Generalized Two-Hop Relay for Flexible Delay Control in MANETs

Jiajia Liu, *Student-Member, IEEE,* Xiaohong Jiang, *Senior Member, IEEE,* Hiroki Nishiyama, *Member, IEEE,* and Nei Kato, *Senior Member, IEEE*

***Abstract*—The available two-hop relay protocols with out-of-order or strictly in-order reception can not provide a flexible control for the packet delivery delay, which may significantly limit their applications to the future MANETs with different delay requirements. This paper extends the conventional two-hop relay and proposes a general group-based two-hop relay algorithm with packet redundancy. In such an algorithm with packet redundancy limit f and group size g (2HR-(f,g) for short), each packet is delivered to at most f distinct relay nodes and can be accepted by its destination if it is a fresh packet to the destination and also it is among g packets of the group the destination is currently requesting. The 2HR-(f,g) covers the available two-hop relay protocols as special cases, like the in-order reception ones ($f \geq 1$, $g = 1$), the out-of-order reception ones with redundancy ($f > 1$, $g = \infty$) or without redundancy ($f = 1$, $g = \infty$). A Markov chain-based theoretical framework is further developed to analyze how the mean value and variance of packet delivery delay vary with the parameters f and g, where the important medium contention, interference and traffic contention issues are carefully incorporated into the analysis. Extensive simulation and theoretical results are provided to illustrate the performance of the 2HR-(f,g) algorithm and the corresponding theoretical framework, which indicate that the theoretical framework is efficient in delay analysis and the new 2HR-(f,g) algorithm actually enables both the mean value and variance of packet delivery delay to be flexibly controlled in a large region.**

***Index Terms*—Mobile ad hoc networks, two-hop relay, packet redundancy, delivery delay.**

I. Introduction

Two-hop relay and its variants have been a class of attractive routing protocols for mobile ad hoc networks (MANETs) due to their efficiency and simplicity [1], [2]. In the two-hop relay routing, the source first transmits packets to the mobiles (relays) it encounters; relays then transmit the packets only if they come in contact with the destination. Thus, each packet travels at most two hops to reach its destination.

The available two-hop relay algorithms adopt either out-of-order or strictly in-order reception, which are the two extreme cases of reception mode. In the in-order two-hop relay algorithms, like the ones proposed in [3]–[5], each packet should be received in-order at its destination. The algorithms in [1], [6], [7] can be regarded as the out-of-order two-hop without redundancy, where a packet has at most one copy and gets accepted by its destination if it is "fresh" (never received before). The out-of-order two-hop relay with redundancy has also been explored recently [8], [9], where each packet may have multiple copies in the transmission process.

For the two-hop relay with in-order reception, lot of reception opportunities may be wasted as the destination only accepts packets according to their sequence orders, resulting in an increase in the packet delivery delay. The out-of-order two-hop relay, on the other hand, can take the full advantage of each reception opportunity but each mobile node there should potentially carry a very big (if not infinite) buffer to accommodate all possible arrivals, which is not really practical for the MANETs. Also, the early arrived packets there may need to wait a long time for the arrivals of other related packets, which may make the early arrived packets become expired and thus useless. The packet delay for two-hop relay MANETs has been extensively studied in the literature, in terms of its order sense scaling laws with network size or its closed-form analytical models (see Section VI for related works). These delay results indicate that the available out-of-order or strictly in-order two-hop relay protocols, although simple and easy to operate, can not provide a flexible control for the packet delivery delay. The lack of a flexible delay control in available two-hop relay protocols may significantly limit their ability to support various delay sensitive applications in the future MANETs, like VoIP [10]–[12], video streaming [13], [14], real-time monitoring and networked control [15], [16], etc.

This paper extends the conventional two-hop relay to a group-based two-hop relay with packet redundancy to enable the packet delivery delay to be flexibly controlled in a large region. The main contributions of this paper are as follows.

- This paper proposes a new 2HR-(f,g) algorithm, where each packet is delivered to at most f distinct relay nodes and can be accepted by its destination if it is a fresh packet to the destination and also it is among g packets of the group the destination is currently requesting. This algorithm is general and covers all the available two-hop routing protocols as special cases, like the in-order ones [3]–[5] ($f \geq 1$, $g = 1$), the out-of-order ones with redundancy [8], [9] ($f > 1$, $g = \infty$) or without redundancy [1], [6], [7] ($f = 1$, $g = \infty$).
- To capture the complex packet delivery process in a MANET with 2HR-(f,g), we further develop a general theoretical framework based on the multi-dimensional Markov chain, which covers the available frameworks for conventional two-hop relay analysis as special cases [17]–[20]. The theoretical framework is powerful in the sense

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.

it enables not only the mean value but also the variance of packet delivery delay to be derived analytically with a careful consideration of the important medium contention, interference and traffic contention issues.
- Extensive simulation and theoretical results are provided to validate the 2HR-(f, g) algorithm and the Markov chain theoretical framework. These results indicate that the theoretical framework is very efficient in packet delay analysis, and more importantly, the new 2HR-(f, g) algorithm makes it possible for us to flexibly control the packet delivery delay (and its variance) in a large region through the proper settings of f and g.

The rest of this paper is outlined as follows. Section II introduces the system models, the 2HR-(f, g) algorithm and the corresponding transmission scheduling scheme. In Section III, we develop the Markov chain-based theoretical framework and provide some basic results. In Section IV, we analytically derive the expected value and standard derivation for packet delivery delay. Section V presents the numerical results to validate the theoretical framework and the 2HR-(f, g) algorithm. Finally we introduce the related works in Section VI and conclude this paper in Section VII.

II. 2HR-(f, g) Algorithm and Transmission Scheduling

A. System Models

The concerned network consists of n mobile nodes inside a unit square, which is evenly divided into $m \times m$ cells. We focus on a slotted system and a fast mobility scenario [21], where only one-hop transmissions are possible within each time slot, and the total number of bits transmitted per slot is fixed and normalized to 1 packet. The nodes independently roam from cell to cell, following the bi-dimensional i.i.d. mobility model [3]. At the beginning of each time slot, each node independently and uniformly selects a cell among all m^2 cells and stays in it for the whole time slot. The protocol model with guarding factor Δ in [22] is adopted as the interference model here. We further assume a permutation traffic pattern in the saturated case [21], where each node is a source and at the same time a destination of some other node, and each source node always has packets waiting for delivery. For a given source-destination pair, we call the traffic between them as a flow.

B. 2HR-(f, g) Algorithm

Without loss of generality, we focus on a tagged flow and denote its source node and destination node as S and D, respectively. As illustrated in Fig. 1 that with the 2HR-(f, g) algorithm, the source node S will deliver at most f copies of a packet P to distinct relay nodes, while the destination D may finally receive the packet from one relay node R^*.

Notice that each node can be a potential relay for other $n-2$ flows (except the two flows originated from and destined for itself), thus, to support the operation of the 2HR-(f, g) algorithm, we assume that each node maintains n individual queues at its buffer: one local-queue for storing the packets that

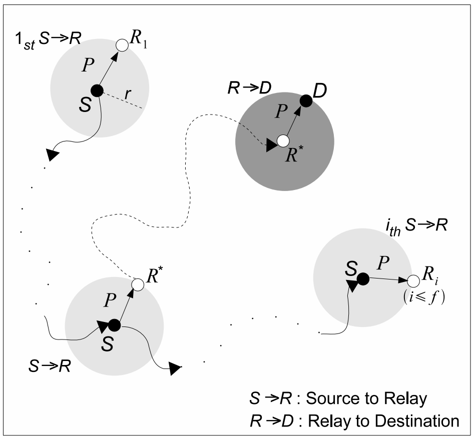


Fig. 1. Illustration of the 2HR-(f, g) relay algorithm for a tagged flow, whose source node S is transmitting packet P to the destination node D.

are locally generated at the node and waiting for their copies (up to f copies for each packet) to be dispatched, one already-sent-queue for storing packets whose f copies have already been dispatched but their reception status are not confirmed yet (from destination node), and $n-2$ parallel relay-queues for storing packets of other flows (one queue per flow).

To support the group based transmission in the 2HR-(f, g) algorithm, the source node S divides packets waiting at its local-queue into consecutive groups, g packets per group, and labels each packet P with a *send group number* $SG(P)$ and a *sequence number* $SN(P)$ ($1 \le SN(P) \le g$). Similarly, the node D also maintains a *request group number* $RG(D)$ and an *indicator vector* $IN(D)$. The $IN(D)$ is a g-bit binary vector that records the reception status of current requesting group at D, where the i_{th} bit $IN_i(D)$ is set as 0 (resp. 1) if the i_{th} packet of the current requesting group has (resp. has not) been received. To simplify the analysis, we assume that each relay node will carry at most one packet for any particular group. We further introduce the following definitions:

- **Fresh packet and non-fresh packet:** A packet is called a fresh packet if it has not been received yet by its destination; a non-fresh packet, otherwise.
- **Fresh node and non-fresh node:** For a tagged packet group, a node (except the source S and the destination D) is called a fresh node if it is carrying a fresh packet for the group; otherwise, if the node is either carrying a non-fresh packet or carrying no packet for the tagged group, it is called a non-fresh node.

Based on the above definitions, the 2HR-(f, g) algorithm can be summarized as follows.

2HR-(f, g) Algorithm: For the tagged flow, every time the node S gets a transmission opportunity, it operates as follows:

Step 1: (Source-to-Destination) If the node D is among its one-hop neighbors, it initiates a handshake with D to get its $RG(D)$ and $IN(D)$. Then it tries to transmit a fresh packet directly to D, where the packet to be transmitted is selected as follows: it first checks its local-queue, starting from its head-of-line packet P_h, to find a fresh packet; if it fails, then it tries to retrieve a fresh packet from the already-sent-queue.

Step 2: Otherwise, if the node D is not among the one-

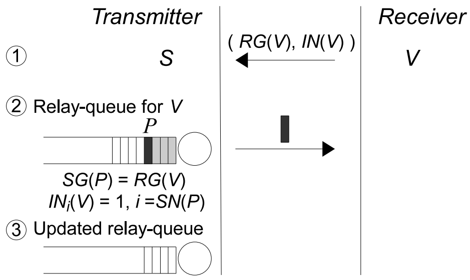


Fig. 2. Illustration of the Relay-to-Destination mode, where the node S (acting as a relay) transmits a fresh packet P destined for the node V.

hop neighbors of S, the node S randomly chooses one of the following two operations to perform:

- (Source-to-Relay) It first randomly selects one node (say R) from its current one-hop neighbors, then initiates a handshake with R to check whether the node R is a non-fresh node. If so, it delivers a new copy of P_h to R; otherwise it remains idle for this time slot.
- (Relay-to-Destination) It acts as a relay and randomly selects one node (say V) as the receiver from its one-hop neighbors. As indicated in Fig. 2 that it first initiates a handshake with V to get the $RG(V)$ and $IN(V)$, then checks its relay-queue specified for V whether there exists a fresh packet of group $RG(V)$. If so, it delivers this packet to V and deletes all packets with $SG \leq RG(V)$ from its relay-queue for V; otherwise it remains idle for this time slot.

Notice that in the above source-to-relay transmission, every time S sends out a copy of P_h it checks whether f copies of P_h have already been delivered. If yes, it puts P_h to the end of the already-sent-queue and then moves ahead the remaining packets in the local-queue. At the relay node R, P_h is put at the end of its relay-queue dedicated to the node D. Thus, each packet may have at most $f+1$ copies in the network (including the one in the already-sent-queue of its source node).

Remark 1: In the 2HR-(f,g) algorithm, if the node D is currently requesting for packets of group i, then any fresh packet belonging to the group i is eligible for reception at the node. The node D will start to receive packets of the next group $i+1$ only after all packets of the group i have been received. Thus, the 2HR-(f,g) algorithm ensures that the inter-group packet reception is strictly in-group-order while the intra-group packet reception is totally out-of-order.

Remark 2: The 2HR-(f,g) algorithm is flexible and general, since its packet delivery process can be flexibly controlled by a proper setting of the redundancy f and group size g. Actually, the new algorithm covers all the available two-hop relays as special cases, like the out-of-order ones with redundancy [8], [9] ($f > 1$, $g = \infty$) or without redundancy [1], [6], [7] ($f = 1$, $g = \infty$), and the strictly in-order ones [3]–[5] ($f \geq 1$, $g = 1$).

C. Transmission Scheduling

To support as many simultaneous transmissions as possible, similar to the "equivalence class" in [23] we define here the "concurrent-set" for transmission scheduling.

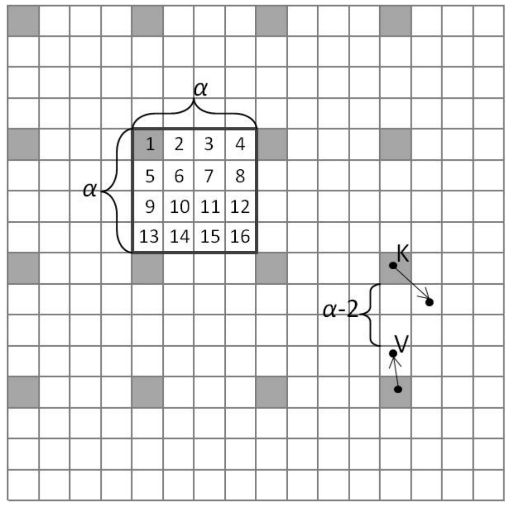


Fig. 3. An example of a concurrent-set of cells with $\alpha = 4$. The cells are divided into 16 different concurrent-sets and all the shaded cells belong to the same concurrent-set. The distribution of all the remaining nodes in the unit square is not shown for simplicity.

Concurrent-set: As illustrated by the shaded cells in Fig. 3, a concurrent-set is a subset of cells in which any two cells have a vertical and horizontal distance of some multiple of α cells, and all the cells there can transmit simultaneously without interfering with each other.

To guarantee the simultaneous transmissions in a concurrent-set without interfering with each other, the parameter α should be set properly. We consider a local transmission scenario, in which a node in some cell can only send packets to the nodes in the same cell or its eight adjacent cells. Two cells are called adjacent if they share a common point. Thus, the maximum distance between a transmitting node (transmitter) and a receiving node (receiver) is $\sqrt{8}/m$, so we set the communication range as $r = \sqrt{8}/m$. Due to the wireless interference, only cells that are sufficiently far away could simultaneously transmit without interfering with each other. As shown in Fig. 3, suppose that during some time slot, the node V is scheduled to receive a packet. According to the definition of "concurrent-set", we know that except the transmitting node of V, another transmitting node (say node K) in the same concurrent-set is at least $(\alpha-2)/m$ away from V. The condition that K will not interfere with the reception at V is that, $(\alpha-2)/m \geq (1+\Delta)\cdot r$. By substituting $r = \sqrt{8}/m$, we obtain that $\alpha \geq (1+\Delta)\sqrt{8}+2$. As α is an integer and $\alpha \leq m$, we set

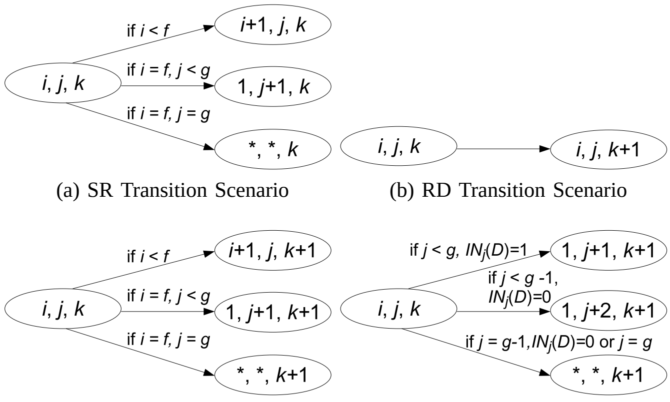
$$\alpha = \min\left\{\lceil (1+\Delta)\sqrt{8}\rceil + 2, m\right\},$$

where $\lceil x \rceil$ returns the smallest integer not less than x.

Notice that each cell will become active (i.e., get transmission opportunity) in every α^2 time slots. If there are more than one nodes inside an active cell, a transmitting node is selected randomly from them, and the selected node then follows the 2HR-(f,g) algorithm for packet transmission.

III. Markov Chain-based Framework

In this section, we develop a Markov chain-based theoretical framework to model the overall behavior of the 2HR-(f,g)



(a) SR Transition Scenario (b) RD Transition Scenario

(c) SR+RD Transition Scenario (d) SD Transition Scenario

Fig. 4. Transition scenarios of a state (i, j, k), where $1 \leq i \leq f$, $1 \leq j \leq g$ and $0 \leq k < g$, $k \leq j$.

algorithm, and then provide some basic results derived from it. This framework and related basic results will help us to perform packet delivery delay analysis in Section IV.

A. Markov Chain-based Theoretical Framework

For a tagged packet group at the source node S, we use a three-tuple (i, j, k) to denote the transient state that S is delivering the i_{th} ($1 \leq i \leq f$) copy for the j_{th} ($1 \leq j \leq g$) packet while the destination node D has already received any k ($0 \leq k < g$, $k \leq j$) of the g packets. We further use $(*, *, k)$ to denote the transient state that S has already finished dispatching the copies of all packets in the tagged group while D has only received k ($0 \leq k < g$) of them. From the operation of the 2HR-(f, g) algorithm we know that if a node pair (S, D) is in state (i, j, k) at the current time slot, then only one of the following four transmission scenarios illustrated in Fig. 4 may happen in the next time slot:

- SR Scenario: source-to-relay transmission only, i.e., S successfully delivers the i_{th} copy to a new relay while none of the relays delivers a fresh packet to D. As shown in Fig. 4a that under such a transition scenario, the state (i, j, k) may transit to three different neighboring states depending on the current copy index i and also the sequence number j of the current packet.
- RD Scenario: relay-to-destination transmission only, i.e., some relay node successfully delivers a fresh packet to node D while S fails to deliver out the i_{th} copy to a new relay node. As shown in Fig. 4b that there is only one target state $(i, j, k + 1)$ under the RD transition scenario.
- SR+RD Scenario: both source-to-relay and relay-to-destination transmissions, i.e., these two transmissions happen simultaneously. We can see from Fig. 4c that depending on both the values of i and j, there are three possible target states under the SR+RD transition scenario, similar to that under the SR scenario.
- SD Scenario: source-to-destination transmission, i.e., S successfully delivers out a fresh packet to D. As shown in Fig. 4d that under the SD transition scenario, the state (i, j, k) may transit to $(1, j + 1, k + 1)$, $(1, j + 2, k + 1)$ or $(*, *, k + 1)$, depending on the sequence number j of the current packet and also its reception status $IN_j(D)$.

If we use A to denote the absorbing state that the destination node D has received all the g packets of the tagged group, then the transition diagrams in Fig. 4 indicate that the packet delivery process in a 2HR-(f, g)-based network can be modeled as a discrete-time finite-state absorbing Markov chain illustrated in Fig. 5, where Figs. 5a, 5b and 5c each represents some cases of the full chain. Specifically, Fig. 5a defines the transitions among neighboring states when no more than one packet is received by D, i.e., $k = 0$; Fig. 5b represents the cases that D may receive at most one more fresh packet of the tagged group given that it has already received k packets of the group, $1 \leq k \leq g-2$; Fig. 5c shows the transition diagrams of how D may receive the last packet. The transitions of SD, SR, RD and SR+RD in Fig. 5 correspond to the 2HR-(f, g) transmissions of source-to-destination, source-to-relay, relay-to-destination, and both source-to-relay and relay-to-destination, respectively.

Remark 3: The Markov chain model in Fig. 5 covers the available models for conventional two-hop relay analysis as special cases when we set $g = 1$ there [17]–[20].

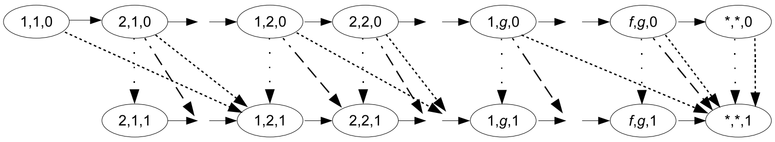
Although the Markov chain framework in Fig. 5 is general enough to model the packet delivery process in a 2HR-(f, g)-based network, it is difficult to directly apply such framework for an accurate packet delay analysis, even for the simple scenario of $f > 1$ and $g = 1$ [17], [19]. This is mainly due to the complicated transitions that may happen among transient states. As shown in Fig. 4 that for a state (i, j, k), its next state may vary significantly with the transition scenarios (SR, RD, SR+RD or SD), with the values of i, j and k, and also with the reception status $IN(D)$.

To simplify the analysis of packet delivery delay and enable both its mean value and variance to be derived analytically, we introduce the following assumption regarding the complex SD transition scenario:

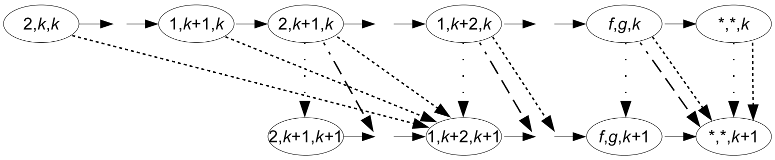
Assumption 1: Under the SD transition scenario in Fig. 4d, the transient state (i, j, k) will always transit to state $(1, j + 1, k + 1)$ whenever $k < j < g$.

The Assumption 1 indicates that for a transient state (i, j, k) under the SD transition scenario, if $k < j < g$, i.e., the source node S is currently delivering the j_{th} packet while only less than j packets have been received at the destination node D by now, then we assume that this j_{th} packet has not been received yet by D.

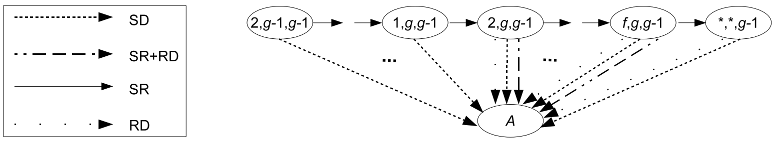
This assumption is due to the following observations: 1) To incorporate all the reception details of the j_{th} packet into analysis, a much complex Markov chain with bigger state space and more complex transitions among neighboring states should be adopted. We must be careful to avoid arriving at intractably complex models, so the case that the j_{th} packet has been received by D is neglected here; 2) Notice that the source node S always delivers out packets sequentially, so the packet delivered out earlier will be likely received early at the destination. Thus, the j_{th} packet currently being delivered at S is very likely not received yet by D given that only $k < j$ packets have been received at D by now. Notice also that this assumption only applies to the special SD transition scenario, which in general happens with negligible probability in comparison with that of the SR or RD transition scenario in a large MANET. Therefore, the simplification introduced



(a) States transition diagram for $k = 0$



(b) States transition diagram for $1 \leq k \leq g-2$



(c) States transition diagram for $k = g-1$

Fig. 5. Transition diagram of the Markov chain for the general 2HR-(f, g) algorithm. For each transient state, the transition back to itself is not shown for simplicity.

by the Assumption 1 will not cause a significant error in the overall packet delay analysis, as to be validated in Section V.

B. Some Basic Results

We can easily see from Fig. 5 that in the Markov-chain model, the total number of transient states β is determined as

$$\beta = \frac{f}{2}(g^2 + 3g - 2) + 1. \tag{1}$$

Actually, these β transient states are arranged into g rows, where the number of transient states L_k in the k_{th} row ($0 \leq k \leq g-1$) is given by

$$L_k = \begin{cases} (g+1-k)f & \text{if } 1 \leq k \leq g-1, \\ g \cdot f + 1 & \text{if } k = 0. \end{cases} \tag{2}$$

We now establish the following lemmas regarding some basic results of the Markov-chain model in Fig. 5, which will help us for packet delay analysis in Section IV.

Lemma 1: For the Markov-chain model in Fig. 5, the number of fresh nodes u_r and the number of non-fresh nodes u_o in the t_{th} transient state of k_{th} row, $t \in [1, L_k]$, $k \in [1, g-1]$, can be determined as

$$u_r \approx t - f, \tag{3}$$

$$u_o \approx n - 2 - t + k - (k-1)f. \tag{4}$$

Lemma 2: For a given time slot and a tagged flow, we use p_1 and p_2 to denote the probability that S conducts a source-to-destination transmission and the probability that S conducts a source-to-relay or relay-to-destination transmission, respectively. Then we have

$$p_1 = \frac{1}{\alpha^2}\left(\frac{9n - m^2}{n(n-1)} - \left(1 - \frac{1}{m^2}\right)^{n-1} \frac{8n + 1 - m^2}{n(n-1)}\right), \tag{5}$$

$$p_2 = \frac{1}{\alpha^2}\left(\frac{m^2 - 9}{n-1}\left(1 - \left(1 - \frac{1}{m^2}\right)^{n-1}\right) - \left(1 - \frac{9}{m^2}\right)^{n-1}\right). \tag{6}$$

Lemma 3: For a tagged flow, suppose that the source node S is delivering copies for some packet group i in the current time slot, the destination node D is requesting the packets of the group i, and there are currently t_1 fresh nodes and t_2 non-fresh nodes for the group i in the network. For the next time slot, we use $P_r(t_1)$ to denote the probability that D will receive a fresh packet, use $P_d(t_2)$ to denote the probability that S will successfully deliver out a copy to some new relay, and use $P_s(t_1, t_2)$ to denote the probability that both a successful source-to-relay transmission and a relay-to-destination transmission will be performed. Then we have

$$P_r(t_1) = p_1 + \frac{t_1}{2(n-2)} p_2, \tag{7}$$

$$P_d(t_2) = \frac{t_2}{2(n-2)} p_2, \tag{8}$$

$$P_s(t_1, t_2) = \frac{t_1 t_2 (m^2 - \alpha^2)}{4m^2\alpha^4} \sum_{k=0}^{n-5} \binom{n-5}{k} h(k) \cdot \left\{ \sum_{t=0}^{n-4-k} \binom{n-4-k}{t} h(t) \left(1 - \frac{18}{m^2}\right)^{n-4-k-t} \right\}, \tag{9}$$

where

$$h(x) = \frac{9\left(\frac{9}{m^2}\right)^{x+1} - 8\left(\frac{8}{m^2}\right)^{x+1}}{(x+1)(x+2)}. \tag{10}$$

Remark 4: The proofs of above lemmas can be found in Appendix A, where the important medium contention, interference and traffic contention issues have been carefully incorporated into the analysis of the probabilities p_1, p_2, $P_r(t_1)$, $P_d(t_2)$ and $P_s(t_1, t_2)$. The basic idea behind the proof for Lemma 1 is to consider a general transient state (i, j, k) and find its mapping relationship with the t_{th} state of k_{th} row in the Markov chain model of Fig. 5, then the (3) follows by assuming that the current j_{th} packet is not among the k received packets, and the (4) follows by assuming that all the

k packets are received from relay nodes. The basic idea of proofs for Lemmas 2 and 3 is to first properly divide the target events, like the event that S conducts a source-to-destination transmission for a given time slot in Lemma 2 and the event that D will receive a fresh packet when there are currently t_1 fresh nodes in Lemma 3, into mutually exclusive cases, represent each case by several simultaneous and independent sub-events, and then derive the target probability based on the laws of multiplication and addition.

IV. Packet Delivery Delay Analysis

With the help of the Markov-chain framework and related basic results in Section III, this section provides the study of both expected value and standard deviation of packet delivery delay under the 2HR-(f, g) algorithm. We first introduce the following definition about the delivery delay of a packet group.

Definition 1: For a packet group at a source node S, the delivery delay of the group is the time elapsed between the time slot S moves the first packet of the group into the head-of-line at its local-queue and the time slot when the destination node D receives the last packet of the group.

For the 2HR-(f, g) relay algorithm, if we denote by $T(f, g)$ the delivery delay of a packet group and denote by T_p the average delivery delay of one packet, then we have

$$T_p = \frac{T(f,g)}{g}. \tag{11}$$

Remark 5: Under the 2HR-(f, g) algorithm, the destination node D queues up the packets of a group until it receives all packets of that group, and then considers the packets of the group delivered. Thus, as the (11) shows that the per packet delivery delay T_p is calculated on a group basis.

A. Expected Packet Delivery Delay

As illustrated in Fig. 5, all β transient states in the Markov chain model are arranged into g rows. We number these transient states sequentially as 1, 2, ..., β, in a left-to-right and top-to-down way. For these transient states, let q_{ij} denote the transition probability from transient sate i to transient state j, then we can define a matrix $\mathbf{Q} = (q_{ij})_{\beta\times\beta}$ of transition probabilities among transient states there. From the theory of Markov chain [24] we know that the fundamental matrix $\mathbf{N}$ of the Markov chain in Fig. 5 is given by

$$\mathbf{N} = (\mathbf{I} - \mathbf{Q})^{-1}, \tag{12}$$

where $\mathbf{N} = (a_{ij})_{\beta\times\beta}$ and the entry a_{ij} denotes the expected number of times in the j_{th} transient state until absorption given that the chain starts from the i_{th} transient state.

Based on the Markov chain structure in Fig. 5, we can actually partition the matrix $\mathbf{N}$ into g-by-g blocks as $\mathbf{N} = (\mathbf{N}_{tk})_{g\times g}$, where the block (i.e., sub-matrix) $\mathbf{N}_{tk}$ corresponds to the expected number of times in the transient states of $(k-1)_{th}$ row of the Markov chain structure given that the Markov chain starts from the transient states of $(t-1)_{th}$ row there. If we use $\mathbf{N}_{tk}(i, j)$ to denote the ij-entry of a block $\mathbf{N}_{tk}$, and further use b_i to denote the time the Markov chain takes to become absorbed given that the chain starts from the i_{th} transient state ($1 \le i \le \beta$), then we have

$$T(f,g) = b_1, \tag{13}$$

where the expected value $\mathbb{E}\{b_1\}$ of b_1 is given by

$$\mathbb{E}\{b_1\} = \sum_{k=1}^{g}\sum_{j=1}^{L_{k-1}} \mathbf{N}_{1k}(1,j). \tag{14}$$

Then the expected packet delivery delay $\mathbb{E}\{T_p\}$ can be determined as

$$\mathbb{E}\{T_p\} = \frac{1}{g}\sum_{k=1}^{g}\sum_{j=1}^{L_{k-1}} \mathbf{N}_{1k}(1,j). \tag{15}$$

B. Standard Deviation

From (11) and (13) we can easily see that the variance of packet delivery delay $Var\{T_p\}$ can be determined as

$$Var\{T_p\} = \frac{1}{g^2} Var\{b_1\}. \tag{16}$$

Since $Var\{b_1\} = \mathbb{E}\{{b_1}^2\} - (\mathbb{E}\{b_1\})^2$ and $\mathbb{E}\{b_1\}$ can be determined by (14), we only need to derive the $\mathbb{E}\{{b_1}^2\}$ here.

Based on the definition of b_i we can see that the expected value $\mathbb{E}\{{b_i}^2\}$ is given by

$$\begin{aligned}\mathbb{E}\{{b_i}^2\} &= \sum_{j=1}^{\beta} q_{ij}\mathbb{E}\{(1+b_j)^2\} \\ &= 1 + 2\sum_{j=1}^{\beta} q_{ij}\cdot\mathbb{E}\{b_j\} + \sum_{j=1}^{\beta} q_{ij}\cdot\mathbb{E}\{{b_j}^2\}.\end{aligned} \tag{17}$$

Let $\mathbf{b}^{(j)} = (\mathbb{E}\{{b_1}^j\}, \mathbb{E}\{{b_2}^j\}, \ldots, \mathbb{E}\{{b_\beta}^j\})^T$, then we can rearrange (17) as

$$\mathbf{I}\cdot\mathbf{b}^{(2)} = \mathbf{c} + 2\mathbf{Q}\cdot\mathbf{b}^{(1)} + \mathbf{Q}\cdot\mathbf{b}^{(2)}, \tag{18}$$

where $\mathbf{c}$ is the $\beta \times 1$ column vector with all entries being 1, i.e., $\mathbf{c} = \{1, 1, \ldots, 1\}^T$.

Then, according to [25], we have

$$\mathbf{b}^{(1)} = \mathbf{N}\cdot\mathbf{c}, \tag{19}$$

$$\mathbf{b}^{(2)} = \mathbf{N}(\mathbf{I} + 2\mathbf{Q}\cdot\mathbf{N})\mathbf{c}. \tag{20}$$

Since $\mathbb{E}\{{b_1}^2\} = \mathbf{e}\cdot\mathbf{b}^{(2)}$, where $\mathbf{e} = \{1, 0, \ldots, 0\}$, the $\mathbb{E}\{{b_1}^2\}$ and $Var\{T_p\}$ can be derived based on $\mathbf{Q}$ and $\mathbf{N}$.

The above results indicate clearly that the only remaining issue for the calculation of both $\mathbb{E}\{T_p\}$ and $Var\{T_p\}$ is the derivation of matrices $\mathbf{Q}$ and $\mathbf{N}$, as discussed in the follows.

C. Derivation of Matrix $\mathbf{Q}$

Notice that for the Markov chain in Fig. 5, the transitions happen only among the transient states of the same row or neighboring rows, so the matrix $\mathbf{Q}$ there can be defined as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_0 & \mathbf{Q}_0^{'} & & & & & & \\ & \mathbf{Q}_1 & \mathbf{Q}_1^{'} & & & & & \\ & & \ddots & \ddots & & & & \\ & & & \mathbf{Q}_k & \mathbf{Q}_k^{'} & & & \\ & & & & \ddots & \ddots & & \\ & & & & & \mathbf{Q}_{g-2} & \mathbf{Q}_{g-2}^{'} \\ & & & & & & \mathbf{Q}_{g-1} \end{bmatrix}, \quad (21)$$

where the block (sub-matrix) $\mathbf{Q}_k$ of size $L_k \times L_k$ corresponds to the transition probabilities among the transient states of k_{th} row in the Markov chain, while the block $\mathbf{Q}_k^{'}$ of size $L_k \times L_{k+1}$ corresponds to the transition probabilities from the transient states of the k_{th} row to that of the $(k+1)_{th}$ row. All the other blocks are zero matrices here and thus omitted for simplicity. The (21) indicates that to derive the matrix $\mathbf{Q}$, we just need to calculate the sub-matrices $\mathbf{Q}_k$ and $\mathbf{Q}_k^{'}$ there.

Calculation of $\mathbf{Q}_k$: Let $\mathbf{Q}_k(i,j)$ denote the ij-entry of the sub-matrix $\mathbf{Q}_k$, $i,j \in [1, L_k]$, then the non-zero entries of $\mathbf{Q}_k$ can be determined as

- when $1 \le k \le g-1$,

$$\mathbf{Q}_k(i,i+1) = \begin{cases} P_d(u_o) & \text{if } 1 \le i \le f, \\ P_d(u_o) - P_s(u_r,u_o) & \\ & \text{if } f+1 \le i < L_k, \end{cases} \quad (22)$$

$$\mathbf{Q}_k(i,i) = \begin{cases} 1 - p_1 - P_d(u_o) & \text{if } 1 \le i \le f, \\ 1 - P_d(u_o) - P_r(u_r) + P_s(u_r,u_o) & \\ & \text{if } f+1 \le i < L_k, \\ 1 - P_r(u_r) & \text{if } i = L_k. \end{cases} \quad (23)$$

- when $k=0$,

$$\mathbf{Q}_0(i,i+1) = P_d(n-1-i) - P_s(i-1,n-1-i) \quad \text{if } 1 \le i < L_0, \quad (24)$$

$$\mathbf{Q}_0(i,i) = \begin{cases} 1 - P_d(n-1-i) - P_r(i-1) & \\ \quad +P_s(i-1,n-1-i) & \text{if } 1 \le i < L_0, \\ 1 - P_r(g \cdot f) & \text{if } i = L_0. \end{cases} \quad (25)$$

Calculation of $\mathbf{Q}_k^{'}$: The sub-matrix $\mathbf{Q}_k^{'}$ is of size $L_k \times L_{k+1}$, where its non-zero ij-entry $\mathbf{Q}_k^{'}(i,j)$ is determined as

- when $1 \le k \le g-2$,

$$\mathbf{Q}_k^{'}(i, j(i)\cdot f) = p_1 \quad \text{if } 1 \le i < L_k, \quad (26)$$

where

$$j(i) = \begin{cases} 1 & \text{if } 1 \le i \le f, \\ \lfloor \frac{i}{f} \rfloor & \text{if } f+1 \le i < L_k, \end{cases} \quad (27)$$

and $\lfloor x \rfloor$ returns the largest integer not greater than x.

$$\mathbf{Q}_k^{'}(i,i-f) = \begin{cases} P_r(u_r) - p_1 - P_s(u_r,u_o) & \\ & \text{if } f+1 \le i < L_k, \\ P_r(u_r) & \text{if } i = L_k, \end{cases} \quad (28)$$

$$\mathbf{Q}_k^{'}(i,i-f+1) = P_s(u_r,u_o) \quad \text{if } f+1 \le i < L_k. \quad (29)$$

- when $k=0$,

$$\mathbf{Q}_0^{'}(i, \lceil \frac{i}{f} \rceil \cdot f) = p_1 \quad \text{if } 1 \le i < L_0, \quad (30)$$

and $\lceil x \rceil$ returns the smallest integer not less than x.

$$\mathbf{Q}_0^{'}(i,i-1) = \begin{cases} P_r(i-1) - p_1 - P_s(i-1,n-1-i) & \\ & \text{if } 2 \le i < L_0, \\ P_r(g \cdot f) & \text{if } i = L_0, \end{cases} \quad (31)$$

$$\mathbf{Q}_0^{'}(i,i) = P_s(i-1,n-1-i) \quad \text{if } 2 \le i < L_0. \quad (32)$$

D. Derivation of Matrix $\mathbf{N}$

We denote the matrix $\mathbf{I} - \mathbf{Q}$ as $\mathbf{G}$, so $\mathbf{N} = \mathbf{G}^{-1}$. Based on the structure of $\mathbf{Q}$ we can see that the matrix $\mathbf{G}$ can also be defined in the block structure. Let $\{\mathbf{G}_k\}$ and $\{\mathbf{G}_k^{'}\}$ denote the main diagonal and upper diagonal blocks of $\mathbf{G}$, then we have

$$\mathbf{G}_k(i,j) = \begin{cases} 1 - \mathbf{Q}_k(i,j) & \text{if } i = j, \\ -\mathbf{Q}_k(i,j) & \text{otherwise,} \end{cases} \quad (33)$$

$$\mathbf{G}_k^{'}(i,j) = -\mathbf{Q}_k^{'}(i,j). \quad (34)$$

The following lemma indicates that the matrix $\mathbf{N}$ can be calculate based on $\{\mathbf{G}_k^{-1}\}$ and $\{\mathbf{G}_k^{'}\}$.

Lemma 4: The fundamental matrix $\mathbf{N} = (\mathbf{N}_{ij})_{g \times g}$ of the Markov chain in Fig. 5 can be determined as

$$\mathbf{N}_{ij} = \begin{cases} \mathbf{0} & \text{if } i > j, \\ \mathbf{G}_{i-1}^{-1} & \text{if } i = j, \\ (-1)^{j-i}\big(\prod_{k=i-1}^{j-2} \mathbf{G}_k^{-1}\mathbf{G}_k^{'}\big)\mathbf{G}_{j-1}^{-1} & \text{if } i < j, \end{cases} \quad (35)$$

where $i,j \in [1,g]$.

Proof: We can easily see that the matrix $\mathbf{G} = \mathbf{I} - \mathbf{Q}$ also has a block partition similar to that of $\mathbf{Q}$ in (21), with sub-matrices (blocks) in the main diagonal and upper diagonal, while other blocks are zero matrices. As to be proved in Lemma 5, each main diagonal block $\mathbf{G}_k$ of $\mathbf{G}$ is invertible (i.e., its inverse matrix $\mathbf{G}_k^{-1}$ exists), $k \in [0, g-1]$. Similarly, we can easily see that $\mathbf{G}$ is invertible and its inverse matrix $\mathbf{G}^{-1}$ (i.e., $\mathbf{N}$) is an upper block triangular matrix. Thus, the formula (35) follows after some basic row operations. ∎

Since $\{\mathbf{G}_k^{'}\}$ can be calculated based on (34), the only remaining issue for evaluating $\mathbf{N}$ is to determine $\{\mathbf{G}_k^{-1}\}$.

Lemma 5: Each sub-matrix $\mathbf{G}_k$ of $\mathbf{G}$ has an inverse $\mathbf{G}_k^{-1}$, where ij-entry $\mathbf{G}_k^{-1}(i,j)$ of $\mathbf{G}_k^{-1}$ can be determined as

$$\mathbf{G}_k^{-1}(i,j) = \begin{cases} 0 & \text{if } i > j, \\ \frac{1}{\mathbf{G}_k(i,i)} & \text{if } i = j, \\ (-1)^{j-i}\big(\prod_{t=i}^{j-1} \frac{\mathbf{G}_k(t,t+1)}{\mathbf{G}_k(t,t)}\big)\frac{1}{\mathbf{G}_k(j,j)} & \text{if } i < j, \end{cases} \quad (36)$$

where $k \in [0, g-1]$, $i,j \in [1, L_k]$.

Proof: As indicated in (33) that for each $\mathbf{G}_k$, we have $\mathbf{G}_k = \mathbf{I}_k - \mathbf{Q}_k$. Obviously, $\mathbf{G}_k$ is a square matrix of size

$L_k \times L_k$. Combining the definitions of $\mathbf{Q}_k$ in the (22), (23), (24) and (25) with the fact that $0 < \mathbf{Q}_k(i,i) < 1$ and $\mathbf{Q}_k(i,i+1) > 0$, we know that $0 < \mathbf{G}_k(i,i) < 1$, $\mathbf{G}_k(i,i+1) < 0$ and all other off-diagonal entries are zero. Thus, $\mathbf{G}_k$ is invertible and its inverse matrix $\mathbf{G}_k^{-1}$ is an upper triangular matrix. After some basic row operations, the (36) follows. ■

V. Numerical Results

In this section, we first verify the efficiency of the Markov chain-based framework through simulation, then apply it to explore how the parameters f and g would affect the packet delivery delay in a 2HR-(f,g) MANET.

A. Simulation Setting

A simulator was developed to simulate the packet delivery process in a 2HR-(f,g) MANET, which is now available at [26]. Similar to the settings adopted in [27], [28], the guard factor here is fixed as $\Delta = 1$, and hence the concurrent-set is defined with $\alpha = \min\{8, m\}$. Besides the bi-dimensional i.i.d. mobility model considered in this paper, we also implemented the simulator for the popular random walk model and random waypoint model, which are defined as follows:

- **Random Walk Model [29]:** At the beginning of each time slot, each node independently makes a decision regarding its mobility action, either staying inside its current cell or moving to one of its eight adjacent cells. Each action happens with the same probability of $1/9$.
- **Random Waypoint Model [30]:** At the beginning of each time slot, each node independently and randomly generates a two-dimensional vector $v = [v_x, v_y]$, where the values of v_x and v_y are uniformly drawn from $[1/m, 3/m]$. The node then moves a distance of v_x along the horizontal direction and a distance of v_y along the vertical direction.

The simulated expected delivery delay (SE) is calculated as the average value of 10^2 batches of simulation results, where each batch consists of 10^4 random and independent simulations. The simulated standard deviation (SSD) is the sample standard deviation, which is calculated as

$$SSD = \sqrt{\frac{1}{w-1}\sum_{i=1}^{w}(x_i - SE)^2}, \quad (37)$$

where $w = 10^6$, and x_i is the observed delivery delay in the i_{th} simulation. Notice that all the simulation results of the expected delivery delay are reported with the 95% confidence intervals.

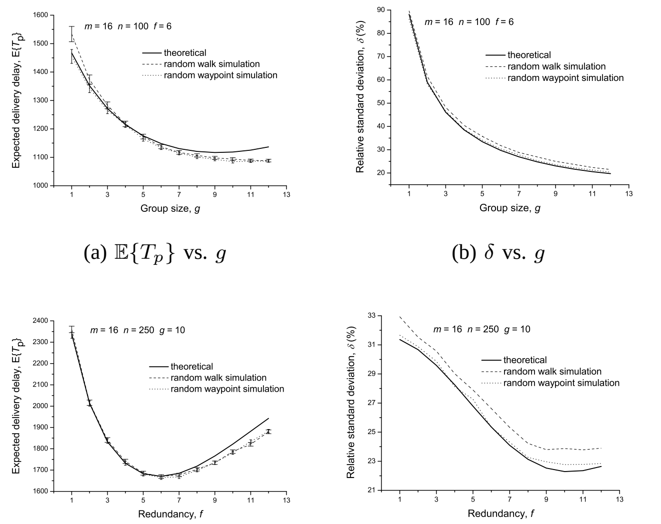
B. Model Validation

Extensive simulations have been conducted to verify the Markov chain-based theoretical framework. For the fixed setting of $m = 16$, we considered three different network scenarios of $n = 100$, 250 and 600, which correspond to the sparse network (with node density 0.39), ordinary network (with node density 0.98) and dense network (with node density 2.34), respectively. For each network scenario, three different settings of parameters f and g have been examined, i.e., $(g = 1, f = 2)$, $(g = 5, f = 4)$ and $(g = 10, f = 6)$. The corresponding simulation results and theoretical results are summarized in Table I.

Table I indicates clearly that the simulation results match nicely with the theoretical ones for both the expected value and standard deviation of packet delivery delay, so our theoretical framework can be used to efficiently model the packet delivery process. A further careful observation of Table I shows that there is still a very small gap ($\leq 5\%$) between the simulation results and theoretical ones. For example, for the case that $n = 100$, $g = 10$ and $f = 6$, the simulated value for $\mathbb{E}\{T_p\}$ is 1086.6 while the theoretical value is 1118.7. Regarding the standard deviation, when $n = 600$, $g = 1$ and $f = 2$, the simulated and theoretical results are 4671.6 and 4665, respectively. This small gap is mainly due to the following two reasons. The first one is that the simplification adopted in the Assumption 1 slightly "slows" down the absorbing speed of the Markov chain, and thus results in a higher absorption time (i.e., delivery delay). The other reason is that we adopted approximations (3) and (4) for the fresh nodes and non-fresh nodes in the theoretical delay analysis, which made the theoretical results shift slightly from simulation ones.

TABLE I
Comparison between simulated and theoretical results for model validation, $m = 16$, Simulated / Theoretical

		$\mathbb{E}\{T_p\}$	$\sqrt{Var\{T_p\}}$
$n = 100$	$g = 1, f = 2$	$1850.3 \pm 3.5/1849.7$	$1805.5/1803.9$
	$g = 5, f = 4$	$1196.7 \pm 0.9/1198.8$	$456.62/456.81$
	$g = 10, f = 6$	$1086.6 \pm 0.5/1118.7$	$242.19/242.68$
$n = 250$	$g = 1, f = 2$	$2703.1 \pm 5.3/2702.5$	$2676.3/2675.1$
	$g = 5, f = 4$	$2043.1 \pm 1.7/2043.2$	$845.23/844.89$
	$g = 10, f = 6$	$1662.8 \pm 0.8/1670.7$	$424.09/423.55$
$n = 600$	$g = 1, f = 2$	$4686.2 \pm 9.2/4685$	$4671.6/4665$
	$g = 5, f = 4$	$4045.3 \pm 3.4/4047.6$	$1741.1/1739.6$
	$g = 10, f = 6$	$3444.5 \pm 1.9/3446$	$967.89/967.94$



(a) $\mathbb{E}\{T_p\}$ vs. g (b) δ vs. g

(c) $\mathbb{E}\{T_p\}$ vs. f (d) δ vs. f

Fig. 6. Delivery delay vs. group size g and redundancy f under random walk and random waypoint mobility models.

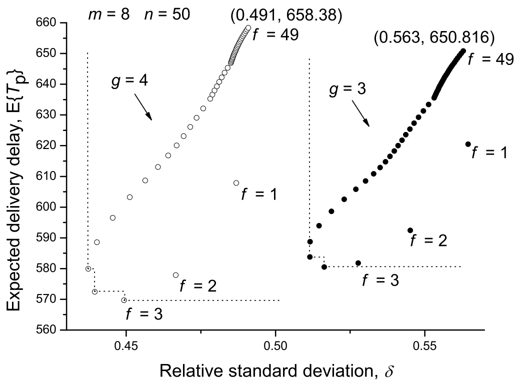


Fig. 7. Achievable delay performance region of a 2HR-(f, g) MANET for the cases of $m = 8$, $n = 50$ and $g = \{3, 4\}$.

To further illustrate the applicability of our theoretical framework to other mobility models, we show in Fig. 6 the $\mathbb{E}\{T_p\}$ and relative standard deviation δ of packet deliver delay under the i.i.d., random walk and random waypoint mobility models, where δ is defined as

$$\delta = \frac{\sqrt{Var\{T_p\}}}{\mathbb{E}\{T_p\}}. \tag{38}$$

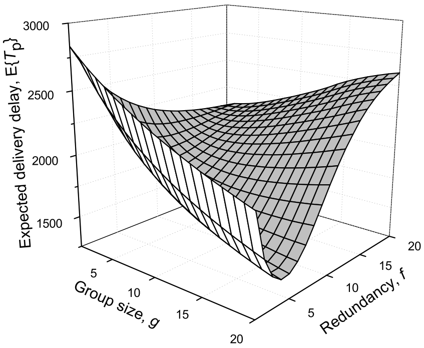
It's interesting to observe from Fig. 6 that the analytical models of $\mathbb{E}\{T_p\}$ and δ, although were developed under the i.i.d. mobility model, can also well approximate the general trends of $\mathbb{E}\{T_p\}$ and δ under the other two mobility models.

Regarding the $\mathbb{E}\{T_p\}$ performance, the results in Fig. 6a and Fig. 6c indicate that the behavior of $\mathbb{E}\{T_p\}$ vs. g under the i.i.d model is slightly different from that of under other two models, but the behavior of $\mathbb{E}\{T_p\}$ vs. f is similar for all the three mobility models. In particular, from Fig. 6a we can see that for the concerned network scenario the minimum $\mathbb{E}\{T_p\}$ of the i.i.d model is reached at $g = 9$, while the minimum $\mathbb{E}\{T_p\}$ of the random walk and random waypoint are reached at $g = 12$ and $g = 10$, respectively. For the $\mathbb{E}\{T_p\}$ vs. f results in Fig. 6c, however, the minimum $\mathbb{E}\{T_p\}$ is reached at the same setting of $f = 6$ for all three mobility models. Different from that of the $\mathbb{E}\{T_p\}$ performance, the results of δ in Fig. 6b and Fig. 6d show that the behavior of δ vs. g is very similar for all three mobility models, while the δ vs. f behavior of the random walk is a little different from that of other two models.

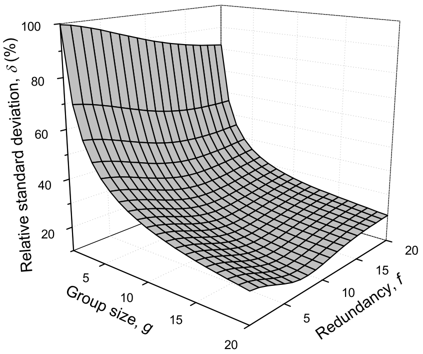
C. Achievable Delay Region

Based on the new Markov chain theoretical framework, we now explore the achievable delay performance region of the 2HR-(f, g) algorithm in terms of its (δ, $\mathbb{E}\{T_p\}$) under the i.i.d. model. For the scenario of $m = 8$, $n = 50$ and $g = \{3, 4\}$, Fig. 7 shows the region of (δ, $\mathbb{E}\{T_p\}$) that the 2HR-(f, g) can achieve by varying the parameter f. Notice that each curve in Fig. 7 consists of multiple discrete points and each point corresponds to a specific value of f, so it may happen that two distinct settings of f achieve different $\mathbb{E}\{T_p\}$ but very similar (even the same) δ. For example, for the curve of $g = 4$, the settings of $f = 1$ and $f = 36$ achieve the same $\delta = 0.487$ but different expected delivery delay of 607.875 and 651.632, respectively. The results in Fig. 7 indicate that the 2HR-(f, g) algorithm actually enables the delay performance (δ, $\mathbb{E}\{T_p\}$) to be flexibly controlled in a large region to adapt to various applications with different delay (and variance) requirements.

It is interesting to notice from Fig. 7 that for a specified group size g, the achievable delay performance region is actually defined by some vertical and horizontal lines determined by several key points, i.e., the Pareto optimal points [31]. For example, when $g = 3$, the achievable delay performance region is determined by the point $(0.516, 580.506)$ ($f = 4$) that results in the minimum $\mathbb{E}\{T_p\}$ of 580.506 and the point $(0.511, 583.765)$ ($f = 5$) that results in the minimum δ of 0.511. For the case that $g = 4$, the achievable delay region is co-determined by three points, i.e., the point $(0.449, 569.695)$ ($f = 3$), point $(0.439, 572.447)$ ($f = 4$) and point $(0.437, 579.933)$ ($f = 5$). Thus, for a specified group size g, any delay performance requirement in terms of (δ, $\mathbb{E}\{T_p\}$) can be supported by the 2HR-(f, g) algorithm as long as the point (δ, $\mathbb{E}\{T_p\}$) falls within the corresponding performance region defined by the group size.



(a) $\mathbb{E}\{T_p\}$ vs. (g, f)

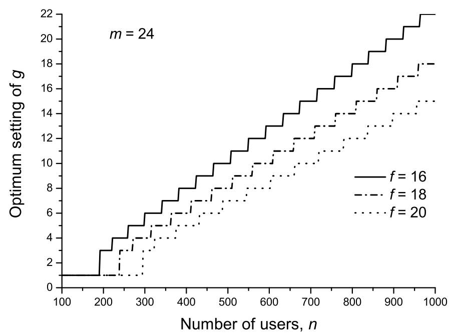


(b) δ vs. (g, f)

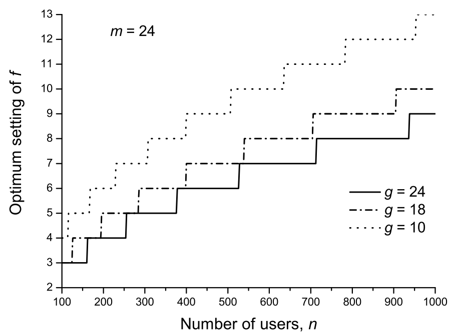
Fig. 8. Delivery delay for a specific network $m = 16$, $n = 250$.

D. Delay Control

To see how the delay can be controlled according to a specified delay target, we now apply our framework to a

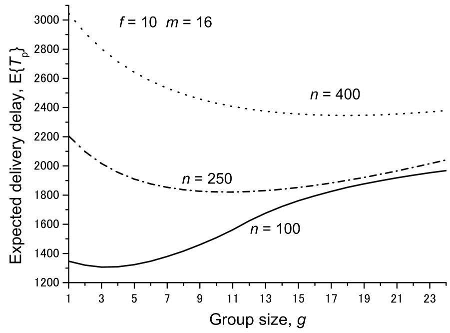


(a) Optimum setting of g vs. n

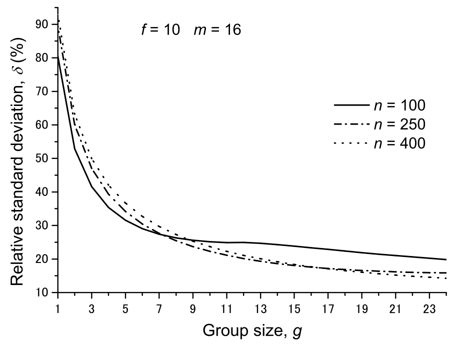


(b) Optimum setting of f vs. n

Fig. 9. Optimum parameter settings vs. number of nodes n



(a) $\mathbb{E}\{T_p\}$ vs. g



(b) δ vs. g

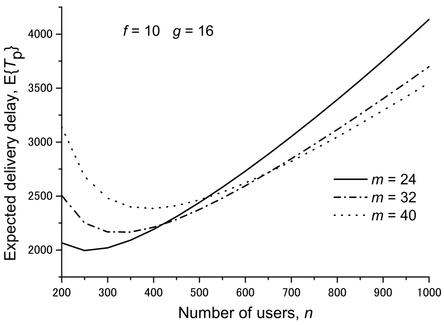
Fig. 10. Delivery delay vs. group size g

network scenario of ($m = 16$, $n = 250$) and show in Fig. 8 how its delay performance $(\delta, \mathbb{E}\{T_p\})$ varies with both g and f there. As shown in Fig. 8a (resp. Fig. 8b) that for a specified target t_p of the mean delay value (resp. a target δ_0 of the relative standard deviation), we can accordingly define a target plane intersecting the z-axis orthogonally at the point $(1, 1, t_p)$ (resp. at the point $(1, 1, \delta_0)$), and thus can get a set of (g, f)-pairs corresponding to the surface below the defined target plane there. By finding the intersection of these two sets of (g, f)-pairs, we can determine the set of (g, f)-pairs to achieve the specified delay target in terms of t_p and δ_0. Fig. 8a also shows that for the network scenario there, although the $\mathbb{E}\{T_p\}$ has different varying trends with g and f but a minimum delivery delay can always be identified. For example, we can see that when $g = 1$, the $\mathbb{E}\{T_p\}$ monotonically decreases with f; for any fixed $g \geq 2$, we can find an optimum setting of f to achieve the corresponding minimum delivery delay. Similarly, when $f \leq 6$ the $\mathbb{E}\{T_p\}$ monotonically decreases with g; for any fixed $f \geq 7$, there also exists an optimum setting of g to achieve the minimum delivery delay. For the network scenario here and all $g, f \in [1, 20]$, the global minimum delivery delay of 1426.75 is achieved at the setting of ($g = 20, f = 4$).

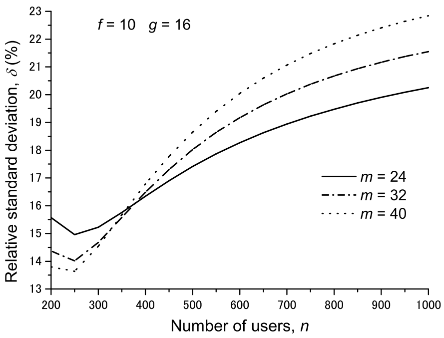
Fig. 8a indicates that for a network scenario with fixed g (resp. f), there exists a corresponding optimum setting of f (resp. optimum setting of g) to achieve the minimal delivery delay. We show in Fig. 9 how such optimum setting of g (resp. f) varies under different network scenarios. One can observe from Fig. 9 that there does not exist a particular optimal value of g (or f) which applies to all networks of different size n. Actually, the optimal setting of g (or f) is a piecewise function of n and one optimal setting only applies to a small range of n. A further careful observation of Fig. 9 indicates that, as n scales up, the optimum setting of f in Fig. 9b becomes less sensitive to the variation of n (i.e., as n increases up, an optimal setting of f applies to a wider range of n), but this is not the case for the optimum setting of g in Fig. 9a. Thus, compared with the optimum setting of f (under a given g), the optimum setting of g (under a given f) depends more heavily on n.

E. Performance Analysis

We now explore how the performance (δ, $\mathbb{E}\{T_p\}$) of the 2HR-(f, g) algorithm varies with different parameters. For the scenarios of $n = \{100, 250, 400\}$ and the fixed setting of $f = 10$ and $m = 16$, Fig. 10 illustrates how $\mathbb{E}\{T_p\}$ and δ vary with group size g. It is interesting to see from Fig. 10 that for a given network with a fixed value of f, the δ always monotonously decreases as g increases, but this is not the case for $\mathbb{E}\{T_p\}$. As shown in Fig. 10b that when $g \leq 2$, the δ is quite high (larger than 50%). It is notable, however, that for most of the MANET applications, the destination node will allow a certain degree of packet out of order determined by the parameter g, so a moderate value of g is usually acceptable. As we can see from Fig. 10b that as g increases beyond 2, the δ drops dramatically to a low level for all



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



(b) δ vs. n

Fig. 11. Delivery delay vs. number of nodes n

three network scenarios considered here, which indicates that our algorithm can stably control the delivery delay for most interested settings of g. Notice also that the affordable group size is limited by the buffer size at each mobile node, and a large group may unavoidably force the early arrived packets at the destination node to wait a long time for other packets (of the same group), which may make the early arrived packets become expired before the arrival of the last packet of the same group. This indicates that using a large group size in the 2HR-(f, g) algorithm may significantly limit its applications to support the delay-sensitive applications in the MANETs. Thus, the group size g should be carefully dimensioned with the considerations of $\mathbb{E}\{T_p\}$, δ and buffer limitation in each node.

Finally, we examine in Fig. 11 how metrics δ and $\mathbb{E}\{T_p\}$ vary with network size n, given that $f = 10$, $g = 16$, and $m = \{24, 32, 40\}$. We can see from Fig. 11a that for a given m (determined by communication range r as $m = \sqrt{8}/r$), we can find a most suitable network size n^* to achieve the minimum expected packet delay $\mathbb{E}\{T_p\}$. A further careful observation of Fig. 11a indicates that the most suitable network size for a minimum expected packet delay varies with m and can be roughly determined as $n^* \approx m \cdot f$. For example, for the setting of $m = 24$, 32 and 40, the corresponding n^* are roughly 240, 320 and 400, respectively. Regarding the performance of δ, the results in Fig. 11b indicate that for a given m, there also exists a most suitable network size to achieve the minimum δ. However, it is interesting to see from Fig. 11b that the most suitable network size for a minimum δ is always 250 for the scenario of $f = 10$ and $g = 16$ here, which actually does not change as m varies.

VI. Related Works

A significant amount of work has been done on the delay performance of the two-hop relay algorithms. These works mainly focus on closed-form analysis or order-sense scaling law study of expected packet delay in a two-hop relay network.

A. Closed-form Delay Analysis

Liu *et al.* [5] considered a two-hop relay algorithm with redundancy and in-order reception in a time-slotted system, and derived closed-form results for the expected end-to-end per packet delay. The expected delivery delay analysis under continuous system models was conducted in [17]–[20], where the inter-meeting time between two nodes, i.e., the time elapsed between two consecutive encounters for a given pair of nodes, is assumed to be exponentially distributed. The network scenarios considered in [17]–[20] were relatively simple, where the network has only one source-destination pair, and the source node has only one single packet to deliver to the destination. It is notable that the Markov chain model in [17]–[20] can be regarded as a special case (i.e., $g = 1$) of our general theoretical framework developed in this paper.

B. Order-sense Delay Scaling Laws

The delay scaling law of the two-hop relay with out-of-order reception but without redundancy has been extensively examined in the regime of ad hoc mobile networks. Gamal *et al.* [32] showed that under the random walk model, the two-hop relay results in a $\Theta(n \log n)$ delay and achieves a $\Theta(1)$ throughput. Later, Mammen *et al.* [7] proved that the same delay and throughput scalings are also achievable even with a variant of the two-hop relay and a restricted mobility model. Gamal *et al.* [33] showed that under the two-dimensional Brownian motion on a torus of size $\sqrt{n} \times \sqrt{n}$, the delay scales as $\Theta(n^{1/2}/v(n))$, where $v(n)$ is the velocity of mobile nodes. Lin *et al.* [34] also considered the Brownian mobility model, and showed that the two-hop relay results in an expected delay of $\Omega(\log n/\sigma_n^2)$, where σ_n^2 is the variance parameter of the Brownian motion model. Sharma *et al.* [35] further showed that when the network is divided into $n^\alpha \times n^\alpha$ cells, the two-hop delay is $\Theta(n)$ for $0 \leq \alpha < 1/2$ and $\Theta(n \log n)$ for $\alpha = 1/2$ under a family of discrete random direction models, while the delay becomes $\Theta(n)$ for $\alpha < 1/2$ and $\Theta(n \log n)$ for $\alpha = 1/2$ when a family of hybrid random walk models are considered. Recently, the delay performance of a variant of two-hop relay has been examined under a correlated mobility model [21], where nodes are partitioned into different groups and all nodes of the same group have to reside concurrently within a circular region around the group center.

In the case of allowing packet redundancy and in-order reception, Neely and Modiano [3] considered a modified version of the two-hop relay algorithm for ad hoc mobile networks, and proved that under the i.i.d. mobility model it achieves $O(\sqrt{n})$ delay with exact $\sqrt{n}$ redundancy for each packet. Sharma and Mazumdar explored the order-sense delay results in ad hoc mobile networks with multiple redundancy for each packet, and proved that it achieves $\Theta(T_p(n)\sqrt{n})$ delay under the random way-point mobility model [36] and achieves $O(T_p(n)\sqrt{n\log n})$ delay under the Brownian mobility model [4], where $T_p(n)$ is the packet transmission time.

VII. Conclusion

This paper proposed a general 2HR-(f,g) algorithm for MANETs, and also developed a Markov chain-based theoretical framework for corresponding performance modeling. We proved that the 2HR-(f,g) algorithm has the capability of flexibly controlling packet delay and its variance in a large region, an important property for future MANETs to support various applications of different delay (and delay variance) requirements. The results in this paper indicate that the control parameters f and g of the 2HR-(f,g) algorithm may affect the packet delay and its variance in very different ways, and a target packet delay (and delay variance) requirement can be actually achieved through various combinations between f and g. Thus, a careful trade-off among packet delay (and delay variance) requirement, packet redundancy (f) and node buffer limitation (related to g) should be examined for the efficient support of a target application.

Notice that in the proposed 2HR-(f,g) algorithm, we considered a very simple scenario where only one node is randomly selected from the one-hop neighbors for the source-to-relay transmission or relay-to-destination transmission, which may cause a waste of the transmission opportunity if a wrong node is selected. Therefore, one future work is to further explore the performance of 2HR-(f,g) under a more flexible scenario, where not only one but many (even all) one-hop neighbors will be considered for the source-to-relay or relay-to-destination transmission to take the full advantage of each transmission opportunity. Another interesting future direction is to extend the theoretical models in this paper to analytically determine the combinations of group size g and redundancy f for the proposed 2HR-(f,g) algorithm to satisfy a given delay requirement and further derive the optimum combination of g and f to achieve the minimum delivery delay under a specific network scenario.

Appendix A
Proof of the Lemmas 1,2 and 3

Proof of Lemma 1: From the Assumption 1 we can easily see that for a transient state (i,j,k) in the k_{th} row of the Markov chain model in Fig. 5, the number of corresponding fresh nodes u_r can be approximated as

$$u_r \approx i-1+(j-1-k)f. \tag{39}$$

Notice that in a large MANET, the probability of direct source-to-destination transmission is negligible in comparison with that of the source-to-relay or relay-to-destination transmissions, so with high probability the destination will receive each of the k packets from a relay node rather than the source node. Thus, the number of corresponding non-fresh nodes u_o can be approximated as

$$u_o \approx n-2-(i-1)-(j-1)f+k. \tag{40}$$

Suppose that the state (i,j,k) is the t_{th} transient state in the k_{th} row of the Markov chain model in Fig. 5, $k\in[1,g-1]$, $t<L_k$, then we have

$$t=(j-k)f+i-1. \tag{41}$$

By combining (41) with (39) and (40), the formulas (3) and (4) then follow.

Proof of Lemma 2: Consider a tagged active cell, the node S can conduct a source-to-destination transmission with D only under the following two mutually exclusive cases: both S and D are in this cell; or S is in this cell while D is in the eight adjacent cells of this cell. If we further assume that aside from S and D, there are k other nodes in this cell, $k\in[0,n-2]$, then the probability that S is selected as the transmitter is $\frac{1}{k+2}$ (resp. $\frac{1}{k+1}$) under the former case (resp. under the latter case). Summing up the probabilities under these two cases, then we have

$$\begin{aligned}
p_1 &= \frac{1}{\alpha^2}\Bigg(\sum_{k=0}^{n-2}\binom{n-2}{k}\Big(\frac{1}{m^2}\Big)^k\Big(\frac{m^2-1}{m^2}\Big)^{n-2-k}\frac{1}{m^2(k+2)} \\
&\quad+\sum_{k=0}^{n-2}\binom{n-2}{k}\Big(\frac{1}{m^2}\Big)^k\Big(\frac{m^2-1}{m^2}\Big)^{n-2-k}\frac{8}{m^2(k+1)}\Bigg) \\
&= \frac{1}{\alpha^2}\Bigg(\sum_{k=0}^{n-2}\binom{n-1}{k+1}\Big(\frac{1}{m^2}\Big)^{k+1}\Big(\frac{m^2-1}{m^2}\Big)^{n-2-k}\frac{1}{k+2} \\
&\quad-\sum_{k=0}^{n-2}\binom{n-2}{k+1}\Big(\frac{1}{m^2}\Big)^{k+1}\Big(\frac{m^2-1}{m^2}\Big)^{n-2-k}\frac{1}{k+2} \\
&\quad+\sum_{k=0}^{n-2}\binom{n-2}{k}\Big(\frac{1}{m^2}\Big)^{k+1}\Big(\frac{m^2-1}{m^2}\Big)^{n-2-k}\frac{8}{k+1}\Bigg) \\
&= \frac{1}{\alpha^2}\Bigg(\frac{9-m^2}{n-1}+\frac{m^2}{n}-\frac{8}{n-1}\Big(\frac{m^2-1}{m^2}\Big)^{n-1} \\
&\quad+\Big(\frac{m^2}{n-1}-\frac{m^2}{n}\Big)\Big(\frac{m^2-1}{m^2}\Big)^{n}\Bigg).
\end{aligned} \tag{42}$$

The formula (5) can then be easily derived from (42) after some basic algebraic operations.

Similarly, S conducts a source-to-relay or relay-to-destination transmission iff the following four events happen simultaneously: S is in an active cell, S is selected as the transmitter, there is at least one other node (except S and D) in the same cell of S or its eight adjacent cells, and the node D is in one of the other m^2-9 cells. Thus, we have

$$\begin{aligned}
p_2 &= \frac{m^2-9}{m^2\alpha^2}\Bigg(\sum_{k=1}^{n-2}\binom{n-2}{k}\Big(\frac{1}{m^2}\Big)^k\Big(\frac{m^2-1}{m^2}\Big)^{n-2-k}\frac{1}{k+1} \\
&\quad+\sum_{k=1}^{n-2}\binom{n-2}{k}\Big(\frac{8}{m^2}\Big)^k\Big(\frac{m^2-9}{m^2}\Big)^{n-2-k}\Bigg) \\
&= \frac{1}{\alpha^2}\Bigg(\frac{m^2-9}{n-1}\Bigg(1-\Big(1-\frac{1}{m^2}\Big)^{n-1}\Bigg)-\Bigg(1-\frac{9}{m^2}\Bigg)^{n-1}\Bigg).
\end{aligned}$$

Proof of Lemma 3: In the next time slot, the destination node D may receive a fresh packet either from the source node S or from one of the t_1 fresh nodes. Notice that these t_1+1 events are mutually exclusive, the probability that D receives a fresh packet from S is p_1, and the probability that D receives a fresh packet from a single fresh node is $\frac{p_2}{2(n-2)}$. By summing the probabilities of these t_1+1 events, the formula (7) follows.

Similarly, given that there are t_2 non-fresh nodes, in the next time slot the node S may deliver out a new copy to any one of them. Notice that these t_2 events are also exclusive, and the probability that S delivers out a copy to a single non-fresh node is $\frac{p_2}{2(n-2)}$, so the formula (8) follows.

To derive $P_s(t_1, t_2)$, let's focus on a specific fresh node R and a specific non-fresh node V, and use $P(S \to V, R \to D)$ to denote the probability that a source-to-relay transmission from S to V and a relay-to-destination transmission from R to D happen simultaneously in the next time slot. Thus, the $P_s(t_1, t_2)$ can be determined as

$$P_s(t_1, t_2) = t_1 t_2 \cdot P(S \to V, R \to D). \tag{43}$$

The basic idea is to treat the event $(S \to V, R \to D)$ as two simultaneous but mutually independent transmissions, i.e., the relay-to-destination transmission $R \to D$ and the source-to-relay transmission $S \to V$, then divide each transmission into multiple mutually exclusive cases, and finally represent each case with several simultaneous but independent sub-events. First consider the active cell with node R. The node R can conduct a relay-to-destination transmission with D only under the following two mutually exclusive cases: D is in this cell or D is in one of the eight adjacent cells. If we further assume that except the S, D, R, V, and the destination node of R's local traffic,there are in total k other nodes in the one-hop neighborhood of R, $k \in [0, n-5]$, among them i nodes are in the same cell as R, $i \in [0, k]$, and the other $k-i$ nodes are in the eight adjacent cells. Then the probability that R and D are selected as the transmitter and the receiver, respectively, is $\frac{1}{(i+2)(k+1)}$ (resp. $\frac{1}{(i+1)(k+1)}$) under the former case (resp. under the latter case). Summing up the probabilities under these two cases, then we get the corresponding probability of the relay-to-destination transmission $R \to D$. Similarly, we can also get the probability of the source-to-relay transmission $S \to V$. By the law of multiplication, then we have

$$\begin{aligned} P(S \to V, R \to D) &= \frac{m^2-\alpha^2}{4m^2\alpha^4} \sum_{k=0}^{n-5} \binom{n-5}{k} \\ &\cdot \Big(\sum_{i=0}^{k} \binom{k}{i} \frac{1}{k+1} \Big(\frac{1}{i+2} + \frac{8}{i+1} \Big) \Big(\frac{1}{m^2} \Big)^{i+1} \Big(\frac{8}{m^2} \Big)^{k-i} \Big) \\ &\cdot \sum_{t=0}^{n-4-k} \binom{n-4-k}{t} \Big(\sum_{j=0}^{t} \binom{t}{j} \frac{1}{t+1} \Big(\frac{1}{j+2} + \frac{8}{j+1} \Big) \\ &\cdot \Big(\frac{1}{m^2} \Big)^{j+1} \Big(\frac{8}{m^2} \Big)^{t-j} \Big) \Big(\frac{m^2-18}{m^2} \Big)^{n-4-k-t}. \end{aligned} \tag{44}$$

Notice that

$$\begin{aligned} &\sum_{i=0}^{k} \binom{k}{i} \frac{1}{k+1} \Big(\frac{1}{i+2} + \frac{8}{i+1} \Big) \Big(\frac{1}{m^2} \Big)^{i+1} \Big(\frac{8}{m^2} \Big)^{k-i} \\ &= \sum_{i=0}^{k} \binom{k+1}{i+1} \frac{1}{k+1} \frac{1}{i+2} \Big(\frac{1}{m^2} \Big)^{i+1} \Big(\frac{8}{m^2} \Big)^{k-i} \\ &\quad - \sum_{i=0}^{k} \binom{k}{i+1} \frac{1}{k+1} \frac{1}{i+2} \Big(\frac{1}{m^2} \Big)^{i+1} \Big(\frac{8}{m^2} \Big)^{k-i} \\ &\quad + \sum_{i=0}^{k} \binom{k}{i} \frac{1}{k+1} \frac{8}{i+1} \Big(\frac{1}{m^2} \Big)^{i+1} \Big(\frac{8}{m^2} \Big)^{k-i} \\ &= \frac{1}{(k+1)(k+2)} \Big(9 \Big(\frac{9}{m^2} \Big)^{k+1} - (k+10) \Big(\frac{8}{m^2} \Big)^{k+1} \Big) \\ &\quad - \frac{1}{(k+1)^2} \Big(8 \Big(\frac{9}{m^2} \Big)^{k+1} - (k+9) \Big(\frac{8}{m^2} \Big)^{k+1} \Big) \\ &\quad + \frac{8}{(k+1)^2} \Big(\Big(\frac{9}{m^2} \Big)^{k+1} - \Big(\frac{8}{m^2} \Big)^{k+1} \Big) \\ &= \frac{9 \big(\frac{9}{m^2} \big)^{k+1} - 8 \big(\frac{8}{m^2} \big)^{k+1}}{(k+1)(k+2)}. \end{aligned} \tag{45}$$

The (9) then follows by combining (43), (44) and (45).

REFERENCES

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

[2] A. A. Hanbali, M. Ibrahim, V. Simon, E. Varga, and I. Carreras, "A survey of message diffusion protocols in mobile ad hoc networks," in *Inter-Perf*, 2008.

[3] 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.

[4] G. Sharma and R. Mazumdar, "On achievable delay/capacity trade-offs in mobile ad hoc networks," in *Wiopt*, 2004.

[5] J. Liu, X. Jiang, H. Nishiyama, and N. Kato, "Delay and capacity in ad hoc mobile networks with f-cast relay algorithms," in *ICC*, 2011.

[6] E. Perevalov and R. S. Blum, "Delay-limited throughput of ad hoc networks," *IEEE Transactions on Communications*, vol. 52, no. 11, pp. 1957–1968, November 2004.

[7] J. Mammen and D. Shah, "Throughput and delay in random wireless networks with restricted mobility," *IEEE Transactions on Information Theory*, vol. 53, no. 3, pp. 1108–1116, 2007.

[8] 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.

[9] T. Small and Z. Hass, "Resource and performance tradeoffs in delay-tolerant wireless networks," in *ACM SIGCOMM Workshop on Delay Tolerant Networks (WDTN)*, 2005.

[10] A. Kashyap, S. Ganguly, S. R. Das, and S. Banerjee, "Voip on wireless meshes: models, algorithms and evaluation," in *INFOCOM*, 2007.

[11] S. Ganguly, V. Navda, K. Kim, A. Kashyap, D. Niculescu, R. Izmailov, S. Hong, and S. R. Das, "Performance optimizations for deploying voip services in mesh networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 11, pp. 2147–2158, November 2006.

[12] J. Cano, J.-C. Cano, C. K. Toh, C. T. Calafate, and P. Manzoni, "Easymanet: an extensible and configurable platform for service provisioning in manet environments," *IEEE Communications Magazine*, vol. 48, no. 12, pp. 159–167, December 2010.

[13] S. Khan, Y. Peng, and E. Steinbach, "Application-driven cross-layer optimization for video streaming over wireless networks," *IEEE Communications Magazine*, vol. 44, no. 1, pp. 122–130, January 2006.

[14] L. Zhou, X. Wang, W. Tu, G.-M. Muntean, and B. Geller, "Distributed scheduling scheme for video streaming over multi-channel multi-radio multi-hop wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 3, pp. 409–420, April 2010.

[15] I.-H. Hou, V. Borkar, and P. R. Kumar, "A theory of qos for wireless," in *INFOCOM*, 2009.

[16] I.-H. Hou and P. R. Kumar, "Scheduling heterogeneous real-time traffic over fading wireless channels," in *INFOCOM*, 2010.

[17] 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.

[18] A. A. Hanbali, P. Nain, and E. Altman, "Performance of ad hoc networks with two-hop relay routing and limited packet lifetime-extended version," *Performance Evaluation*, vol. 65, no. 6-7, pp. 463–483, June 2008.

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

[20] 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.

[21] 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.

[22] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, March 2000.

[23] S. R. Kulkarni and P. Viswanath, "A deterministic approach to throughput scaling in wireless networks," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1041–1049, June 2004.

[24] C. M. Grinstead and J. L. Snell, *Introduction to Probability: Second Revised Edition*. American Mathematical Society, 1997.

[25] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. D. Van Nostrand, 1963.

[26] C++ programs for the 2hr-(f, g) manet. [Online]. Available: http://distplat.blogspot.com

[27] The network simulator ns-2. [Online]. Available: http://www.isi.edu/nsnam/ns/

[28] Qualnet. [Online]. Available: http://www.scalable-networks.com/products/qualnet/

[29] 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.

[30] S. Zhou and L. Ying, "On delay constrained multicast capacity of large-scale mobile ad-hoc networks," in *INFOCOM*, 2010.

[31] H. Boche, S. Naik, and M. Schubert, "Pareto boundary of utility sets for multiuser wireless systems," *IEEE/ACM Transactions on Networking*, vol. 19, no. 2, pp. 589–601, April 2011.

[32] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Optimal throughput-delay scaling in wireless networks-part i: The fluid model," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2568–2592, June 2006.

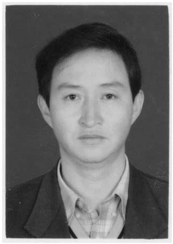
[33] ——, "Throughput-delay trade-off in wireless networks," in *INFOCOM*, 2004.

[34] X. Lin, G. Sharma, R. R. Mazumdar, and N. B. Shroff, "Degenerate delay-capacity tradeoffs in ad-hoc networks with brownian mobility," *IEEE/ACM Transactions on Networking, Special Issue on Networking and Information Theory*, vol. 52, no. 6, pp. 2777–2784, June 2006.

[35] G. Sharma, R. Mazumdar, and N. B. Shroff, "Delay and capacity trade-offs for mobile ad hoc networks: A global perspective," *IEEE/ACM Transactions on Networking*, vol. 15, no. 5, pp. 981–992, October 2007.

[36] G. Sharma and R. Mazumdar, "Delay and capacity trade-off in wireless ad hoc networks with random way-point mobility," in *Dept. Elect. Comput. Eng., Purdue Univ.,West Lafayette, IN*, 2005. [Online]. Available: http://ece.purdue.edu/∼gsharma/

Jiajia Liu received his B.S. and M.S. Degrees both in Computer Science from Harbin Institute of Technology in 2004 and from Xidian University in 2009, respectively. He is currently a PhD candidate at the Graduate School of Information Sciences at Tohoku University. His research interests include performance modeling and evaluation, scaling laws of wireless networks, stochastic network optimization, and optimal control.



Xiaohong Jiang received his B. S, M. S and Ph D degrees from Xidian University, Xian, China, in 1989, 1992 and 1999, respectively. Dr. Jiang is currently a full professor of Future University Hakodate, Japan. Before joining Future University Hakodate, he was an associate professor in Tohoku University, Japan, was an assistant professor in the Graduate School of Information Science, Japan Advanced Institute of Science and Technology (JAIST) . His current research interests include wireless sensor networks, optical networks, and network coding, etc. Dr. Jiang has authored and coauthored more than 160 publications in journals, books and international conference proceedings, which include IEEE/ACM Transactions on networking, IEEE Transactions on Communications, and IEEE Journal of Selected Area on Communications. Dr. Jiang was also the winner of the Best Paper Award of WCNC08 and the ICC2005-Optical Networking Symposium. He is a senior member of IEEE. (Email: jiang@fun.ac.jp)

Hiroki Nishiyama received his M.S. and Ph.D. in Information Science from Tohoku University, Japan, in 2007 and 2008, respectively. He was a Research Fellow of the Japan Society for the Promotion of Science (JSPS) until finishing his Ph.D, when he then went on to become an Assistant Professor at the Graduate School of Information Sciences at Tohoku University. He has received Best Paper Awards from the IEEE Global Communications Conference 2010 (GLOBECOM 2010) as well as the 2009 IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC 2009). He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and an IEEE member. He was also a recipient of the 2009 FUNAI Foundation's Research Incentive Award for Information Technology. His active areas of research include, traffic engineering, congestion control, satellite communications, ad hoc and sensor networks, and network security.

Nei Kato received his M.S. and Ph.D. Degrees in Information Science from Tohoku University, Japan, in 1988 and 1991, respectively. He joined the Computer Center of Tohoku University in 1991 and has been a full professor at the Graduate School of Information Sciences since 2003. He has been engaged in research on computer networking, wireless mobile communications, image processing and neural networks, and has published more than 200 papers in journals and peer-reviewed conference proceedings.

He currently serves as the chair of the IEEE Satellite and Space Communications Technical Community (TC), the secretary for the IEEE Ad Hoc & Sensor Networks TC, the vice chair of the IEICE Satellite Communications TC, a technical editor for IEEE Wireless Communications (since 2006), an editor for IEEE Transactions on Wireless Communications (since 2008), and as an associate editor for IEEE Transactions on Vehicular Technology (since 2009). He also served as a co-guest-editor for IEEE Wireless Communications Magazine SI on "Wireless Communications for E-healthcare", a symposium co-chair of GLOBECOM'07, ICC'10, ICC'11, ChinaCom'08, ChinaCom'09, and the WCNC2010-2011 TPC Vice Chair.

Besides his academic activities, he also serves as a member on the Telecommunications Council expert committee, the special commissioner of the Telecommunications Business Dispute Settlement Commission for the Ministry of Internal Affairs and Communications in Japan, and as the chairperson of ITU-R SG4 in Japan. Nei Kato is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and a senior member of IEEE.