**An Analytical Model for Broadcasting by Self Pruning in Wireless Ad hoc Networks**

Yu Huang\(^1,2\), Bo Liu\(^2\), Xianping Tao\(^1,2\), Jiannong Cao\(^3\), Beihong Jin\(^4\)

\(^1\)State Key Laboratory for Novel Software Technology
Nanjing University, Nanjing China
\(^2\)Department of Computer Science and Technology
Nanjing University, Nanjing China
\(^3\)Internet and Mobile Computing Lab, Department of Computing
Hong Kong Polytechnic University, Hong Kong China
\(^4\)Technology Center of Software Engineering
Institute of Software, Chinese Academy of Sciences, Beijing China

{yuhuang,txp,lj}@nju.edu.cn
wy90021@126.com
csjcao@comp.polyu.edu.hk
jbh@otcaix.iscas.ac.cn

**Abstract**

The broadcast operation is a fundamental service in wireless ad hoc networks. The naïve flooding mechanism may trigger a great number of data retransmissions at the same time and congest the networks, resulting in the broadcast storm problem. Self pruning is a simple heuristic, which reduces redundant data transmission of flooding, making use of the neighbor knowledge. However, exiting evaluation on self pruning is mainly based on experiments and greatly lacks detailed theoretical analysis. To this end, we propose in this paper an analytical model, which estimates the expected cost and the scalability of self pruning. To evaluate the accuracy of the proposed analytical model, we compare the analytical results with results obtained from the experiments. The experimental evaluation shows that the proposed analytical model accurately estimates the expected cost and scalability of self pruning, especially when the density of nodes is high.

**Keywords** wireless ad hoc networks, broadcasting, flooding, self pruning, analytical model

**1. Introduction**

A wireless ad hoc network consists of a collection of autonomous hosts which communicate with each other via multi-hop wireless links. It does not require any pre-existing infrastructure and is relatively easy to deploy and reconfigure [2,3,4,5,13]. The broadcast operation is a fundamental service in wireless ad hoc networks. For example, it is used by a variety of routing protocols, such as DSR [6] and AODV[11], to set up the routing. A naïve broadcast mechanism is flooding, in which each node forwards the packet once and only once [10]. However, flooding may impose excessive data transmission and is prone to the broadcast storm problem [1] if the forwarding nodes are not carefully designed. A variety of broadcast mechanisms have been proposed to reduce redundant packet forwarding imposed by flooding. Such mechanisms can be classified as: probability-based, area-based and neighbor knowledge-based [14].

Self pruning is a simple heuristic, which reduces redundant data transmission of flooding, making use of the neighbor knowledge [8]. Using self pruning, when a node B receives a broadcast packet from node A, but finds that all its neighbors are covered by A, it prunes itself from the set of forwarding nodes. The broadcast cost by self pruning greatly varies in dynamic wireless ad hoc networks. Existing performance evaluation
mainly relies on simulations, and lacks theoretical analysis, as far as we know [8]. To this end, we propose in this paper an analytical model, which estimates the expected broadcast cost of self pruning. We also study the scalability of self pruning based on the analytical model. To evaluate the proposed analytical model, we compare the analytical results with results obtained from the experiments. The experimental evaluation shows that our model accurately estimates the broadcast cost and the scalability of self pruning, especially when the density of nodes is high.

The rest of this paper is organized as follows. Section 2 describes the preliminaries. Section 3 presents the analytical model. In Section 4, we evaluate the analytical model by experiments. Section 5 concludes the paper with a summary and the future work.

2. Preliminaries

We describe a wireless ad hoc network as a unit disk graph \( G = (V, E) \). The node set \( V \) represents a set of nodes in the network and the edge set \( E \) represents a set of bidirectional links between the neighboring nodes. Two nodes are considered neighbors if and only if their geographic distance is less than the transmission range.

Self pruning is a neighbor knowledge-based heuristic. Each node \( A \) attaches the set of its neighbors \( N(A) \) to the broadcast packet. When a node \( B \) receives the broadcast packet from node \( A \) and this is the first time node \( B \) receives the packet, it checks the set \( N(B) - N(A) \). Node \( B \) relays the packet only if set \( N(B) - N(A) \) is not empty, i.e., the relay of node \( B \) can cover new nodes, as shown in Fig. 1.

However, in wireless ad hoc networks, this metric is not appropriate, due to the intrinsic broadcasting nature of radio transmission. When a node in the wireless ad hoc network broadcasts, every neighbor of it receives the packet. Thus, the number of nodes which relay the packet is used to measure the broadcast cost in wireless ad hoc networks [8, 9]. Following this metric, we have that as for the broadcast cost in wireless ad hoc networks:

- The upper bound of broadcast cost is that of flooding, which is the total number of nodes in the network.
- The lower bound is the number of nodes in the minimum connected domination set (MCDS), while constructing the MCDS in a wireless ad hoc network is an NP-hard problem[9].

3. Analytical Model for Self Pruning

In this section, we present the analytical model for estimating the cost and study the scalability of broadcasting by self pruning in wireless ad hoc networks. We first describe the assumptions of our model. Then we present the analytical results.

3.1. Assumptions

Assumptions adopted in our analytical model are listed as follows.

1. All nodes in the network are scattered to the network area uniformly at random, and the resulting network is connected (if it is not, we only consider the connected component).
2. Width and length of the network area is much larger than the transmission range of wireless hosts, thus enabling us to ignore the edge effect.
3. Each node has the same probability to initiate the broadcast.
4. Packet loss and node failure are not considered.

Observe that, given the assumptions above, broadcasting by self pruning can cover all nodes in the network. Notations used in the analytical model are listed in Table 1.

<table>
<thead>
<tr>
<th>Table 1. Notations used in the analytical model</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
</tr>
<tr>
<td>( N_1 )</td>
</tr>
<tr>
<td>( N_2 )</td>
</tr>
</tbody>
</table>
3.2. Analysis on the Broadcast Cost of Self Pruning

According to the design of self pruning, when a node $B$ receives the broadcast packet from node $A$, it decides whether to relay the packet according to the status of set $N(B)-N(A)$, as shown in Fig. 1. If set $N(B)-N(A)$ is empty (node $B$ cannot cover new neighbors), it does not need to relay the packet.

To estimate the status of set $N(B)-N(A)$, we study the distribution of all the nodes in the network, excluding $A$ and $B$. Since node $B$ receives the packet from node $A$, we have:

$$ d(A, B) \leq r $$

We divide all nodes in the network, excluding $A$ and $B$, into 3 categories:

1) **Neighbors of node $A$, excluding node $B$.** Such nodes appear in the region within the network area defined as:

$$ N_1 = \{i \mid d(i, A) \leq r, i \neq A, i \neq B\} $$

$N_1$ is denoted as *Area I* in Fig. 1.

2) **Neighbors of Node $B$, which are not neighbors of node $A$.** Such nodes appear in the region defined as:

$$ N_2 = \{i \mid d(i, B) \leq r, d(i, A) > r\} $$

$N_2$ is denoted as *Area II* in Fig. 1.

3) **Nodes which are neither node $A$’s neighbors nor node $B$’s neighbors.** Such nodes appear in the region defined as:

$$ N_3 = \{i \mid d(i, A) > r, d(i, B) > r\} $$

$N_3$ is denoted as *Area III* in Fig. 1.

According to the design of self pruning, the probability that node $B$ does not need to relay the broadcast packet is:

$$ P(x) = \Pr\{N(B) - N(A) = \Phi\} = \Pr\{\text{All nodes in the network, excluding } A \text{ and } B, \text{ are in region } H-N_2\}. $$

Here, $x$ stands for the distance between $A$ and $B$, $0 < x \leq r$.

Since each node is scattered to the network area uniformly at random, we can approximately measure the probability that a node appear in region $H-N_2$ by the ratio between the area of two regions:

$$ \Pr\{i \mid i \in H-N_2\} = \frac{S(H) - S(N_2)}{S(H)}. $$

Since

$$ S(N_2) = \pi r^2 - 2(2r \arccos \frac{x}{2r} - \frac{x}{2} \sqrt{4r^2 - x^2}) $$

we have that

$$ \Pr\{i \mid i \in H-N_2\} = \frac{S(H) - \pi r^2 + 2(2r \arccos \frac{x}{2r} - \frac{x}{2} \sqrt{4r^2 - x^2})}{S(H)}. $$

Since each node appears in the network area independently, the probability that no nodes appear in region $N_2$ is:

$$ \overline{P(x)} = \Pr\{i \mid i \in H-N_2\} = \left(1 - \frac{S(H) - \pi r^2 + 2(2r \arccos \frac{x}{2r} - \frac{x}{2} \sqrt{4r^2 - x^2})}{S(H)}\right)^{n-2} $$

Thus, the probability that node $B$ needs to relay the packet is:

$$ P(x) = 1 - \overline{P(x)} $$

$$ = 1 - \left(1 - \frac{S(H) - \pi r^2 + 2(2r \arccos \frac{x}{2r} - \frac{x}{2} \sqrt{4r^2 - x^2})}{S(H)}\right)^{n-2} $$

According to the discussion above, we obtain the probability $P(x)$ for a given pair of nodes with distance $x$. We now calculate the average probability of relaying the packet for any possible pair of neighboring nodes.

For all possible distributions of nodes in the network area, the average probability of relaying the packet is:

$$ P^{(n)} = \int_0^x P(x) \Pr\{d(A, B) = x\} dx $$
Here, $Pr\{d(A,B) = x\}$ is the probability that the distance between $A$ and $B$ is $x$. Since $B$ receives the packet from $A$, we have that $x \leq r$. Node $B$ appears in circle $C(A)$ uniformly at random. Thus, we have that:

$$Pr\{d(A,B) = x\} = \frac{d}{dx} Pr\{d(A,B) \leq x\}.$$

Note that:

$$Pr\{d(A,B) \leq x\} = \frac{\pi x^2}{2r^2} = \frac{x^2}{2r^2}.$$

Thus, we have:

$$Pr\{d(A,B) \leq x\} = \frac{d}{dx} Pr\{d(A,B) \leq x\} = \frac{d}{dx} \frac{x^2}{2r^2} = \frac{2x}{r^2}.$$

According to the discussion above, we finally obtain:

$$P^{(o)} = \int_0^r \left(1 - \frac{(S(H) - S(N_2))^{n-2} 2x}{S(H)} \right)^{n-2} \frac{2x}{r^2} dx.$$

$P^{(o)}$ estimates the expected broadcast cost using self pruning in given network area. Based on $P^{(o)}$, we can study the performance of self pruning in a variety of network conditions. In the following Section 3.4, we study the scalability of self pruning based on equation (1).

### 3.3. Analysis on the Scalability of Self Pruning

One essential issue in evaluating self pruning is its scalability. We can analyze the scalability of self pruning based on equation (1). Looking into the expression of $P^{(o)}$, we first have:

**Lemma 1** $\lim_{n \to \infty} P^{(o)} = 1$.

**Proof:** Since $\int_0^r \frac{2x}{r^2} dx = 1$, we have that:

$$\lim_{n \to \infty} P^{(o)} = 1 \iff \lim_{n \to \infty} \int_0^r \left( \frac{(S(H) - S(N_2))^{n-2} 2x}{S(H)} \right)^{n-2} \frac{2x}{r^2} dx = 0.$$

For any sufficiently small value $\varepsilon > 0$, we choose $\delta \in (0, r\sqrt{\varepsilon})$. Thus, we have:

$$\lim_{n \to \infty} \int_0^r \left( \frac{(S(H) - S(N_2))^{n-2} 2x}{S(H)} \right)^{n-2} \frac{2x}{r^2} dx = f_1 + f_2.$$

Here,

$$f_1 = \lim_{n \to \infty} \int_0^\delta \left( \frac{(S(H) - S(N_2))^{n-2} 2x}{S(H)} \right)^{n-2} \frac{2x}{r^2} dx,$$

$$f_2 = \lim_{n \to \infty} \int_\delta^r \left( \frac{(S(H) - S(N_2))^{n-2} 2x}{S(H)} \right)^{n-2} \frac{2x}{r^2} dx.$$

First we study the value of $f_1$. Since

$$\frac{S(H) - S(N_2)}{S(H)} < 1,$$

we have

$$f_1 < \lim_{n \to \infty} \int_0^\delta \frac{2x}{r^2} dx = \frac{\delta^2}{r^2} (3)$$

Since $0 < \delta < r\sqrt{\varepsilon}$, we have $f_1 < \varepsilon$.

Then we study the value of $f_2$. Since in $f_2$, the minimum value of $x$ is $\delta > 0$, we have that

$$\frac{S(H) - S(N_2)}{S(H)} < c < 1$$

Here, $c$ is a constant. So we have:

$$f_2 \leq \lim_{n \to \infty} c^{n-2} \int_\delta^r \frac{2x}{r^2} dx = 0$$

Thus we have $f_1 + f_2 < \varepsilon$. Finally we obtain that:

$$\lim_{n \to \infty} \int_0^r \left( \frac{(S(H) - S(N_2))^{n-2} 2x}{S(H)} \right)^{n-2} \frac{2x}{r^2} dx = 0$$

This proves that $\lim_{n \to \infty} P^{(o)} = 1$. ■

According to Lemma 1, we have that when there are sufficiently large number of nodes in given network area, all nodes will definitely relay the broadcast packet. This shows that:

**Theorem 1.** Self pruning transforms to flooding when the node density is sufficiently large in given network area.

### 4. Experimental Evaluation

In this section, we conduct simulations to evaluate the analytical model we propose. We first present the experiment methodology and configurations. Then we discuss the evaluation results.

#### 4.1. Experiment Methodology and Configurations

In the experiment, we study the accuracy of our analytical model in estimating the expected broadcast cost of self pruning. Results from the analytical model and those from the experiments are compared. We first study the performance of
self pruning in networks with normal size. Then we increase the number of nodes in the network to study the scalability. In the experiment, all nodes in the network are scattered to a rectangular area of size $100 \times 100$ m$^2$. The transmission range of each host is set to 15 m. The number of nodes are varied from 70 to 1000 in the experiments. The results are averaged over 1000 experiments. The number of forwarding nodes is used to evaluate the cost of self pruning. The experiment configurations are listed in Table 2.

<table>
<thead>
<tr>
<th>Table 2. Experiment Configurations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network area</td>
</tr>
<tr>
<td>Transmission range</td>
</tr>
<tr>
<td>Network size (cost estimation)</td>
</tr>
<tr>
<td>Network size (scalability analysis)</td>
</tr>
<tr>
<td>Times of experiments</td>
</tr>
</tbody>
</table>

### 4.2. Estimation of Broadcast Cost

In this experiment, we study the broadcast cost of self pruning with both our analytical model and experiments, in networks with different sizes (70-200). We find that our analytical model is quite accurate, especially when the density of nodes is high (# of nodes > 140), as shown in Fig. 2 and 3. In all network conditions, the discrepancy between the analytical results and the experiment results is within 4 percents. The maximum discrepancy (3.9%) is reached when the network size is 90.

#### Fig. 2. Broadcast cost of self pruning

![Fig. 2. Broadcast cost of self pruning](image)

### 4.3. Analysis on Scalability

We find from experiment results in Section 4.2 (Fig. 2 & 3) that the broadcast cost and the ratio of relaying nodes increase when the number of nodes in the network area increases. In this experiment, we keep increasing the number of nodes until the ratio of relaying nodes is close to 100%, as shown in Fig. 4. We find in this experiment that the ratio of relaying nodes quickly increases to 100% when the network size increases from 70 to 200. Then the ratio slows down and gradually increases to 100%, which is in accordance to our analytical result that self pruning transforms to flooding when there are sufficiently larger number of nodes in the network area in Section 3.3.

#### Fig. 3. Ratio of relaying nodes

![Fig. 3. Ratio of relaying nodes](image)

### 5. Conclusion and Future Work

Self pruning is a neighbor knowledge-based heuristic for broadcasting in wireless ad hoc networks. It reduces redundant data transmission of flooding by pruning, from the set of forwarding nodes, the nodes whose relay cannot cover new nodes. The evaluation of self pruning is mainly based on experiments and lacks detailed theoretical analysis. In this paper, we propose an analytical model for self pruning, which estimates the expected broadcast cost and the scalability.
Experiments are conducted to evaluate the accuracy of the proposed analytical model. The evaluation results show that the analytical model accurately estimates the performance of self pruning, especially when the node density is high.

In our future work, we will extend our analytical model to analyze more broadcast algorithms. We will also study how to design new heuristics for broadcasting in wireless ad hoc networks based on our analytical results.

Acknowledgements

This work is supported by the National Natural Science Foundation of China(No. 60736015, 60721002) and the National High-Tech Research and Development 863 Program of China(No. 2007AA01Z178).

References


