Throughput Capacity of MANETs with Power Control and Packet Redundancy

Throughput Capacity of MANETs with Power Control and Packet Redundancy

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

***Abstract*—This paper studies the exact per node throughput capacity of a MANET, where the transmission power of each node can be controlled to adapt to a specified transmission range υ and a generalized two-hop relay with limited packet redundancy f is adopted for packet routing. Based on the concept of automatic feedback control and the Markov chain model, we first develop a general theoretical framework to fully depict the complicated packet delivery process in the challenging MANET. With the help of the framework, we are then able to derive the exact per node throughput capacity for a fixed setting of both υ and f. Based on the new throughput result, we further explore the optimal throughput capacity for any f but a fixed υ and also determine the corresponding optimum setting of f to achieve it. This result helps us to understand how such optimal capacity varies with υ (and thus transmission power) and to find the maximum possible throughput capacity of such a network for any f and υ. Interestingly, our results show that increasing the transmission power of the nodes improves the capacity, which is the same as that proved in fixed networks.**

***Index Terms*—Capacity, mobile ad hoc networks, power control, packet redundancy.**

I. Introduction

Mobile ad hoc network (MANET), a flexible and self-autonomous wireless network architecture, is very promising to find many important applications in the daily information exchange, disaster relief, military troop communication, etc. By now, the lack of a general Shannon limit-like network capacity theory is still a challenging roadblock stunting the development and commercialization of MANETs [1], [2]. It is expected such a theory can help us to understand the fundamental network throughput limit and thus serves as an instruction guideline for the network design, performance optimization and engineering of future MANETs [3]–[5].

Since the seminal work of Grossglauser and Tse (2001) [6], a lot of research efforts have been devoted to a better understanding of the MANET throughput capacity (i.e., the maximum achievable input rate that can be supported by a spatial and temporal scheduling algorithm) under various mobility models. Grossglauser and Tse [6] showed that under the i.i.d. mobility model, it is possible to achieve a $\Theta(1)$ per node throughput by employing a two-hop relay scheme. Following this line, it was later proved that the $\Theta(1)$ per node throughput can also be achieved under other mobility models, like the random walk model [7], the two-dimensional Brownian motions model [8] and the restricted mobility model [9]. Moraes *et al.* further showed that under uniform mobility model, we can still have the $\Theta(1)$ throughput even with a variant of the two-hop relay scheme, where each packet is only broadcasted once by its source and all nodes that receive the packet will act as its relays [10]. Shila *et al.* in [11] showed that by deploying m base stations in MANETs, the $\Theta(1)$ capacity can also be achieved while improving the average end-to-end delay by a factor of m.

Recently, the trade-off between the throughput capacity and delay performance in MANETs has also been extensively explored. Perevalov *et al.* [12] studied the delay-limited throughput of MANETs and reported that under the i.i.d. mobility model, the achievable throughput is of order $\Theta(n^{-1/3})$ for a fixed delay value d and the throughput increases as $d^{2/3}$ when the delay d is a moderate value. Lin *et al.* [13] considered the Brownian motion model and showed that the two-hop relay scheme proposed by Grossglauser and Tse, while capable of achieving a per node throughput of $\Theta(1)$, incurs an expected packet delay of $\Omega(\log n/\sigma_n^2)$, where σ_n^2 is the variance parameter of the Brownian motion model. Neely *et al.* [14] proved that under the i.i.d. mobility model, it is able to achieve $O(1/\sqrt{n})$ throughput and $O(\sqrt{n})$ delay by introducing exact $\sqrt{n}$ redundancy for each packet. More recently, the per node throughput capacity and delay trade-off has also been studied under the random waypoint model [15].

It is notable that the above works mainly focus on deriving the order sense results of throughput capacity in MANETs. Although the order sense results are helpful for us to understand the general scaling law and thus the growth rate of throughput capacity with network size n, they tell us little about the exact achievable per node throughput. In practice, however, such exact achievable throughput is of great interest for network designers. Some initial works are now available on the exact capacity study of MANETs. Neely *et al.* [14], [16] established the exact capacity of cell partitioned MANETs under both the i.i.d. mobility model and the more general Markovian mobility models, where it was assumed that the transmission power (and thus transmission range) of each node is fixed and the interference among simultaneous link transmissions can be avoided by using orthogonal channels in adjacent cells. Gao *et al.* [17] later extended the above work to a class of MANETs where the group-based scheduling is adopted to

This work was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2012 proceedings.

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.

schedule simultaneous link transmissions. Recently, Liu *et al.* [18] explored the exact capacity for the MANETs based on a specific two-hop relay routing algorithm with limited packet redundancy, i.e., a limited number of copies can be dispatched for each packet, and further extended capacity analysis to the scenario where each transmitter is allowed to conduct multiple rounds of probing for identifying a possible receiver [19]. Closed-form models has also been developed for achievable throughput analysis in a directional antenna-based MANET [20]. It is noticed that the capacity results in [18]–[20] hold only when the packet redundancy is smaller than a specific value (i.e., falls within a restricted range); while in [14], [16], the capacity was derived without considering the important interference, medium contention and traffic contention issues. In this paper, we explore the exact throughput capacity with a careful consideration of these issues and also for a general setting of packet redundancy.

Another limit of available works is that the impact of node transmission range on the throughput capacity of MANETs has been largely neglected. Since it is generally believed that the local transmission mode could result in the maximum per node throughput capacity, these work generally adopt the local transmission mode in their analysis, where either each node has a small transmission range of $\Theta(1/\sqrt{n})$ [6], [9], [10], [12], [13], [15], [21], or it can only transmit to some other node(s) in the same cell [7], [8], [14]. Therefore, the throughput capacity under a general setting of node transmission range remains unknown by now. To address this issue, we study the exact per node throughput capacity of a MANET where the transmission power of each node can be flexibly controlled such that the transmission range can be adapted to a specific value.

The main contributions of this paper are summarized as follows:

- By modeling the packet dispatching at the source and the packet receiving at the destination as Markov chains and applying the concept of automatic feedback control to characterize the service rate adaptation between the source and the destination, we first develop a general theoretical framework to depict the complicated packet delivery process in the challenging MANET.
- With the help of the theoretical framework, we then develop the exact per node throughput capacity $\mu(v, f)$ for any specified setting of transmission range v and packet redundancy limit f. Simulation results are also provided to validate the throughput capacity result.
- Based on the new throughput result, we further explore the optimal capacity $\max_f\{\mu(v, f)\}$ for a fixed v and also determine the corresponding optimum setting of f to achieve it. This result helps us to understand how such optimal capacity varies with v and to find a suitable v (and also f) to achieve the possible maximum throughput capacity $\max_{v,f}\{\mu(v, f)\}$ of such a network.

The rest of this paper is outlined as follows. Section II provides the system assumptions and definitions, and Section III discusses the issues of transmission scheduling and packet routing. In Section IV, we develop the theoretical framework for achievable per node throughput and present numerical results to validate it. We study in Section V the throughput maximization problem and explore the impact of transmission range on the maximum per node throughput capacity, and finally conclude the paper in Section VI.

II. System Assumptions and Definitions

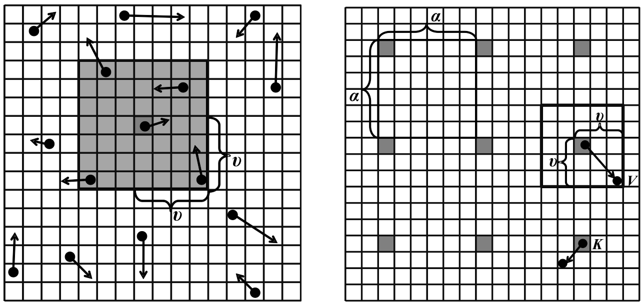
Network Model: Similar to previous works [8], [14], [18], in this paper we consider a two-dimensional cell-partitioned unit torus with n independent mobile nodes, as illustrated in Fig. 1(a). Time is slotted, and in order to exclusively explore and thus clearly illustrate the impact of transmission range on per node throughput capacity, we assume a $\sqrt{n} \times \sqrt{n}$ cell partition[1]. We assume a limited channel bandwidth such that the total number of bits that can be transmitted per time slot is fixed and normalized to one packet. We further assume that during each time slot each node has the knowledge about which cell it resides in based on its location information (For node localization, please refer to [22], [23]).

The cell-partitioned network model is popular and has been widely adopted in literature [7], [8], [14]. The motivations behind assuming such a network model are two folds: first, it could establish a mapping from the node behavior (like node movements and data transmissions) in a given network area to that among finite network cells. Such mapping can enable a flexible control of both node moving velocity and transmission range by accordingly adjusting the cell size. Second, it enables closed-form expressions to be derived for various performance metrics and is very helpful for network performance modeling and analysis.

Mobility Model: This paper focuses on the independent and identically distributed (i.i.d.) mobility model [24]–[26], which is defined as follows: at time slot $t = 0$, a node is uniformly placed in one of the n cells; at the beginning of time slot $t = 1$, the node independently and randomly selects a cell among the n cells with equal probability of $1/n$, then moves to the selected cell and stays inside for the whole time slot. This process is repeated by the node in every subsequent time slot. One can easily see that as the node movements are independent of each other, the n nodes are uniformly and randomly distributed in the n cells at each time slot. Furthermore, since the node movements are also independent from time slot to time slot, the node positions are totally reshuffled at each time slot.

Note that the time a node takes to move from one cell to another cell actually depends on the distance between two cells and the node moving velocity. Specifically, if at the beginning of a time slot a node selects a cell which is too far away from its current cell, the node may even take multiple time slots to move to the selected cell. For a general transmission range and a given node pair, it is extremely difficult to formulate the exact time when two nodes will move into reciprocal communication range, and the available time that can be utilized for data transmission. It becomes much more difficult

[1]In the case that $\sqrt{n}$ is not an integer, i.e., when the network is divided into $m \times m$ equal cells where $m = \lceil n \rceil$ or $m = \lfloor n \rfloor$, the corresponding per node throughput capacity can also be easily derived based on the theoretical framework developed in this paper.



(a) Cell-partitioned network.

(b) An illustration of transmission-group with $\alpha = 6$.

Fig. 1. Network cell partition and transmission-group.

(if not impossible) to analytically explore the probability of a node conducting successful data transmissions in a time slot, especially when taking into account the important issues of interference and medium contentions in the challenging MANETs.

It is noticed that the main focus of this paper is to develop a theoretical framework for deriving closed-form expressions for per node throughput capacity under the general settings of transmission range and packet redundancy. In order to simplify the analysis and thus keep the theoretical framework tractable, we assume that the time a node takes to move from one cell to another cell is neglected, similar to that in previous studies [7], [8], [14], [16], [27]. It is further noticed that under this assumption, the per node throughput capacity derived in this paper could actually serve as a meaningful bound in the limit of infinite mobility. The results in [14] indicate that the network capacity derived under the i.i.d. mobility model is actually identical to the one derived under other non-i.i.d. mobility models (like the Markovian random walk model and random waypoint model) if they follow the same steady state channel distribution.

Communication Model: To account for the interference among simultaneous transmissions, the Protocol model introduced in [8], [28] is adopted here. For a link i at time slot t, we use $T_i(t)$ and $R_i(t)$ to denote the positions of the corresponding transmitter and receiver, respectively. Based on the Protocol model, the transmission of the link i can be successful at the time slot t if for any other link j with simultaneous transmission we have

$$|T_j(t) - R_i(t)| \geq (1 + \Delta)|T_i(t) - R_i(t)|$$

here Δ is a protocol specified guard factor for interference control. In order to explore the impact of power control on per node throughput capacity, similar to [29] we assume that each node employs a power level so as to cover a set of cells with horizontal and vertical distances no more than $\upsilon - 1$ cells away from its current cell, where $1 \leq \upsilon \leq \lfloor \frac{\sqrt{n}+1}{2} \rfloor$, and $\lfloor x \rfloor$ is the floor function. With such power control, a node could transmit to any other node in a square area centered at the cell of the node and of side length $(2\upsilon - 1)$, as illustrated in Fig. 1(a).

Traffic Model: This paper considers the permutation traffic pattern widely adopted in previous studies [6], [14], [18], [24], [27], [30], [31]. Under such traffic model, there will be in total n distinct flows, where each node is the source of its locally generated traffic flow and at the same time the destination of a flow originated from another node. For the traffic flow originated at each node, we assume it has an average rate of λ (packets/slot). The packet arrival process at each node is independent of the mobility process and packets arrive at the beginning of a time slot. For the purpose of throughput capacity analysis, we assume that no lifetime is associated with each packet and the buffer size at each node is large enough (or infinite) such that the packet loss due to buffer overflow will never happen.

Throughput Capacity: We call a traffic input rate λ (packets/slot) feasible or achievable if there exists a spatial and temporal scheduling algorithm such that under this input rate the queue length at each node will never increase to infinity as the time goes to infinity. The per node throughput capacity is then defined as the maximum feasible input rate λ. Without incurring any ambiguity, hereafter we call such capacity as throughput capacity for brevity.

Remark 1: The cell-partitioned network model, i.i.d. mobility model, Protocol interference model, and permutation traffic model, which have been widely adopted in literature for MANET performance modeling and analysis, are actually abstracted from the actual environment to capture some major features in the real MANETs. Specifically, the cell-partitioned network model is mainly used for a MANET with discrete (slotted) time system to establish an efficient mapping from the given network area to a finite number of network cells. It enables a flexible control of both the node movement speed and the node transmission range to be made for the challenging MANET, by accordingly setting the cell side length. The practical meanings of the i.i.d. mobility are two folds: firstly, it can be used to represent a large class of node mobility which has homogeneous mobility patterns and exhibits uniform node distributions at each time slot; secondly, since with the i.i.d. mobility model the network topology varies dramatically and the network behavior can never be predicted, the network performance observed under such model can be regarded as the limiting performance in the regime of infinite mobility. Given a transmitter-receiver pair, by defining an exclusive region around the receiver via the guard factor Δ, the Protocol interference model can effectively prevent other nodes which are not sufficiently far away from the receiver from transmitting simultaneously. Furthermore, it can also enable a flexible control of simultaneous transmissions by adjusting the value of Δ according to the specified system requirement. The permutation traffic model actually represents the worst-case scenario of uni-cast in the real-world MANETs. In light of the fact that in the actual MANETs some nodes may have no traffic to deliver or receive, the per node throughput capacity derived under the permutation traffic model may serve as a meaningful achievable lower bound.

III. Transmission Scheduling and Routing

This section introduces the transmission scheduling and routing schemes to be adopted in this paper.

A. Transmission-Group Based Scheduling

It is noticed that under the CSMA/CA-based 802.11 DCF (distributed coordination function), packets can be transmitted at any time, and that whether the packets are successfully received or not depends on the actual SINR at the receiver, which is totally different from the time slotted system and the Protocol interference model considered in this paper. According to the Protocol interference model, multiple links could simultaneously transmit if they are sufficiently far away from each other. In order to support as many simultaneous link transmissions as possible under the time slotted system and the Protocol interference model, we consider in this paper a simple transmission-group based scheduling scheme, where at the beginning of each time slot each node is assumed to be able to judge whether it is inside an active cell or not. Actually, such scheduling scheme has also been widely adopted in literature, see, for example [18], [27], [32], [33].

Transmission-group: A transmission-group is a subset of cells, where any two of them have a vertical and horizontal distance of some multiple of α cells and all of them could conduct transmissions simultaneously.

An example of transmission-group is illustrated in Fig. 1(b), where all the shaded cells are of the same transmission-group and each of them can simultaneously support a transmitting node in it without interfering with each other. It is notable that for the transmission-group based scheduling with parameter α, there will be in total α^2 distinct transmission-groups, where each cell belongs to one distinct transmission-group. If all transmission-groups alternatively become active (i.e., get the transmission opportunity), then each transmission-group (and thus each cell) becomes active in every α^2 time slots.

Setting of Parameter α: To support as many simultaneous transmissions as possible, we need to properly tune the parameter α according to the parameters υ and Δ. As illustrated in Fig. 1(b), suppose that node V is scheduled to receive from a transmitting node, while node K in other active cells of the same transmission-group is transmitting to another node. Notice that in this paper we consider a network scenario where each node employs a power level so as to cover a set of cells which have a horizontal and vertical distance of no greater than $\upsilon - 1$ cells away from its current cell, $1 \le \upsilon \le \lfloor \frac{\sqrt{n}+1}{2} \rfloor$. Thus, we assume that the node V is at a distance of (x, y) ($x, y \in [-\upsilon+1, \upsilon-1]$) cells away from its transmitting node, where the x and y denote the horizontal distance and vertical distance, respectively. It is trivial to see that we only need to consider the cases that $x \in [0, \upsilon-1], y = \upsilon-1$. We can easily see that the distance from node V to its transmitting node is at most $\frac{1}{\sqrt{n}}\sqrt{\upsilon^2 + (x+1)^2}$, while another simultaneous transmitting node (say the node K in Fig. 1(b)) is at least $\frac{1}{\sqrt{n}}\sqrt{(\alpha-\upsilon)^2 + x^2}$ away from the node V. According to the interference model, the condition that K will not interfere with the reception at V is that for any $x \in [0, \upsilon - 1]$,

$$\frac{1}{\sqrt{n}}\sqrt{(\alpha-\upsilon)^2 + x^2} \ge (1+\Delta)\frac{1}{\sqrt{n}}\sqrt{\upsilon^2 + (x+1)^2}$$

To ensure above inequality for each $x \in [0, \upsilon - 1]$, we have

$$\alpha \ge \upsilon + \sqrt{2(\Delta+1)^2\upsilon^2 - (\upsilon-1)^2}$$

Since α is an integer and $\alpha \le \sqrt{n}$, we can set α as follows

$$\alpha = \min\{\upsilon + \lceil \sqrt{2(\Delta+1)^2\upsilon^2 - (\upsilon-1)^2} \rceil, \lfloor \sqrt{n} \rfloor\} \quad (1)$$

Selection of Transmitting Node: As shown in [18], there exists a non-negligible probability (which approaches $1-2e^{-1}$ as n goes to infinity) that there are at least two nodes falling within an an active cell. When an active cell has more than one node, the selection of transmitting node can be implemented by a mechanism similar to the DCF. At the beginning of each time slot, each node independently judges whether it is inside an active cell or not. If not, it remains silent and will not contend for the transmitting opportunity. Otherwise, it starts its back-off counter with a seed randomly selected from $[0, CW]$ (CW represents the contention window), and then overhears the channel until its back-off counter becomes 0 or it hears a broadcasting message from a transmitter. If no broadcasting message is heard during the back-off counting process, it broadcasts out a message denoting itself as the transmitter. Based on the back-off counting mechanism, each node of an active cell has the same probability to win the transmission opportunity and thus become the transmitting node.

Remark 2: The transmission-group based scheduling scheme is adopted due to the following advantages: first, it can be easily implemented for the considered MANETs since it is fully distributed without any centralized management; second, closed-form expressions can be derived for the probability of successful transmissions during each time slot under the transmission-group based scheduling scheme, based on which theoretical analysis can be further conducted for per node throughput capacity. It is also noticed that the parameter α in equation (1) is determined according to the extreme case where the distance between a transmitter-receiver pair is maximized. However, one can easily see that such extreme case rarely happen, i.e., with high probability the distance between the transmitter-receiver pair is smaller. Therefore, the transmission-group based scheduling scheme may unavoidably result in an inefficient spatial reuse due to the conservative setting of α.

B. 2HR-f Routing Scheme

In this paper, we consider a generalization of the classic two-hop routing scheme with f-cast (2HR-f) [14], [18], [34], $f \in [1, n-2]$, where each packet waiting at the source is delivered to at most f distinct relay nodes (i.e., each packet has a limited redundancy f) and should be received in order at its destination. For the permutation traffic pattern considered in this paper, there are in total n distinct traffic flows. Without loss of generality, we focus on a tagged flow in our discussion and use S and D to denote the source node and the destination node, respectively.

As a common complication for the designing of relay algorithm with packet redundancy, there may exist some residual packets (copies) lingering for a long time in the network, even after they have already been received by the destination. Obviously, such remnant packets may create excess congestion, waste network buffer, and must somehow be removed. Some approaches have been proposed to address this issue, such as the Time-To-Live (TTL) based mechanism [35] and the broadcasting based feedback mechanism [36]. These mechanisms work fine in lots of scenarios, however, due to the packet lifetime limit and specific communication resources taken for feedback broadcasting, both of them may unavoidably affect the per node throughput capacity. Towards this end, we adopt the sequence number based mechanism instead in this paper [14]. Specifically, the source S labels each packet P of the tagged flow with a *sequence number* $SN(P)$, while the destination D maintains a *request number* $RN(D)$ to indicate the sequence number of the packet it is currently requesting. Every time S or relay nodes meet D, they can efficiently update their buffers by deleting all packets with sequence numbers less than the current request number obtained from D. Note that the sequence number mechanism ensures that every packet is received in order at the destination. For the more general case where a certain level of out-of-order reception is allowed at the destination, an efficient group number based mechanism was proposed in [37] to remove the remnant packet copies, and please refer to [37] for details.

The 2HR-f scheme is summarized as follows.

2HR-f Routing Scheme: When S wins the transmission opportunity at a time slot, S first overhears the channel for a specified interval of time to check whether D is inside the one-hop transmission range.

1) If S hears the reply from D within the specified time interval, it initiates a handshake with D and then transmits a packet directly to D ("Source-to-Destination" transmission);
2) If no broadcasting reply is overheard during the specified time interval, a receiving node (say R) is randomly selected among the nodes within the one-hop transmission range of S based on a mechanism similar to the selection of transmitting node. With probability $1/2$, S and R then perform either the "Source-to-Relay" or "Relay-to-Destination" transmission:
   - Source-to-Relay: Suppose that packet P is the locally generated packet for which S is currently delivering copies. S first initiates a handshake with R to check whether R has already received a copy of P. If not, S delivers out a new copy of P to R if less than f copies of P have been delivered by now; otherwise, S remains idle for this time slot.
   - Relay-to-Destination: S initiates a handshake with R to check if S carries a packet P^* destined for R with $SN(P^*) = RN(R)$. If so, S delivers the packet P^* to R; otherwise, S remains idle for this time slot.

IV. THROUGHPUT CAPACITY

In this section, we first introduce some basic results related to transmission probabilities and service times under the 2HR-f routing, and then use them to derive the per node throughput capacity.

A. Some Basic Results

We present the following lemmas first, i.e., Lemmas 1, 2, 3, 4, and 5. The basic ideas of proofs are similar to that in

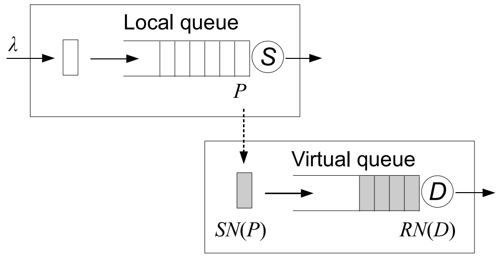


Fig. 2. An illustration of the local queue at S and the virtual queue at D.

[18], [37], and please refer to [38] for the proofs of these lemmas. Note that Lemmas 1, 2, 3, 4, and 5 further generalize the results established in [18], [37] to the case in which each node could transmit to the cells with horizontal and vertical distance of no more than $\upsilon - 1$ cells away from its current cell, $1 \leq \upsilon \leq \lfloor \frac{\sqrt{n}+1}{2} \rfloor$. Actually, with these lemmas one can easily recover the results in [18], [37] by setting $\upsilon = 2$.

Lemma 1: For the tagged flow and a given time slot, 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. By setting $m = (2\upsilon - 1)^2$, we have

$$p_1 = \frac{1}{\alpha^2}\Big\{\frac{m-1}{n-1}\Big(1-\Big(\frac{n-1}{n}\Big)^{n-1}\Big) + \frac{1}{n}\Big(\frac{n-1}{n}\Big)^{n-1}\Big\} \quad (2)$$

$$p_2 = \frac{1}{\alpha^2}\Big\{\frac{n-m}{n-1}\Big(1-\Big(\frac{n-1}{n}\Big)^{n-1}\Big) - \Big(\frac{n-m}{n}\Big)^{n-1}\Big\} \quad (3)$$

Lemma 2: Suppose that S is delivering copies for packet P in the current time slot, D is also requesting for P, i.e., $SN(P) = RN(D)$, and there are already j ($1 \leq j \leq f+1$) copies of P in the network (including the original one at S). For the next time slot, we use $P_r(j)$ to denote the probability that D will receive P, use $P_d(j)$ to denote the probability that S will deliver out a copy of P to a new relay (if $j \leq f$), and use $P_s(j)$ to denote the probability of simultaneous source-to-relay and relay-to-destination transmissions. By setting $m = (2\upsilon - 1)^2$, we have

$$P_r(j) = p_1 + \frac{j-1}{2(n-2)}p_2 \quad (4)$$

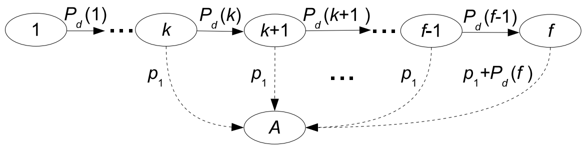
$$P_d(j) = \frac{n-j-1}{2(n-2)}p_2 \quad (5)$$

$$P_s(j) = \frac{(j-1)(n-j-1)(n-\alpha^2)}{4n\alpha^4}\sum_{k=0}^{n-5}\binom{n-5}{k}h(k) \cdot\Big\{\sum_{t=0}^{n-4-k}\binom{n-4-k}{t}h(t)\Big(\frac{n-2m}{n}\Big)^{n-4-k-t}\Big\} \quad (6)$$

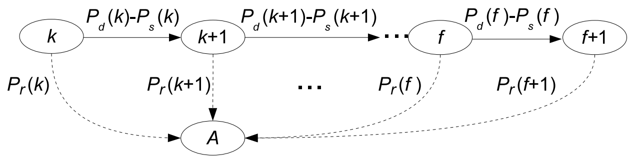
where

$$h(x) = \frac{m\big(\frac{m}{n}\big)^{x+1} - (m-1)\big(\frac{m-1}{n}\big)^{x+1}}{(x+1)(x+2)}$$

In order to derive the per node throughput capacity, we now define the following two queues for S and D. As shown in Fig. 2, the local queue at S stores the packets locally generated at S, while the virtual queue at D stores the sequence numbers



(a) Absorbing Markov chain for the packet distribution process at S.



(b) Absorbing Markov chain for the packet reception process at D.

Fig. 3. Absorbing Markov chains for a packet P of the tagged flow, given that D starts to request for P when there are already k copies of P in the network. For each transient state, the transition back to itself is not shown for simplicity.

of the packets not received yet by D. Every time a packet is locally generated at S, it is put to the end of the local queue; every time a packet, say P, is moved to the head-of-line of the local queue, its sequence number $SN(P)$ is put to the end of the virtual queue. The head-of-line entry of the local queue represents the packet that S is currently distributing copies, while the head-of-line entry of the virtual queue represents the sequence number of the packet that D is currently requesting, i.e., $RN(D)$. When S finishes copy distribution for the head-of-line packet at the local queue, S moves ahead the remaining packets waiting behind; when D receives the packet whose sequence number equals the head-of-line entry at the virtual queue, D moves ahead the remaining entries.

Remark 3: The local queue defined at S characterizes the local packet generation and copy dispatching at the source node, which corresponds to the "Source-to-Relay" transmission of the 2HR-f scheme; while the virtual queue introduced to D depicts the packet delivery from the intermediate relay nodes to the destination, which corresponds to the "Relay-to-Destination" transmission.

The service time of the local queue at S and the service time of the virtual queue at D can then be defined as follows:

Definition 1: For a packet P of the tagged flow, its service time at the local queue is defined as the time elapsed between the time slot when S starts to deliver copies for P and the time slot when S stops distributing copies for P.

Definition 2: For a packet P of the tagged flow, its service time at the virtual queue is defined as the time elapsed between the time slot when D starts to request for P and the time slot when D receives P.

Actually, the local queue characterizes the packet arrival and copy distribution at S, while the virtual queue defines the packet delivery from relay nodes to D. Note that each entry waiting in the virtual queue represents the sequence number of a packet that has not been received yet by D. Without loss of generality, consider a packet P and suppose that there are k copies of P in the network when D starts to request for P, $1 \le k \le f+1$. If we denote by A the state that S has finished the copy distribution for P or the state that D has received a copy of P, then the service process of P at the local queue and the service process of P at the virtual queue can be defined by two finite-state absorbing Markov chains, as defined in Fig. 3(a) and Fig. 3(b), respectively. We further denote by $X_S(k)$ the corresponding service time at the local queue and denote by $X_D(k)$ the corresponding service time at the virtual queue. Then according to the Markov chain theory [39], $X_S(k)$ (resp. $X_D(k)$) can be regarded as the time it takes the Markov chain in Fig. 3(a) (resp. Fig. 3(b)) to become absorbed given that the chain starts from the state 1 (resp. the state k).

Lemma 3: For a packet P of the tagged flow, suppose that there are k copies of P in the network when D starts to request for P, $1 \le k \le f+1$, then we have

$$\mathbb{E}\{X_S(k)\} = \begin{cases} \sum_{i=1}^{k-1} \frac{1}{P_d(i)} + \frac{1}{p_1+P_d(k)} \\ \quad \cdot \big(1+\sum_{j=1}^{f-k} \phi_1(k,j)\big) \text{ if } 1 \le k \le f, \\ \sum_{i=1}^{f} \frac{1}{P_d(i)} \qquad \text{if } k = f+1. \end{cases} \tag{7}$$

$$\mathbb{E}\{X_D(k)\} = \begin{cases} \frac{1}{P_r(k)+P_d(k)-P_s(k)} \big(1+\sum_{j=1}^{f-k} \phi_2(k,j) \\ \quad + \frac{P_d(f)-P_s(f)}{P_r(f+1)} \phi_2(k,f-k)\big) \\ \qquad \text{if } 1 \le k \le f-1, \\ \frac{1}{P_r(f)+P_d(f)-P_s(f)} \big(1+\frac{P_d(f)-P_s(f)}{P_r(f+1)}\big) \\ \qquad \text{if } k = f, \\ \frac{1}{P_r(f+1)} \qquad \text{if } k = f+1. \end{cases} \tag{8}$$

where

$$\phi_1(k,j) = \prod_{t=1}^{j} \frac{P_d(k+t-1)}{p_1+P_d(k+t)}$$

$$\phi_2(k,j) = \prod_{t=1}^{j} \frac{P_d(k+t-1)-P_s(k+t-1)}{P_r(k+t)+P_d(k+t)-P_s(k+t)}$$

Lemma 4: For any $1 \le k \le f$, we have

$$\mathbb{E}\{X_S(k)\} < \mathbb{E}\{X_S(k+1)\} \tag{9}$$

$$\mathbb{E}\{X_D(k)\} > \mathbb{E}\{X_D(k+1)\} \tag{10}$$

For the tagged flow, if we further denote by $\overline{X}_S$ the average service time of all packets in the local queue served by S, and denote by $\overline{X}_D$ the average service time of all entries in the virtual queue served by D, then we can establish the following result based on Lemma 4.

Lemma 5: For any given transmission range parameter v and packet redundancy limit f, $1 \le v \le \lfloor \frac{\sqrt{n}+1}{2} \rfloor$, $1 \le f \le n-2$, we have

$$\mathbb{E}\{X_S(1)\} \le \overline{X}_S \le \mathbb{E}\{X_S(f+1)\} \tag{11}$$

$$\mathbb{E}\{X_D(f+1)\} \le \overline{X}_D \le \mathbb{E}\{X_D(1)\} \tag{12}$$

B. Per Node Throughput Capacity

For the tagged flow, suppose that currently packet P is the head-of-line packet at the local queue of S, D just receives the last packet before P and there are already k ($1 \le k \le f$) copies of P in the network. We further assume that the packet waiting right behind P in the local queue is packet P', and

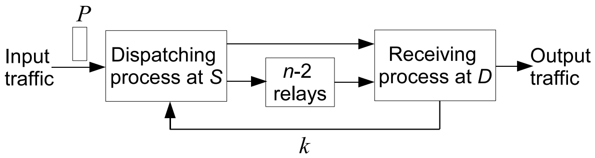


Fig. 4. An illustration of the automatic feedback control system defined for the packet delivery process of the tagged flow, where the parameter k is automatically updated to adjust to the service rates at the S and D.

D will start to request for P' when there are k' copies of P' in the network. Then we have the following two cases:

- If $\mathbb{E}\{X_S(k)\} \leq \mathbb{E}\{X_D(k)\}$, the copy dispatching of P at S is faster than the packet receiving at D and thus we have $k' \geq k$ for packet P' in the average case. According to (9) and (10), it follows that $\mathbb{E}\{X_S(k')\} \geq \mathbb{E}\{X_S(k)\}$ and $\mathbb{E}\{X_D(k')\} \leq \mathbb{E}\{X_D(k)\}$. Since $\mathbb{E}\{X_S(k)\} \leq \mathbb{E}\{X_D(k)\}$, then we have
$$\mathbb{E}\{X_D(k) - X_S(k)\} \geq \mathbb{E}\{X_D(k') - X_S(k')\} \quad (13)$$
The above inequality indicates that from P to P', the expected gap between the service time at the destination and the service time at the source tends to become smaller. Since (13) also holds for the packets (if any) waiting behind P', we can see that the $\overline{X}_S$ and $\overline{X}_D$ (and thus the average service rates at S and D) will gradually approach each other until a balance is achieved.
- If $\mathbb{E}\{X_S(k)\} \geq \mathbb{E}\{X_D(k)\}$, the packet receiving of P at D is faster than the copy dispatching at S and thus we have $k' \leq k$ for packet P' in the average case. Similarly, we have $\mathbb{E}\{X_S(k')\} \leq \mathbb{E}\{X_S(k)\}$ and $\mathbb{E}\{X_D(k')\} \geq \mathbb{E}\{X_D(k)\}$. Together with $\mathbb{E}\{X_S(k)\} \geq \mathbb{E}\{X_D(k)\}$, it follows
$$\mathbb{E}\{X_S(k) - X_D(k)\} \geq \mathbb{E}\{X_S(k') - X_D(k')\} \quad (14)$$
which means that the expected gap between the service time at the source and the service time at the destination also tends to become smaller from P to P'.

Note that every time a relay receives a packet P from the source node, it puts P into the end of its relay-queue specified for the destination; on the other hand, every time it delivers a packet P to the destination, it deletes from the relay-queue all packets with sequence numbers less than $SN(P)$ (since the destination has already received all the packets before P). Therefore, from (13) and (14) one can see that, when averaged over all packets of the tagged flow, the average rate of buffer occupation and the average rate of buffer flush (i.e., the average rate of deleting packets from the relay-queues) at the $n-2$ relay nodes tends to being equal and thus the network system will gradually evolve towards being stable.

The above analysis indicates that under the 2HR-f routing scheme, the parameter k is automatically updated from packet to packet so as to adjust the service rates at S and D. Based on this intrinsic feature of automatic updating for parameter k, we can model the packet delivery process of the tagged flow as an automatic feedback control system shown in Fig. 4, where the packet dispatching process at S and the packet receiving process at D can be defined by the two absorbing Markov chains in Fig. 3(a) and Fig. 3(b), respectively[2].

Now we are ready to derive the throughput capacity for the tagged flow. We denote by V_S the long-term average packet dispatching rate at S and denote by V_D the long-term average packet receiving rate at D, where

$$V_S = \lim_{t\to\infty} \frac{\text{the number of dispatched packets at } S}{t} \quad (15)$$

$$V_D = \lim_{t\to\infty} \frac{\text{the number of received packets at } D}{t} \quad (16)$$

We then have the following result.

Lemma 6: For the tagged flow and any given parameters of υ and f, we have

$$V_S \leq \frac{1}{\mathbb{E}\{X_S(1)\}} \quad (17)$$

$$V_D \leq \frac{1}{\mathbb{E}\{X_D(f+1)\}} \quad (18)$$

Proof: We first prove (17). For the local queue at the source S, suppose that during some time interval t node S has successfully served $N_S(t)$ locally generated packets (i.e., S has distributed copies for $N_S(t)$ local packets). According to the definition in (15), we have

$$V_S = \lim_{t\to\infty} \frac{N_S(t)}{t} \quad (19)$$

Notice that during the time interval t, the local queue may be empty and thus the queue server S may become idle. Denote by $I_S(t)$ the accumulated vacancy time at S during the time interval t, then we have

$$\overline{X}_S = \lim_{t\to\infty} \frac{t - I_S(t)}{N_S(t)} \quad (20)$$

Since $I_S(t) \geq 0$, combining the (19) and (20), we have

$$V_S \leq \frac{1}{\overline{X}_S} \quad (21)$$

Substituting (11) into (21), (17) then follows. After a similar derivation, (18) also follows. ■

We now can establish the following main result on per node throughput capacity.

Theorem 1: Consider a cell-partitioned MANET where each node could transmit to the cells with horizontal and vertical distance of no more than $\upsilon - 1$ cells away from its current cell, $1 \leq \upsilon \leq \lfloor \frac{\sqrt{n}+1}{2} \rfloor$, and each packet follows the 2HR-f routing scheme, $1 \leq f \leq n-2$. If we denote by $\mu(\upsilon, f)$ the per node (flow) throughput capacity, i.e., the network can stably support any input rate $\lambda < \mu(\upsilon, f)$, then we have

$$\mu(\upsilon, f) = \min\left\{ p_1 + \frac{f}{2(n-2)} p_2, \frac{p_1 + p_2/2}{1 + \sum_{j=1}^{f-1} \prod_{t=1}^{j} \frac{(n-t-1)p_2}{2(n-2)p_1 + (n-t-2)p_2}} \right\} \quad (22)$$

[2]Notice that for a packet P of the tagged flow, its corresponding parameter k depends only on the delivery process of the last packet received just before P.

Proof: For the tagged flow, as its packet delivery process can be defined by the automatic feedback control system in Fig. 4, we can see that if the network is stable (i.e., the queue length at the source and at relay nodes will not go to infinity) under the input rate λ, then we have

$$\lambda = V_S = V_D \tag{23}$$

This is due to the fact that in a stable control system, the long-term average rate of the input traffic is equal to that of the output one.

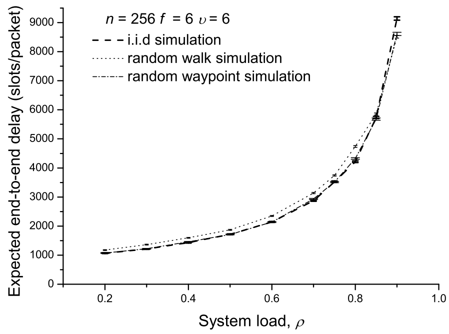
Based on (17) and (18), we then have

$$\mu(\upsilon, f) = \min\left\{\frac{1}{\mathbb{E}\{X_S(1)\}}, \frac{1}{\mathbb{E}\{X_D(f+1)\}}\right\} \tag{24}$$

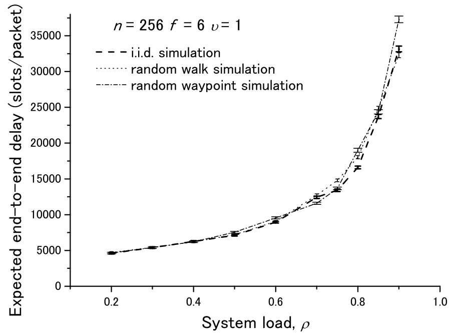
After substituting (7) and (8) into (24), (22) follows. ■

In this paper, we assume transmission coverage of a square area so as to approximate the actual circular transmission area and simplify the theoretical analysis. It is noticed that the adoption of other general cases of transmission coverage area, like the circle or hexagon transmission coverage area, will affect only the setting of transmission scheduling scheme (i.e., the distance parameter α in Section III-A), and the probabilities of a node conducting source-to-destination, source-to-relay, and relay-to-destination transmissions in each time slot. However, it will never change the fundamental properties or the basic features of the end-to-end packet delivery process, as characterized by our automatic feedback control and Markov chain model based framework. Therefore, the theoretical framework developed in this paper can also be applied to analyze the per node throughput capacity under the general circle or hexagon transmission coverage area, after accordingly updating the parameter α in (1), and probabilities p_1 and p_2 in Lemma 1.

Remark 4: In the 2HR-f routing scheme, the mechanism of sequence number and request number is adopted to ensure all packets are received in order at the destination. This is actually very strict since according to the buffer space, the destination may receive and keep somewhat out-of-order packets, rather than accept all incoming packets strictly in their sequence order. The generalized 2HR-f routing scheme which allows a certain level of out-of-order packet reception, i.e., the group-based two-hop relay algorithm with packet redundancy, was first proposed in [37]. 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 has never been received by the destination before and also it is among g packets of the group the destination is currently requesting. Packets are divided into consecutive groups at the source node, g packet per group, and the level of out-of-order packet reception at the destination can be flexibly controlled by adjusting the value of parameter g. Actually, the theoretical framework developed in this paper has been further extended to analyze the per node throughput capacity under the general 2HR-(f, g) algorithm in [40], and please kindly refer to [40] for details.



(a) Network scenario ($n = 256$, $f = 6$, $\upsilon = 6$) with $\mu(6, 6) = 1.17 \times 10^{-3}$ (packets/slot).



(b) Network scenario ($n = 256$, $f = 6$, $\upsilon = 1$) with $\mu(1, 6) = 2.84 \times 10^{-4}$ (packets/slot).

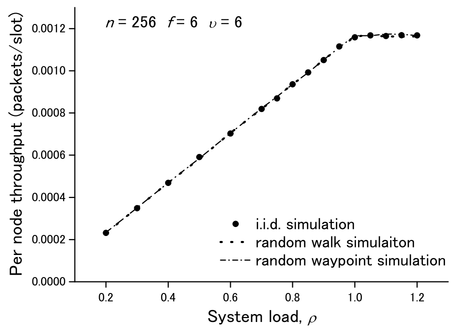
Fig. 5. The expected end-to-end packet delay for $n = 256$.

C. Model Validation

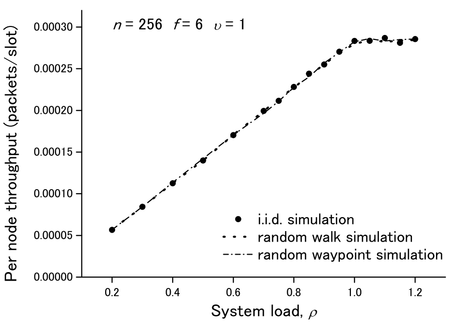
To validate the derived per node throughput capacity, a dedicated simulator in C++ was developed, which is now available at [41]. In our simulation, the traffic flow originated from each node is assumed to be a Poisson stream. Similar to the settings in [42], [43], the guard factor Δ is fixed as $\Delta = 1$ here. In addition to the i.i.d. mobility model considered in this paper, we also implemented our simulator for the popular random walk and random waypoint models:

- **Random Walk Model [7]:** 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 [25]:** At the beginning of each time slot, each node independently and randomly generates a two-dimensional vector $\boldsymbol{\nu} = [\nu_x, \nu_y]$, where the values of ν_x and ν_y are uniformly drawn from $[1/\sqrt{n}, 3/\sqrt{n}]$. The node then moves a distance of ν_x along the horizontal direction and a distance of ν_y along the vertical direction. The pause time is zero.

Extensive simulations have been conducted to validate the per node throughput capacity. We included here the simulation results of two network scenarios ($n = 256$, $f = 6$, $\upsilon = 6$) and ($n = 256$, $f = 6$, $\upsilon = 1$), and the simulation results

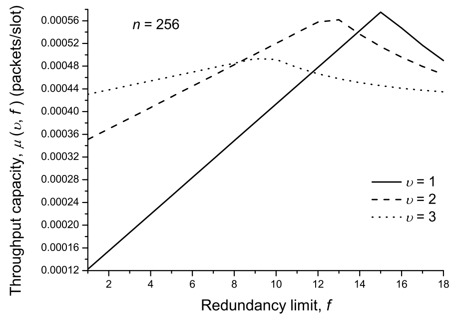


(a) The actual per node throughput for network scenario ($n = 256$, $f = 6$, $\upsilon = 6$).

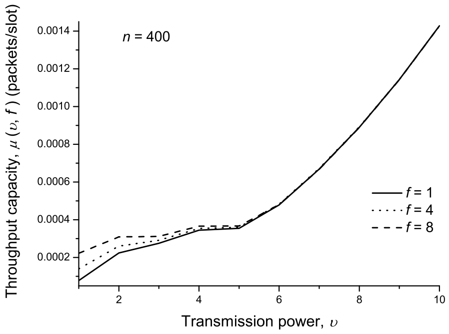


(b) The actual per node throughput for network scenario ($n = 256$, $f = 6$, $\upsilon = 1$).

Fig. 6. The actual per node throughput for $n = 256$.



(a) $\mu(\upsilon, f)$ vs. f



(b) $\mu(\upsilon, f)$ vs. υ

Fig. 7. Throughput capacity $\mu(\upsilon, f)$ vs. redundancy limit f and transmission power υ.

of other scenarios can be easily obtained by our simulator [41] as well. Fig. 5 shows how the expected end-to-end delay varies with the system load $\rho = \lambda/\mu$ under the two scenarios. Note that all the simulation results of expected end-to-end delay are reported with the 95% confidence interval. Fig. 5 indicates clearly that our theoretical framework can be used to nicely characterize the throughput capacity of the considered MANETs. As observed from Fig. 5 that when the system load ρ approaches 1 (i.e., when the traffic input rate λ approaches the throughput capacity μ), the packet delay rises up sharply and becomes extremely sensitive to the variations of ρ. Such skyrocketing behavior of packet delay as ρ approaches 1 serves as an intuitive validation for the throughput capacity determined by our theoretical framework.

In order to further validate the throughput capacity result, we examined the actual per node throughput and summarized the corresponding simulation results in Fig. 6. As observed from Fig. 6 that for both the network scenarios there, the actual per node throughput first monotonically increases with ρ and becomes flattened as ρ beyond 1. Specifically, as the system load ρ increases beyond 1 (i.e., the traffic input rate λ increases beyond the throughput capacity μ), the actual per node throughput becomes flattened at 1.16×10^{-3} and 2.83×10^{-4} in Figs. 6(a) and 6(b), respectively, which means that the throughput capacity determined in Theorem 1 is tight and achievable.

From Figs. 5 and 6, it is also interesting to note that the network we consider here actually exhibits very similar behaviors in terms of per node throughput under either the i.i.d. model, the random walk model or the random waypoint model. In this sense, our theoretical framework, although was developed for the throughput capacity under the i.i.d. mobility model, can also be used to nicely analyze that under the random walk and the random waypoint models as well.

We now explore how the per node throughput capacity $\mu(\upsilon, f)$ of the 2HR-f routing scheme varies with parameters υ and f. For the scenarios of $\upsilon = \{1, 2, 3\}$ and $n = 256$, Fig. 7(a) illustrates how $\mu(\upsilon, f)$ varies with packet redundancy limit f. We can see from Fig. 7(a) that for a given υ, we can find an optimum setting of f to achieve the maximum throughput capacity $\mu(\upsilon, f)$ (as to be analytically derived in Section V). For example, for the setting of $\upsilon = 1, 2$ and 3, the corresponding optimum setting of f are $15, 13$ and 9, respectively. It is also interesting to notice that when $f \leq 8$, the $\mu(\upsilon, f)$ of the case $\upsilon = 3$ is always the greatest one among that of all three cases, while the $\mu(\upsilon, f)$ of the case $\upsilon = 1$ becomes the greatest one when $f \geq 14$. However, this is not the case for the relationship between $\mu(\upsilon, f)$ and transmission power υ illustrated in Fig. 7(b). One can observe from Fig. 7(b) that, the $\mu(\upsilon, f)$ of all the three cases $f = \{1, 4, 8\}$ have very similar varying tendency with υ. Specifically, when $\upsilon \geq 5$ the three curves of $\mu(\upsilon, f)$ there almost coincide with each other. The impact of f on $\mu(\upsilon, f)$, therefore, can be almost neglected as the transmission power υ increases beyond some specific value.

V. Throughput Optimization

As shown in Fig. 7(a), for a given υ, there exists an optimum setting of f to achieve the maximum $\mu(\upsilon, f)$. We now proceed to explore the following optimization problem.

Throughput Optimization Problem: For a 2HR-f-based MANET with a fixed transmission range υ for each node, calculate its maximum per node throughput capacity for any value of f.

For a fixed transmission range υ, if we denote by μ^* the corresponding maximum per node throughput capacity, then the throughput optimization problem can be formulated as

$$\begin{aligned}\mu^* &= \max_f\{\mu(\upsilon, f)\}\\ &= \max\ \min\left\{\frac{1}{\mathbb{E}\{X_S(1)\}}, \frac{1}{\mathbb{E}\{X_D(f+1)\}}\right\} \end{aligned} \quad (25)$$

subject to:

$$1 \le f \le n-2, 1 \le \upsilon \le \lfloor \frac{\sqrt{n}+1}{2} \rfloor$$

where the $\mathbb{E}\{X_S(1)\}$ and $\mathbb{E}\{X_D(f+1)\}$ are defined in (7) and (8), respectively.

Regarding the solution of this optimization problem, we have the following result.

Lemma 7: For any given $\upsilon \in [1, \lfloor \frac{\sqrt{n}+1}{2} \rfloor]$, we have

$$\mu^* = max\left\{\frac{1}{\mathbb{E}\{X_D(f+1)\}\big|_{f=f_0}}, \frac{1}{\mathbb{E}\{X_S(1)\}\big|_{f=f_1}}\right\} \quad (26)$$

where

$$f_0 = \max\Big\{f \mid \mathbb{E}\{X_S(1)\} \le \mathbb{E}\{X_D(f+1)\}\Big\} \quad (27)$$

$$f_1 = \min\Big\{f \mid \mathbb{E}\{X_D(f+1)\} \le \mathbb{E}\{X_S(1)\}\Big\} \quad (28)$$

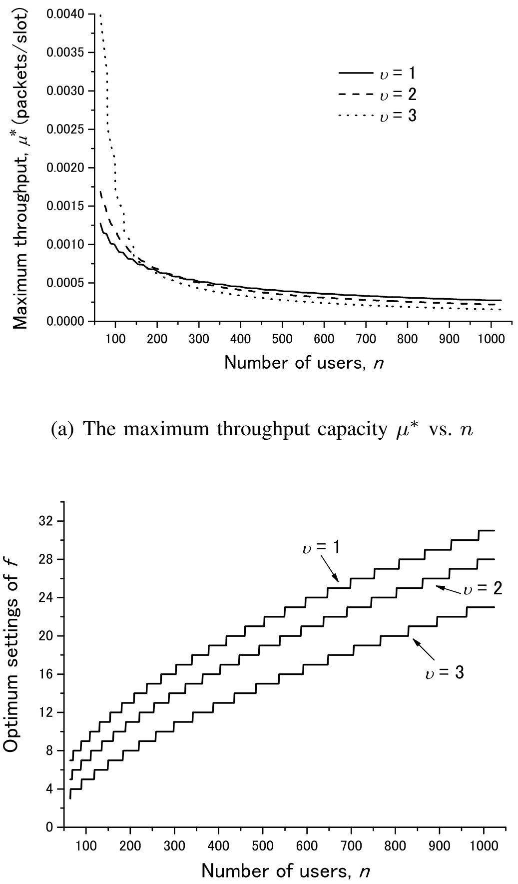
Proof: We first prove that f_0 and f_1 defined above do exist. According to (7) and (8), we have

$$\begin{aligned}\mathbb{E}\{X_S(1)\}|_{f=1} &= \frac{1}{p_1 + P_d(1)}\\ &\le \frac{1}{p_1 + \frac{p_2}{2(n-2)}} = \mathbb{E}\{X_D(f+1)\}|_{f=1} \end{aligned} \quad (29)$$

Notice that

$$\begin{aligned}&\mathbb{E}\{X_S(1)\}|_{f=n-2}\\ &= \frac{1}{p_1+p_2/2}\Big(1+\sum_{j=1}^{n-3}\prod_{t=1}^{j}\frac{(n-t-1)p_2}{2(n-2)p_1+(n-t-2)p_2}\Big)\\ &\ge \frac{1}{p_1+p_2/2} = \mathbb{E}\{X_D(f+1)\}|_{f=n-2} \end{aligned} \quad (30)$$

It is easy to see from (7) and (8) that as f increases, the $\mathbb{E}\{X_S(1)\}$ monotonically increases while the $\mathbb{E}\{X_D(f+1)\}$ monotonically decreases. Combining with the results in (29) and (30), we can see that the f_0 and f_1 defined above do exist.



(a) The maximum throughput capacity μ^* vs. n

(b) The optimum setting of f vs. n

Fig. 8. The maximum throughput capacity μ^* and the corresponding optimum setting of f for networks with n varying from 64 to 1024.

From (24) we know that

$$\mu(\upsilon, f) = \begin{cases} 1/\mathbb{E}\{X_D(f+1)\} & \text{if } 1 \le f \le f_0, \\ 1/\mathbb{E}\{X_S(1)\} & \text{if } f_1 \le f \le n-2. \end{cases} \quad (31)$$

Thus, (26) can be derived directly based on (31). The above results indicate that for a MANET with a fixed υ, there exists an optimum setting of f (f_0 or f_1) to achieve the optimal per node throughput capacity μ^*. ■

To illustrate the optimal throughput capacity μ^*, we show in Fig. 8(a) and Fig. 8(b) that for $\upsilon = \{1, 2, 3\}$, how μ^* and the corresponding optimum setting of f vary with network size n. Fig. 8(a) shows clearly that for all the three settings of υ here, although the corresponding μ^*'s all decrease quickly as the network size increases, their varying tendencies with n are actually different. A careful observation of Fig. 8(a) indicates that among the three settings of υ here, the μ^* of the case $\upsilon = 3$ decreases most dramatically with n while the one of the case $\upsilon = 1$ decreases least significantly with n. It is also interesting to note that when n is in different ranges, the impact of υ on μ^* can also be different. For example, when $n \le 143$, the μ^* of the case $\upsilon = 3$ is always the greatest one among that of all three cases, while the μ^* of the case $\upsilon = 1$ becomes the greatest one when n is larger than 270. The results in Fig. 8(b)

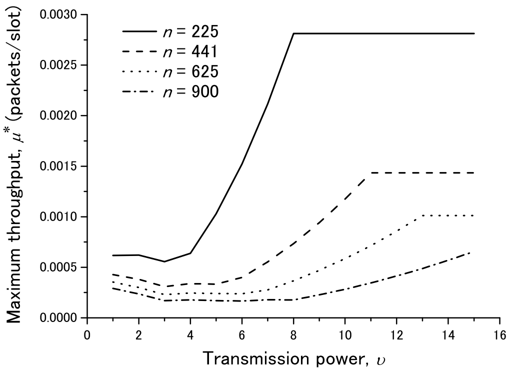


Fig. 9. The throughput capacity μ^* vs. node transmission region υ.

show that for a given υ, the corresponding optimum setting of f is actually a piecewise function of n. We can also see from the figure that for each network size n, the optimum f of case $\upsilon = 3$ is the smallest one among that of all three cases. This can be intuitively interpreted as follows. For one given network, if a bigger υ (and thus a larger transmission range) is adopted, a node will have higher probability to meet its destination or relay nodes and thus can deliver packets more fast, resulting in a fewer number of redundant copies for each packet before it arrives at its destination.

To further explore the relationship between υ and μ^*, we summarize in Fig. 9 how μ^* varies with υ for scenarios of $n = \{225, 441, 625, 900\}$. It is interesting to see that, for each network scenario there, as υ increases μ^* always first decreases and then increases. Actually, the effect of increasing υ is two-fold: on one hand, it increases the probability that a node meets the destination or relay nodes, but on the other hand it decreases the number of simultaneous transmissions in network. As illustrated in Fig. 9 that when υ is small, the latter negative effect dominates, while as υ gradually increases, the positive effect becomes the dominant one. It is further noticed that for the cases $n = 225$, $n = 441$ and $n = 625$, μ^* does not increase any more (i.e., saturates to a value) when υ increases beyond some threshold ($\upsilon = 8, 11$ and 13, respectively). This is because that with such settings of υ, a node is able to cover the whole network region and thus the destination node receives each packet directly from the source. According to (2), the saturated value of μ^* can be determined as

$$\mu^* = \frac{1}{\alpha^2}\Big(1 - \big(\frac{n-1}{n}\big)^n\Big)$$

It is notable that the results in both Fig. 9 and Fig. 8(a) actually imply a fact that for the mobile networks considered in this paper, we may have a significant improvement of per node throughput capacity and even achieve its possible maximum value through adopting a bigger υ (and thus a larger transmission range) for each node, which is the same as that proved in fixed networks [44], [45].

VI. Conclusion

Distinguished from the available works which mainly focus on deriving the order sense results and exploring the scaling laws of throughput capacity in MANETs, this paper addressed another basic problem: for a MANET with general node transmission range control and packet redundancy control, what is the exact achievable per node throughput. We found that for the network scenarios considered in this paper, it may not be always true that adopting the minimum transmission range can achieve the maximum per node throughput capacity, as what is generally believed in literature. This finding indicates that further deliberate studies are necessary to reveal the real achievable throughput of MANETs. Another interesting finding of our work is that the MANETs considered in this paper actually exhibit very similar behaviors in terms of per node throughput under different node mobility models, like the i.i.d., random walk and random waypoint. The results of this work can provide a clear guideline to network engineers by indicating whether a given vector of traffic input rate can be supported or not by the network, what is the possible optimal throughput capacity that can be achieved via all combinations of transmission range and packet redundancy. Furthermore, as our very initial step towards thoroughly understanding the fundamental capacity of general MANETs, this paper considered only some simple scenarios, and the results of this work are expected to lay a foundation for future analysis under other more general scenarios.

It is noticed that in this paper, we consider a simple network scenario where there is no lifetime associated with each packet and the buffer size at each node is assumed to be large enough. Note that the per node throughput capacity is actually a monotonically non-decreasing function of packet lifetime and buffer size at each node, and there may exist some thresholds of packet lifetime and also buffer size, beyond which the per node throughput capacity cannot increase anymore. Therefore, one interesting future research is to derive the per node throughput capacity of the 2HR-f under a more general scenario in which each packet is of limited lifetime and each node has constrained buffer space, and further explore the possible thresholds of packet lifetime and buffer size there. Note also that the closed-form throughput capacity results presented in this paper only hold for MANETs with homogeneous node mobility patterns, so our another future research direction is to extend the current theoretical models and closed-form results to other more general scenarios with heterogeneous mobility patterns.

References

[1] J. Andrews, S. Shakkottai, R. Heath, N. Jindal, M. Haenggi, R. Berry, D. Guo, M. Neely, S. Weber, S. Jafar, and A. Yener, "Rethinking information theory for mobile ad hoc networks," *IEEE Communications Magazine*, vol. 46, no. 12, pp. 94–101, December 2008.

[2] A. Goldsmith, M. Effros, R. Koetter, M. Medard, A. Ozdaglar, and L. Zheng, "Beyond shannon: The quest for fundamental performance limits of wireless ad hoc networks," *IEEE Communications Magazine*, vol. 49, no. 5, pp. 195–205, May 2011.

[3] L. X. Cai, X. S. Shen, J. W. Mark, and L. Cai, "Capacity analysis and mac enhancement for uwb broadband wireless access networks," *Elsevier Computer Networks*, vol. 51, no. 11, pp. 3265–3277, August 2007.

[4] J. Zhang, J. Chen, and Y. Sun, "Transmission power adjustment of wireless sensor networks using fuzzy control algorithm," *Wireless Communication and Mobile Computing*, vol. 9, no. 6, pp. 805–818, June 2009.

[5] Y. Zhuang, J. Pan, and L. Cai, “Minimizing energy consumption with probabilistic distance distributions in wireless sensor networks,” in *INFOCOM*, 2010.

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

[7] 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.

[8] ——, “Throughput-delay trade-off in wireless networks,” in *INFOCOM*, 2004.

[9] 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.

[10] R. M. de Moraes, H. R. Sadjadpour, and J. Garcia-Luna-Aceves, “Throughput-delay analysis of mobile ad-hoc networks with a multi-copy relaying strategy,” in *SECON*, 2004.

[11] D. M. Shila and Y. Cheng, “Ad hoc wireless networks meet the infrastructure: Mobility, capacity and delay,” in *INFOCOM*, 2012.

[12] 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.

[13] 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.

[14] 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.

[15] 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/

[16] R. Urgaonkar and M. J. Neely, “Network capacity region and minimum energy function for a delay-tolerant mobile ad hoc network,” *IEEE/ACM Transactions on Networking*, vol. 19, no. 4, pp. 1137–1150, August 2011.

[17] J. Gao, J. Liu, X. Jiang, O. Takahashi, and N. Shiratori, “Exact capacity study for a class of manets,” in *ICCC*, 2012.

[18] 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.

[19] 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.

[20] Y. Chen, J. Liu, X. Jiang, and O. Takahashi, “Throughput analysis in mobile ad hoc networks with directional antennas,” ad hoc networks, 2012, to be published.

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

[22] M. Li and Y. Liu, “Rendered path: Range-free localization in anisotropic sensor networks with holes,” in *MobiCom*, 2007.

[23] Z. Zhu, A. M.-C. So, and Y. Ye, “Universal rigidity: Towards accurate and efficient localization of wireless networks,” in *INFOCOM*, 2010.

[24] 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.

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

[26] Y. Wang, X. Chu, X. Wang, and Y. Cheng, “Optimal multicast capacity and delay tradeoffs in manets: A global perspective,” in *INFOCOM*, 2011.

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

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

[29] P. Gupta and P. R. Kumar, “Critical power for asymptotic connectivity in wireless networks,” Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming, W. M. McEneaney, G. Yin, and Q. Zhang, pp. 547–566, 1998.

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

[31] 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.

[32] 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.

[33] C. Zhang, Y. Fang, and X. Zhu, “Throughput-delay tradeoffs in large-scale manets with network coding,” in *INFOCOM*, 2010.

[34] 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.

[35] 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.

[36] Z. Haas and T. Small, “A new networking model for biological applications of ad hoc sensor networks,” *IEEE/ACM Transactions on Networking*, vol. 14, no. 1, pp. 27–40, February 2006.

[37] 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.

[38] ——, “On throughput capacity of manets with power control and packet redundancy,” Graduate school of information sciences, Tohoku university, 2013, technical report 201302.01. [Online]. Available: http://distplat.blogspot.com

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

[40] J. Liu, X. Jiang, H. Nishiyama, and N. Kato, “Throughput capacity of the group-based two-hop relay algorithm in manets,” in *GLOBECOM*, 2012.

[41] C++ simulator for the 2hr-f manet with power control. [Online]. Available: http://distplat.blogspot.com

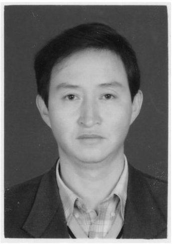
[42] The network simulator ns-2. [Online]. Available: http://www.isi.edu/nsnam/ns/

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

[44] A. Karnik, A. Iyer, and C. Rosenberg, “Throughput-optimal configuration of fixed wireless networks,” *IEEE/ACM Transactions on Networking*, vol. 16, no. 5, pp. 1161–1174, October 2008.

[45] A. Behzad and I. Rubin, “High transmission power increases the capacity of ad hoc wireless networks,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 1, pp. 156–165, January 2006.

Jiajia Liu (S’11-M’12) 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, and received his Ph.D. degree in information sciences from Tohoku University in 2012. He is currently a JSPS special research fellow in Tohoku University. He was a recipient of the Chinese Government Award for Outstanding Ph.D. Students Abroad in 2011, the Tohoku University RIEC Student Award and the Tohoku University Professor Genkuro Fujino Award in 2012. He also received the Yasujiro Niwa Outstanding Paper Award in 2012, the Best Paper Award of IEEE WCNC 2012, the Best Student Paper Award of CHINACOM 2012, and IEEE STG Award of IEEE GLOBECOM 2011 and ICC 2011. He was also the awardee of the prestigious Dean Award and President Award of Tohoku University. His research interests include performance modeling and evaluation, scaling laws of wireless networks, stochastic network optimization, and optimal control.



Xiaohong Jiang is currently a full professor of Future University Hakodate, Japan. Before joining Future University, Dr. Jiang was an Associate professor, Tohoku University, from Feb. 2005 to Mar. 2010. Dr. Jiang's research interests include computer communications networks, mainly wireless networks and optical networks, interconnection networks for massive parallel computing systems, routers/switches design for high performance networks, network coding for wireless networks, VoIP over wireless networks, network security, VLSI/WSI systems, etc. He has published over 200 technical papers at premium international journals and conferences, which include over 30 papers published in top IEEE journals and top IEEE conferences, like IEEE/ACM Transactions on Networking, IEEE Journal on Selected Areas on Communications, IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Communications, IEEE INFOCOM. Dr. Jiang was the winner of the Best Paper Award and Outstanding Paper Award of IEEE WCNC 2012, IEEE WCNC 2008, IEEE ICC 2005-Optical Networking Symposium, and IEEE/IEICE HPSR 2002. He is a Senior Member of IEEE, a Member of IEICE.

Hiroki Nishiyama (M'08) 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 prestigious Japan Society for the Promotion of Science (JSPS) until the completion of his PhD, following which he went on to become an Assistant Professor at the Graduate School of Information Sciences (GSIS) at Tohoku University. He was promoted to his current position of an Associate Professor at GSIS in 2012, when he was just 29 years old. He was acclaimed with the Best Paper Awards in many international conferences including IEEE's flagship events, namely the IEEE Wireless Communications and Networking Conference in 2012 (WCNC'12) and the IEEE Global Communications Conference in 2010 (GLOBECOM'10). He is a young yet already prominent researcher in his field as evident from his valuable contributions in terms of many quality publications in prestigious IEEE journals and conferences. He was also a recipient of the IEICE Communications Society Academic Encouragement Award 2011 and the 2009 FUNAI Foundation's Research Incentive Award for Information Technology. He received the Best Student Award and Excellent Research Award from Tohoku University for his phenomenal performance during the undergraduate and master course study, respectively. His research covers a wide range of areas including traffic engineering, congestion control, satellite communications, ad hoc and sensor networks, and network security. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Nei Kato (A'03-M'04-SM'05-F'13) received his Bachelor Degree from Polytechnic University, Japan in 1986, M.S. and Ph.D. Degrees in information engineering from Tohoku University, Japan, in 1988 and 1991, respectively. He joined Computer Center of Tohoku University at 1991, and has been a full professor with the Graduate School of Information Sciences since 2003. He has been engaged in research on satellite communications, computer networking, wireless mobile communications, smart grid, image processing, and pattern recognition. He has published more than 300 papers in peer-reviewed journals and conference proceedings. He currently serves as the Vice Chair of IEEE Ad Hoc & Sensor Networks Technical Committee, the Chair of IEEE ComSoc Sendai Chapter, the steering committee member of WCNC and voting member of GITC, an editor of IEEE Wireless Communications (2006-), IEEE Network Magazine (2012-), IEEE Transactions on Wireless Communications (2008-), IEEE Transactions on Vehicular Technology (2010-), IEEE Trans. on Parallel and Distributed Systems. He has served as the Chair of IEEE Satellite and Space Communications Technical Committee (2010-2012), a co-guest-editor of several Special Issues of IEEE Wireless Communications Magazine, a symposium co-chair of GLOBECOM'07, ICC'10, ICC'11, ICC'12, Vice Chair of IEEE WCNC'10, WCNC'11, ChinaCom'08, ChinaCom'09, Symposia co-chair of GLOBECOM'12, ICC'14, and workshop co-chair of VTC2010. His awards include Minoru Ishida Foundation Research Encouragement Prize (2003), Distinguished Contributions to Satellite Communications Award from the IEEE Communications Society, Satellite and Space Communications Technical Committee (2005), the FUNAI information Science Award (2007), the TELCOM System Technology Award from Foundation for Electrical Communications Diffusion (2008), the IEICE Network System Research Award (2009), the IEICE Satellite Communications Research Award (2011), the KDDI Foundation Excellent Research Award (2012), IEICE Communications Society Distinguished Service Award (2012), IEEE GLOBECOM Best Paper Award (twice), IEEE WCNC Best Paper Award, and IEICE Communications Society Best Paper Award (2012). Besides his academic activities, he also serves on the expert committee of Telecommunications Council, Ministry of Internal Affairs and Communications, and as the chairperson of ITU-R SG4 and SG7, Japan. Nei Kato is a Distinguished Lecturer of IEEE Communications Society (2012-2013) and the co-PI of A3 Foresight Program (2011-2014) funded by Japan Society for the Promotion of Sciences(JSPS), NSFC of China, and NRF of Korea. He is a fellow of IEEE and IEICE.