Performance Modeling for Relay Cooperation in Delay Tolerant Networks

Performance Modeling for Relay Cooperation in Delay Tolerant Networks

Jiajia Liu, Xiaohong Jiang, Hiroki Nishiyama, and Nei Kato

***Abstract*—Delay tolerant networks (DTNs) rely on the mobility of nodes and sequences of their contacts to compensate for lack of continuous connectivity and thus enable messages to be delivered from end to end in a "store-carry-forward" way, where multiple relay nodes are usually employed in the message delivery process. In this paper, we focus on such relay cooperation and analytically explore its impact on the delivery performance in DTNs. Specifically, we first develop a continuous time Markov chain-based theoretical framework to model the complicated message delivery process in delay tolerant networks adopting the two-hop relay algorithm. We then derive closed-form expressions for both the expected delivery delay and the corresponding expected delivery cost, where the important relay behaviors of forwarding traffic for itself or for other nodes are carefully incorporated into the analysis.**

***Index Terms*—Delay tolerant networks, relay cooperation, two-hop relay, delivery delay, delivery cost.**

I. Introduction

Delay tolerant networks (DTNs) are sparsely distributed and highly mobile wireless ad hoc networks, where the transmission opportunities come up and down from time to time, and no contemporaneous end-to-end path may ever exist at any given time instant [1]–[3]. The traditional route-based routing algorithms proposed for mobile ad hoc networks, such as DSR [4], AODV [5], etc., fail to work properly, as they require the simultaneous availability of a number of links.

The "store-carry-forward" kind of routing relies on the mobility of nodes and sequences of their contacts to compensate for lack of continuous connectivity and thus enable messages to be delivered from end to end. Therefore, it is believed that the "store-carry-forward" routing is a promising alternative for DTN scenarios and will become a natural routing option for the DTN routing [6]–[9]. Among these "store-carry-forward" routing protocols, the two-hop relay and its variants [3], [10]–[12] have become a class of attractive routing protocols due to its efficiency and simplicity. In the two-hop relay routing, the source transmits copies of its packets to all mobiles (relays) it encounters; relays transmit the packets only if they come in contact with the destination [13]–[18]. Since the source will also transmit its packets directly to the destination every time such transmission opportunity arises, each packet travels at most two hops to reach its destination as shown in Fig. 1.

J. Liu, H. Nishiyama and N. Kato are with the Graduate School of Information Sciences, Tohoku University, Aobayama 6-3-09, Sendai, 980-8579, JAPAN. E-mail: {liu-jia,kato}@it.ecei.tohoku.ac.jp.

X. Jiang is with the School of Systems Information Science, Future University Hakodate, Kamedanakano 116-2, Hakodate, Hokkaido, 041-8655, JAPAN. E-mail: jiang@fun.ac.jp.

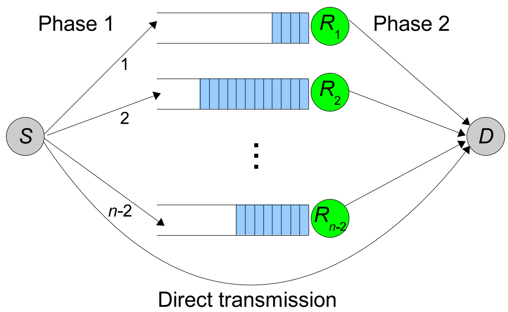


Fig. 1. Illustration of the two-hop relay routing protocol, where the destination node D receives a packet either directly from the source node S or from one of the $n-2$ distinct relay nodes.

It is easy to see that when operating under the two-hop relay routing, multiple relay nodes are usually employed in the packet delivery process of a tagged traffic flow (as shown in Fig. 1, for the tagged flow there all other nodes except the source and the destination may be employed as relays). Actually, the performances of the two-hop relay routing, such as the delivery delay (the time it takes a packet to reach its destination) and the delivery cost (the total number of transmissions until a packet is delivered to the destination) rely heavily on such relay cooperation behaviors. Some initial work [19]–[22] has been done to derive the delivery delay and delivery cost performances for two-hop relay routing, where all other nodes are assumed to be perfect relays in the packet delivery process and will forward the packets in a cooperative and altruistic way.

Consequently, such cooperative and mutually helping routing inflicts significant energy consumption and storage cost on each node. In the real world, however, as mobile nodes are usually not only energy-constrained but also storage-limited, the intermediate relay nodes may act selfishly. It is noticed that recently, some interesting works has been done to address the important node selfishness issue in relay cooperation and explore its impact on the delivery performance of two-hop relaying. Panagakis *et al.* in [23] experimentally examined the effect of node cooperation on the message delivery delay and the transmission overhead incurred until message delivery or the termination of the message spreading process, where a node may choose to probabilistically drop a newly received message or refuse to forward a buffered message. Karaliopoulos *et al.* in [24] further assumed a specific group of selfish relay nodes with such individual selfishness behaviors and analytically evaluated its impact on the expected message delivery delay. Li *et al.* in [25], [26] obtained explicit expressions for the system performances of message delivery delay and delivery cost in a social selfishness scenario where there are two groups of relay nodes, and a relay node has greater incentive to help forward messages from the nodes in the same group, but less interests to forward the messages from nodes of the other group [27], [28].

It is noticed that all the available works in literature [19]–[25] suffered from the same limitation that all of them considered a very simple network scenario with only a single source-destination pair. Under such scenario, all the other nodes (except the source and the destination) act as "pure" relays, and have only one kind of selfish behavior, to either carry and forward messages for the source or not. In the DTNs, however, there may simultaneously co-exist multiple source-destination pairs (traffic flows). Each node may act not only as a relay carrying and forwarding messages for other nodes, but also as a source trying to deliver out its locally generated message. Thus, a node may become more willing to forward its own message rather than that of others when it encounters some node. This kind of selfish behaviors may become much more significant when the nodes are operating under both QoS requirements (e.g., delivery delay requirements) and energy consumption constraints. In this paper, we focus on this kind of node selfishness in relay cooperation and analytically explore how it will influence the delivery performance of the two-hop relay routing in the challenging DTNs.

The main contributions of this paper are summarized as follows.

- We focus on a DTN scenario where each node has a locally generated traffic destined for some node and also an incoming traffic originated from some other node, and develop a continuous time Markov chain-based theoretical framework to model the complicated packet delivery process in such network scenarios.
- With the help of the developed theoretical framework, we further derive closed-form expressions for both the expected delivery delay and the expected delivery cost, where the important relay behaviors of forwarding traffic for itself or for other nodes, are carefully incorporated into the analysis.
- Finally, we provide extensive numerical results to validate our theoretical framework and explore how the node selfishness in relay cooperation and network size will influence both the expected delivery delay and the expected delivery cost of two-hop relay routing.

The rest of this paper is outlined as follows. Section II introduces the system models. In Section III, we develop the continuous time Markov chain-based theoretical framework and derive closed-form expressions for both the expected delivery delay and the expected delivery cost. We provide extensive numerical results in Section IV and conclude this paper in Section V.

II. System Model

Network Model: We consider a delay tolerant network with n mobile nodes. We assume that two nodes are able to communicate with each other only when they are within reciprocal transmission range. Similar to [11]–[13], [15]–[22], we consider a limited channel bandwidth and assume that the number of bits that can be successfully transmitted during each contact duration between any node pair is fixed and normalized to one packet here, such that a packet can be successfully transmitted during a contact. Every time two nodes meet each other, we assume that either node has the same probability of $1/2$ to become the transmitting node.

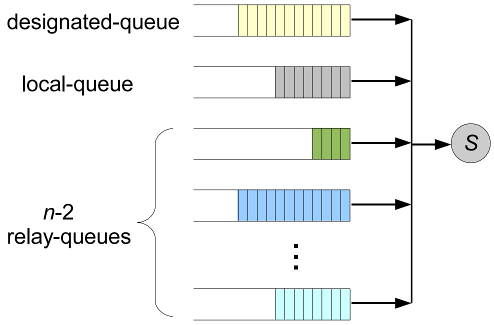


Fig. 2. Illustration of the queue structure at the buffer of node S, which contains one designated-queue for packets destined for itself, one local-queue for its locally generated traffic and $n-2$ parallel relay-queues for traffic of other flows.

Traffic Pattern: In order to fully capture the node selfishness behavior of forwarding traffic for itself or for other nodes, we assume here the permutation traffic pattern widely adopted in previous studies [10], [15], [29]–[33]. Under such a traffic pattern, each node has a locally generated traffic destined for some node and also an incoming traffic originated from some other node, i.e., each node is not only the source of its own traffic flow but also the destination of some other traffic flow. Thus, there are in total n distinct traffic flows inside the whole network.

Queue Structure: Without loss of generality, we focus on a tagged flow and denote its source and destination by node S and node D, respectively. According to the two-hop relay [10], [14], [15], [29], S can also be a potential relay for other $n-2$ flows (except the two flows originated from and destined for itself). As illustrated in Fig. 2, we assume that S maintains n individual queues at its buffer, one designated-queue for storing the packets that are destined for itself, one local-queue for storing the packets that are locally generated and destined for node D, and $n-2$ parallel relay-queues for storing packets destined for other $n-2$ nodes (excluding S and D).

It is noticed that for a general packet P generated at S, there may exist remnant copies of P carried by relay nodes after P has been received by D. In order to remove such remnant copies of already received packets, we adopt a packet sequence number based mechanism for the two-hop relaying [15], [29]. For the tagged flow, S labels each packet P waiting in the local-queue with a *send number* $SN(P)$, such that P can be easily retrieved from the queue buffers by other relay nodes. Similarly, D maintains a *request number* $RN(D)$ to indicate the send number of the packet it is currently requesting, such that all packets are received in order at D.

Mobility Model: We assume that the node inter-meeting times, i.e., the time elapsed between two consecutive contacts of a given node pair, are exponentially distributed with inter-meeting intensity λ. The validity of this assumption has been discussed in [34], and it has been demonstrated to be fairly accurate for a number of mobility models, like Random Direction, Random Waypoint and Random Walker, [11], [13], [22]. As shown in [34], the inter-meeting intensity λ can be determined by

$$\lambda = \frac{8\omega R\nu}{\pi L^2} \tag{1}$$

where R refers to the transmission range of each node and is small enough with respect to the length L of square area, i.e., $R \ll L$, ν is the scalar velocity of nodes and the constant $\omega = 1$ (resp. $\omega = 1.368$) for the Random Direction (resp. Random Waypoint) mobility model.

Node Behavior: For the tagged flow, every time the source S encounters some node R (rather than the destination D), if S wins the opportunity to transmit a packet then S has two choices: to either act as a source delivering to R a copy of its local-queue packet, or act as a relay delivering to R a packet in the relay-queue specified for R. The former choice expedites the delivery process of its own traffic flow (destined for D), while the latter improves the information delivery process of other flow (destined for node R). It is notable that a node may become more willing to forward its own traffic rather than that of others when it encounters some node. This kind of selfish behaviors may become much more significant when mobile nodes are operating under both QoS requirements (e.g., delivery delay requirements) and energy consumption constraints. To analytically characterize the impact of such behavior on the delivery performance of two-hop relaying, we assume that S will deliver its own packet to R with probability p, and deliver to R a packet (if available) destined for R with probability $1-p$, $0 \le p \le 1$.

III. Markovian Analysis

In this section, we first develop a continuous time Markov chain-based theoretical framework to model the packet delivery process of the two-hop relay routing in DTNs and derive some related basic results, then proceed to derive closed-form expressions for the expected delivery delay and the expected delivery cost.

A. Markov Chain Framework and Related Basic Results

Without loss of generality, we focus on the head-of-line (HoL) packet P_h of the tagged flow. Recall that when two nodes meet each other, either one has the same probability to become the transmitting node. Given that the destination node D is requesting P_h, it is easy to see that the source node S will deliver out a copy of P_h with probability $p/2$ when encountering a relay node, and a relay node carrying P_h will forward P_h with probability $(1-p)/2$ when encountering D.

If we use the total copy number of packet P_h in the network (including the original one at the source node S) to denote a transient state, the whole packet delivery process of P_h can be modeled with an absorbing CTMC (Continuous Time Markov Chain). Since the source node S can deliver copies of P_h to at most $n-2$ distinct relay nodes, the corresponding CTMC is a finite-state absorbing CTMC. We illustrate the transition diagram of the Markov chain in Fig. 3, where the state A denotes the absorbing state, i.e., the destination node D successfully receives the packet P_h.

From Fig. 3, we can see that for each transient state k there, $1 \le k < n-1$, it may have at most three different transitions: the transition to its neighboring state $k+1$, the transition to the absorbing state A and the transition back to itself. We assume that when in state k, the Markov chain either transits to state $k+1$ after time $S_1(k)$ or transits to state A after time $S_2(k)$.

It is noticed that state k will transit to state $k+1$ if and only if the source S successfully delivers out a new copy of packet P_h. As there are $n-1-k$ relay nodes without carrying a copy of P_h in state k, say $R_1, R_2, \ldots, R_{n-1-k}$, we denote by T_i the time it takes S to deliver a copy of P_h to relay R_i, $1 \le i \le n-1-k$. As the node inter-meeting times between any node pair are assumed to be exponentially distributed with inter-meeting intensity λ, it is easy to see that T_i follows the exponential distribution with mean $\frac{2}{p\lambda}$. Since $S_1(k) = \min\{T_1, T_2, \ldots, T_{n-1-k}\}$, $S_1(k)$ follows an exponential distribution with mean $\frac{2}{(n-k-1)p\lambda}$. Similarly, we can see that $S_2(k)$ follows an exponential distribution with mean $\frac{2}{(k-kp+p)\lambda}$. If we denote by $b_1(k)$ the rate of state k transiting to state $k+1$ and denote by $b_2(k)$ the rate of state k transiting to state A, then we have

$$Pr(S_1(k) < x) = 1 - \mathrm{e}^{-b_1(k)x} \tag{2}$$
$$Pr(S_2(k) < x) = 1 - \mathrm{e}^{-b_2(k)x} \tag{3}$$

where

$$b_1(k) = \frac{1}{2}(n-k-1)p\lambda \tag{4}$$
$$b_2(k) = \frac{1}{2}(k-kp+p)\lambda \tag{5}$$

If we further denote by $S(k)$ the overall sojourn time inside a general transient state k and denote by $b(k)$ the rate of state k transiting back to itself, $1 \le k \le n-1$, then we have the following lemma.

Lemma 1: For a general transient state k, $1 \le k \le n-1$, the sojourn time $S(k)$ follows an exponential distribution with mean $\frac{1}{b(k)}$, i.e.,

$$P_r(S(k) < x) = 1 - \mathrm{e}^{-b(k)x} \tag{6}$$

where

$$b(k) = \frac{1}{2}(np - 2kp + k)\lambda \tag{7}$$

Proof: Since there are two outgoing transitions from state k in the Markov chain of Fig. 3, $1 \le k < n-1$, i.e., the transition to neighboring state $k+1$ and the transition to absorbing state A, the overall sojourn time $S(k)$ can be determined as

$$S(k) = \min\{S_1(k), S_2(k)\} \tag{8}$$

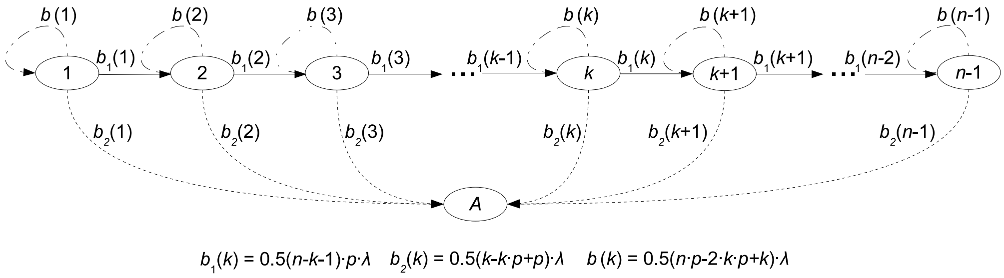


Fig. 3. Transition diagram of the finite-state absorbing CTMC for the HoL packet P_h of the local-queue at the source node S, given that the destination node D is also requesting P_h. For each transient state k, the corresponding transition rate $b_1(k)$ to its neighboring state, the transition rate $b_2(k)$ to the absorbing state A and the transition rate $b(k)$ to itself are listed at the bottom part of the figure.

Together with (2) and (3), we have

$$\begin{aligned} Pr(S(k) < x) &= Pr(S_1(k) < x \mid S_1(k) < S_2(k)) \\ &= \frac{P_r(S_1(k) < x, S_1(k) < S_2(k))}{P_r(S_1(k) < S_2(k))} \end{aligned} \tag{9}$$

Since

$$\begin{aligned} &P_r(S_1(k) < x, S_1(k) < S_2(k)) \\ &= \int_0^x b_1(k)\mathrm{e}^{-b_1(k)t}\mathrm{d}t \int_t^\infty b_2(k)\mathrm{e}^{-b_2(k)u}\mathrm{d}u \\ &= \int_0^x b_1(k)\mathrm{e}^{-(b_1(k)+b_2(k))t}\mathrm{d}t \\ &= \frac{b_1(k)}{b_1(k)+b_2(k)}\left(1-\mathrm{e}^{-(b_1(k)+b_2(k))x}\right) \end{aligned} \tag{10}$$

and

$$P_r(S_1(k) < S_2(k)) = \frac{b_1(k)}{b_1(k)+b_2(k)} \tag{11}$$

Substituting (10) and (11) into (9), it follows (6) and (7). It's easy to further verify that (6) and (7) also hold for the case $k = n-1$. Therefore, we finish the proof for Lemma 1. ■

For a general transient state k in the CTMC of Fig. 3, if we denote by $p_1(k)$ and $p_2(k)$ the transition probability from state k to state $k+1$ and the transition probability from state k to state A, respectively, combining (4), (5) and (7) we can see that

$$p_1(k) = \frac{b_1(k)}{b(k)} = \frac{(n-k-1)p}{np-2kp+k} \tag{12}$$

$$p_2(k) = \frac{b_2(k)}{b(k)} = \frac{k-kp+p}{np-2kp+k} \tag{13}$$

We further assume that when the CTMC of Fig. 3 enters the absorbing state A there are in total N_d message copies in the network, i.e., the Markov chain in Fig. 3 gets absorbed from state N_d, $1 \le N_d \le n-1$. Notice that these N_d message copies include (resp. exclude) the copy at the source node S (resp. at the destination node D). Thus, we have the following lemma.

Lemma 2: The pdf (probability distribution function) of N_d can be given by

$$P_r(N_d = k) = \frac{(n-2)! \cdot p^{k-1}(k-kp+p)}{(n-k-1)! \cdot \prod_{j=1}^{k}(np-2jp+j)} \tag{14}$$

where $1 \le k \le n-1$.

Proof: Given $N_d = k$, we can see that the last transient state before the Markov chain gets absorbed is the state k, i.e, the Markov chain becomes absorbed along the path $1 \to 2 \to 3 \to \cdots \to k \to A$. Thus, we have

$$P_r(N_d = k) = \prod_{j=1}^{k-1} p_1(j) \cdot p_2(k) \tag{15}$$

Substituting (12) and (13) into (15), it follows (14) after some basic algebraic operations. ■

B. Expected Delivery Delay and Expected Delivery Cost

With the help of the developed Markov chain framework and the related basic results, we proceed to derive closed-form expressions for the expected packet delivery delay and the expected packet delivery cost. We first introduce here the following definitions for the packet delivery delay and the packet delivery cost.

Definition 1: For the HoL packet P_h at the local-queue of the source node S, the packet delivery delay is the time elapsed between the time when S starts to transmit P_h and the time when the destination node D receives P_h.

Definition 2: For the HoL packet P_h at the local-queue of the source node S, the packet delivery cost is the total number of transmissions for packet P_h between the time when S starts to transmit P_h and the time when the destination node D receives P_h.

Notice that in Definition 2, the packet delivery cost includes the last transmission from the source node S (or some relay) to the destination D. If we denote by T_d and C_d the delivery delay and delivery cost, respectively, then we have the following theorems about $\mathbb{E}\{T_d\}$ and $\mathbb{E}\{C_d\}$.

Theorem 1: The expected delivery delay $\mathbb{E}\{T_d\}$ can be determined as

$$\mathbb{E}\{T_d\} = \sum_{k=1}^{n-1} P_r(N_d = k) \cdot \sum_{j=1}^{k} \frac{1}{b(j)} \tag{16}$$

where the $P_r(N_d = k)$ and $b(j)$ are given by (14) and (7), respectively.

Proof: We denote by $L_d(s)$ the Laplace-Stieltjes transform of T_d, $s \ge 0$, thus we have

$$\mathbb{E}\{T_d\} = -\frac{\mathrm{d}L_d(s)}{\mathrm{d}s}\Big|_{s=0} \tag{17}$$

Since

$$
\begin{aligned}
L_d(s) &= \mathbb{E}\{\mathrm{e}^{-T_d \cdot s}\} \\
&= \sum_{k=1}^{n-1} \mathbb{E}\{\mathrm{e}^{-T_d \cdot s} \mid N_d = k\} \cdot P_r(N_d = k) \qquad (18) \\
&= \sum_{k=1}^{n-1} \mathbb{E}\{\mathrm{e}^{-\sum_{j=1}^{k} S(j)\cdot s} \mid N_d = k\} \cdot P_r(N_d = k) \qquad (19) \\
&= \sum_{k=1}^{n-1} f(s,k) \cdot P_r(N_d = k) \qquad (20)
\end{aligned}
$$

where

$$
f(s,k) = \mathbb{E}\{\mathrm{e}^{-\sum_{j=1}^{k} S(j)\cdot s} \mid N_d = k\} \qquad (21)
$$

(18) follows by conditioning on the N_d, and (19) follows after substituting the $(T_d \mid_{N_d=k}) = \sum_{j=1}^{k} S(j)$. Notice that in (20), as $P_r(N_d = k)$ is given by (14) in Lemma 2, the only remaining issue for derivation of $L_d(s)$ is to derive $f(s,k)$.

Since $S(1)$, $S(2)$, $\ldots$, $S(k)$ in (21) are mutually independent, we have

$$
f(s,k) = \prod_{j=1}^{k} \mathbb{E}\{\mathrm{e}^{-S(j)\cdot s}\} \qquad (22)
$$

where

$$
\begin{aligned}
&\mathbb{E}\{\mathrm{e}^{-S(j)\cdot s}\} \\
&= \int_0^{\infty} \mathrm{e}^{-s\cdot x} \cdot b(j)\mathrm{e}^{-b(j)x}\mathrm{d}x \qquad (23) \\
&= \frac{b(j)}{s+b(j)} \qquad (24)
\end{aligned}
$$

and (23) follows after substituting (6) in Lemma 1. Substituting (24) into (22), we have

$$
f(s,k) = \prod_{j=1}^{k} \Big(1 + \frac{s}{b(j)}\Big)^{-1} \qquad (25)
$$

Combining (17) and (20), we can see that

$$
\mathbb{E}\{T_d\} = \sum_{k=1}^{n-1} P_r(N_d = k) \cdot \Big(- \frac{\mathrm{d}f(s,k)}{\mathrm{d}s}\Big|_{s=0}\Big) \qquad (26)
$$

where

$$
\begin{aligned}
&\Big(- \frac{\mathrm{d}f(s,k)}{\mathrm{d}s}\Big|_{s=0}\Big) \\
&= \Big(\sum_{j=1}^{k} \Big(1+\frac{s}{b(j)}\Big)^{-2} \frac{1}{b(j)} \cdot \prod_{i=1, i\neq j}^{k} \Big(1+\frac{s}{b(i)}\Big)^{-1}\Big)\Big|_{s=0} \qquad (27) \\
&= \sum_{j=1}^{k} \frac{1}{b(j)} \qquad (28)
\end{aligned}
$$

and (27) follows after substituting (25).

Substituting (28) into (26), it follows (16). Then we finish the proof for Theorem 1. ■

Theorem 2: The expected delivery cost $\mathbb{E}\{C_d\}$ can be determined as

$$
\mathbb{E}\{C_d\} = \sum_{k=1}^{n-1} \frac{(n-2)! \cdot p^{k-1}(k^2 - k^2 p + kp)}{(n-k-1)! \cdot \prod_{j=1}^{k}(np - 2jp + j)} \qquad (29)
$$

Proof: As indicated by Lemma 2, the CTMC in Fig. 3 will become absorbed from state k with probability $P_r(N_d = k)$. Notice that when the Markov chain arrives at the state k, $k-1$ transmissions in total are taken for the packet P_h. Plus the last transmission taken from state k to the absorbing state A, we can see that when the Markov chain gets absorbed from state k, the corresponding delivery cost is also k. Thus, the expected delivery cost $\mathbb{E}\{C_d\}$ can be given by

$$
\mathbb{E}\{C_d\} = \sum_{k=1}^{n-1} k \cdot P_r(N_d = k) \qquad (30)
$$

after substituting (14), it follows (29) after some basic algebraic operations. ■

IV. Numerical Results

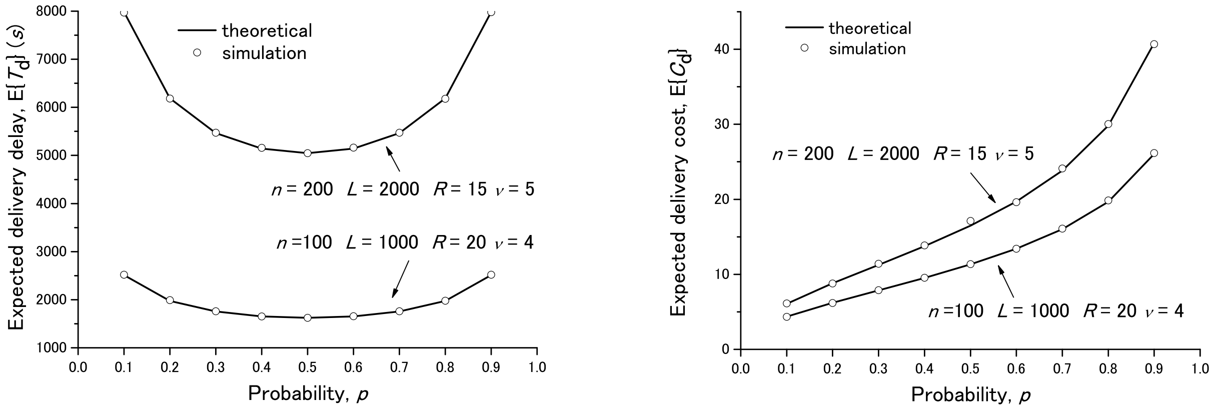
A. Simulation Setting

Our simulations were based on the node contact traces generated by OMNeT++, and the packet delivery was simulated by a C++ program which receives pre-recorded contact traces as input. The simulation scenario was a delay tolerant network with n nodes moving in a square area with side length L. Each mobile node adopted a uniform communication range R and a constant moving speed ν. The mobility model was the Random Waypoint model with no pause time, and in order to avoid the transient effects, the initial distribution of nodes was drawn from the stationary distribution [35].

B. Theoretical Vs. Simulation Results

To verify the theoretical framework, extensive simulation studies were conducted with the settings of $(n = 200, L = 2000\ m, R = 15\ m, \nu = 5\ m/s)$ and $(n = 100, L = 1000\ m, R = 20\ m, \nu = 4\ m/s)$. Notice that under these two network settings, the node inter-meeting intensity can be determined as $\lambda = 6.5332 \times 10^{-5}\ s^{-1}$ and $\lambda = 2.7875 \times 10^{-4}\ s^{-1}$, respectively, according to (1). The corresponding theoretical and simulation results were summarized in Fig. 4. Notice that all the simulated expected delivery delay and simulated expected delivery cost were calculated as the average value of 10^4 random and independent simulations.

Fig. 4 shows clearly that the simulation results match nicely with the theoretical ones for both $\mathbb{E}\{T_d\}$ and $\mathbb{E}\{C_d\}$, so our framework can be used to efficiently model the packet delivery process under the two-hop relay routing in DTNs. A further careful observation of Fig. 4a indicates that for both the network scenarios there, as p increases from 0.1 to 0.9, there exists an optimum setting of probability p, i.e., $p = 0.50$, which minimizes the expected delivery delay $\mathbb{E}\{T_d\}$. Specifically, for the network scenario $(n = 200, L = 2000\ m, R = 15\ m, \nu = 5\ m/s)$ (resp. $(n = 100, L = 1000\ m, R = 20\ m, \nu = 4\ m/s)$), a minimum $\mathbb{E}\{T_d\}$ of $5045.26\ s$ (resp. $1624.85\ s$) is achieved at the setting $p = 0.50$.



(a) Theoretical and simulation $\mathbb{E}\{T_d\}$ vs. p

(b) Theoretical and simulation $\mathbb{E}\{C_d\}$ vs. p

Fig. 4. Comparisons between theoretical and simulation results for model validation under the settings of $(n = 200, L = 2000\ m, R = 15\ m, \nu = 5\ m/s)$ and $(n = 100, L = 1000\ m, R = 20\ m, \nu = 4\ m/s)$.

It is also interesting to notice that in Fig. 4a both the curves of $\mathbb{E}\{T_d\}$ are symmetric with the line $p = 0.50$, which means that $\mathbb{E}\{T_d\}$ achieved at the setting p is the same as that of the setting $1 - p$. Therefore, we can see that under the two-hop relay which relies heavily on node cooperations for packet routing, the selfish behavior that each node adopts a higher probability to distribute its own packets may not help improve the delivery delay performance. The generally unselfish behavior, i.e., the simple fair setting of $p = 0.5$, achieves the best delivery delay performance for each node.

It is easy to observe from Fig. 4b that for both the network scenarios there, the expected delivery cost $\mathbb{E}\{C_d\}$ monotonically increases as p varies from 0.1 to 0.9. A further careful observation of Fig. 4b indicates that the sensitivity (the slope) of $\mathbb{E}\{C_d\}$ also increases with p. Specifically, $\mathbb{E}\{C_d\}$ increases almost linearly with p when $0.1 \leq p \leq 0.5$; while as p approaches 0.9, $\mathbb{E}\{C_d\}$ rises up sharply. In light of the symmetric behavior of $\mathbb{E}\{T_d\}$ observed from Fig. 4a, we can see that when operating under the two-hop relay routing, each mobile node may select a probability $p \in [0, 0.5]$, where a higher p value achieves a smaller delivery delay but unavoidably results in a higher delivery cost. Thus, a careful trade-off needs to be made according to the specified delivery delay requirement and energy consumption constraints.

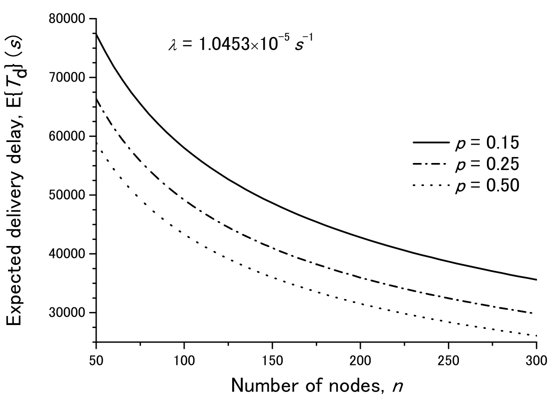
C. $\mathbb{E}\{T_d\}$ and $\mathbb{E}\{C_d\}$ Vs. n

Based on the theoretical framework, we further proceed to explore how the number of nodes n, will affect the expected delivery delay $\mathbb{E}\{T_d\}$ and expected delivery cost $\mathbb{E}\{C_d\}$. The node inter-meeting intensity was fixed as $\lambda = 1.0453 \times 10^{-5}\ s^{-1}$, which corresponds to the network setting of $(L = 5000\ m, R = 15\ m, \nu = 5\ m/s)$. We consider three settings of p ($p = 0.15$, 0.25 and 0.50) and let n varies from 50 to 300 to ensure that the resulted network is sparsely distributed and in line with a DTN scenario. As shown in Figs. 5a and 5b, for all three settings of p there, $\mathbb{E}\{T_d\}$ monotonically decreases with n while $\mathbb{E}\{C_d\}$ monotonically increases with n. Actually, such behaviors can be interpreted as follows: as the number of mobile nodes n increases up, there will be more chances for the source node S to meet other nodes and thus deliver out copies for its HoL packet. Since more relay nodes will be employed to help forward the packet, more transmissions will be conducted and thus the packet delivery cost is increased. On the other hand, the increasing of relay nodes will also improve the packet propagation speed and thus shorten the packet delivery delay.

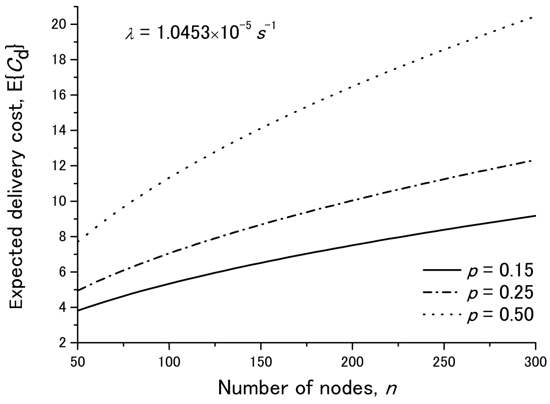
V. Conclusion

In this paper, we investigated the impact of relay cooperation on the delivery performance of two-hop relay routing in delay tolerant networks. A continuous time Markov chain-based theoretical framework was developed to model the packet delivery process under such routing scheme. Closed-form expressions were further derived for both the expected delivery delay and the expected delivery cost, with a general setting of the relay forwarding behavior. Our results show that for a given DTN the $\mathbb{E}\{T_d\}$ achieved at the setting p is the same as that of the setting $1 - p$, and the generally unselfish behavior, i.e., the simple fair setting of $p = 0.5$, achieves the best delivery delay performance for each node. Therefore, each mobile node may select a probability $p \in [0, 0.5]$, where a higher p value achieves a smaller delivery delay but unavoidably results in a higher delivery cost. Thus, a careful trade-off needs to be made according to the specified delivery delay requirement and energy consumption constraints.

Notice that the theoretical models and closed-form expressions in this paper were developed mainly under the assumptions of homogeneous packet size and permutation traffic pattern. Therefore, one of our future research directions is to develop theoretical models for other more general network scenarios, like the heterogeneous packet sizes and hybrid traffic patterns.



(a) $\mathbb{E}\{T_d\}$ vs. n



(b) $\mathbb{E}\{C_d\}$ vs. n

Fig. 5. The expected delivery delay $\mathbb{E}\{T_d\}$ and expected delivery cost $\mathbb{E}\{C_d\}$ vs. the number of nodes n.

REFERENCES

[1] S. Burleigh, A. Hooke, L. Torgerson, K. Fall, V. Cerf, B. Durst, K. Scott, and H. Weiss, "Delay-tolerant networking: an approach to interplanetary internet," *IEEE Communications Magazine*, vol. 41, no. 6, pp. 128–136, June 2003.

[2] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on opportunistic forwarding algorithms," *IEEE Transactions on Mobile Computing*, vol. 6, no. 6, pp. 606–620, June 2007.

[3] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Efficient routing in intermittently connected mobile networks: The multiple-copy case," *IEEE/ACM Transactions on Networking*, vol. 16, no. 1, pp. 77–90, February 2008.

[4] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*, 1996.

[5] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *WMCSA*, 1999.

[6] S. Jain, K. Fall, and R. Patra, "Routing in a delay tolerant network," in *SIGCOMM*, 2004.

[7] Z. Zhang, "Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges," *IEEE Communications Surveys & Tutorials*, vol. 8, no. 1, pp. 24–37, March 2006.

[8] K. Fall and S. Farrell, "Dtn: an architectural retrospective," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 5, pp. 828–836, June 2008.

[9] T. Matsuda and T. Takine, "(p, q)-epidemic routing for sparsely populated mobile ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 5, pp. 783–793, June 2008.

[10] M. Grossglauser and D. N. Tse, "Mobility increases the capacity of ad hoc wireless networks," in *INFOCOM*, 2001.

[11] E. Altman, F. D. Pellegrini, and L. Sassatelli, "Dynamic control of coding in delay tolerant networks," in *INFOCOM*, 2010.

[12] E. Altman, T. Basar, and F. D. Pellegrini, "Optimal control in two-hop relay routing," *IEEE Transactions on Automatic Control*, vol. 56, no. 3, pp. 670–675, March 2011.

[13] E. Altman and F. D. Pellegrini, "Forward correction and fountain codes in delay tolerant networks," in *INFOCOM*, 2009.

[14] J. Liu, X. Jiang, H. Nishiyama, and N. Kato, "Group-based two-hop relay with redundancy in manets," in *HPSR*, 2011.

[15] ——, "Delay and capacity in ad hoc mobile networks with f-cast relay algorithms," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2738–2751, August 2011.

[16] E. Altman, A. P. Azad, T. Basar, and F. D. Pellegrini, "Optimal activation and transmission control in delay tolerant networks," in *INFOCOM*, 2010.

[17] E. Altman, T. Basar, and F. D. Pellegrini, "Optimal monotone forwarding policies in delay tolerant mobile ad-hoc networks," in *Inter-Perf*, 2008.

[18] E. Altman, G. Neglia, F. D. Pellegrini, and D. Miorandi, "Decentralized stochastic control of delay tolerant networks," in *INFOCOM*, 2009.

[19] A. A. Hanbali, P. Nain, and E. Altman, "Performance of ad hoc networks with two-hop relay routing and limited packet lifetime," in *Valuetools*, 2006.

[20] A. A. Hanbali, A. A. Kherani, and P. Nain, "Simple models for the performance evaluation of a class of two-hop relay protocols," in *Proc. IFIP Networking*, 2007.

[21] R. Groenevelt, P. Nain, and G. Koole, "The message delay in mobile ad hoc networks," *Performance Evaluation*, vol. 62, no. 1-4, pp. 210–228, October 2005.

[22] A. Panagakis, A. Vaios, and I. Stavrakakis, "Study of two-hop message spreading in dtns," in *WiOpt*, April 2007.

[23] ——, "On the effects of cooperation in dtns," in *Comsware*, January 2007.

[24] M. Karaliopoulos, "Assessing the vulnerability of dtn data relaying schemes to node selfishness," *IEEE Communications Letters*, vol. 13, no. 12, pp. 923–925, December 2009.

[25] Y. Li, P. Hui, D. Jin, L. Su, and L. Zeng, "Evaluating the impact of social selfishness on the epidemic routing in delay tolerant networks," *IEEE Communications Letters*, vol. 14, no. 11, pp. 1026–1028, November 2010.

[26] Y. Li, G. Su, D. O. Wu, D. Jin, L. Su, and L. Zeng, "The impact of node selfishness on multicasting in delay tolerant networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 5, pp. 2224–2238, June 2011.

[27] P. Hui, K. Xu, V. Li, J. Crowcroft, V. Latora, and P. Lio, "Selfishness, altruism and message spreading in mobile social networks," in *NetSciCom*, 2009.

[28] Q. Li, S. Zhu, and G. Cao, "Routing in socially selfish delay tolerant networks," in *INFOCOM*, 2010.

[29] M. J. Neely and E. Modiano, "Capacity and delay tradeoffs for ad-hoc mobile networks," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1917–1936, June 2005.

[30] P. Li, Y. Fang, and J. Li, "Throughput, delay, and mobility in wireless ad-hoc networks," in *INFOCOM*, 2010.

[31] L. Ying, S. Yang, and R. Srikant, "Optimal delay-throughput trade-offs in mobile ad hoc networks," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4119–4143, September 2008.

[32] M. Garetto, P. Giaccone, and E. Leonardi, "Capacity scaling in ad hoc networks with heterogeneous mobile nodes: The subcritical regime," *IEEE/ACM Transactions on Networking*, vol. 17, no. 6, pp. 1888–1901, December 2009.

[33] D. Ciullo, V. Martina, M. Garetto, and E. Leonardi, "Impact of correlated mobility on delay-throughput performance in mobile ad-hoc networks," in *INFOCOM*, 2010.

[34] R. Groenevelt, "Stochastic models in mobile ad hoc networks," Ph.D. dissertation, University of Nice Sophia Antipolis, April 2005.

[35] J. yves Le Boudec and M. Vojnovic, "Perfect simulation and stationary of a class of mobility models," in *INFOCOM*, 2005.