

# Locating Randomly Selected Destinations in Large Multi-hop Wireless Networks

Jing Deng

Department of Computer Science  
University of New Orleans  
New Orleans, LA 70148, USA  
E-mail: jing@cs.uno.edu

**Abstract.** We focus on the problem of locating randomly selected destinations in large multi-hop wireless networks where potential destinations are distributed randomly in a circular region. In order to reduce the number of rebroadcasts, it may be beneficial to broadcast the inquiry packet with a certain Time-To-Live (TTL) to search those nodes within a limited hop-count from the source node, hoping to locate the destination without using the network-wide broadcast. Such a technique is termed Expanding Ring Search (ERS), which has been shown to be generally ineffective by other researchers when the TTL value is incremented by one after each failure. Therefore, it is interesting to identify the best search strategy, in terms of which set of TTLs should be used, for ERS. In this work, we estimate the expected number of network nodes that are within a fixed hop from the source node located at the center of the circular region. We argue that, when the maximum hop-count of the network is known, a search strategy with three steps are the optimum. Closed-form search sets are derived through our analysis. Numerical and simulation results are presented to support our claim. It is shown that the search strategies based on the search sets perform relatively well in a number of different network regions, even when the source node is not located at the center of the region.

**Keywords:** Expanding Ring Search, Time-To-Live, Optimum Strategy

## 1 Introduction

In wireless communication networks such as wireless ad hoc networks [1] and wireless sensor networks [2], communication nodes may need to inquire destinations that are unknown to themselves thus far. Since such an inquiry may take place before a routing path is found, routing information is usually unavailable in the context. One way to send the inquiry packet is to use the flooding technique, in which the packet is broadcasted and all neighboring nodes forward them until the destination is reached. Such a flooding scheme is usually termed as pure flooding. Pure flooding is rather expensive since it involves all nodes. In fact, the “broadcast storm” problem [3] may appear when inappropriate rebroadcasts are performed and when packet collisions take place frequently, resulting in more rebroadcasts. These collisions degrade the overall network performance and should

be avoided. Advanced techniques have been developed to reduce the number of redundant rebroadcasts while maintaining the reachability of the flooding process [3].

Since the destination may reside in an area that is relatively close to the source node, an Expanding Ring Search (ERS) technique has been proposed and used in a number of networking protocols, such as routing [4, 5]. In the ERS scheme, the query packet is broadcasted with a Time-To-Live (TTL) value. When the packet is received by other nodes, the TTL value on the packet is decremented. Then the packet will be rebroadcasted only if the TTL value is positive. The exact implementations of the ERS scheme vary in different protocols. For example, in some ERS techniques, the initial TTL value is set to 1 for the first (ring) search. If such a search does not find the destination, a new search is initiated with an incremented TTL value, 2. The process continues until the initial TTL value reaches a threshold,  $L$ . Then a network-wide flooding is initiated [6]. The ERS scheme in DSR [4] uses the California Split rule [7], where the TTL value is doubled every time the previous TTL value fails. In AODV [5], an ERS scheme is implemented to start with a TTL value of `TTL_START` and to increase the TTL by `TTL_INCREMENT`.

Recently, several research papers have been published with the focus of analyzing search strategies such as the ERS schemes under various network settings [6, 8, 9]. It has been shown that the ERS schemes that increment the TTL value by one after each failure are generally ineffective. Therefore, it is interesting to identify the best search strategy, in terms of which set of TTLs should be used, for ERS. In this work, we estimate the expected number of network nodes that are within a fixed hop from the source node located at the center of the circular region. Our estimation is based on the uniform distribution of nodes on wireless networks. Through our analysis, we argue that, when the maximum hop-count of the network is known, a search strategy with three steps are the optimum. Closed-form search sets are derived through our analysis.

The rest of this paper is organized as follows. In Section 2, we summarize the related work. In Section 3, we present the network model of our analysis and our estimation of number of nodes that are within  $i$  hops from the source node. Our main analytical results are presented in Section 4. Section 5 presents numerical and simulation results in our study. Concluding remarks are given in Section 6.

## 2 Related Work

In [8], the authors studied geography-based and hop-based flooding control methods. It has been proven that two-tier and three-tier hop-based flooding control methods can reduce the cost of broadcast. Both of the cost and the latency have been studied in [8]. A general formula to determine good parameters for two-tier and three-tier schemes has been provided and investigated. Small network and the role of caching in flooding have been studied as well. Different to [8], we concentrate on the cost of broadcast and provide a general framework to prove the optimality of the three-step (three-tier) techniques.

Reference [9] presents a series of theoretical results on TTL broadcasts and methods to derive search strategies, such as sequences of TTL values. The authors have proposed a dynamic programming formulation with which optimal search strategies can be derived when the node distribution is known. However, only the optimal strategies of a linear network with uniformly distributed location can be derived from the dynamic programming

technique. The paper gives the necessary and sufficient conditions on location distribution under which some commonly used search strategies are optimal. The authors have further shown that, when the destination distribution is unknown, a worst-case search cost is minimized when randomized strategies are used.

In [6], the authors studied the optimum  $L$  for one class of ERS schemes with a search set of  $\{1, 2, \dots, L\}$ . The authors claimed that there existed an optimum value of  $L$ , which minimizes the expected cost of broadcasts. However, it is generally non-optimal to use a consecutive sequence of integers as the search set, as shown in [8] and [9].

In [7], the authors studied the California Split rule in a network that reduces to a path. In the California Split rule, the set of TTLs is  $\mathbf{u} = \{x_1, x_2, \dots\}$  where  $x_i = 2^{i-1}$  for all  $i \geq 1$ . In [10], alternative query algorithms of the Gnutella network have been explored through simulations. A new query algorithm based on multiple random walks have been proposed and evaluated.

### 3 Network Model and Preliminaries

In this section, we firstly present our network model and notations. Then we give some preliminary studies on hop-counts, which will be used in Section 4.

#### 3.1 Network Model

We make the following assumptions in the network that we study:

- We assume that the source node locates at the center of a circular network region, in which nodes are distributed randomly. The destination is chosen from these nodes randomly;
- The query packet carries a TTL value and a sequence number. When a node overhears a packet with sequence number that is larger than what it has seen from the source node so far, the packet is processed. First, it decrements the TTL value on the packet. If the decremented TTL value is positive, the packet is then rebroadcasted;
- We define the “search cost”, or “cost”, as the number of expected (re)broadcasts before the source node receives a cost-free acknowledgment from the intended destination (receiver) of the inquiry packet;
- We assume the broadcasts are collision-free, i.e., all broadcasts are carefully scheduled so that all neighbors of the sender will be able to overhear the message successfully. We argue that the inclusion of packet collisions into our model only increases the number of broadcasts of all schemes proportionally. By eliminating such collisions, we simplify the analysis significantly without sacrificing its correctness.

We generalize a Ring Search (RS) scheme with a search set of  $\mathcal{R} = \{r_1, r_2, \dots, r_k\}$ . In this scheme, a broadcast with TTL= $r_1$  is sent first. If the search does not locate the destination, a new broadcast with TTL= $r_2$  is sent. Such a process continues until a broadcast with TTL= $r_k$  fails to locate the destination, at which time a network-wide flooding is initiated. Note that  $RS(\Phi)$  is the pure flooding technique and  $RS(\{1, 2, \dots, L\})$  is the class of ERS schemes studied in [6].

Throughout the paper, we use the following notations:

- $M$ : number of network nodes;
- $H$ : the maximum hop that  $M$  nodes may spread from the source node at the center of the network;
- $r$ : radio transmission range, which is set to 1 unit distance;
- $n(i)$ : number of nodes that are exactly  $i$  hops away from the source node,  $i = 0, 1, \dots, H$ . Thus,  $n(0) = 1$ ;
- $N(i)$ : number of nodes that are within (less than or equal to)  $i$  hops from the source node,  $i = 0, 1, \dots, H$ . Therefore,  $N(i) = \sum_{j=0}^i n(j)$ ;
- $p(i)$ : the probability that the destination is chosen within  $i$  hops from the source node. Neglecting the effect of the source node,<sup>1</sup>

$$p(i) = \frac{N(i)}{M} ; \tag{1}$$
- $\mathcal{R} = \{r_1, r_2, \dots, r_k\}$ : the search set, where  $1 \leq r_1 < r_2 < \dots < r_k \leq H$ ;
- $k$ : the size of the search set,  $0 \leq k \leq H$ ;
- $C(\mathcal{R})$ : the cost of the RS scheme when a search set  $\mathcal{R}$  is used.

### 3.2 Preliminaries

We use the random distribution of nodes to estimate  $N(i)$  and  $n(i)$ , the number of nodes within  $i$  hops from the source node and the number of nodes on the  $i$ -th hop “ring”, respectively,  $0 \leq i \leq H$ . The approximate number of nodes in the circle of radius  $ir$  and centered at the source node is  $\rho\pi(ir)^2$ , where  $\rho$  is the nodal density. Since some of these nodes may not be within  $i$  hops from the source node,  $N(i)$  may be approximated by

$$N(i) \approx \gamma\rho\pi(ir)^2 = m \cdot i^2 , \tag{2}$$

where  $\gamma$  is the ratio of nodes in the circle that are actually within  $i$  hops from the source node and  $m = \gamma\rho\pi r^2$ . Based on (2), an estimate of  $n(i)$  is given by

$$n(i) \approx N(i) - N(i-1) = m \cdot (2i-1) . \tag{3}$$

Based on the definitions of  $H$  and  $M$  and (2), we have:<sup>2</sup>

$$M = N_H \approx m \cdot H^2 . \tag{4}$$

## 4 Theoretical Analysis

### 4.1 Problem Statement

In this paper, we focus on the following problem:

*Problem 1.* When a source node is located at the center of a circular region with maximum hop-count from the center as  $H$  and the destination is chosen randomly from the network, what is the best  $\mathcal{R} = \{r_1, r_2, \dots, r_k\}$ ,  $1 \leq r_1 < r_2 < \dots < r_k \leq H$ ,  $0 \leq k \leq H$ , with which the RS( $\mathcal{R}$ ) scheme has the lowest expected search cost  $C(\mathcal{R})$ ?

<sup>1</sup> The total number of nodes to be chosen from is  $M - 1$ , as the source node is excluded from the random selection. An accurate calculation is  $p(i) = \frac{N(i)-1}{M-1}$ .

<sup>2</sup> Therefore, we have assumed that the network is “full” with nodes within  $H$ -hops. Our results will be studied, through simulations, in networks where this assumption is not satisfied.

## 4.2 The $RS(\Phi)$ Scheme

The  $RS(\Phi)$  Scheme is equivalent to the pure flooding scheme. We have the following lemma in regard to the cost of the  $RS(\Phi)$  scheme:

**Lemma 1.** *The cost of the  $RS(\Phi)$  scheme is given by*

$$C(\Phi) = M . \quad (5)$$

*Proof.* Since all nodes (including the source node and the destination node<sup>3</sup>) in the network must broadcast the message exactly once, the number of rebroadcasts is the total number of nodes,  $M$ .

## 4.3 The $RS(\{r_1\})$ Schemes

When  $k = 1$ , we have the following theorem in regard to the cost of the  $RS(\{r_1\})$  scheme:

**Theorem 1.** *The cost of the  $RS(\{r_1\})$  scheme is*

$$C(\{r_1\}) = C(\Phi) + [N(r_1 - 1) - N(r_1)] \leq C(\Phi) . \quad (6)$$

*Therefore, when  $n(r_1) = N(r_1) - N(r_1 - 1) > 0$ , the  $RS(\{r_1\})$  scheme out-performs the  $RS(\Phi)$  scheme.*

*Proof.* The cost of the  $RS(\{r_1\})$  scheme depends on whether the chosen destination resides within  $r_1$  hops from the source node. If so, the search cost is  $N(r_1 - 1)$ . Otherwise, the search cost is  $N(r_1 - 1) + M$ . The expected search cost of the  $RS(\{r_1\})$  scheme is then:

$$C(\{r_1\}) = p(r_1) \cdot N(r_1 - 1) + [1 - p(r_1)] \cdot [N(r_1 - 1) + M] . \quad (7)$$

Based on (1) and Lemma 1, (7) becomes

$$C(\{r_1\}) = M + N(r_1 - 1) - p(r_1) \cdot M = M + N(r_1 - 1) - N(r_1) = C(\Phi) - n(r_1) .$$

Therefore, when  $n(r_1) > 0$ , the  $RS(\{r_1\})$  scheme out-performs the  $RS(\Phi)$  scheme.

## 4.4 The $RS(\{\mathcal{R}\})$ Schemes

In this subsection, we investigate the  $RS(\{\mathcal{R}\})$  schemes with  $\mathcal{R}$  in a more general form. The proofs of the following theorems and lemmas are omitted due to the page limit.

**Theorem 2.** *The relation between the cost of the  $RS(\{\mathcal{R}\})$  scheme,  $\mathcal{R} = \{r_1, r_2, \dots, r_k\}$ ,  $1 \leq r_1 < r_2 < \dots < r_k \leq H$ ,  $0 \leq k \leq H$ , and that of the  $RS(\Phi)$  scheme is the following:*

$$C(\{r_1, r_2, \dots, r_k\}) = C(\Phi) + \sum_{i=1}^k \left[ N(r_i - 1) - \frac{N(r_i) \cdot N(r_{i+1} - 1)}{M} \right] , \quad (8)$$

where we have defined  $r_{k+1} \equiv H + 1$  for presentation purpose.

Now we develop the optimum search set  $\mathcal{R}$  for the  $RS(\mathcal{R})$  schemes.

---

<sup>3</sup> We have also counted the destination's rebroadcast, which does not take place.

**Lemma 2.** *The cost of the  $RS(\{r_1, r_2, \dots, r_k\})$  scheme decreases with the increase of  $r_k \leq H$ , i.e.,*

$$C(\{r_1, r_2, \dots, r_{k-1}, r_k\}) > C(\{r_1, r_2, \dots, r_{k-1}, r'_k\}), \quad (9)$$

when  $r_k < r'_k \leq H$ .

Therefore, an optimum  $RS(\{r_1, r_2, \dots, r_k\})$  scheme must satisfy  $r_k = r_k^* = H$ .

**Theorem 3.** *When  $k = 2$  in the  $RS(\{r_1, r_2, \dots, r_k\})$  schemes, the optimum values for  $r_1$  and  $r_2$  are*

$$\begin{cases} r_1^* = \frac{H^2}{2H-1} \\ r_2^* = H \end{cases}. \quad (10)$$

The integer value of  $r_1^*$  is  $\lceil \frac{H}{2} \rceil$ , where  $\lceil x \rceil$  is the ceiling function giving the minimum integer that is larger than  $x$ .

Before we present our results on the search set with  $k = 3$ , we develop a lemma.

**Lemma 3.** *For a large positive integer  $H$  and a real-valued variable  $z$ ,  $0 < z \leq \frac{H-2}{H}$ , the function  $f(z)$*

$$f(z) = -\frac{1}{1-z^2} + H^2 z^2 - \left[ \frac{(zH+1)(H-1)}{H} \right]^2, \quad (11)$$

achieves minimum when  $z$  takes a value close to

$$z^* = \frac{(H-1)^2}{(2H-1)H}. \quad (12)$$

**Theorem 4.** *When  $k = 3$  in the  $RS(\{r_1, r_2, \dots, r_k\})$  schemes, the optimum values for  $r_1$ ,  $r_2$ , and  $r_3$  are*

$$\begin{cases} r_1^* = 1 \\ r_2^* = \frac{(H-1)^2}{2H-1} \\ r_3^* = H \end{cases}. \quad (13)$$

The integer value of  $r_2^*$  is  $\lfloor \frac{(H-1)^2}{2H-1} \rfloor$ , where  $\lfloor x \rfloor$  is the floor function giving the maximum integer that is no greater than  $x$ .

In our simulation results, we will see that the  $\mathcal{R}$  set given by (13) in Theorem 4 is indeed the optimum search set for all integer  $k$ . Therefore, we have the following conjecture:

*Conjecture 1.* The  $RS(\{1, \lfloor \frac{(H-1)^2}{2H-1} \rfloor, H\})$  scheme is the optimum scheme that achieves lowest cost among all RS schemes.

## 5 Performance Evaluations

We present our numerical and simulation results in this section. We study the Number of Reduced Broadcasts (NRB) by the RS scheme over pure flooding. The NRB of an RS scheme is defined as the number of (re)broadcasts of pure flooding minus the expected number of (re)broadcasts of the RS scheme,  $NRB = C(\Phi) - C(\mathcal{R})$ . The efficiency,  $\eta$ , of an RS scheme is defined as its Normalized NRB (NNRB),  $\eta = \frac{C(\Phi) - C(\mathcal{R})}{C(\Phi)}$ . The higher the value of  $\eta$  is, the better the RS scheme. It can be proven that  $\eta < 1$ .

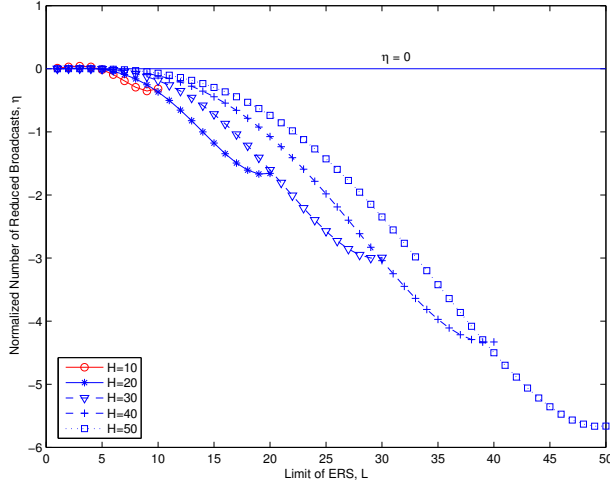


Fig. 1. Performance of the ERS schemes with limit as  $L \leq H$ .

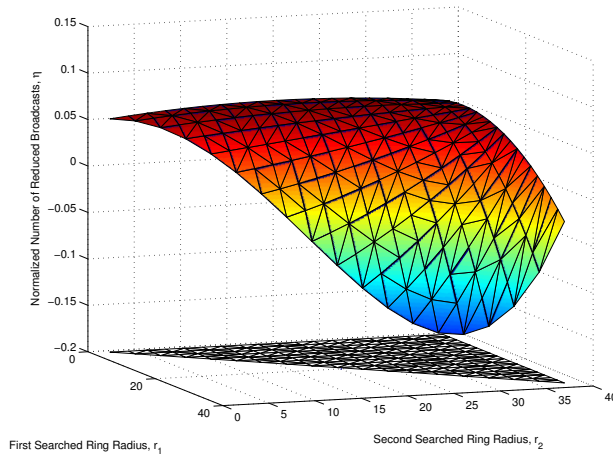


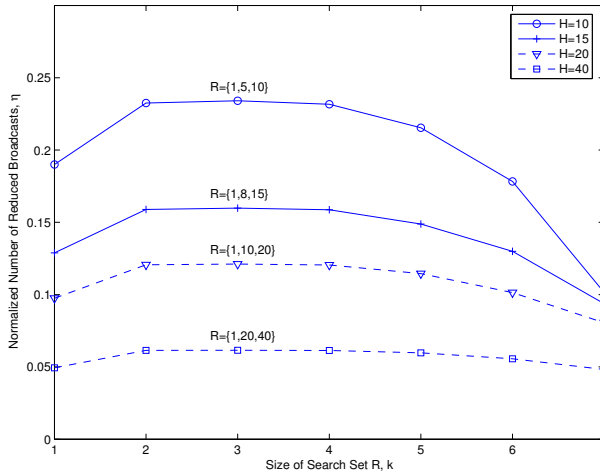
Fig. 2. Performance of the RS schemes with  $k = 3, H = 40, r_3 = 40$ .

## 5.1 Numerical Results

Our numerical results are calculated based on (8). Note that we do not need the exact value of  $m$  because the efficiency is a normalized value.

In Fig. 1, the performance of a class of ERS schemes, the  $RS(\mathcal{R})$  scheme with  $\mathcal{R} = \{1, 2, \dots, L\}$  and a limit of  $L \leq H$  is shown for different maximum hop-count of the network,  $H$ . From Fig. 1, we can observe that the benefit of using these ERS schemes is extremely limited. In most of the scenarios that we have shown, the ERS schemes have higher search cost than pure flooding ( $\eta < 0$ ).

We show the performance of the  $RS(\{r_1, r_2, r_3 = H\})$  schemes with  $H = 40$  in Fig. 2. The shaded plane at  $\eta = -0.2$  level is the region of  $(r_1, r_2)$  that we have used to calculate and show their efficiency. As observed from Fig. 2, the  $RS(\{r_1, r_2, r_3 = H\})$  scheme is more efficient when  $r_1$  is closer to 1, except in the small region where  $r_2$  is close to  $H$ . The most efficient point of  $(r_1, r_2)$  is when  $r_2$  is in the middle of  $[0, 40]$ ,  $(r_1, r_2) = (1, 20)$ , with



**Fig. 3.** Optimum search sets of the RS schemes.

an efficiency of  $\eta = 0.06$ . Therefore, when  $H = 40$ , the best of the  $RS(\{r_1, r_2, r_3 = H\})$  schemes achieves approximately 6% search cost saving.

In Fig. 3, the performance of optimum search sets with different set size,  $k$ , is shown for different  $H$ . We have gone through all possible combinations for each  $k$  in  $\mathcal{R}$  and identified the best search set with  $k = 1, 2, \dots, 7$  for  $H = 10, 15, 20$ , and  $40$ . The difference between optimum 2-, 3-, and 4-step schemes are rather small. However, the optimum  $k$  is 3 with the optimum search set of different  $H$  agreeing with Theorem 4. Therefore, the use of larger  $k$  should be discouraged. We also observe from Fig. 3 that the gain of using an RS scheme decreases with the increase of  $H$ . However, when the values of  $H$  in wireless networks are small, the benefit of using an optimum RS scheme is clearly shown. For example, the cost saving is higher than 25% when  $H \leq 10$ .

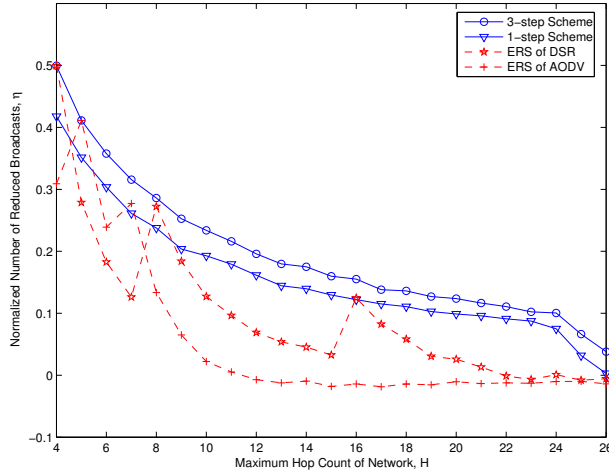
## 5.2 Simulation Results

Simulations have been performed to support our numerical results. The value of  $\rho$  is 10 and the radio transmission range is set to 1 unit distance in our simulations. Therefore, there are approximately 10 nodes in each unit square area. We have simulated different shapes of region (circular or rectangular) and different locations of the source node (center or corner). The simulations were constructed in Matlab.

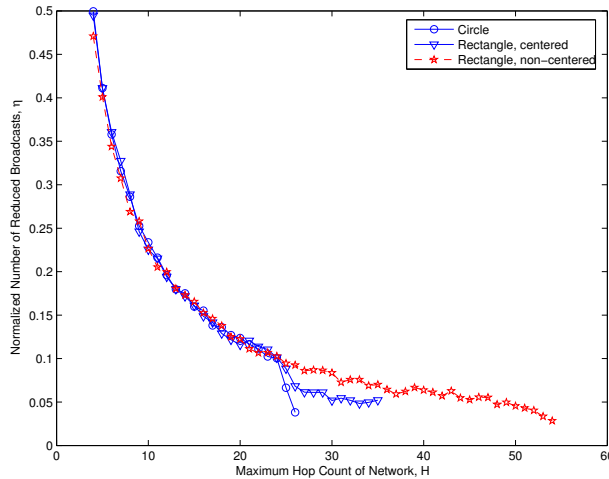
The simulation results of a circular region are presented in Fig. 4. The circular region has a radius of 20. The source node is located at the center of the region. We varied the maximum hop-count of the network, from which the source node chooses its intended destination. Therefore, when  $H$  gets larger, border effects may become more noticeable. It can be seen that the 3-step RS scheme based on Theorem 4 out-performs all other schemes including the 1-step RS scheme based on Lemma 2. The ERS scheme implemented in DSR [4] is equivalent to an  $RS(\{1, 2, 4, 8, 16, \dots\})$  scheme. On the other hand, AODV [5] implements an  $RS(\{1, 3, 5, 7\})$  scheme. The peaks of these two schemes are due to the incidental match between these search sets and the optimum search sets.

In order to show how the 3-step RS scheme performs under various network settings, we have simulated the RS scheme based on Theorem 4 in different regions and different





**Fig. 4.** Simulation results of a circular network region (source at center).



**Fig. 5.** Simulation results of different regions and source locations (3-step RS schemes).

source locations and shown its efficiency in Fig. 5. The curve related to “Circle” is the same as the one shown in Fig. 4, i.e., circular region with the source node located at the center. The curve termed as “Rectangle, centered” represents the performance of the 3-step RS scheme in a rectangle area of  $40 \times 40$  and the source node located at the center of the area, i.e.,  $(20, 20)$ . The third curve, with a legend of “Rectangle, non-centered”, shows the results in the same rectangle area, but with the source node located at one of the corners, e.g.,  $(0, 0)$ . It is interesting to see that the efficiency of the 3-step RS scheme are almost identical in most values of  $H$  under three network settings, except that the possible maximum hop-count of the network,  $H$ , is larger under the last two settings. The results shown in Fig. 5 suggest that the 3-step RS schemes are rather robust.

## 6 Concluding Remarks

The wide-spread use of flooding to locate services or destinations in wireless networks underlines the importance of such a technique. In this paper, we focus on the problem of locating a randomly chosen destination in a large multi-hop wireless network with Expanding Ring Search (ERS). We have generalized pure flooding and Expanding Ring Search (ERS) into Ring Search (RS) schemes, in which a search set ( $\mathcal{R}$ ) is used to set the Time-To-Live (TTL) field of the inquiry packet sequentially before a network-wide flooding is initiated. Based on the random distribution of nodes in the network region, we have developed a theoretical framework to prove that a special 3-step RS scheme is optimum among all 3-step RS schemes, in terms of lowering the number of (re)broadcasts. Numerical and simulation results have been presented to support our claim.

Our simulation results further suggest that this 3-step RS scheme is actually optimum among all RS schemes. Unfortunately, we have not been able to prove such an optimality. In our future work, we will try to prove the global optimality of the 3-step RS scheme. The analysis of the RS schemes in irregular network region is interesting as well.

## References

1. Haas, Z.J., Deng, J., Liang, B., Papadimitratos, P., Sajama, S.: Wireless ad hoc networks. In Proakis, J.G., ed.: Wiley Encyclopedia of Telecommunications, John Wiley & Sons (2002)
2. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E.: A survey on sensor networks. *IEEE Communications Magazine* **40** (2002) 102–114
3. Ni, S., Tseng, Y., Sheu, J.: The broadcast storm problem in a mobile ad hoc network. In: Proc. of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '99), Seattle, WA, USA (1999) 152–162
4. Johnson, D.B., Maltz, D., Hu, Y.C.: The dynamic source routing protocol for mobile ad hoc networks (DSR) (2003) Internet Draft.
5. Perkins, C.E.: Ad hoc on-demand distance vector (AODV) routing (2003) Internet Draft.
6. Hassan, J., Jha, S.: Optimising expanding ring search for multi-hop wireless networks. In: Proc. of IEEE Global Telecommunications Conference / Wireless Communications, Networks, and Systems (GLOBECOM '04), Dallas, TX, USA (2004)
7. Baryshnikov, Y., Coffman, E., Jelenkovic, P., Momcilovic, P., Rubenstein, D.: Flood search under the california split rule. *Operations Research Letters* **32** (2004)
8. Cheng, Z., Heinzelman, W.B.: Flooding strategy for target discovery in wireless networks. In: Proc. of ACM (MSWiM '03), San Diego, CA, USA (2003)
9. Chang, N., Liu, M.: Revisiting the TTL-based controlled flooding search: Optimality and randomization. In: Proc. of the 10th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '04), Philadelphia, PA, USA (2004) 85–99
10. Lv, Q., Cao, P., Cohen, E., Li, K., Shenker, S.: Search and replication in unstructured peer-to-peer networks. In: Proc. of the ACM International Conference on Supercomputing (ICS '02), New York, NY, USA (2002) 84–95