Hop Count Distribution of Multihop Paths in Wireless Networks With Arbitrary Node Density: Modeling and Its Applications

Jia-Chun Kuo, *Member, IEEE*, and Wanjiun Liao, *Member, IEEE*

***Abstract*—The significance of hop counts on the throughput and delay performance in multihop wireless networks has been well demonstrated in the literature. However, so far, there has been very little analytical work on determining the expected hop count for packet forwarding in multihop wireless networks. In this paper, we develop an analytical framework for the hop count distribution in a multihop wireless network with an arbitrary node density. We derive the average progress per hop and obtain the path connectivity probability in a network. Together with the derived per-hop progress and the path connectivity probability, we can express the probability distribution for the expected hop count in multihop wireless networks. We also demonstrate that, based on our analytical result, many network design guidelines can be provided. Specifically, the average packet delivery ratio in the network under a hop count limitation can be estimated accurately, and the tradeoff between the flooding cost and the search latency for target location discovery that is commonly used in many *ad hoc* routing protocols can also be evaluated.**

***Index Terms*—Hop count distribution, multihop wireless networks, path connectivity.**

I. Introduction

Wireless multihop networks have received much attention in recent years. In such a network, packets are transmitted via multiple hops to reach their destinations. The multihop path for each flow is determined by the routing protocol. For each node in most of the existing multihop routing protocols, the neighbor node with the shortest distance to the destination is selected as the relaying node for the next hop. Such selection results in a longer per-hop progress and less hop counts and, thus, smaller end-to-end delay for packet delivery.

The impact of hop counts on the network performance has been widely studied in the literature. In [1], Li *et al.* show by simulations that the end-to-end throughput is degraded inversely with the number of required hop counts for the path due to mutual interferences between adjacent links. In [2], the significance of hop counts on the network capacity is analytically demonstrated. The impact of hop counts on the tradeoff between throughput and end-to-end delay in multihop wireless networks is studied in [3]. Hop count also affects the target searching cost and latency in most existing *ad hoc* routing protocols [4]. When the hop count distribution is known *a priori*, the optimal flooding strategy can be determined accordingly [5]. Although the significance of hop count distribution of multihop paths is well recognized, there is no analytical work on determining the required hop count for packet transmissions in multihop wireless networks. Such an analytical study is important because it can provide many useful insights in the network performance and design guidelines for multihop wireless networks [6].

The hop count distribution of multihop paths is jointly determined by many factors, such as the hop progress, the node density, the routing mechanism, and other network parameters. In this paper, we attempt to develop the hop count distribution for multihop wireless networks with an arbitrary node density. The derivation can be decomposed into two parts: to obtain the average per-hop progress and to calculate the path connectivity. The expected per-hop progress in multihop wireless networks has been discussed in the literature under certain assumptions [7]–[9], but the path connectivity probability has never been developed. In [10], the behavior of packet forwarding via a multihop path in mobile *ad hoc* networks under a high node density is modeled as the ripples generated by a stone being dropped into a lake. The results, however, are based on the assumptions of high node density and the progress per hop being equal to the transmission range and, thus, cannot be applied to other scenarios.

In this paper, we develop the probability distribution of expected hop counts for packet transmissions in multihop wireless networks with an arbitrary node density. To simplify the problem, we assume that 1) for each node, the neighbor node with the shortest distance to the destination is selected as the relaying node for the next hop, and 2) each node is statistically identical. With a given node density, we calculate the average per-hop progress and the path connectivity probability. Then, we derive the probability distribution of the expected number of hops for each packet to traverse from the source to the destination. We demonstrate that based on the analytical results, the flooding cost and search latency in target location discovery can be evaluated. The tradeoff between cost and latency for different flooding schemes is discussed, based on which guidelines are provided for determining the most appropriate flooding scheme to use when the node density is given.

Manuscript received November 22, 2005; revised July 29, 2006 and October 13, 2006. This work was supported by the National Science Council, Taiwan, R.O.C., under Center Excellence Grant NSC95-2752-E-002-006-PAE and Grant NSC95-2221-E-002-066. The review of this paper was coordinated by Dr. J. Misic.

The authors are with the Department of Electrical Engineering and the Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan, R.O.C. (e-mail: wjliao@ntu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2007.897663

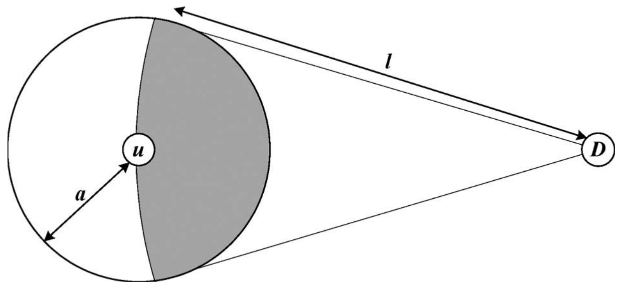

Fig. 1. Effective relaying area of node u to the destination D.

The rest of the paper is organized as follows. In Section II, the analytical model is derived and is verified via simulations. In Section IV, the applications based on the analytical results are demonstrated. They include the estimation of packet delivery ratio with a hop count limitation and the evaluation of different flooding schemes for target location discovery. Finally, the paper is concluded in Section V.

II. Modeling Hop Count Distribution

A. System Model and Definition

In this analysis, each node is assumed to have the same transmission capability, and packets are transmitted toward the destination in a multihop manner. The neighbor node that is closest to the destination is selected as the next relaying node in the path. We consider static nodes only, and it can also be regarded as a snapshot of mobile nodes.

Let $Cov_T(u, a)$ denote the transmission coverage of node u with radius a. The set of nodes located in $Cov_T(u, a)$ is defined as the *neighbor nodes* of node u. The number of neighbor nodes for each node is referred to as the *node density* in this paper. The effective relaying area of node u to the destination D, which is denoted by $Cov_E(u, D)$, is defined as $Cov_T(u, a) \cap Cov_R(D, l)$, where l is the distance between u and D, and $Cov_R(D, l)$ is the coverage area centered at D with radius l. The set of neighbor nodes located in $Cov_E(u, D)$ is said to be the *effective neighbor nodes* of node u toward the destination D. For example, the shaded region in Fig. 1 is $Cov_E(u, D)$. Since the set of nodes located in $Cov_E(u, D)$ are all closer to the destination D than node u, they are potential candidates for the next hop in the forwarding path.

The actual hop distance traversed by each packet is referred to as the *per-hop progress*. The per-hop progress ranges from 0 to the maximum transmission range of each node. The connection from node u toward the destination D on a multihop path is disconnected[1] if there is no neighbor node located in $Cov_E(u, D)$. This may often happen when the node density is low. To ensure the path connectivity, the node density must exceed a certain threshold. Intuitively, each node needs at least two neighbor nodes to keep the path connected: One is for the previous hop, and the other is for the next hop.

The spatial distribution of neighbor nodes for each node significantly affects the link connectivity probability and the per-hop progress. If node u has no neighbor nodes in $Cov_E(u, D)$, the link is disconnected, and the progress is 0. If there is at least one node in $Cov_E(u, D)$, the link is connected, and the average progress equals the distance between nodes u and D minus the distance between the next relaying node and D.

[1]We do not consider rerouting for disconnected links.

TABLE I
Notation Used in the Analysis

A	Node's maximum transmission range
N	Node density, i.e., the number of neighbor nodes in one node's maximum transmission coverage
n'	The number of neighbor nodes in one node's feasible region
B	Effective node density, i.e., the number of effective neighbor nodes in one node's transmission range
$L_{b,m}$	Random variable representing the remaining distance to the destination after traversing m hops given that the effective node density is b, $m = 0, 1, 2, \ldots$, where $L_{b,0}$ is the initial distance to the destination.
P_R	The probability that one neighbor node falls in the effective relaying area of the considered node
$f_{L_{b,m}}(l_0)$	Probability density function of $L_{b,m}$
$F_{X_b}(x)$	Cumulative distribution function of the distance from the closest effective relaying node to the destination given that there are b effective neighbor nodes
$f_{X_b}(x)$	Probability density function of the distance from the closest effective relaying node to the destination given that b effective neighbor nodes
$D_{b,m}$	Expected progress at the mth hop given that the effective node density is b
$R_{n,m}$	Expected progress at the mth hop with node density n
P_i	The probability that the destination node can be reached at i hops
$P_{n,k}$	The probability that a k-hop path is connected, given that the node density is n

The number of hops required for a multihop path is determined by 1) the distance between the source and the destination nodes and 2) the per-hop progress. The longer the distance (or the smaller the per-hop progress), the larger the number of hops that each packet must traverse. The path connectivity should also be taken into account. We will develop the probability distribution of hop counts for multihop paths in the rest of this section. For ease of derivation, we assume that the distance between the source and the destination nodes is uniformly distributed. The notations used in the analysis are summarized in Table I.

B. Per-Hop Progress

When the destination node is located within the transmission range of the transmitting node, packets can be transmitted directly. This happens at the last hop in the path, and the hop length is equal to the distance between the transmitting node and the destination. Hence, the average progress for this hop is given by

$$\int_0^a y \cdot \frac{1}{a} \cdot dy = \frac{a}{2} \tag{1}$$

where a is the node's maximum transmission range.

When the destination node cannot be reached directly, each transmitting node in the path always selects the neighbor node that is closest to the destination as the next relaying node. Denote the expected progress at the mth hop under node density n by $R_{n,m}$. To derive the expected progress at each hop, we need to take the effective node density into consideration. Here, the effective node density means the number of neighbor nodes that can be the candidates for the relaying nodes in the forwarding path. In most cases, the effective node density is not equal to the node density because some of the neighbor nodes will not be chosen for relaying in the forwarding path. Thus, we start the derivation with a given effective node density, and then extend the analytical result to the case with a given node density.

1) Per-Hop Progress With a Given Effective Node Density: We first assume that all neighbor nodes can be the candidates for the relaying node in the forwarding path. Let $L_{b,m}$ denote the remaining distance to the destination node after m hops given that the number of effective neighbor nodes is b. Thus, $L_{b,0}$ is the initial (remaining) distance to the destination node. Since the initial distance is independent of the number of effective neighbor nodes, for convenience, we let $L_0 = L_{b,0}$, $b = 0, 1, 2, \cdots$, in the following discussion.

Fig. 2 gives three arcs centered at the destination node D with a radius of l_0, x, and $l_0 - a$, respectively, where l_0 is the initial distance between nodes S and D. These arcs represent all possible locations at which the next relaying node may be located, and the radius of each arc corresponds to the remaining distance to node D. For example, the arc with a radius of x represents those locations that the remaining distance from the next relaying node N_1 to the destination node D is x, where $l_0 \geq x \geq l_0 - a$. The probability that the next relaying node is located at a specific arc is proportional to the length of the arc, which is given by the product of the radius and the angle of the arc.

Let $F_{X_1}(x)$ denote the conditional cumulative distribution function for a random variable $L_{1,1}$, given that the initial distance $L_0 = l_0$. $L_{1,1}$ is defined as the remaining distance to the destination node, given that the effective node density is one. Hence

$$F_{X_1}(x) = \frac{K(S,a) \cap Cov_S(\theta_x, x)}{Cov_E(S,D)} = \frac{\int_0^{\theta_x} \int_{l_0-a}^{x} r \cdot \theta \cdot dr d\theta}{\int_0^{\theta} \int_{l_0-a}^{l_0} r' \cdot \theta' \cdot dr' d\theta'} = \frac{(x - l_0 + a)^2 \left[x^2 - (l_0 - a)^2\right]}{a^2 \left(2 l_0 a - a^2\right)}. \quad (2)$$

Here, we approximate the area coverage of $K(S,a) \cap Cov_S(\theta_x, x)$ and $Cov_E(S,G)$ by $\int_0^{\theta_x} \int_{l_0-a}^{x} r \cdot \theta \cdot dr d\theta$ and $\int_0^{\theta} \int_{l_0-a}^{l_0} r' \cdot \theta' \cdot dr' d\theta'$, respectively, where θ_x is the angle formed by the arc with radius x, and θ_x is approximated by $\sqrt{(x-(l_0-a))/a} \cdot \theta$; $Cov_S(\theta_x, x)$ is the sector formed by

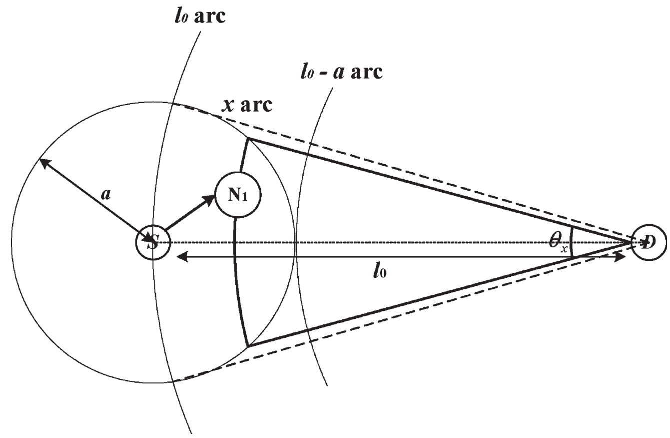

Fig. 2. Per-hop progress.

θ_x. Thus, the conditional probability density function $f_{X_1}(x)$ of $L_{1,1}$, given the initial distance $L_0 = l_0$, can be expressed by

$$f_{X_1}(x) = \frac{d}{dx} F_{X_1}(x) = \frac{2(a - l_0 + x)^2}{a^2 \left(-a^2 + 2 l_0 a\right)} + \frac{2(a - l_0 + x)\left(-(-a + l_0)^2 + x^2\right)}{a^2 \left(-a^2 + 2 l_0 a\right)}. \quad (3)$$

Taking the expectation of $L_{1,1}$, we have

$$E[L_{1,1}] = \int_a^{l_{\max}} E[L_{1,1} | L_0 = l_0] \cdot f_{L_0}(l_0) \cdot dl_0 \quad (4)$$

where

$$E[L_{1,1} | L_0 = l_0] = \int_{l_0-a}^{l_0} x \cdot \left(\frac{2(a - l_0 + x)^2}{a^2 \left(-a^2 + 2 l_0 a\right)} + \frac{2(a - l_0 + x)\left(-(-a + l_0)^2 + x^2\right)}{a^2 \left(-a^2 + 2 l_0 a\right)} \right) \cdot dx$$

and $f_{L_0}(l_0)$ is the probability density function of L_0, and the value taken by L_0 ranges from a to $l_{\max}$ (i.e., $l_{\max}$ is the maximum distance between source and destination nodes).

Accordingly, the average progress at the first hop under the condition that there is only one effective neighbor node can be expressed by

$$D_{1,1} = E[L_0] - E[L_{1,1}].$$

To derive $E[L_{1,2}]$, which can be derived from $L_{1,1}$, the range of the remaining distance should be adjusted accordingly because the remaining distance to the destination node is decreased by $R_{n,1}$ on average, where $R_{n,1}$ is the expected progress at the first hop and will be derived in Section II-B2b. Generally, we can express the ranges of $L_{1,m}$ and derive the values of $D_{1,m}$ after m hops, $m = 1, 2, 3, \cdots$.

Next, we consider a general case for the expected progress at the mth hop, given that the number of effective neighbor nodes is b. Recall that all nodes are distributed independently and identically, and the neighbor node closest to the destination is chosen as the next relaying node. Thus, the cumulative probability distribution function $F_{X_b}(x)$ of $L_{b,m}$ is given by

$$\begin{aligned} F_{X_b}(x) &= \Pr\{d(N_1) \le x \cup d(N_2) \le x \cup \cdots \cup d(N_b) \le x\} \\ &= C_1^b \cdot \Pr\{d(N_1) \le x\} - C_2^b \\ &\quad \cdot \Pr\{d(N_1) \le x \cap d(N_2) \le x\} + \cdots \\ &\quad + (-1)^{b-1} C_b^b \cdot \Pr\{d(N_1) \le x \cap d(N_2) \\ &\qquad \le x \cap \cdots \cap d(N_b) \le x\} \\ &= 1 - \left[1 - \frac{(x - l_{m-1} + a)^2 \left(x^2 - (l_{m-1} - a)^2\right)}{a^2 \left(2 l_{m-1} a - a^2\right)}\right]^b \end{aligned} \tag{5}$$

where $N_1 \cdots N_b$ represent the b effective neighbor nodes, $d(N_i)$ represents the remaining distance from the effective neighbor node i to the destination node, $C_i^j = j!/i!(j-i)!$, and l_{m-1} is the remaining distance to the destination node at the mth hop.

Accordingly, we have $f_{X_b}(x) = (d/dx)F_{X_b}(x)$, and

$$\begin{aligned} E[L_{b,m}] &= \int E[L_{b,m} | L_{b,m-1} = l_{m-1}] \cdot f_{L_{b,m-1}}(l_{m-1}) \cdot dl_{m-1} \\ &= \int_a^{l_{\max,m-1}} \int_{l_{m-1}-a}^{l_{m-1}} x \cdot f_{X_b}(x) \cdot dx \cdot f_{L_{b,m-1}}(l_{m-1}) \cdot dl_{m-1} \end{aligned} \tag{6}$$

where $l_{\max,m-1}$ is the maximum value of $L_{b,m-1}$.

Finally, we obtain

$$D_{b,m} = E[L_{b,m-1}] - E[L_{b,m}]. \tag{7}$$

2) Impact of Node Density on the Per-Hop Progress: The expected per-hop progress $D_{b,m}$ in Section II-B1 is derived under the assumption that each node has b effective neighbor nodes in a multihop path. We now generalize to the case that the node density is given (i.e., not all neighbor nodes can be candidates for the relaying nodes in the forwarding path) and examine the impact of the node density on the expected per-hop progress.

a) Feasible region for packet forwarding: Consider a node, e.g., X, on the routing path in a network with node density n. Not all neighbor nodes of node X can be the candidates for the relaying node to the previous hop or to the next hop. Let $Cov_{NF}(X,a)$ denote the area within $Cov_T(X,a)$ that the set of neighbor nodes for node X cannot be the candidates for the relaying nodes. $Cov_{NF}(X,a)$ is referred to as the nonforwarding area, and the area $Cov_T(X,a) - Cov_{NF}(X,a)$, the feasible region for node X. For example, the shaded areas in Fig. 3 are the nonforwarding areas for node N_2.

In what follows, we attempt to determine the average number (or the percentage) of neighbor nodes that are potentially involved (i.e., being the candidates for the relaying node) in packet forwarding. We focus only on the case with $n > 2$, since

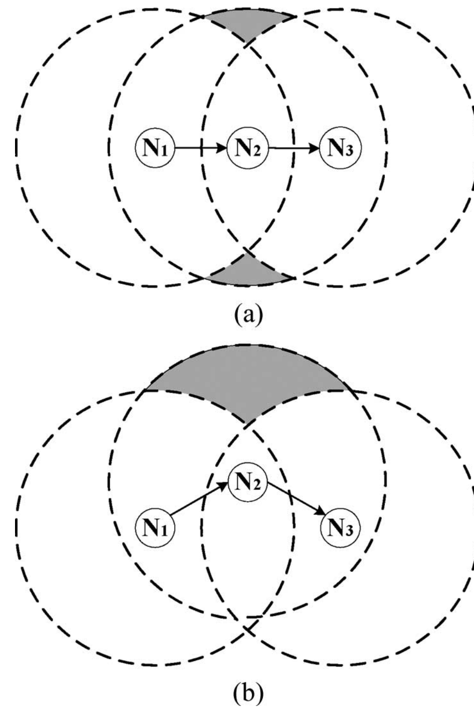

Fig. 3. Two examples for the nonforwarding region of node N_2.

when n equals 2, all neighbors of each relaying node must be involved in the packet forwarding to maintain the connectivity (i.e., one is for the previous hop, and the other is for the next hop).

The coverage area of a nonforwarding region determines the number of neighbor nodes that are not involved in packet transmissions. From [11], this average coverage area can be approximated by $0.19\pi a^2$, and the intersection area of two adjacent nodes X and Y can be approximated by $Cov_T(X,a) \cap Cov_T(Y,a) \approx 0.59\pi a^2$. Thus, we obtain

$$Cov_T(Y,a) - (Cov_T(X,a) \cap Cov_T(Y,a)) \approx 0.41\pi a^2.$$

For a 2-hop connected path, e.g., $N_1 \to N_2 \to N_3$, as shown in Fig. 3, we have

$$\begin{aligned} &Cov_T(N_2,a) - Cov_T(N_1,a) \\ &\qquad \cap Cov_T(N_2,a) - Cov_{NF}(N_2,a) \approx 0.22\pi a^2 \\ &Cov_T(N_2,a) - Cov_T(N_3,a) \\ &\qquad \cap Cov_T(N_2,a) - Cov_{NF}(N_2,a) \approx 0.22\pi a^2 \end{aligned}$$

thus, leading to

$$Cov_T(N_1,a) \cap Cov_T(N_2,a) \cap Cov_T(N_3,a) \approx 0.37\pi a^2.$$

The intersection area of N_1, N_2, and N_3 in a connected 2-hop path can be approximately two times the excluded region of the relaying node N_2. The excluded region of a relaying node (e.g., N_2) refers to the region that is inside the transmission range of the relaying node and in which no neighbor nodes exist. For example, Fig. 4(a) gives the excluded region of node N_2 for the 2-hop connected path shown in Fig. 3(a). If there were a node in this excluded region, N_2 would not be

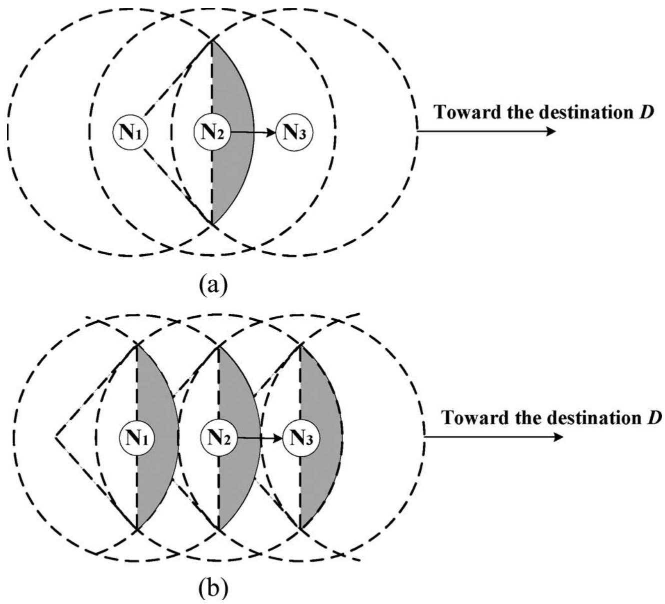

Fig. 4. Excluded region based on the example shown in Fig. 3(a). (a) Excluded region of N_2. (b) Excluded regions of N_1, N_2, and N_3 in $Cov_T(N_2, a)$.

selected as the relaying node by N_1 according to the selection strategy for the next relaying node. Since each relaying node in a multihop path has an excluded region, there must be three excluded regions in the transmission range of each node (i.e., for the node itself, plus the two nodes at the previous and the next hops), as shown in Fig. 4(b), for example. Thus, about 55.5%[2] of the transmission range per node is in the excluded region, and all neighbor nodes can only be located in 44.5% of a node's transmission range.

Of this region (i.e., $0.445\pi a^2$), where neighbor nodes can be located, the nonforwarding region takes $0.19\pi a^2$, which is about half (i.e., $0.19\pi a^2/0.445\pi a^2 = 42.7\%$) of this region for each node. In fact, the ratio 42.7% may be slightly underestimated because when both nodes N_1 and N_3 are not located in each other's transmission range,[3] the nonforwarding region of node N_2 in Fig. 3 should be larger than $0.19\pi a^2$, as mentioned in [12]. For simplicity, we assume that about one half of neighbor nodes are located in the nonforwarding region, and they are not involved in the multihop path. Later, we will show via simulation that such approximation is acceptable.

According to the discussion above, for a connected multihop path, each relaying node must have at least two neighbor nodes to be involved in packet forwarding. Except for the two neighbor nodes that maintain the path connectivity, each node has only about half of the neighbor nodes located in the feasible region. Thus, we conclude that for each relaying node in a network with node density n, the number of neighbor nodes located in the feasible region n' is given by

$$n' = \begin{cases} n, & \text{if } n = 2 \\ 2 + \frac{n-2}{2}, & \text{if } n > 2 \end{cases}. \tag{8}$$

[2]Each excluded region is about $0.37/2 = 0.185 = 18.5\%$ of a transmission range. Thus, the total area coverage of the three excluded regions is 55.5%.

[3]On average, it makes N_3 become farther away from N_2 and increases the additional coverage area by N_2's transmissions.

b) Per-hop progress with a given node density: Given a node density, the number of neighbor nodes for each relaying node falling in its feasible region is given by (8). Based on this result, we further derive the expected per-hop progress under a given node density. We start the derivations with a simple case with $n' = 3$, as shown in Fig. 5. In both subfigures of Fig. 5, node S is the source, and nodes N_1, N_2, and N_3 are the relaying nodes for the first three hops. For source node S, at least one of its neighbor nodes [e.g., N_1 in Fig. 5(b)] must be located in its effective relaying area [i.e., $Cov_E(S, D)$] to maintain the path connectivity, or the link is broken. The other two neighbor nodes can be distributed in either way, as shown in Fig. 5. Fig. 5(a) shows that one of these two nodes (i.e., N_4) is located in $Cov_E(S, D)$ but is not chosen as the next relaying node, and the other node (i.e., N_5) is located on the other side of the destination node. In this case, node N_1 has only one choice for the next hop (i.e., N_2) because n' is fixed (i.e., $n' = 3$), and the two neighbor nodes of N_1 (i.e., S and N_4) are already located on the side opposite to the destination node. Similarly, N_2 (N_3) has two (one) neighbor nodes to choose as the next relaying node. We can conclude that the distribution of the neighbor nodes around the source node S determines the distributions of neighbor nodes at the successive hops in a connected multihop path, and each distribution has an occurrence probability. For example, the occurrence probability of the case in Fig. 5(a) is given by the probability that either N_4 or N_5 is located within $Cov_E(S, D)$, which is $C_1^2 \cdot (P_R)^1(1 - P_R)^1$, where P_R is the probability for any neighbor nodes of node S falling within $Cov_E(S, D)$, and

$$P_R = \frac{Cov_E(S, D)}{Cov_T(S, a)}.$$

The second case [Fig. 5(b)] states that the two neighbor nodes N_4 and N_5 are not located within $Cov_E(S, D)$. Thus, the occurrence probability is given by $C_0^2 \cdot (P_R)^0(1 - P_R)^2$. From the occurrence probability of each case in a connected multihop path in Fig. 5 and the given node density, the expected per-hop progress can be calculated accordingly. Let $R_{n,m,}$ denote the expected progress at the mth hop in a connected multihop path with node density n. The progresses at the first two hops are then given by

$$\begin{aligned} R_{3,1} = {} & C_1^2 \cdot (P_R)^1(1 - P_R)^1 \cdot D_{2,1} + C_0^2 \\ & \cdot (P_R)^0(1 - P_R)^2 \cdot D_{1,1} \\ R_{3,2} = {} & C_1^2 \cdot (P_R)^1(1 - P_R)^1 \cdot D_{1,2} + C_0^2 \\ & \cdot (P_R)^0(1 - P_R)^2 \cdot D_{2,2}. \end{aligned}$$

Note that for a connected multihop path, it is unlikely that all neighbor nodes of source S are located in $Cov_E(S, D)$. If this were the case, the next relaying node (i.e., N_1 in this case) would not have a neighbor to choose for the next hop. This leads to a disconnected path, contradicting the requirement that the considered multihop path is connected. Thus, we conclude that in a connected multihop path and with a node density n, the

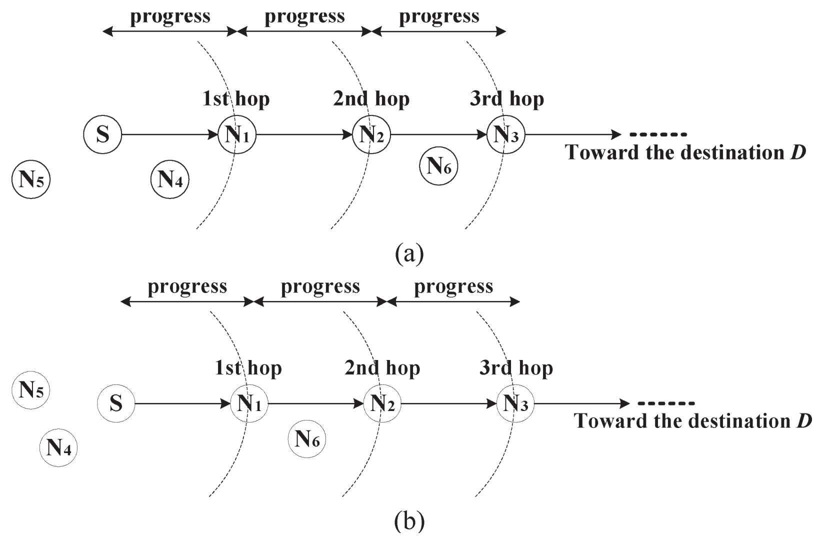

Fig. 5. Two cases of node distribution in a connected path.

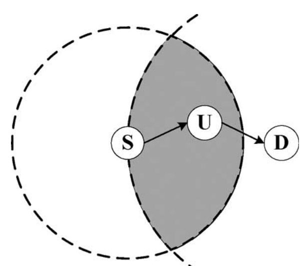

Fig. 6. Two-hop transmission path.

expected progress at the mth hop is given by (9), shown at the bottom of the page, where

$$n' = \begin{cases} n, & \text{if } n = 2 \\ 2 + \frac{n-2}{2}, & \text{if } n > 2 \end{cases}.$$

C. Connectivity Probability of a Multihop Path

The derivations so far have been based on the assumption that the multihop path is connected. Next, we derive the connectivity probability of a k-hop path for a source–destination (S–D) pair. Note that since a 1-hop path (i.e., the destination node is located in the transmission range of the source node) is connected by default, we focus only on a k-hop ($k \geq 2$) path, which is discussed in the following.

Consider a 2-hop path for an S–D pair shown in Fig. 6. Intuitively, the connectivity probability of this path is equal to the probability that at least one of node S's neighbor nodes (e.g., node U) is located within node S's effective relaying area. Since the node density is n, the occurrence probability is $1-(1-P_R)^n$. Similarly, for an intermediate node i on a multihop path, the connectivity probability for node i to its next hop is equal to the probability that one of its neighbor nodes is located within its effective relaying area. Thus, the connectivity probability of a k-hop path with node density n, i.e., $P_{n,k}$, is given by $(1-(1-P_R)^n)^{k-1}$.

D. Probability Distribution of Required Hop Counts

Finally, the probability that the destination node can be reached at a certain hop count is derived. This probability is determined by the expected per-hop progress and the connectivity probability of the multihop path. In Section II-B, the expected per-hop progress is derived under a given node density in a connected multihop path. In Section II-C, the probability that a multihop path is connected is also obtained. From both results, we can obtain the probability that the destination node can be reached at a certain hop count.

Let P_i denote the probability that the destination node is reached at the ith hop. Given the initial distance L_0 and the node density, the probability that the destination node can be reached at the first hop is equal to the probability that the initial distance L_0 is equal to or less than the transmission range a. Thus, we have $P_1 = \int_0^a f_{L_0}(l_0)dl_0$.

However, if the destination node cannot be reached directly, the probability that the destination node can be reached at the second hop is equal to the probability that the destination node is away from the source node in a distance ranging over $(a, a + R_{n,1}]$, given that the multihop path is connected. This is because

$$R_{n,m} = \begin{cases} \sum_{k=0}^{n'-2} C_k^{n'-1} \cdot (P_R)^k (1-P_R)^{n'-1-k} \cdot D_{k+1,m}, & \text{if } m = 1,3,5,7\ldots \\ \sum_{k=0}^{n'-2} C_k^{n'-1} \cdot (P_R)^k (1-P_R)^{n'-1-k} \cdot D_{n'-1-k,m}, & \text{if } m = 2,4,6,8\ldots \end{cases} \quad (9)$$

the expected progress at one hop is equal to $R_{n,1}$. Thus, $P_2 = \int_a^{a+R_{n,1}} f_{L_0}(l_0)dl_0 \cdot P_{n,2}$.

Generally, the probability that the destination node can be reached at the ith hop is equal to the probability that the destination node is away from the source node in the distance ranging over $(a + \sum_{l=0}^{i-2} R_{n,l}, a + \sum_{l=0}^{i-1} R_{n,l}]$, given that the multihop path is connected. Thus, we have

$$P_i = \begin{cases} \int_0^a f_{L_0}(l_0)dl_0, & \text{for } i = 1 \\ \int_{a+\sum_{l=0}^{i-2} R_{n,l}}^{a+\sum_{l=0}^{i-1} R_{n,l}} f_{L_0}(l_0)dl_0 \cdot P_{n,i}, & \text{for } i \geq 1 \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

III. Simulation Results

In this section, we validate the analytical model in the previous section via simulations. In the simulation, nodes are uniformly distributed in a $200\sqrt{\pi} \times 200\sqrt{\pi}$ unit square area. The maximum transmission range of each node is set to 10 units, and the initial distance L_0 between the source and destination nodes is uniformly distributed over [0, 60]. We vary the total number of nodes in the network to emulate the change in the node density. For example, the scenario with 1200 nodes distributed in the square area corresponds to the case that each node has two neighbor nodes (i.e., node density $= 2$).

A. Expected Progress in One Hop

We first measure the expected per-hop progress as a function of the number of effective neighbor nodes in the simulation. Fig. 7 shows that more effective neighbor nodes result in a longer expected progress, which fits our expectation. Fig. 7 also plots the analytical curve of $D_{b,1}$[4] and shows that the analytical result matches the simulation one very well.

B. Probability That the Destination is Reached With a Certain Hop Count

The next simulation is conducted to plot the hop count distribution. We first consider a sparse network with node density $n = 2$. Fig. 8 shows that the longer the distance (i.e., more hop counts) between the source and the destination nodes, the smaller the probability that packets can reach the destination node. This is reasonable because for a k-hop S–D pair, due to the possible link disconnection, a larger k (i.e., the longer path length) results in a smaller value of P_i. This figure plots the analytical curves of both the cumulative distribution function (cdf) and probability density function (pdf) in comparison to the simulation results. It also verifies our assumption that all neighbor nodes are involved in packet forwarding when $n = 2$.

We then simulate a denser network with $n = 4$, 6, and 8. This time, the value of n in $R_{n,m}$ is replaced with n' in (8), i.e., $n' = 2 + (n-2)/2$. Again, as shown in Fig. 9, both the simulation and analytical curves match very well, verifying the analysis in Section II. Based on these derived probability

[4]When b is fixed, the expected progress at each hop is approximately identical, i.e., $D_{b,1} \approx D_{b,2} \approx D_{b,3} \approx \ldots$. Here, we only consider the first hop with different numbers of effective neighbor nodes.

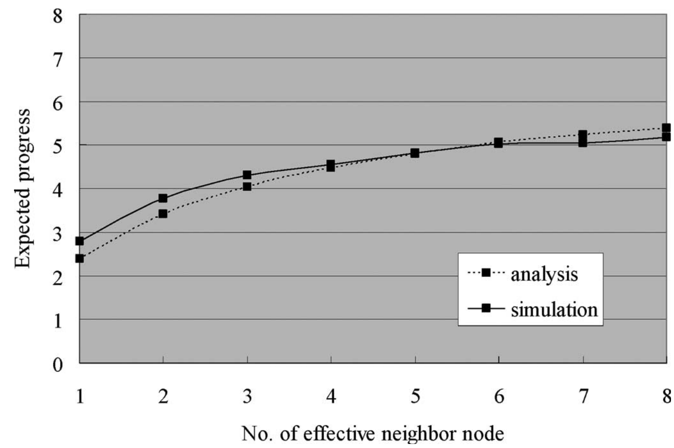

Fig. 7. Expected progress per hop.

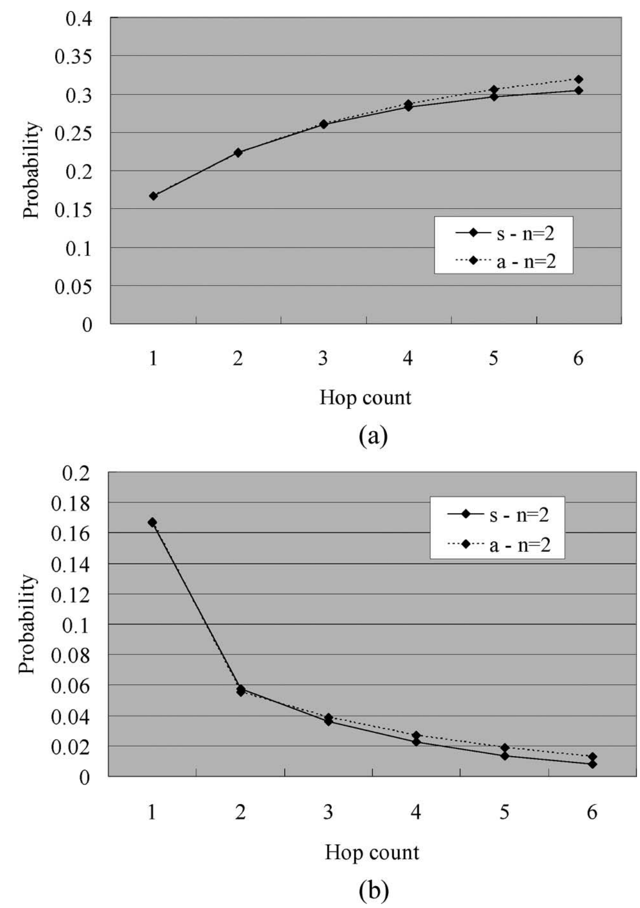

Fig. 8. Analytical and simulation results with $n = 2$. (a) cdf. (b) pdf.

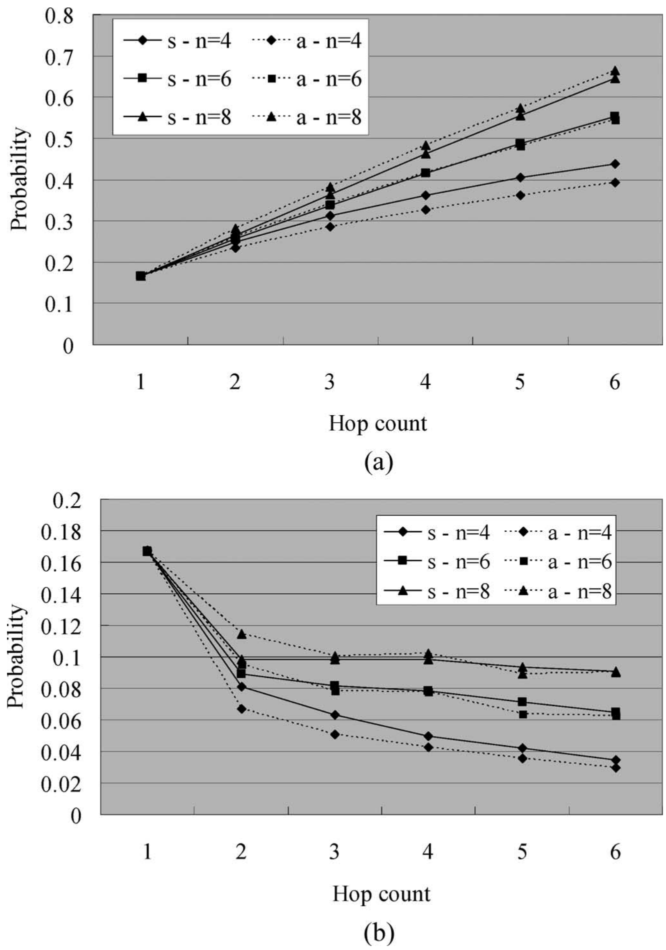

Fig. 9. Analytical and simulation results with $n = 4, 6, 8$. (a) cdf. (b) pdf.

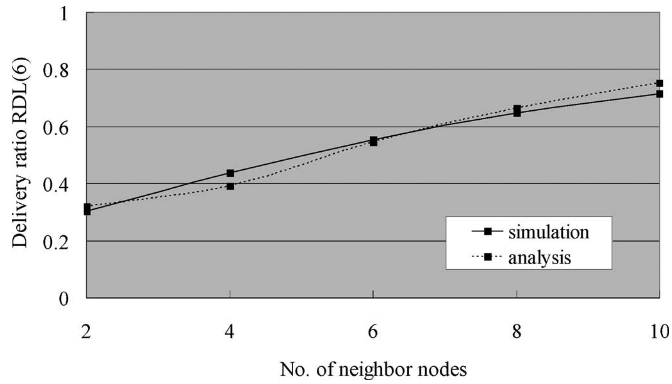

Fig. 10. Delivery ratio at six hops with different node densities.

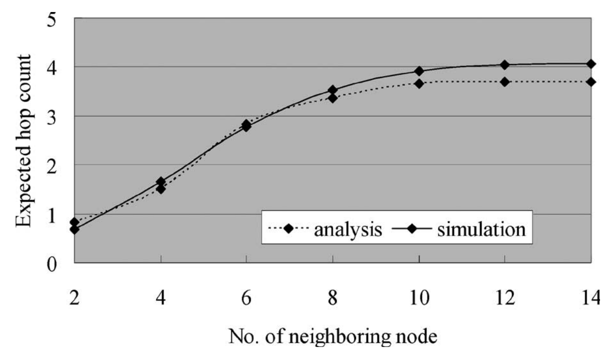

Fig. 11. Expected hop count for packet transmissions.

distributions, we can determine the optimal flooding search strategy, as mentioned in [5].

C. Delivery Ratio With Different Numbers of Neighbor Nodes

This simulation is to measure the delivery ratio, which is defined as the probability that packets can be transmitted successfully to the destination node, within a given hop count. The value of this ratio is determined jointly by many system parameters, in particular, the node density. Fig. 10 plots the delivery ratio within six hops under different node densities. It shows that the higher the node density, the larger the ratio. This is reasonable because when there are more neighbors for each node, the expected per-hop progress is larger, which, in turn, results in a larger search area covered by six hops. Therefore, the probability of the destinations being located within the 6-hop coverage area increases, and the delivery ratio also increases. In addition, a higher node density decreases the probability of link disconnection. Again, the analytical curve included in Fig. 10 is to demonstrate the applicability of this model. Let $R_{DL}(k)$ denote this ratio within k hops, and

$$R_{DL}(k) = \sum_{i=1}^{k} P_i.$$

Fig. 10 shows that the analytical delivery ratio within six hops fits the simulation result well, proving the accuracy of our analytical model. This result indicates the relationship between the node density and the delivery ratio and provides guidelines to decide what the proper node density is for the network in order to maintain the delivery ratio above a certain threshold.

D. Expected Hop Count for Packet Transmissions

Finally, we examine the expected hop count for packet transmissions in the network. As shown in Fig. 11, when the number of neighbor node is lower, the expected hop count is smaller because a shorter path is likely to be connected. This result matches the observation made in [12]. In this figure, the analytical result also fits the simulation result well.

IV. Evaluation of Different Flooding Schemes

A. Different Flooding Schemes on Target Discovery

In this section, we further show that based on our analytical model, different flooding schemes that are commonly adopted in on-demand *ad hoc* routing protocols for target location discovery can be evaluated in terms of flooding cost and search latency. According to [4], the flooding cost is referred to as the total number of copies a packet is transmitted for target discovery because packet transmission is related to energy consumption. For a flooding to k hops, which is referred to as k-hop limited, the cost equals the number of nodes whose distance is less than k hops. The nodes that are exactly k hops away are not counted because they will no longer forward the packet. The search latency is defined as $2lT$, where T is the per-hop delay, and l is the total number of hops traversed by a flooded packet to reach the destination. The latency is doubled because the source node has to wait for an acknowledgement from the destination. We assume that if no acknowledgement is heard by the source after a timeout period (i.e., the destination is unreachable), the packet will be retransmitted.

The performance of three different types of flooding schemes is evaluated.

1) *Blind flooding:* The entire network is flooded (e.g., [13]).
2) *Two-tier flooding:* The neighbor nodes within finite hops are searched first, and if the target is not found, the entire network is flooded. The searching packet in the target discovery phase of dynamic source routing [14] is an example.
3) *Expansion-ring flooding:* The source node incrementally enlarges the search range from an initial value to a predefined threshold. If the target is still not found, the entire network is flooded. The searching packet for target location in *ad hoc* on demand distance vector routing [15] is an example.

The cost and latency for each type of flooding schemes in a network with diameter h hops are defined as follows.

1) For blind flooding, since the packet is flooded to the entire network, the cost of blind flooding, which is denoted by C_h, is equal to the number of nodes located within $(h-1)$ hops, and the latency is given by $\sum_{l=1}^{h} P_l \cdot 2 \cdot l \cdot T$.
2) For two-tier flooding with an initial j-hop search, the cost is given by $(\sum_{i=1}^{j} P_i) \cdot C_j + (\sum_{l=j+1}^{h} P_l) \cdot C_h$, and the latency is $\sum_{i=1}^{j} P_i \cdot 2iT + \sum_{l=j+1}^{h} P_l \cdot (2lT + 2jT)$.
3) For expansion-ring flooding with an initial j-hop search, the cost is given by $(\sum_{i=1}^{j} P_i) \cdot C_j + \sum_{l=j+1}^{h} P_l \cdot C_l$, and the latency is $\sum_{i=1}^{j} P_i \cdot 2iT + \sum_{l=j+1}^{h} P_l \cdot (2lT + \sum_{k=1}^{l-1} 2kT)$.

We compare six flooding schemes in this paper. Scheme 1 is the blind flooding, and Schemes 2–4 are the two-tier flooding with the first searching packet set to 1-hop, 2-hop and 2-hop limited, respectively. Schemes 5 and 6 are the expansion-ring flooding with the first searching packet set to 1-hop and 2-hop limited, respectively.

Fig. 12 depicts the flooding cost [Fig. 12(a)] and the search latency [Fig. 12(b)] under different node densities for the six schemes. In this evaluation, the timeout value is set to $12T$, and T is 1 s for all schemes. It shows that more neighbor nodes lead to higher cost but lower latency for all six schemes. The reason is that when the node density increases, the number of nodes involved in packet flooding increases, and thus, the cost also increases. Moreover, an increase in the node density leads to an increase in the expected per-hop progress, and thus, the destination may be reached at fewer hops. This results in a decrease in search latency. Fig. 12 also demonstrates that with blind flooding (i.e., Scheme 1), since the packet is flooded to the entire network, it results in the highest flooding cost but the smallest latency due to the unlimited search range, irrespective of the number of neighbor nodes; with expansion-ring flooding with initial hop count $= 1$ (i.e., Scheme 5), due to more careful flooding by enlarging its search range gradually, it has the lowest flooding cost but incurs the highest latency, again, irrespective of the number of neighbor nodes.

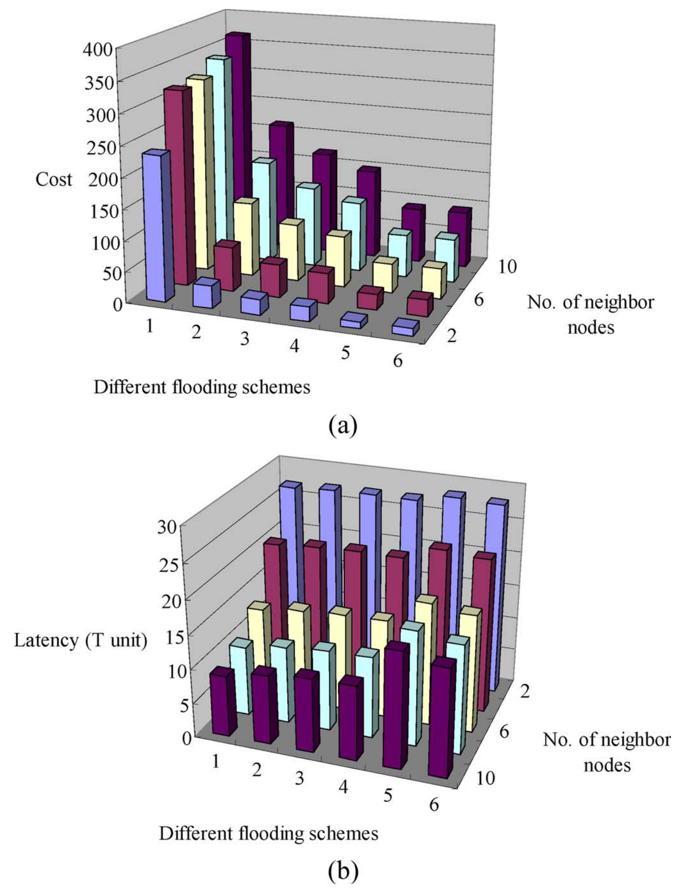

Fig. 12. Performance of different flooding schemes. (a) Cost. (b) Latency.

B. Tradeoff Between Search Cost and Latency

To better understand the tradeoff between the search cost and the latency for different types of flooding schemes, we plot the search cost and the latency for the three flooding schemes in Fig. 13: blind flooding, two-tier flooding, and expansion-ring flooding (i.e., the first searching packets for both two-tier flooding and expansion-ring flooding are 1-hop limited). The subfigures show that for each type of scheme, a higher node density results in a higher search cost and a lower latency. The search cost and latency curves intersect at a point of the best choice in terms of both search cost and latency with respect to the node density. These points indicate that when both search cost and latency performances are jointly taken into account, a network with a different node density should choose a different flooding scheme for target discovery. In this example, the point of the best choice for blind flooding is at $n = 3.5$, two-tier flooding at $n = 6$, and expansion-ring flooding at $n = 7$. This result shows that when the node density is small (i.e., n is about 3–4), blind flooding is a better choice; however, when the density is moderate (i.e., $n = 6$), two-tier flooding is preferable. When the density becomes large (i.e., $n \geq 7$), expansion-ring flooding performs best. In addition, as

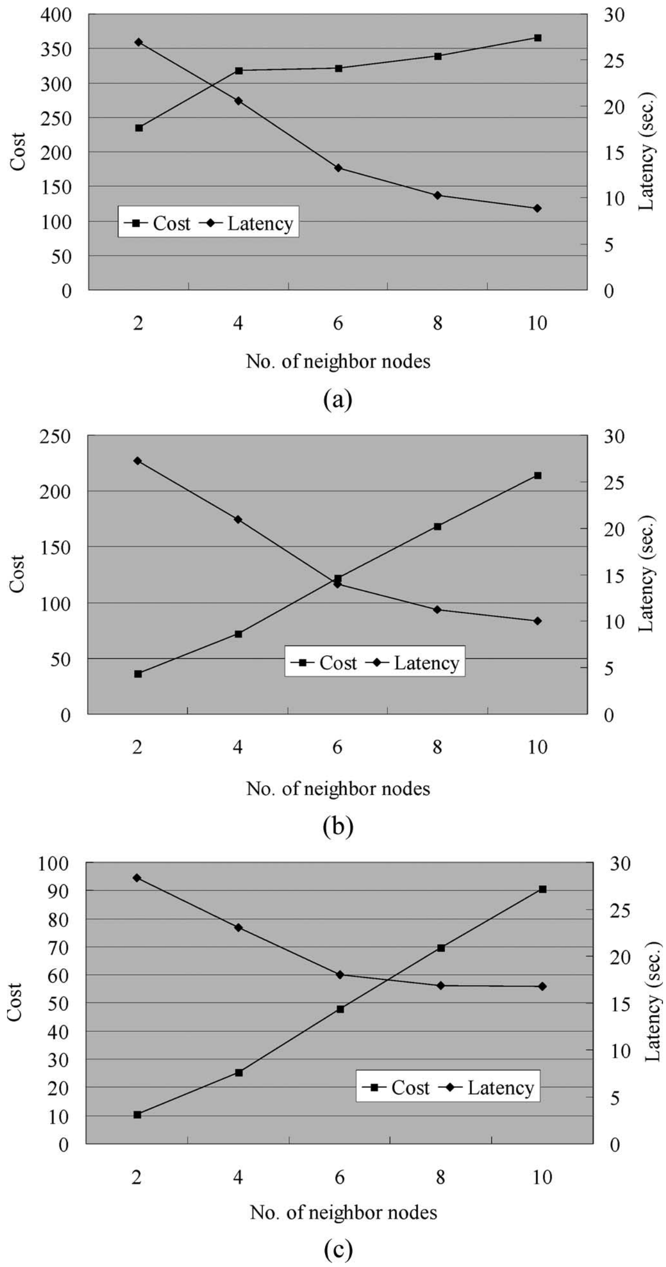

Fig. 13. Effect of different flooding schemes. (a) Blind flooding. (b) Two-tier flooding $(\text{initial hop count} = 1)$. (c) Expansion-ring flooding $(\text{initial hop count} = 1)$.

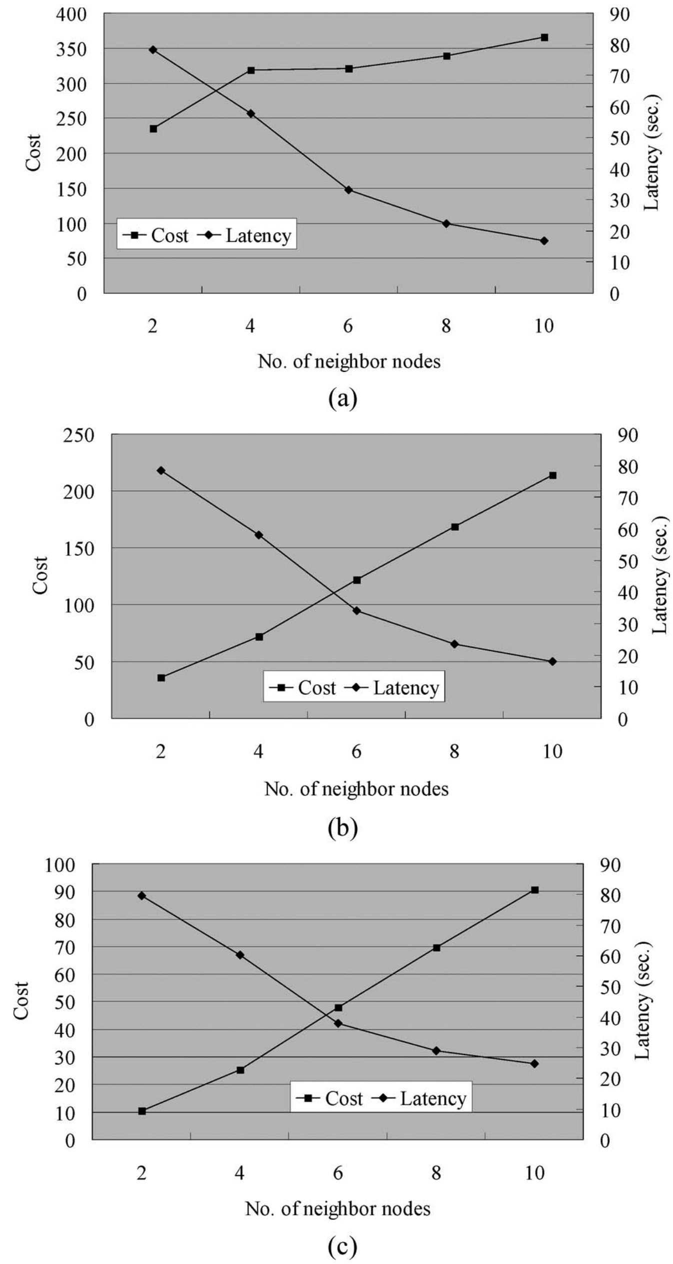

Fig. 14. Performance of different flooding schemes with timeout $= 36T$. (a) Blind flooding. (b) Two-tier flooding $(\text{initial hop count} = 1)$. (c) Expansion-ring flooding $(\text{initial hop count} = 1)$.

the first searching packets in the two-tier and expansion-ring flooding schemes are set to a different value, the point will also be changed.

We next examine the impacts of the timeout value $2lT$ and the per-node delay T on the best choice of node density for the three schemes. The results of different settings are plotted in Figs. 14 and 15. In Fig. 14, the value of T is set to 1 s as in Fig. 13, but the timeout value is now $36T$; in Fig. 15, the timeout value stays at $12T$, but the value of T now becomes 2 s. Both figures show a similar result to Fig. 13, except that, as the timeout value becomes larger (i.e., $36T$), the performance of two-tier flooding and expansion-ring flooding is very close, showing that these two parameters do not play an important role on the choice of the best flooding schemes.

V. Conclusion

In this paper, we develop an analytical model to describe the hop count distribution for each S–D pair in multihop wireless networks with an arbitrary node density. We demonstrate that based on the proposed analytical model, the delivery ratio can be estimated accurately, and the tradeoff between the flooding cost and the search latency for target location discovery that is commonly used in most existing *ad hoc* routing protocols can be evaluated. More importantly, the analytical result can provide guidelines to determine the most appropriate flooding scheme to use for any given node density and can also be used to design optimal search strategies that minimize the average search cost for target discovery in multihop wireless networks such as wireless mesh and *ad hoc* networks.

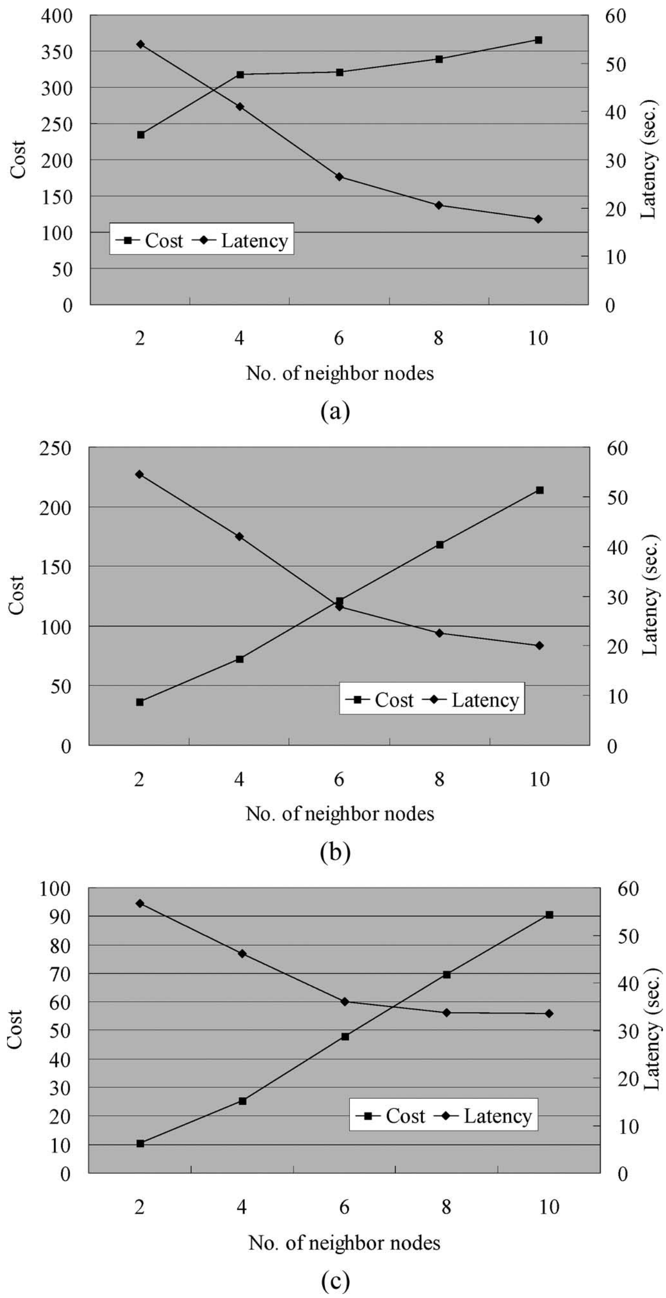

Fig. 15. Performance of different flooding schemes with $T = 2$ s. (a) Blind flooding. (b) Two-tier flooding (initial hop count $= 1$). (c) Expansion-ring flooding (initial hop count $= 1$).

REFERENCES

[1] J. Li, C. Blake, D. Couto, H. Lee, and R. Morris, "Capacity of ad hoc wireless networks," in *Proc. ACM MobiCom*, 2001, pp. 61–69.

[2] J. Jun and M. Sichitiu, "The nominal capacity of wireless mesh networks," *IEEE Wireless Commun.*, vol. 10, no. 5, pp. 8–14, Oct. 2003.

[3] A. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput–delay trade-off in wireless networks," in *Proc. IEEE Infocom*, 2004, pp. 464–475.

[4] Z. Cheng and W. B. Heinzelman, "Flooding strategy for target discovery in wireless networks," in *Proc. ACM MSWiM*, Sep. 2003, pp. 33–41.

[5] N. Chang and M. Liu, "Revisiting the TTL-based controlled flooding search: Optimality and randomization," in *Proc. ACM MobiCom*, Sep. 2004, pp. 85–99.

[6] R. Bruno, M. Conti, and E. Gregori, "Mesh networks: Commodity multihop ad hoc networks," *IEEE Commun. Mag.*, vol. 43, no. 3, pp. 123–131, Mar. 2005.

[7] L. Kleinrock and J. Silvester, "Optimum transmission radii for packet radio networks or why six is a magic number," in *Proc. IEEE Nat. Telecommun. Conf.*, Birmingham, AL, Dec. 1978, pp. 4.3.1–4.3.5.

[8] T. C. Hou and V. O. K. Li, "Transmission range control in multihop packet radio networks," *IEEE Trans. Commun.*, vol. COM-34, no. 1, pp. 38–44, Jan. 1986.

[9] R. Nagpal, H. Shrobe, and J. Bachrach, "Organizing a global coordinate system from local information on an ad hoc sensor network," in *Proc. Int. Workshop IPSN*, 2003, pp. 333–348.

[10] J.-C. Kuo and W. Liao, "Modeling the behavior of flooding on target location discovery in mobile ad hoc networks," in *Proc. IEEE ICC*, 2005, pp. 3015–3019.

[11] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, and J.-P. Sheu, "The broadcast storm problem in a mobile ad hoc network," in *Proc. ACM MobiCom*, Aug. 1999, pp. 152–162.

[12] E. Royer, P. Melliar-Smith, and L. Moser, "An analysis of the optimum node density for ad hoc mobile networks," in *Proc. IEEE ICC*, 2001, pp. 857–861.

[13] C. Ho, K. Obraczka, G. Tsudik, and K. Viswanath, "Flooding for reliable multicast in multi-hop ad hoc networks," in *Proc. Int. Workshop DIALM*, 1999, pp. 64–71.

[14] D. Johnson and D. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*, T. Imielinski and H. Korth, Eds. Norwell, MA: Kluwer, 1996, ch. 5.

[15] C. E. Perkins, E. M. Royer, and S. R. Das, *Ad hoc on Demand Distance Vector (AODV) Routing*. IETF RFC 3561.

Jia-Chun Kuo (M'03) received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, R.O.C., in 2001, where he is currently working toward the Ph.D. degree with the Department of Electrical Engineering.

His research interests include modeling and performance analysis in wireless *ad hoc*, sensor, and mesh networks.

Wanjiun Liao (S'96–M'97) received the B.S. and M.S. degrees from National Chiao Tung University, Hsinchu, Taiwan, R.O.C., in 1990 and 1992, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1997.

She joined the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, as an Assistant Professor in 1997, where she has been a Full Professor since August 2005. Her research interests include wireless networks, multimedia networks, and broadband access networks.

Dr. Liao is an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IEEE TRANSACTIONS ON MULTIMEDIA. She was a tutorial Cochair of IEEE INFOCOM 2004 and the TPC Vice Chair of IEEE Globecom 2005 (Autonomous Network Symposium). She has received many research awards. Papers she coauthored with her students received the Best Student Paper Award at the First IEEE International Conferences on Multimedia and Expo (ICME) in 2000 and the Best Paper Award of the First International Conference on Communication, Circuits, and Systems in 2002. She received the K. T. Li Young Research Award from the Association for Computing Machinery Taipei chapter in 2003 for her research achievement. She was elected as one of the Ten Distinguished Young Women in Taiwan in 2000 and was listed in the *Marquis Who's Who* and the *Contemporary Who's Who*.