Delay Modeling for Broadcast-Based Two-Hop Relay MANETs

Juntao Gao† and Xiaohong Jiang†
† School of Systems Information Science
Future University Hakodate, Japan
Email: gaojuntao223@gmail.com, jiang@fun.ac.jp

***Abstract*—Understanding the delay performance in mobile ad hoc networks (MANETs) is of fundamental importance for supporting Quality of Service (QoS) guaranteed applications in such networks. Despite lot of research efforts in last several decades, the important end-to-end delay modeling in MANETs remains a challenging issue. This is partially due to the highly dynamical behaviors of MANETs but also due to the lack of an efficient theoretical framework to depict the complicated network state transitions under such dynamics. This paper demonstrates the potential application of the Quasi-Birth-and-Death process (QBD) theory in MANETs delay analysis by applying it to the end-to-end delay modeling in broadcast-based two-hop relay MANETs. We first demonstrate that the QBD theory actually enables a novel and powerful theoretical framework to be developed to efficiently capture the complicated network state transitions in the concerned MANETs. We then show that with the help of the theoretical framework, we are able to analytically model the exact expected end-to-end delay and also the exact per node throughput capacity in such MANETs. Extensive simulations are further provided to validate the efficiency of our QBD theory-based models.**

I. Introduction

Mobile ad hoc networks (MANETs) represent a class of important wireless ad hoc networks with mobile nodes. The flexible and distributed MANETs are robust and rapidly deployable/reconfigurable, so they are highly appealing for a lot critical applications, like deep space communication, disaster relief, battlefield communication, outdoor mining, etc. To facilitate the application of MANETs in providing Quality of Service (QoS) guaranteed services, understanding delay performance of such networks is of fundamental importance [1]. Notice the end-to-end delay, i.e. the time elapsed between the time a packet is generated at its source node and the time it's delivered to its destination, is the most fundamental delay performance for MANETs. However, the analytical modeling of the practical end-to-end delay in MANETs remains a technical challenge. This is due to the highly dynamical behaviors of MANETs in terms of topology changing, wireless medium contention, interference and traffic contention, the complicated queueing process of a packet at its source node and the delivery process of the packet among mobile nodes.

To simplify delay analysis in MANETs, lot of preliminary work has been done to analyze the packet delivery delay, which composes only a part of packet end-to-end delay and is the time duration from the time a packet starts to be distributed by its source node to the time the packet is delivered to its destination. For the simple network scenarios where there is one single packet to be delivered without any traffic contention, the distribution of packet delivery delay was derived in [2], [3], the closed-form Laplace-Stieltjes transform of packet delivery delay was obtained in [4], while the modeling of expected packet delivery delay was further discussed in [5], [6] and [7]. For more practical network scenarios with multiple packets and traffic contention issue, the expected packet delivery delay was studied in [8], where the absorbing markov chain theory was applied to conduct delay analysis which is however not suitable for packet end-to-end delay analysis. Recently, some attempts have been observed on end-to-end delay analysis in MANETs. Analytical approximations of expected packet end-to-end delay were provided [9]–[11], while a closed-form upper bound on the expected packet end-to-end delay was derived in [12], [13]. However, the results in [9]–[13] indicate clearly that approximations or upper bound on packet end-to-end delay, although can reflect the general trend of end-to-end delay, may introduce significant errors in packet end-to-end delay analysis and thus is far from satisfactory.

In summary, despite much research activity on the delay performance study of MANETs in the last several decades, the important issue of end-to-end delay modeling in MANETs remains a technical challenge. This is partially due to the highly complicated network dynamics, but also due to the lack of an efficient theoretical framework to depict the complex network state transitions in such networks. In this paper, we demonstrate the potential application of the QBD theory in MANETs delay analysis by applying it to the end-to-end delay modeling in a broadcast-based two-hop relay (2HR-B for short) MANET[1]. The main contributions of this paper are summarized as follows:

- We first demonstrate that the QBD theory actually enables a novel and powerful theoretical framework to be developed to efficiently capture the complex network state transitions in MANETs.
- With the help of the theoretical framework, we then show that we are able to analytically model the exact expected end-to-end delay and also exact per node throughput capacity in a broadcast-based two-hop relay MANET.
- Extensive simulations are further provided to validate the efficiency of our QBD theory-based modeling for end-to-end delay and per node throughput capacity in MANETs.

The rest of this paper is organized as follows. We introduce

[1]The wireless broadcast feature is also exploited in [14], [15].

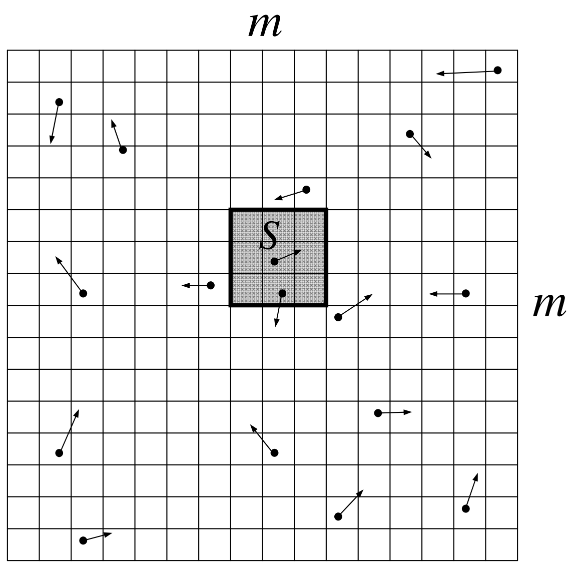

Fig. 1. A snapshot of a cell partitioned MANET with $m = 16$.

the system models in Section II. Section III is dedicated to the QBD-based modeling of exact expected end-to-end delay and per node throughput capacity. Section IV provides simulation results to validate the efficiency of our theoretical framework and the derived delay and capacity results, and we conclude this paper in Section V.

II. System Models

In this section, we introduce first the network models concerning node mobility, wireless channel, radio and traffic pattern issues in MANET, and then the Equivalent-Class based Medium Access Control (MAC-EC) protocol for transmission scheduling. Finally, the 2HR-B routing scheme for packet delivery is described.

A. Network Models

Mobility and channel models: As shown in Fig. 1 that we consider a unit torus MANET partitioned evenly into $m \times m$ cells [16]–[18]. In the concerned MANET, there are n nodes roaming according to the i.i.d. mobility model [8], [12], [13]. We consider the time slotted system, where each node randomly chooses one cell to stay in at the beginning of every time slot. We assume that a common bandwidth limited wireless channel is shared by all nodes for data transmissions. At any time slot, the data transmitted between any two nodes through the wireless channel is normalized to one packet.

Radio model: Every transmitting node (transmitter) employs the same radio power to transmit data through the common wireless channel. To enable a transmitter (say S in Fig. 1) to cover its own cell and also its 8 neighbor cells (called coverage cells of the transmitter hereafter), the corresponding radio range r of the transmitter should be set as $r = \sqrt{8}/m$. Based on the widely used protocol model [8], [17]–[20], data transmission from a transmitter i to a receiving node (receiver) j can be conducted only if the Euclidean distance d_{ij} between them is less than r (i.e., $d_{ij} \leq r$), while the data can be successfully received by receiver j only if $d_{kj} \geq (1+\Delta)\cdot r$ holds for any other concurrent transmitting node k, here $k \neq i$ and Δ is a specified guard-factor for interference prevention.

Traffic model: Similar to previous works [8], [13], [17], we consider the permutation traffic pattern, in which each node acts as the source of a traffic flow and at the same time the destination of another traffic flow. Thus, there are in total n distinct traffic flows in the MANET. Each source node exogenously generates packets for its destination according to an i.i.d. process with average rate λ (packets/slot) [12] (a packet is generated with probability λ at the beginning of each time slot).

B. MAC-EC Protocol

For a fair wireless medium access among all network nodes, we consider here a MAC protocol MAC-EC for transmission scheduling based on the idea of Equivalent-Class (EC).

Active cell and equivalent-class: At a time slot, one cell is called an active cell if node(s) in this cell are allowed to contend for wireless medium access and thus packet transmission. An EC is then defined as a set of cells such that when the EC is active (i.e., all cells in it are active), all transmitters of this set, each from one distinct cell there, can transmit without interfering each other[2].

As illustrated in Fig 2 that we consider here the EC consisting of the set of cells with vertical and horizontal distance between any two of them being some integer multiple of α cells, $\alpha \leq m$. Thus, we have in total α^2 ECs in the MANET, which are scheduled to be active alternatively as time evolves.

To enable as many number of concurrent transmissions to be conducted as possible while avoiding interference between these transmissions, the parameter α should be set appropriately. As illustrated in Fig 2 that for the transmission from a transmitter S to its possible receiver R to be successful, the distance between R and another possible closest concurrent transmitter W, i.e., $(\alpha-2)/m$, should satisfy the following condition according to the protocol model:

$$(\alpha-2)/m \geq (1+\Delta)\cdot r \qquad (1)$$

Notice that $r = \sqrt{8}/m$ and $\alpha \leq m$, the parameter α should be determined as

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

where $\lceil x \rceil$ takes the least integer value greater than x.

With the help of ECs, the MAC-EC protocol can be easily implemented as follows. At each time slot, a node can easily know which cell (and thus which EC) it belongs to based on the Global Positioning System (GPS). If the cell is currently not active, the node simply keeps silent; otherwise, if the cell is currently active (i.e., the EC the cell belongs to is scheduled to be active at current time slot), the node will contend fairly with other nodes in the cell to become transmitter and thus get fair access to the common wireless channel.

[2]Such set is called "concurrent-set" in [8], similar schemes were also presented in [17], [18], [20]

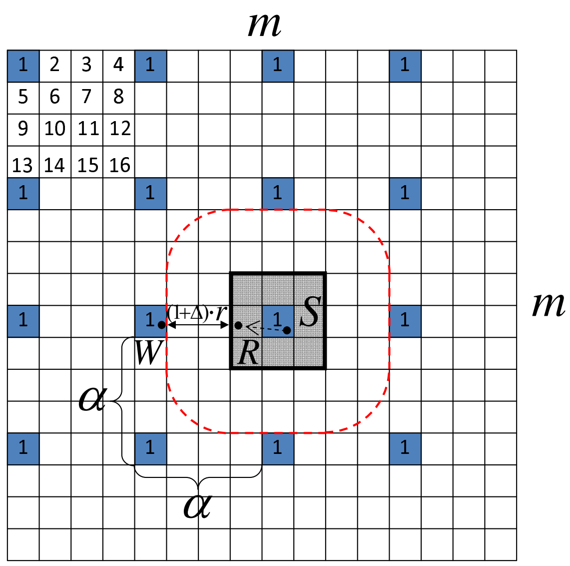

Fig. 2. Illustration of ECs in a MANET with $m = 16$. There are 16 ECs in this MANET with $\alpha = 4$. All shaded cells belong to EC 1.

C. 2HR-B Routing

Once a node, say S in Fig. 2, succeeds in wireless medium access contention and becomes a transmitter, it executes the 2HR-B routing protocol defined in Algorithm 1 for packet delivery.

To facilitate the operation of the 2HR-B routing protocol, each node is equipped with three types of First In First Served (FIFS) queues: one source-queue, one broadcast-queue and $n-2$ parallel relay-queues (no relay-queue is needed for the node itself and its destination node). For a source node S with destination node D, its three types of queues are defined as:

Source-queue: Source-queue stores packets exogenously generated at S and destined for D. These exogenous packets will be distributed out to relay nodes later in the FIFS way.

Broadcast-queue: Broadcast-queue stores packets from source-queue that have already been distributed out by S but have not been acknowledged yet by D the reception of them.

Relay-queue: For each destination node (excluding S and D), one relay-queue is allocated to store redundant copies of packets distributed out by the source of that destination node.

To ensure the in-order packet reception at D, similar to previous works [8], [12], [13] that S will label an exogenously generated packet with a unique identification number $ID(S)$, which increases by 1 every time a packet is generated; destination D also maintains an acknowledgment number $ACK(D)$ indicating that D is currently requesting the packet with $ID(S) = ACK(D)+1$ (i.e, the packets with $ID(S) \leq ACK(D)$ have already been received by D).

III. End-to-End Delay Modeling

In this section, we first present some preliminaries, and then develop a novel theoretical framework based on QBD theory to efficiently capture the complex network state transitions. Finally, with the help of the theoretical framework, we derive the exact expected end-to-end delay and also exact per node throughput capacity for the considered MANET.

Algorithm 1 2HR-B Routing Protocol

Transmitter S selects to do packet-broadcast with probability q and to do packet-delivery with probability $1-q$;
if S selects packet-broadcast **then**
  S executes Procedure 2;
else
  S executes Procedure 3;
end if

Procedure 2 packet-broadcast

S distributes out the head-of-line (HoL) packet of source-queue in a wireless broadcast fashion to all nodes in the coverage cells of S;
Any node, say R, locating in the coverage cells of S reserves a copy of the packet received and inserts it into the end of $R's$ relay-queue associated with D (if D is in S's coverage cells and D is currently requesting the packet S is distributing out, then D receives that packet, otherwise discards it);
S moves that HoL packet out of soure-queue and inserts it into the end of its broadcast-queue;
S moves ahead the remaining packets in its source-queue;

Procedure 3 packet-delivery

S randomly selects a node U (denote U's source by V) as its receiver from nodes in its coverage cells;
S initiates a handshake with U to acquire the packet number $ACK(U)+1$ and know which packet U is currently requesting;
S checks its corresponding relay-queue/broadcast-queue whether it bears a packet with $ID(V) = ACK(U)+1$;
if S bears such packet **then**
  S delivers that packet to U;
  S clears all packets with $ID(V) \leq ACK(U)$ from its corresponding relay-queue/broadcast-queue;
  S moves ahead the remaining packets in its corresponding relay-queue/broadcast-queue;
end if

A. Preliminaries

Due to symmetry of all traffic flows, we only focus on one specific traffic flow from source S to destination D in our analysis. We first define the packet end-to-end delay and per node throughput capacity.

Definition 1: The end-to-end delay T_e of a packet is the time elapsed between the time slot the packet is generated at its source and the time slot it is delivered to its destination.

Definition 2: Per node throughput capacity μ is defined as the maximum packet generation rate λ every node in the concerned MANET can stably support.

Notice that once a generic packet is generated at S, it first experiences a queueing process in the source-queue of S before it is distributed out (served), and then experiences a network delivery process after it is distributed out into network by S before it is successfully received by D. Since D requests packets in order according to $ACK(D)$, all packets distributed out by S will be also delivered (served) in order. Thus, we can treat the network delivery process as a queueing process of one virtual network-queue.

To model packet end-to-end delay, we need to model the two queueing processes in source-queue and network-queue, where the departure process of source-queue is just the arrival process of network-queue. The following lemmas provide the calculation of some probabilities and also show that the source-queue and network-queue can be decoupled and thus can be analyzed separately in the considered MANET.

Lemma 1: In a time slot, let p_b denote the probability that S selects to do packet-broadcast to distribute out a packet, let $p_c(j)$ denote the probability that j copies of a packet exist in the MANET after it being distributed out by S, $1 \le j \le n-1$, and let $p_r(j)$ denote the probability that D receives the packet it is currently requesting given that there are j copies of the packet in the network. Then we have

$$p_b = \frac{qm^2}{\alpha^2 n}\left\{1 - \left(\frac{m^2-1}{m^2}\right)^n\right\} \tag{3}$$

$$p_c(j) = \frac{n\binom{n-2}{j-1}(m^2-9)^{n-1-j}}{m^{2n}-(m^2-1)^n}\left\{(m^2-9)f(j) + f(j+1)\right\} \tag{4}$$

$$p_r(j) = \frac{j(1-q)m^2}{\alpha^2 n(n-1)}\left\{1-\left(\frac{m^2-1}{m^2}\right)^n - \frac{n}{m^2}\left(\frac{m^2-9}{m^2}\right)^{n-1}\right\} \tag{5}$$

where

$$f(x) = \frac{9^x - 8^x}{x} \tag{6}$$

Proof: Due to space limit, the proof is omitted here. Please kindly refer to [21] for proof details. ■

Lemma 2: For the considered MANET with MAC-EC protocol for transmission scheduling and 2HR-B protocol for packet delivery, the arrival process of network-queue is a Bernoulli process with probability λ and it is independent of the state of source-queue.

Proof: We know from Section II-A that the arrival process of source-queue in S is a Bernoulli process with probability λ. The service process of source-queue is actually also a Bernoulli process, because in every time slot S gets a chance with constant probability p_b to do packet-broadcast to distribute out a packet in source-queue (or equivalently the source-queue is served with probability p_b in every time slot). Thus, the source-queue in S follows a Bernoulli/Bernoulli queue, and in equilibrium the packet departure process of source-queue is also a Bernoulli process with probability λ, which is independent of the state of source-queue (i.e., the number of packets in source-queue) [22]. Because the arrival process of network-queue is just the departure process of source-queue, this finishes the proof of this Lemma. ■

B. Theoretical Framework

We use $L(t)$ to denote the number of packets already distributed out by S and detained in network to be delivered to D at time slot $t, L(t) \ge 1$, and use $J(t)$ to denote the number of copies of the packet D is currently requesting at time slot t (i.e., the packet with $ID(S) = ACK(D)+1$) in the network, $1 \le J(t) \le n-1$. Then the queueing process of network-queue follows the following markov process as time slot t evolves

$$\big\{\{(0,0)\} \cup \{(L(t), J(t))\}, t = 0,1,2,\cdots\big\}, \tag{7}$$

where $(0,0)$ corresponds to the empty network-queue state. $L(t)$ increases by 1 if S distributes out a packet from its source-queue while D does not receive the packet it is requesting at slot t, $L(t)$ decreases by 1 if S does not distribute out a packet from its source-queue while D receives the packet it is requesting at slot t, and $L(t)$ keeps unchanged, otherwise.

All states in (7) can be divided into following subsets

$$N(0) = \{(0,0)\} \tag{8}$$

$$N(l) = \big\{\{(l,j)\}, 1 \le j \le n-1\big\}, l \ge 1 \tag{9}$$

where subset $N(0)$ is called level 0 and subset $N(l)$ is called level l. It is notable that when network-queue is in some state of one level at a time slot, the next state of one-step state transitions could only be some state in the same level or in its adjacent levels. Thus, we know from [23], [24] that the queueing process of network-queue here forms a QBD with state transition matrix $\mathbf{Q}$ as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{B_1} & \mathbf{B_0} & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{B_2} & \mathbf{A_1} & \mathbf{A_0} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{A_2} & \mathbf{A_1} & \mathbf{A_0} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{A_2} & \mathbf{A_1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{10}$$

Let $\mathbf{I}$ denote an identity matrix of size $(n-1)\times(n-1)$, the matrices $\mathbf{B_1}$, $\mathbf{B_0}$, $\mathbf{B_2}$, $\mathbf{A_1}$, $\mathbf{A_0}$ and $\mathbf{A_2}$ are then defined as

$$\mathbf{B_1} = \left[p^{b_1}\right] \tag{11}$$

$$\mathbf{B_0} = \left[\, p_{01}^{b_0} \quad p_{02}^{b_0} \quad \cdots \quad p_{0j}^{b_0} \quad \cdots \quad p_{0(n-1)}^{b_0} \,\right] \tag{12}$$

$$\mathbf{B_2} = \left[\, p_{10}^{b_2} \quad p_{20}^{b_2} \quad \cdots \quad p_{j0}^{b_2} \quad \cdots \quad p_{(n-1)0}^{b_2} \,\right]^T \tag{13}$$

$$\mathbf{A_0} = \mathrm{diag}\left(p_{11}^{a_0}, p_{22}^{a_0}, \cdots, p_{jj}^{a_0}, \cdots, p_{(n-1)(n-1)}^{a_0}\right) \tag{14}$$

$$\mathbf{A_2} = \mathbf{B_2}\mathbf{v_2} \tag{15}$$

$$\mathbf{v_2} = \left[\, p_c(1) \quad p_c(2) \quad \cdots \quad p_c(j) \quad \cdots \quad p_c(n-1) \,\right] \tag{16}$$

$$\mathbf{A_1} = [p_{ij}^{a_1}]_{(n-1)\times(n-1)} \tag{17}$$

Here p^{b_1} is the probability that the state of network-queue changes from $(0,0)$ to $(0,0)$; $p_{0j}^{b_0}$ is the probability that the state of network-queue changes from $(0,0)$ to $(1,j)$, $1 \le j \le n-1$; $p_{j0}^{b_2}$ is the probability that the state of network-queue changes from $(1,j)$ to $(0,0)$, $1 \le j \le n-1$; $p_{jj}^{a_0}$ is the probability that the state of network-queue changes from (l,j) to $(l+1,j)$, $l \ge 1$, $1 \le j \le n-1$; $p_{ij}^{a_1}$ is the probability that the state of network-queue changes from (l,i) to (l,j), $l \ge 1$, $1 \le i,j \le n-1$.

Lemma 3: The probabilities p^{b_1}, $p^{b_0}_{0j}$, $p^{b_2}_{j0}$, $p^{a_0}_{jj}$, $p^{a_1}_{ij}$ defined in the above QBD process are determined as follows

$$p^{b_1} = 1 - \frac{\lambda \cdot q \cdot (m^2 - 9)}{\alpha^2 (n-1) p_b} \left\{ 1 - \left(\frac{m^2 - 1}{m^2} \right)^{n-1} \right\} \tag{18}$$

$$p^{b_0}_{0j} = \frac{\lambda \cdot q \cdot \binom{n-2}{j-1} (m^2 - 9)^{n-j}}{\alpha^2 m^{2n-2} p_b} f(j) \tag{19}$$

$$p^{b_2}_{j0} = p_r(j) - p_{br}(j-1) \tag{20}$$

$$p^{a_0}_{jj} = \lambda - p_{br}(j-1) \tag{21}$$

$$p^{a_1}_{ij} = \begin{cases} p_{\overline{br}}(i) + p_{br}(i-1) \cdot p_c(i) & \text{if} \quad i = j \\ p_{br}(i-1) \cdot p_c(j) & \text{if} \quad i \neq j \end{cases} \tag{22}$$

where

$$\begin{aligned} p_{br}(x) = x & \frac{\lambda (q - q^2)(m^4 - m^2 \alpha^2)}{\alpha^4 n (n-1)(n-2) p_b} \\ & \cdot \left\{ 1 - 2 \left(\frac{m^2 - 1}{m^2} \right)^n + \left(\frac{m^2 - 2}{m^2} \right)^n \right. \\ & \left. - \frac{n}{m^2} \left(\frac{m^2 - 9}{m^2} \right)^{n-1} + \frac{n}{m^2} \left(\frac{m^2 - 10}{m^2} \right)^{n-1} \right\} \end{aligned} \tag{23}$$

$$p_{\overline{br}}(x) = 1 - p_{br}(x-1) - p^{a_0}_{xx} - p^{b_2}_{x0} \tag{24}$$

Proof: Due to space limit, the proof is omitted here. Please kindly refer to [21] for proof details. ■

Remark 1: Notice that network dynamics in terms of topology change, wireless medium contention, interference and traffic contention have been jointly taken into considerations in the calculations of probabilities in Lemmas 1 and 3 [21].

C. End-to-End Delay and Per Node Throughput Capacity

With the help of the above QBD-based theoretical framework, we are able to present our main results regarding the exact per node throughput capacity μ and exact expected end-to-end delay $\mathbb{E}(T_e)$ for the concerned MANET.

Theorem 1: For the considered MANET, its per node throughput capacity μ is given by

$$\mu = \min \left\{ p_b, \frac{1}{\sum_{j=1}^{n-1} \frac{p_c(j)}{p_r(j)}} \right\} \tag{25}$$

Proof: In equilibrium, the service rate of source-queue μ_s is

$$\mu_s = p_b \tag{26}$$

and the service rate of network-queue μ_d (the rate D receives its requested packets) is

$$\mu_d = \frac{1}{\sum_{j=1}^{n-1} \frac{p_c(j)}{p_r(j)}}. \tag{27}$$

To ensure network stability, packet generation rate λ at S should satisfy

$$\lambda < \min\{\mu_s, \mu_d\} \tag{28}$$

Thus, the per node throughput capacity μ is determined as

$$\mu = \min\{\mu_s, \mu_d\} \tag{29}$$

■

Theorem 2: For the concerned MANET, where each source node exogenously generates packets according to a Bernoulli process with probability λ ($\lambda < \mu$), the expected end-to-end delay $\mathbb{E}(T_e)$ of a generic packet is determined as

$$\mathbb{E}(T_e) = \frac{\overline{L}_1 + \overline{L}_2}{\lambda}, \tag{30}$$

where

$$\overline{L}_1 = \frac{\lambda - \lambda^2}{p_b - \lambda} \tag{31}$$

$$\overline{L}_2 = \frac{\mathbf{y_1}(\mathbf{I} - \mathbf{R})^{-2}\mathbf{1}}{\upsilon}, \tag{32}$$

$$\mathbf{R} = \mathbf{A_0}(\mathbf{I} - \mathbf{A_1} - \mathbf{A_0}\mathbf{1}\mathbf{v_2})^{-1}, \tag{33}$$

$$\upsilon = y_0 + \mathbf{y_1}(\mathbf{I} - \mathbf{R})^{-1}\mathbf{1}. \tag{34}$$

here $\mathbf{1}$ is a column vector of size $(n-1) \times 1$ with all elements being 1, y_0 is a scalar value, $\mathbf{y_1}$ is a row vector of size $1 \times (n-1)$, and y_0 and $\mathbf{y_1}$ are determined as

$$[y_0, \mathbf{y_1}] = [y_0, \mathbf{y_1}] \begin{bmatrix} \mathbf{B_1} & \mathbf{B_0} \\ \mathbf{B_2} & \mathbf{A_1} + \mathbf{R}\mathbf{A_2} \end{bmatrix}, \tag{35}$$

Proof: From Lemma 2 we know that we can analyze queueing processes of source-queue and network-queue separately. Since the source-queue at S follows a Bernoulli/Bernoulli queue, we know from [12], [23] the expected number of packets $\overline{L}_1$ of the source-queue is determined as

$$\overline{L}_1 = \frac{\lambda - \lambda^2}{p_b - \lambda} \tag{36}$$

As network-queue follows a QBD characterized by matrix $\mathbf{Q}$, the queueing process of network-queue can be analyzed based on two related matrices $\mathbf{R}, \mathbf{G}$ defined in [23] and [24], where

$$\mathbf{R} = \mathbf{A_0}(\mathbf{I} - \mathbf{A_1} - \mathbf{A_0}\mathbf{G})^{-1} \tag{37}$$

Due to the special structure of $\mathbf{A_2}$, which is the product of a column vector $\mathbf{B_2}$ by a row vector $\mathbf{v_2}$, matrix $\mathbf{G}$ can be calculated as

$$\mathbf{G} = \mathbf{1}\mathbf{v_2} \tag{38}$$

Based on the results in [23], the expected number of packets $\overline{L}_2$ of network-queue is given by

$$\overline{L}_2 = \frac{\mathbf{y_1}(\mathbf{I} - \mathbf{R})^{-2}\mathbf{1}}{\upsilon}, \tag{39}$$

where $\mathbf{y_1}$ and υ are determined by (35) and (34).

Finally, by applying Little's Theorem, (30) follows. This finishes the proof of Theorem 2. ■

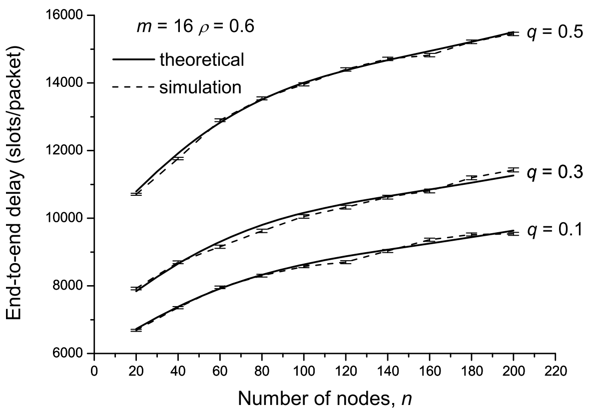

Fig. 3. Average packet end-to-end delay VS. number of nodes n in MANET.

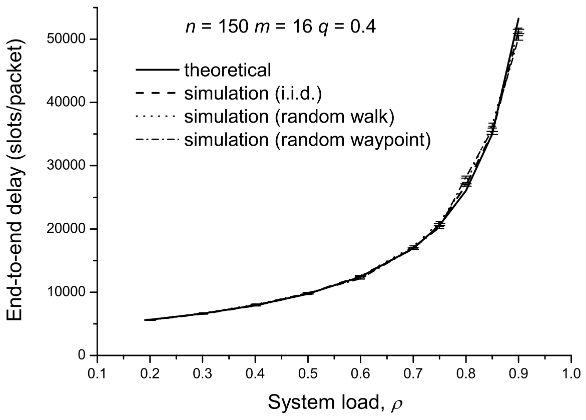

Fig. 4. Average packet end-to-end delay VS. system load ρ in MANET.

IV. Simulation Results

To validate the QBD-based theoretical results on expected end-to-end delay and per node throughput capacity, a customized C++ simulator has been developed to simulate packet generation, distribution and delivery processes in the considered MANET, where not only the i.i.d. node mobility model but also the typical random walk [25] and random waypoint [26] mobility models have been implemented[3].

- **Random Walk Model:** At the beginning of each time slot, each node moves to another cell as follows: it first independently selects a cell with equal probability $1/9$ among its current cell and its 8 neighboring cells; it then moves into that cell and stays in it until the end of that time slot.
- **Random Waypoint Model:** At the beginning of each time slot, each node moves to another cell as follows: it first independently generates a two-element vector $[x, y]$, where both elements x and y are uniformly drawn from $[1/m, 3/m]$; it then moves along the horizontal and vertical direction of distance x and y, respectively.

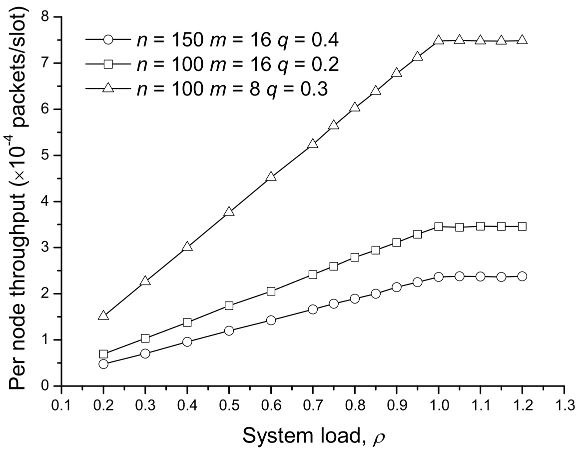

Fig. 5. Per node throughput VS. system load ρ in MANET.

A. End-to-End Delay Validation

For networks of different size n, Fig. 3 shows both theoretical and simulation results on packet end-to-end delay under the settings of $m = 16$, system load $\rho = 0.6$ ($\rho = \lambda/\mu$) and packet-broadcast probability $q = \{0.1, 0.3, 0.5\}$. Notice that all simulation results are reported with small 95% confidence intervals. The results in Fig. 3 show clearly that in the wide range of network scenarios considered here, theoretical results all match very nicely with the simulated ones, indicating that our QBD-based theoretical modeling is really efficient in capturing the expected packet end-to-end delay behavior of concerned MANETs. From Fig. 3 we can also see that as network size n increases, packet end-to-end delay increases as well. This is because that in the concerned MANET with fixed unit area and fixed setting of $m = 16$, as n increases the contention for wireless channel access becomes more intensive, resulting in a lower packet deliver opportunity and thus a longer packet end-to-end delay.

For the setting of $n = 150, m = 16, q = 0.4$, Fig. 4 shows both the theoretical and simulation results on packet end-to-end delay when system load ρ changes from $\rho = 0.2$ to $\rho = 0.9$. In addition to the i.i.d. mobility model considered in this paper, the corresponding simulation results for the random walk and random waypoint mobility models have also been included in Fig. 4 for comparison. Again, we can see from Fig. 4 that our theoretical delay model is very efficient. It is interesting to see from Fig. 4 that although our theoretical framework is developed under the i.i.d. mobility model, it can also nicely capture the general packet end-to-end delay behavior under some more realistic mobility models like random walk and random waypoint.

[3]The program of our simulator is now available online at [27]. Similar to [28], the guard-factor is set as $\Delta = 1$.

B. Throughput Capacity Validation

Another observation of Fig. 4 is that the average packet end-to-end delay increases sharply as system load ρ approaches 1.0 (i.e., as packet generation rate λ approaches per node throughput capacity μ), which serves as an intuitive verification of our theoretical per node throughput capacity result. To further validate the our theoretical model on throughput capacity, Fig. 5 provides the simulation results on the achievable per node throughput (the average rate of packet delivery to destination) under three network scenarios $\{n = 150, m = 16, q = 0.4\}$, $\{n = 100, m = 16, q = 0.2\}$ and $\{n = 100, m = 8, q = 0.3\}$ when system load ρ there increases from $\rho = 0.2$ up to $\rho = 1.2$. We can see from Fig. 5 that for each network scenario there, the corresponding per node throughput first increases monotonously as system load ρ increases from $\rho = 0.2$ to $\rho = 1.0$, and then remains a constant and does not increase anymore when ρ further increases beyond 1.0 (i.e., when packet generation rate λ goes beyond the theoretical per node throughput capacity μ). Thus, our theoretical capacity model is also efficient in depicting the per node throughput capacity behavior of the considered MANET.

V. Conclusion

The main finding of this paper is that the Quasi-Birth-and-Death process (QBD) can be a promising theory to tackle the challenging issue of analytical delay modeling for MANETs. We demonstrated through a broadcast-based two-hop relay MANET that QBD theory can help us: 1) to develop a theoretical framework to capture the complicated queueing state transitions in the highly dynamic MANET, 2) to analytically model the expected end-to-end delay and also the per node throughput capacity of the network, and 3) to enable many important network dynamics like topology changing, wireless medium contention, interference and traffic contention to be jointly considered in the delay modeling process.

References

[1] L. H. II and R. Tafazolli, "A survey of qos routing solutions for mobile ad hoc networks," *IEEE Communications Surveys and Tutorials*, vol. 9, no. 2, pp. 50–70, 2nd Quarter 2007.

[2] T. Small and Z. Hass, "Resource and performance tradeoffs in delay-tolerant wireless networks," in *Proc. ACM SIGCOMM workshop on Delay-tolerant networking (WDTN)*, 2005, pp. 260–267.

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

[4] 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.

[5] 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, pp. 191–202.

[6] 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.

[7] 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.

[8] J. Liu, X. Jiang, H. Nishiyama, and N. Kato, "Generalized two-hop relay for flexible delay control in manets," *IEEE/ACM Transactions on Networking*, vol. 20, no. 6, pp. 1950–1963, December 2012.

[9] A. A. Hanbali, R. de Haan, R. J. Boucherie, and J.-K. van Ommeren, "A tandem queueing model for delay analysis in disconnected ad hoc networks," in *Proceedings of the 15th international conference on Analytical and Stochastic Modeling Techniques and Applications (ASMTA)*, 2008, pp. 189–205.

[10] A. Jindal and K. Psounis, "Contention-aware performance analysis of mobility-assisted routing," *IEEE Transactions on Mobile Computing*, vol. 8, no. 2, pp. 145–161, February 2009.

[11] J. Liu, J. Gao, X. Jiang, H. Nishiyama, and N. Kato, "Capacity and delay of probing-based two-hop relay in manets," *IEEE Transactions on Wireless Communications*, vol. 11, no. 11, pp. 4172–4183, November 2012.

[12] 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.

[13] J. Liu, X. Jiang, H. Nishiyama, and N. Kato, "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.

[14] R. M. de Moraes, H. R. Sadjadpour, and J. Garcia-Luna-Aceves, "Taking full advantage of multiuser diversity in mobile ad hoc networks," *IEEE Transactions on Communications*, vol. 55, no. 6, pp. 1202 – 1211, June 2007.

[15] M. J. Neely and R. Urgaonkar, "Optimal backpressure routing for wireless networks with multi-receiver diversity," *Ad Hoc Networks*, vol. 7, no. 5, pp. 862–881, July 2009.

[16] 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.

[17] D. Ciullo, V. Martina, M. Garetto, and E. Leonardi, "Impact of correlated mobility on delay-throughput performance in mobile ad hoc networks," *IEEE/ACM Transactions on Networking*, vol. 19, no. 6, pp. 1745–1758, December 2011.

[18] P. Li, Y. Fang, J. Li, and X. Huang, "Smooth trade-offs between throughput and delay in mobile ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 11, no. 3, pp. 427–438, March 2012.

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

[20] 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.

[21] J. Gao and X. Jiang, "Delay modeling for broadcast-based two-hop relay manets," School of systems information science, Future university hakodate, Tech. Rep., January 2013, [Online]. Available: http://researchplatform.blogspot.jp/.

[22] H. Daduna, *Queueing Networks with Discrete Time Scale: Explicit Expressions for the Steady State Behavior of Discrete Time Stochastic Networks*. Springer, 2001.

[23] A. S. Alfa, *Queueing Theory for Telecommunications: Discrete Time Modelling of a Single Node System*. Springer, 2010.

[24] G. Latouche and V. Ramaswamy, *Introduction to Matrix Analytic Methods in Stochastic Modeling*. ASA-SIAM Series on Statistics and Applied Probability, 1999.

[25] 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.

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

[27] C++ simulator for the 2HR-B MANETs. [Online]. Available: http://researchplatform.blogspot.jp/.

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