponential timeouts for both the RWP and the RD mobility models. In the RWP model [5] each node is assigned an initial location in a given area (typically a square) and travels at a constant speed to a destination chosen randomly in this area. The speed is chosen randomly in (v_{\min}, v_{\max}) , independently of the initial location and destination. After reaching the destination the node may pause for a random time, after which a new destination and speed are chosen, independently of previous speeds, destinations, and pause times. In the RD model [3] each node is assigned an initial direction, speed and travel time. The node then travels in that direction at the given speed and for the given duration. When the travel time has expired, the node may pause for a random time, after which a new direction, speed and travel time are chosen at random, independently of all previous directions, speeds and travel times. When a node reaches a boundary it is either reflected or the area wraps around so that the node reappears on the other side. In both mobility models nodes move independently of each other.

In our simulation settings, for both the RWP and the RD models the area is a square of side-length $L = 2000m$, the speed is constant and equals to $V = 10m/sec.$, there is no pause time, and the transmission range R is constant and the same for all nodes. In addition, in the RD model the travel time is constant and equals to $30sec.$ and the nodes reflect on reaching the boundaries. It has been experimentally observed in [8] that whenever $R \ll L$ then the node inter-meeting time is exponentially distributed with rate $\lambda = 10.94 \frac{RV}{\pi L^2}$ for RWP and $\lambda = 8 \frac{RV}{\pi L^2}$ for RD.

For different values of the ratio R/L (resp. N), Table 1 (resp. 2) reports the expected delivery delay obtained from the exact result in (1) and by simulations for the RWP model and the RD model, and give relative errors.

From the results in Tables 1-2 we conclude that, for both mobility models, the ratio R/L is a key factor. More accurate results are obtained for the RD model for a given ratio R/L and a given number of nodes (relative error of 6% when the ratio R/L is less than or equal to 1.25% for 20 nodes – see Table 2-A).

R/L (%)	1.25	1	0.5	0.1	1.25	1	0.5	0.1
$E_m[T_d]$ (s)	1216	1529	3154	20102	1678	2116	4416	30264
$E_{sim}[T_d]$ (s)	945	1245	2851	20861	1596	1988	4176	31651
$ 1 - \frac{E_{sim}[T_d]}{E_m[T_d]} $ (%)	22	18	10	4	6	6	5	4

(A)

(B)

Table 1: Expected delivery delay calculated from (1) and by simulations ($\mu = 0.0001$, $N = 20$): (A) RWP model, (B) RD model).

N	10	20	30	40	100	10	20	30	40	100
$E_m[T_d]$ (s)	4344	3154	2596	2257	1436	6116	4416	3622	3141	1987
$E_{sim}[T_d]$ (s)	4093	2851	2237	1839	1068	6208	4141	3297	2867	1512
$ 1 - \frac{E_{sim}[T_d]}{E_m[T_d]} $ (%)	6	9	14	18	26	1	6	9	9	24

(A)

(B)

Table 2: Expected delivery delay calculated from (1) and by simulations ($\mu = 0.0001$, $R/L = 0.5\%$): (A) RWP model, (B) RD model).

6.2 Comparison of two-hop relay and epidemic routing protocols

In this section, we compare the expected delivery delay, $E[T_d]$, and the expected number of packets transmitted, $E[G_d]$, as a function of μ , the timeout intensity, for the two-hop relay and the epidemic routing protocols. The absorbing Markov chain modeling the epidemic routing protocol is the same as the absorbing Markov chain in Section 3, except that the birth rate in state i is now equal to $\lambda i(N - i)$, since in the epidemic routing protocol all nodes are allowed to generate copies of the packet. The death rate (resp. absorption rate) in state i is unchanged and equal to $\mu(i - 1)$ (resp. λi). The computation of the expected delivery delay and the expected number of packets transmitted for the epidemic routing protocol is therefore similar to that carried out for the two-hop relay protocol, except that the fundamental matrix for the epidemic routing protocol cannot be computed in explicit form. This matrix was obtained numerically.

As expected, we observe that the epidemic routing protocol induces a smaller expected delivery delay than the two-hop relay protocol, but at the expense of a much more important overhead in terms of the number of copies generated. We also point out that the conclusions drawn from the results in Figure 6b apply for the energy consumptions P_s and P_d in the case where the energy to transmit (resp. decode) a packet is constant, since we have shown in Section 4 that in this case P_s and P_d are both linear functions of $E[G_d]$.

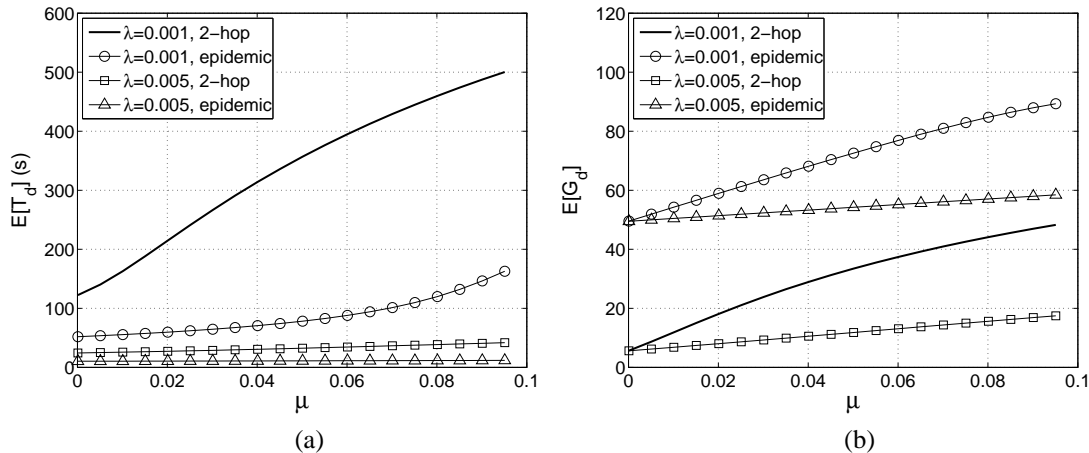


Figure 6: Expected delivery delay (left) and expected number of packet transmitted (right) for two-hop relay and epidemic routing protocols as a function of μ ($N = 100$).

6.3 Limited energy consumption

For different values of λ , the inter-meeting time rate, Figure 7a plots the expected delivery time, $\overline{T_d^K}$, under the K -limited two-hop relay protocol for different values of K , the maximum number of copies of the packet that the source may generate (see Section 4). For each λ , we observe there exists a threshold K_0 such that $\overline{T_d^K}$ is almost constant when $K \geq K_0$ ($K_0 \sim 20$ for $\lambda = 0.001$). Figure 7b plots the expected delivery time, $\overline{G_d^K}$, under the K -limited two-hop relay protocol for different values of K . Similarly to $\overline{T_d^K}$, We observe that there exists a threshold K_0 such that $\overline{T_d^K}$ is almost constant when $K \geq K_0$.

7 Concluding Remarks

In this work, we have evaluated the main QoS metrics of the two-hop relay protocol under the assumption that packets in relay nodes have a limited lifetime. Closed-form expressions have been derived for the probability distribution of the packet delivery delay, the expected number of copies in the system at the delivery instant, and the overall expected number of copies generated by the source at the delivery instant. We have observed that the latter metrics is directly related to the energy needed to transmit the packet to the destination node, in the case when the energy needed to transmit a packet between two nodes and the energy needed to decode a packet are constant. We have also proposed, and evaluated, a modification of the two-hop relay protocol that limits the number of copies of the packet that the source may generate.

In this paper our work has focused on the performance of the two-hop relay protocol before the destination receives the packet for the first time. It would also be interesting to quantify the impact of using an anti-packet mechanism on the total amount of energy consumed by the network during

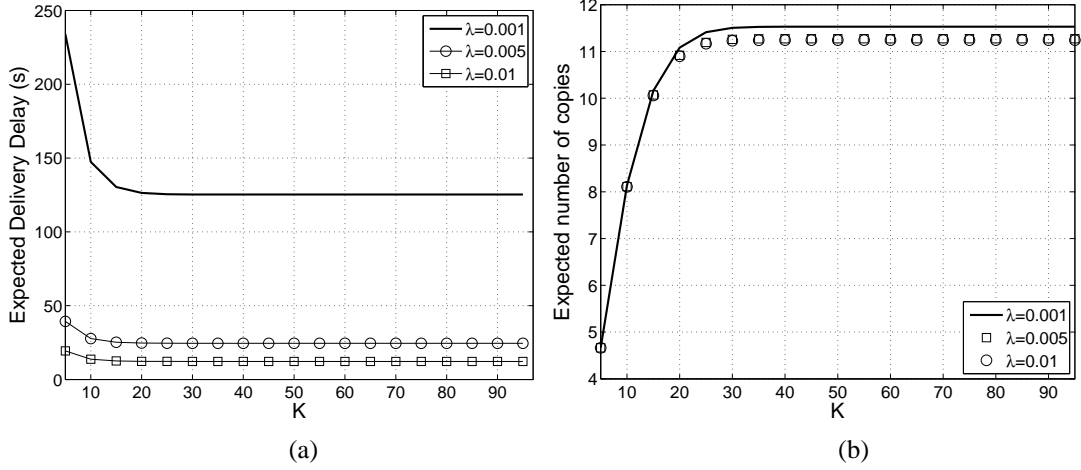


Figure 7: Expected delivery delay (left) and expected number of packet transmitted (right) under K -limited two-hop relay protocol for different values of K ($N=100$, $\mu=0.001$).

the entire lifetime of the packet, including its copies, in the network. Also, we have assumed that there is no timeout on the packet lifetime at the source. This assumption may not be realistic in some applications, and would therefore be worthwhile to relax it.

This study is part of a research effort towards developing *simple* analytical models for quantifying the performance of relay protocols for MANETs and, in particular, for better understanding the delay-energy tradeoff of this class of protocols.

Appendix I: The fundamental matrix

Lemma 2 The matrix $\mathbf{I}-\mathbf{Q}$ has N distinct, real, and strictly negative eigenvalues z_1, \dots, z_N given by

$$z_k = \frac{-N(2\rho + 1) + 1 - (N + 1 - 2k)\sqrt{4\rho + 1}}{2}, \quad k = 1, 2, \dots, N.$$

Therefore the fundamental matrix $\mathbf{M} = (\mathbf{I} - \mathbf{Q})^{-1}$ exists, and its (i,j) -entry is given by

$$m(i, j) = -\frac{N\rho + j - 1}{\binom{N-1}{i-1}\rho^{i-1}} \sum_{k=1}^N \frac{\Psi_i^k \Psi_j^k}{z_k \Psi^k \tau^2 (\Psi^k)^T}, \quad (17)$$

with $\Psi^k = (\Psi_1^k, \dots, \Psi_N^k)$ where

$$\Psi_i^k = \sum_{l=\max(0, i-1-N+k)}^{\min(i-1, k-1)} \binom{k-1}{l} \binom{N-k}{i-1-l} (-1)^{i-1} x_1^{k-1-l} x_2^{N-k-i+1+l},$$

$$x_1 = \frac{-1 - \sqrt{1+4\rho}}{2\rho}, \quad x_2 = \frac{-1 + \sqrt{1+4\rho}}{2\rho},$$

and where $\tau = \text{diag}(\tau_1, \dots, \tau_N)$, with $\tau_i = \left(\binom{N-1}{i-1} \rho^{i-1}\right)^{-1/2}$. \diamond

Proof. To simplify the computation of \mathbf{M} we introduce the matrix \mathbf{A} defined as

$$\mathbf{A} = -\mathbf{B}(\mathbf{I} - \mathbf{Q}), \quad (18)$$

where $\mathbf{B} = \text{diag}(b(1), \dots, b(N))$ with $b(i) = N\rho + (i-1)$. Matrices \mathbf{M} and \mathbf{A} are related through the simple identity $\mathbf{M} = -\mathbf{A}^{-1}\mathbf{B}$. In the following we will compute \mathbf{A}^{-1} . We will follow the approach developed in [1]. We first compute the eigenvalues and left/right eigenvectors of \mathbf{A} .

Eigenvalues of \mathbf{A} .

Let z be some eigenvalue of \mathbf{A} and let $\Psi = (\Psi_1, \dots, \Psi_N)$ be the associated left eigenvector. That is, $\Psi\mathbf{A} = z\Psi$, or equivalently,

$$\rho(N - (i-1))\Psi_{i-1} - (\rho N + i - 1 + z)\Psi_i + i\Psi_{i+1} = 0 \quad (19)$$

for $i = 1, \dots, N$, with $\Psi_0 = \Psi_{N+1} = 0$ by convention. Let $\psi(x) = \sum_{j=1}^N \Psi_j x^j$ denote the generating function of Ψ . Multiplying (19) by x^i and then summing over i yields

$$\frac{\psi'(x)}{\psi(x)} = \frac{\rho N x - (\rho N - 1 + z) - 1/x}{\rho x^2 + x - 1}. \quad (20)$$

Let the zeros of $x^2 + x/\rho - 1/\rho$ be $x_1 = \frac{-1 - \sqrt{1+4\rho}}{2\rho}$ and $x_2 = \frac{-1 + \sqrt{1+4\rho}}{2\rho}$. The unique solution of (20) such that $\Psi_N = 1$ is

$$\psi(x) = x(x_1 - x)^{c_1} (x_2 - x)^{c_2} \quad (21)$$

$$c_1 := \frac{x_1^2 \rho N - x_1(\rho N - 1 + z) - 1}{\rho x_1(x_1 - x_2)}, \quad c_2 := \frac{-x_2^2 \rho N + x_2(\rho N - 1 + z) + 1}{\rho x_2(x_1 - x_2)}.$$

It is easily seen that $c_1 + c_2 = N - 1$ (Hint: use $x_1 x_2 = -1/\rho$), so that (21) also writes

$$\psi(x) = x(x_1 - x)^{c_1} (x_2 - x)^{N-1-c_1}. \quad (22)$$

Because $\psi(x)$ is a polynomial of degree N , we observe from (22) that necessarily c_1 is an integer lying in the set $\{0, 1, \dots, N-1\}$ since x_1 and x_2 are always distinct.

The equations $c_1 = k - 1$ for $k = 1, \dots, N$ give the following N eigenvalues of \mathbf{A} :

$$z_k = \frac{-N(2\rho + 1) + 1 - (N + 1 - 2k)\sqrt{4\rho + 1}}{2}, \quad k = 1, \dots, N. \quad (23)$$

All eigenvalues of \mathbf{A} are distinct (obvious from (23)). Furthermore, z_k increases as k increases, and it is easily seen that $z_N < 0$ for $\rho > 0$. Thus, $z_k < 0$ for all $k = 1, \dots, N$.

Left eigenvectors of \mathbf{A} .

Recall that $\Psi^k = (\Psi_1^k, \dots, \Psi_N^k)$ is the left eigenvector associated with the eigenvalue z_k of \mathbf{A} . The i th component Ψ_i^k of the eigenvector Ψ^k is the coefficient of x^i in the polynomial $x(x_1 - x)^{k-1}(x_2 - x)^{N-k}$ that is

$$\Psi_i^k = (-1)^{i-1} x_1^{k-1} x_2^{N-k-i+1} \sum_{l=\max(0, i-1-N+k)}^{\min(i-1, k-1)} \binom{k-1}{l} \binom{N-k}{i-1-l} \left(\frac{x_2}{x_1}\right)^l.$$

Right eigenvectors of \mathbf{A} .

Recall that $\Phi^k = (\Phi_1^k, \dots, \Phi_N^k)^T$ is the right eigenvector associated with the eigenvalue z_k , for $k = 1, \dots, N$. We proceed like in [1, Section 2.4], that is we look for a diagonal matrix $\tau = \text{diag}(\tau_1, \dots, \tau_N)$ such that

$$\tau^{-1} \mathbf{A} \tau = (\tau^{-1} \mathbf{A} \tau)^T. \quad (24)$$

It is easily found that (Hint: solve $\tau_i^2/\tau_{i+1}^2 = \rho(N-i)/i$ for $i = 1, \dots, N$ with $\tau_1 = 1$)

$$\tau_i = \left(\binom{N}{i} \frac{i}{N} \rho^{i-1} \right)^{-1/2}, \quad i = 1, \dots, N$$

satisfy (24). The identity $\Psi^k \mathbf{A} = z_k \Psi^k$ implies that $\Psi^k \tau (\tau^{-1} \mathbf{A} \tau) = z_k \Psi^k \tau$. Therefore, $\Psi^k \tau$ is a left eigenvector of the matrix $\tau^{-1} \mathbf{A} \tau$ associated with the eigenvalue z_k . Since the matrix $\tau^{-1} \mathbf{A} \tau$ is symmetric, it has identical left and right eigenvectors. Hence, $(\tau^{-1} \mathbf{A} \tau) (\Psi^k \tau)^T = z_k (\Psi^k \tau)^T$ which gives that $\mathbf{A} \tau^2 (\Psi^k)^T = z_k \tau^2 (\Psi^k)^T$. This shows that $\alpha_k \tau^2 (\Psi^k)^T$ is a right eigenvector associated with the eigenvalue z_k for any constant $\alpha_k \neq 0$.

Without loss of generality we select the constants $\alpha_1, \dots, \alpha_N$ so that $\Psi^k \Phi^k = 1$ for every $k = 1, \dots, N$. Hence, $\alpha_k = 1/\Psi^k \tau^2 (\Psi^k)^T$ for $k = 1, \dots, N$. Finally, $\Phi^k = \tau^2 (\Psi^k)^T / \Psi^k \tau^2 (\Psi^k)^T$, or equivalently

$$\Phi_i^k = \frac{1}{\Psi^k \tau^2 (\Psi^k)^T} \left(\binom{N}{i} \frac{i}{N} \rho^{i-1} \right)^{-1} \Psi_i^k, \quad i = 1, \dots, N. \quad (25)$$

The proof is concluded by noting that

$$\hat{a}(i, j) = \sum_{k=1}^N \frac{\Phi_i^k \Psi_j^k}{z_k} = \binom{N-1}{i-1} \rho^{i-1} \sum_{k=1}^N \frac{\Psi_i^k \Psi_j^k}{z_k \Psi^k \tau^2 (\Psi^k)^T}, \quad (26)$$

by using (25). Equation 26 together with $m(i, j) = -(N\rho + (j-1))\hat{a}(i, j)$ (coming from $\mathbf{M} = -\mathbf{A}^{-1}\mathbf{B}$) gives (17). \square

Remark 7.1 The solution to (20) is not unique, since if $\psi(x)$ is a solution then $\alpha\psi(x)$ is also a solution for any constant $\alpha \neq 0$. This is of course related to the fact that eigenvectors are uniquely defined up to a multiplicative constant. Taking $\alpha = 1$ (like we implicitly did in (22)) implies that the N th component, Ψ_N , of the left eigenvector Ψ is taken to be equal to 1 (coefficient of x^N in (22)).

Remark 7.2 Replacing x by 1 in (20) implies the following relation between the eigenvalue z_k and its corresponding left eigenvector Ψ^k

$$\sum_{j=1}^N j\Psi_j^k = -\frac{z_k}{\rho} \sum_{j=1}^N \Psi_j^k, \quad (27)$$

that is true for $1 \leq k \leq N$.

Appendix II: Distribution of delivery delay

The delivery delay, T_d , given that there are i copies in the network at time 0 is the time to absorption of the Markov chain **MC** of Figure 1, given that $I(0) = i$. In order to compute the distribution of T_d , $P_i(T_d \leq t)$, given that $I(0) = i$, first we derive the Laplace Steiltjes Transform (LST) given that $I(0) = i$, $f_i(s) = E_i[e^{-sT_d}]$, and next we invert $f_i(s)$. $f_i(s)$ will actually hold for any complex number s such that $\Re(s) \geq 0$. Starting at state i , if we condition on the next possible transition of the **MC**, $f_i(s)$ reads

$$f_i(s) = \frac{\lambda i}{\lambda N + \mu(i-1) + s} + \frac{\mu(i-1)}{\lambda N + \mu(i-1) + s} f_{i-1}(s) + \frac{\lambda(N-i)}{\lambda N + \mu(i-1) + s} f_{i+1}(s) \quad (28)$$

for $i = 1, \dots, N$ (by convention $f_0(\cdot) = f_{N+1}(\cdot) = 1$).

Multiplying both sides of (28) by $\lambda N + \mu(i-1) + s$ and dividing by μ , yields

$$(i-1)f_{i-1}(s) - \left(\rho N + i - 1 + \frac{s}{\mu}\right) f_i(s) + \rho(N-i)f_{i+1}(s) = -\rho i \quad (29)$$

for $i = 1, \dots, N$, with $\rho := \lambda/\mu$.

Let $\mathbf{f} = (f_1(s), \dots, f_N(s))^T$.

In matrix form (29) writes

$$\left(\mathbf{A} - \frac{s}{\mu}\mathbf{I}\right) \mathbf{f} = \mathbf{b}, \quad s \geq 0, \quad (30)$$

with \mathbf{I} the N -by- N identity matrix, $\mathbf{b} := (-\rho, -2\rho, \dots, -\rho N)^T$, and \mathbf{A} is the N -by- N matrix defined in (18) in Appendix I. It is shown in Appendix I that the matrix \mathbf{A} is invertible and diagonalizable, namely, there exists an invertible matrix \mathbf{F} such that

$$\mathbf{A} = \mathbf{F} \text{diag}(z_1, \dots, z_N) \mathbf{F}^{-1}.$$

where z_1, \dots, z_N are the eigenvalues of \mathbf{A} , the j th right eigenvector of \mathbf{A} , Φ^j , is the j th column of the matrix \mathbf{F} , and the left eigenvector, Ψ^i , is the i th row of the matrix \mathbf{F}^{-1} . Hence

$$\mathbf{A} - (s/\mu)\mathbf{I} = \mathbf{F} \text{diag} (z_1 - s/\mu, \dots, z_N - s/\mu) \mathbf{F}^{-1},$$

and

$$(\mathbf{A} - (s/\mu)\mathbf{I})^{-1} = \mathbf{F} (\text{diag} (1/(z_1 - s/\mu), \dots, 1/(z_N - s/\mu))) \mathbf{F}^{-1} \quad (31)$$

provided that $z_k - s/\mu \neq 0$ for $k = 1, \dots, N$. Since, we know from Appendix I, that the eigenvalues of \mathbf{A} are all strictly negative, so that the right-hand side of (31) is well defined (in particular) for all $s \geq 0$. Therefore, (cf. (30))

$$\mathbf{f} = \mathbf{F} (\text{diag} (1/(z_1 - s/\mu), \dots, 1/(z_N - s/\mu))) \mathbf{F}^{-1} \mathbf{b}, \quad s \geq 0. \quad (32)$$

Since Φ^j is the j th column of the matrix \mathbf{F} , and Ψ^i is the i th row of the matrix \mathbf{F}^{-1} , we see from (32) that the i th component, $f_i(s)$, of the vector \mathbf{f} is given by

$$f_i(s) = -\rho \sum_{j=1}^N \sum_{k=1}^N \frac{j \Phi_i^k \Psi_j^k}{z_k - s/\mu}, \quad s \geq 0. \quad (33)$$

Closed-form expressions for the eigenvalues and right/left eigenvectors of the matrix \mathbf{A} are provided in Lemma 2. Lemma 2 and Remark 7.2 gives that

$$f_i(s) = \left(\binom{N}{i} \frac{i}{N} \rho^{i-1} \right)^{-1} \sum_{k=1}^N \frac{\Psi_i^k (\Psi^k \mathbf{1}^T)}{\Psi^k \tau^2 (\Psi^k)^T} \frac{z_k}{z_k - s/\mu} \quad s \geq 0, \quad (34)$$

where $\mathbf{1}^T$ is the column vector of dimension N whose all components are equal to 1, one. Equation 34 implies that the distribution of T_d given that the MC starts at state 1 reads

$$P_i(T_d < t) = 1 - \left(\binom{N-1}{i-1} \rho^{i-1} \right)^{-1} \sum_{k=1}^N \frac{\Psi_i^k}{\Psi^k \tau^2 (\Psi^k)^T} \Psi^k \mathbf{1}^T e^{z_k \mu t}. \quad (35)$$

Since $z_k < 0$ for $1 \leq k \leq N$ and $\rho > 0$, thus $\lim_{t \rightarrow +\infty} P_i(T_d > t) = 0$. Further, $P_i(T_d > 0) = 1$ (Hint: $\Psi^i \Phi^j = 0$ for $i \neq j$, and $\Psi^k \Phi^k = 1$).

References

- [1] D. Anick, D. Mitra and M. M. Sondhi, Stochastic Theory of a Data-Handling System with Multiple Sources, *Bell System Technical Journal*, Vol. 61, No. 8, October 1982, pp. 1871-1896.
- [2] F. Baccelli and P. Brémaud, *Palm Probabilities and Stationary Queues*. Springer-Verlag, 1987.
- [3] C. Bettstetter, Mobility Modeling in Wireless Networks: Categorization, Smooth Movement, Border Effects, *ACM Mobile Computing and Communications Review*, **5**, 3, pp. 55-67, July 2001.

-
- [4] C. Bettstetter, H. Hartenstein, and X. Pérez-Costa. Stochastic Properties of the Random Way-point Mobility Model, *ACM/Kluwer Wireless Networks, Special Issue on Modeling and Analysis of Mobile Networks*, Vol. 10, No. 5, pp. 555-567, Sept 2004.
- [5] J. Broch, A. D. Maltz, B. D. Johnson, Y.-C.Hu and Jetcheva, A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols *Proc. of ACM MOBICOM 1998*, Dallas, TX, USA, pp. 85-97. October 25-30, 1998.
- [6] A. El Gamal, J. Mammen, B. Prabhakar and D. Shah, Throughput-Delay Trade-off in Wireless Networks, *Proc. of INFOCOM 2004*, Hong Kong, 2004.
- [7] C. Grinstead and J. Snell, *Introduction to Probability*. American Mathematical Society, 1997.
- [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, October 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, August, 2002, pp. 477-486.
- [10] P. Gupta and P. R. Kumar, The Capacity of Wireless Networks, *IEEE/ACM Transactions on Information Theory*, Vol. 46, No. 2, March, 2000, pp. 388-404.
- [11] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
- [12] T. G. Kurtz, Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes, *J. Applied Prob.*, Vol. 7, pp. 49-58, 1970.
- [13] P. Nain, D. Towsley, B. Liu and Z. Liu, Properties of Random Direction Models, *Proc. of INFOCOM 2005*, Miami, FL, USA, March 2005.
- [14] T. Small and Z. J. Haas The Shared Wireless Infostation Model: A New Ad Hoc Networking Paradigm, *Proc. of MOBIHOC*, Anapolis, MD, USA, 2003.
- [15] T. Small and Z. J. Haas, Resource and Performance Tradeoffs in Delay-Tolerant Wireless Networks, *Proc of ACM SIGCOMM'05 Workshop on Delay Tolerant Networking*, Philadelphia, PA, USA, August 2005.
- [16] E. Zhang and G. Neglia and J. Kurose and D. Towsley, Performance Modeling of Epidemic Routing, *Umass Computer Science Technical Report 2005-44*.



Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique que
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)

<http://www.inria.fr>

ISSN 0249-6399