# Analysis on a Localized Pruning Method for Connected Dominating Sets

JOHN SUM[1], JIE WU[2] AND KEVIN HO[3]
[1]*Department of Information Management*
*Chung-Shan Medical University*
*Taichung, 402 Taiwan*
*E-mail: pfsum@csmu.edu.tw*
[2]*Department of Computer Science and Engineering*
*Florida Atlantic University*
*Boca Raton, Florida 33431, U.S.A.*
*E-mail: jie@cse.fau.edu*
[3]*Department of Computer Science and Communication Engineering*
*Providence University*
*Taichung, 433 Taiwan*
*E-mail: ho@pu.edu.tw*

While *restricted rule-k* has been succeeded in generating a connected dominating set (CDS) of small size, not much theoretical analysis on the size has been done. In this paper, an analysis on the expected size of a CDS generated by such algorithm and its relation to different node density is presented. Assume $N$ nodes are deployed uniformly and randomly in a square of size $L_N \times L_N$ (where $N$ and $L_N \to \infty$); three results are obtained. (1) It is proved that the node degree distribution of such a network follows a Poisson distribution. (2) The expected size of a CDS that is derived by the restricted pruning rule-$k$ is a decreasing function with respect to the node density $\hat{n}$. For $\hat{n} \geq 30$. it is found that the expected size is close to $N/\hat{n}$. (3) It is proved that the lower bound on the expected size of a CDS for a Poissonian network of node density $\hat{n}$ is given by $\{\frac{1}{\hat{n}-1} - \frac{\hat{n}}{\hat{n}-1}\exp(-(\hat{n}-1))\}N$. The second result is of paramount importance for practitioners. It provides the information about the expected size of a CDS when the node density $\hat{n}$ is between 6 and 30. The data (expected CDS size) for this range can hardly be provided by simulations.

*Keywords:* connected dominating set (CDS), expected size, lower bound, restricted pruning rule, wireless mobile ad hoc network

## 1. INTRODUCTION

In wireless ad hoc networks, the selecting of a subset of nodes (*i.e.*, construction of a virtual backbone) for efficient message routing is always a crucial problem. In the last decade, much research has been conducted in order to develop a simple and yet efficient algorithm for the construction of such a virtual backbone. Amongst them, distributed algorithms based on the idea of connected dominating set (CDS) [4] have been proposed and have succeeded in generating a virtual backbone of small size [2, 10-13]. In these algorithms, a CDS is constructed by going through two processes, namely the *marking process* and the *pruning process*. In the marking process, a node will mark itself *true* if it

has two unconnected neighbors. Otherwise, it will mark itself *false*. Once the marking process has finished, each *true* node will check if its local condition fulfills the conditions specified by a pruning rule-*k*. With respect to the pruning rule-*k*, a marked node unmarks itself if there exists *a set of connected nodes* whose coverage can cover all its neighbors and, at the same time, the ID of the marked node is smaller than the IDs of the connected nodes. If the connected nodes are restricted to direct neighbors of the marked node, the pruning rule is called *restricted rule-k*.[1] Otherwise, it is called *unrestricted rule*.

Although the Wu & Li decentralized algorithm is simple and efficient in terms of computational complexity, little theoretical work has been done concerning the size of the CDS being generated. Only Dai & Wu in [2] and Hansen *et al.* in [5] have provided analytical studies on this issue. Let $N$ be the total number of nodes and $N \to \infty$, Dai & Wu showed that the size of a pruning rule $k$ CDS is of constant-times-larger than the minimal CDS. Hansen *et al.* considered the situation that the size of the square (say $L_N^2$) grows linearly with $N$. The expected size of a CDS derived by the restricted pruning rule-*k* is in an order linear to $L_N^2$ and lower bounded by $L_N^2 / \pi$.

As observed from the simulated results presented in [2], this lower bound does not fit the cases when the node density is low. In this regard, an in-depth investigation about the expected size of a CDS generated by restricted pruning rule-*k* is inevitable. In particular, we would like to investigate *how the size changes with the node density*, and *when it reaches the lower bound as derived in* [5].

To do so, we need to derive the expected size of a CDS in terms of the node degree distribution, and the probability that a *true* node will turn out to be marked false during the pruning process. We call the latter probability the *unmarked probability*. Here, node degree is defined as the total number of neighbors a node has. Unfortunately, we will state later in the text that this unmarked probability cannot be obtained analytically. The *random sampling* technique is needed to determine these values numerically. Therefore, the contributions of this paper are as follows.

1. For a network of $N$ nodes that are uniformly and randomly generated in a square of size $L_N \times L_N$, the node degree distribution follows a Poisson distribution when $L_N$, $N \to \infty$.

2. A procedure based on the idea of random sampling is proposed and the unmarked probabilities against different node degrees are obtained.

3. It is found that the expected size of a CDS is almost a decreasing function with respect to the node density. The size of the CDS reaches its lower bound when the node density is greater than or equal to 30.

4. The lower bound on the expected size of a CDS for a Poissonian network of node density $\hat{n}$ is given by $\{\frac{1}{\hat{n}-1} - \frac{\hat{n}}{\hat{n}-1}\exp(-(\hat{n}-1))\}N$.

The third result is of paramount importance for practitioners. It provides the information about the expected size of a CDS when the node density $\hat{n}$ is between 6 and 30. The expected CDS size for this range can hardly be provided by simulations.

The rest of the paper is organized in four sections. In the next section, the algorithm of the marking process and the restricted pruning rule will be presented. The node degree distribution of a network of randomly deployed nodes will be analyzed and presented in

---

[1] In this paper, the terms *restricted rule-k* and *restricted pruning rule-k* are used interchangeably.

section 3. An empirical procedure to obtain the unmarked probability and the analysis on the expected size will be elucidated in the same section. Finally, the conclusion is presented in section 4.

## 2. RESTRICTED PRUNING RULE-*k*

Consider a mobile ad hoc network of $N$ nodes that are randomly and uniformly deployed within a two-dimensional square of area $L \times L$. Because of the transmission power of a radio signal, two nodes can communicate with each other if their distance apart is less than an allowable transmission range, say $r$ ($r \ll L$). In other words, two nodes are neighbors if the distance between them is less than $r$.

Once a node has been deployed, (i) it generates a uniformly random ID for itself and broadcasts to other nodes nearby (if any) about its ID. Then, (ii) it waits and listens to the signals from nearby nodes about their IDs and the IDs of their neighbors. In accordance with the received list of IDs, the node can check whether its ID is unique. If the ID already exists, the node will generate another random ID and then repeat steps (i) and (ii). (iii) As long as the IDs have been received, it updates the list of the IDs of its neighbors and broadcasts this neighbor information to its neighbors. The listen-update-broadcast cycle is then repeated a few more times until there are no more updates on the neighbor list. The resultant network graph is denoted by $V$.

When a complete list of neighbor information has been obtained, each node carries out the following algorithm to determine whether it is a gateway node for message routing. Let $id(x)$ and $\mathcal{N}(x)$ be the ID and the set of neighbor nodes of a node located at $x$. The marker of $x$ is denoted by $\mathcal{M}(x)$.

**Wu-Li Marking Process [13]:** A node located at $x$ sets its marker to *True*, *i.e.* $\mathcal{M}(x) = T$, if there exists two unconnected neighbor nodes.

**Dai-Wu Restricted Pruning Rule $k$ [2]:** A marked node unmarks itself if its neighbor nodes can be covered by a set of connected neighbor nodes whose IDs are larger than node $x$. That is to say, $\mathcal{M}(x) = F$, if there exists $x_1, x_2, \ldots, x_k \in \mathcal{N}(x)$ such that

(i)   $\mathcal{M}(x) = T$.
(ii)  $id(x) < id(x_i)$ for all $j = 1, 2, \ldots, k$.
(iii) $\mathcal{N}(x) \subset \mathcal{N}(x_1) \cup \mathcal{N}(x_2) \ldots \cup \mathcal{N}(x_k)$.
(iv) $x_1, x_2, \ldots, x_k$ form a connected graph.

To realize step (ii) in practice, each marked node first sorts the IDs of its neighbors in ascending order. Then, all nodes with IDs larger than $id(x)$ will be selected. The selected neighbor nodes are further checked with respect to their coverage (step (iii)) and connectivity (step (iv)). Finally, $\mathcal{M}(x)$ changes to $F$ if the selected nodes form a connected graph and can cover all the neighbors of $x$.

The beauty of this pruning rule is that the algorithm is completely distributed. Only direct neighbor information is needed. No global information is required. Each node can perform the pruning locally. Plus, the computational complexity is small. For a node of

degree $d$, its complexity is in an order of $\mathcal{O}(d^2)^2$. Consider a graph of finite-mean-node degree (say $\mu$) and variance (say $\sigma^2$). It can be shown by the Chebyshev Inequality that

$$P(|d - \mu| \leq m\sigma) \leq \frac{1}{m^2}.$$

In other words, for a finite $m$, pruning $(1 - m^{-2})$ nodes has a complexity of just $\mathcal{O}(m^2 \sigma^2)$. The cost paid for conducting the marking and pruning process is low.

## 3. ANALYSIS

Assume a network of $N$ nodes, *i.e.* $|V| = N$. Let $P(d)$ be the node degree distribution of $V$. $P(\mathcal{M}(x) = F \mid deg(x) = d)$ is the probability that a node of degree $d$ is unmarked after the pruning step. During the marking process, a node will be marked if there are two neighbor nodes of $x$ that are not neighbors to each other. As nodes are deployed uniformly and randomly, the probability that a node of degree $d$ will be marked in the marking process will be given by

$$P(\text{Node } x \text{ is marked} \mid deg(x) = d) = 1 - \beta^{d(d-1)/2}, \tag{1}$$

where $\beta$ is the probability that the distance of any two random nodes within a unit circle is less than or equal to the radius.

By conducting a computer simulation that generates 10,000 points uniformly and randomly in a circle of radius $r$, and then counts the percentage of pairs of nodes whose separation is less than $r$, it finds that $\beta$ is equal to 0.5852. Then

$$P(\text{Node } x \text{ is marked} \mid deg(x) = d) = 0.995$$

for $d = 5$. Fig. 1 shows the probability that a node of degree $d$ is not marked during the marking process. Clearly, one can assume that this probability vanishes when $d > 6$.
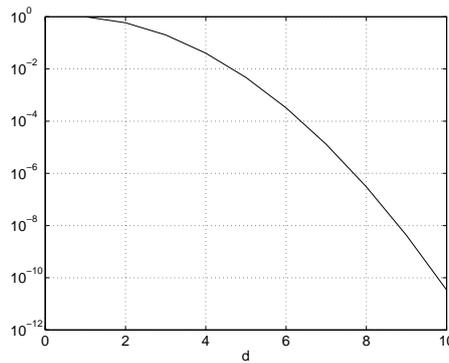


Fig. 1. The probability that a node of degree $d$ is not marked during the marking process, *i.e.* $\beta^{d(d-1)/2}$ versus $d$.

---

[2] Theorem 4 in [2].

Suppose a network graph is of Poisson node degree distribution with large mean node degree. The percentage of nodes of small node degree will be very small. One can thus assume that all nodes are marked after the initial marking process has been performed. The expected size of the CDS can be given by

$$E\left[|V_{cds}|\right] = N(1 - \sum_{x} \sum_{d} P\left(\mathcal{M}(x) = F \,|\, deg(x) = d\right) P(deg(x) = d)). \qquad (2)$$

As $P(deg(x) = d)$ is homogenous for all $x \in V$, the expected size can simply be expressed as follows:

$$E\left[|V_{cds}|\right] = N(1 - \sum_{d} P\left(\mathcal{M}(d) = F \,|\, d\right) P(d)), \qquad (3)$$

where the factor $P(\mathcal{M}(d) = F \,|\, d)$ corresponds with to the probability that a node of degree $d$ is unmarked.

### 3.1 Node Degree Distribution *P(d)*

Suppose the nodes are randomly and uniformly distributed, and let $\hat{n}$ be the average number of nodes within a unit disk. The average node degree $\lambda$ will thus be $\hat{n} - 1$. The node degree distribution of $V$ follows a Poisson distribution.

**Theorem 1**    For a mobile ad hoc network $V$, in which the mobile nodes are randomly and uniformly distributed, the node degree distribution $P(d)$ is given by

$$P(d) = \exp(-\lambda)\frac{\lambda^d}{d!}, \lambda = \hat{n} - 1, \qquad (4)$$

where $\hat{n}$ is the average node density.

***Proof:*** Let $N$ be the total number of nodes of $V$, and the area of deployment is much larger than a unit disk. The number of nodes deployed within a unit disk will then follow a binomial distribution,

$$P\{\text{Exactly } n \text{ nodes in a unit disk}\} = \frac{N!}{n!(N-n)!}\delta^n (1-\delta)^{(N-n)}$$

where

$$\delta = \frac{\text{Area of unit disk}}{\text{Deployment Area}}.$$

For large $N$, $\hat{n} = N\delta$ and

$$P\{\text{Exactly } n \text{ nodes in a unit disk}\} = \exp(-\hat{n})\frac{\hat{n}^n}{n!}.$$

Therefore, the probability of a node having node degree $d$ (*i.e.*, the number of neighbor nodes) is given by a Poisson distribution with average node degree $\lambda = \hat{n} - 1$. ❑
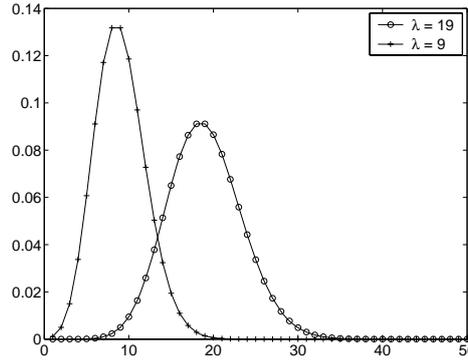
Fig. 2. The node degree distributions of $V$ for which the node densities are 20 ($\lambda = 19$) and 10 ($\lambda = 9$) respectively. The y-axis corresponds to value of $P(d)$, while the x-axis corresponds to the node degree $d$.

For illustration, Fig. 2 shows two examples in which $\hat{n}$ equals 20 and 10 respectively.

### 3.2 Unmarked Probability $P(\mathcal{M}(d) = F \,|\, d)$

Recall that a marked node $x$ unmarks itself if there exists $x_1, x_2, \ldots, x_k \in \mathcal{N}(x)$ such that

(i) $\mathcal{M}(x) = T$.
(ii) $id(x) < id(x_j)$ for all $j = 1, 2, \ldots, k$.
(iii) $\mathcal{N}(x) \subset \mathcal{N}(x_1) \cup \mathcal{N}(x_2) \ldots \cup \mathcal{N}(x_k)$.
(iv) $x_1, x_2, \ldots, x_k$ form a connected graph.

Consider the condition (i). As we have assumed that all the nodes are marked,

$P(\mathcal{M}(x) = T) = 1, \forall x \in V$.

Consider the condition (ii). For a node of degree $d$, it might have 1 neighbor node, 2 neighbor nodes, 3 neighbor nodes and so on, up to $d$ neighbor nodes that have IDs larger than itself. Since all node IDs are uniformly and randomly generated in a constant range, say [0, 1], the probability that $id(x) < id(y)$ for all $y \in \mathcal{N}(x)$ is given by

$$P(id(x) < id(y) \,|\, y \in \mathcal{N}(x)) = \int_0^1 (1-u)du = \frac{1}{2}.$$

As a result, the probability that exactly $k$ ($k \le d$) neighbor nodes that have larger IDs is given by

$$P(\text{Exactly } k \text{ out of } d \text{ neighbors having larger IDs}) = \binom{d}{k}\left(\frac{1}{2}\right)^d, \qquad (5)$$

for all $k = 0, 1, 2, \ldots, d$.

The final question left behind is this: *If there are k neighbor nodes with larger IDs, will these nodes form a connected graph, and simultaneously will the rest of the other d − k nodes be neighbors of these nodes?* Unfortunately, it is not an easy question to answer. It all depends on the locations of these $d$ neighbor nodes. Take a look at the illustrative examples shown in Fig. 3. In both cases, $k = 6$. Even though both sets of neighbor nodes can cover the whole circle, one is connected (Fig. 3 (a) and the other is disconnected (Fig. 3 (b)).
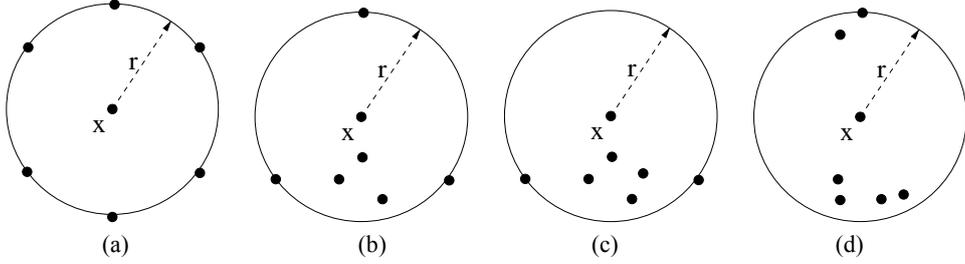


|     (a)     |     (b)     |     (c)     |     (d)     |

Fig. 3. (a) and (b) show two illustrative examples in which the neighbor nodes of *x* can cover the whole circle. (a) Six neighbor nodes are located evenly on the circumference of the circle. They form a connected graph. (b) Three neighbor nodes are located evenly on the circumference and three other nodes are located at the lower half of the circle. The graph being induced from these neighbor nodes is a disconnected graph. (c) and (d) show two illustrative examples in which the neighbor nodes of *x* cannot cover the whole circle. Again, one forms a connected graph (c) and the other does not (d).

Let $\Omega(x)$ be the unit circle centered at location $x$. Let $X = (x_1, x_2, \ldots, x_k) \in \Omega(x)^k$ be an augmented random vector, in which $x_1, x_2, \ldots, x_k \in \Omega(x)$. The graph induced by $X$ is denoted by $G_X$. Furthermore, we let $I(X)$ be an indicator function defined as follows:

$$I(X) = \begin{cases} 1 & \text{if } G_X \text{ is connected,} \\ 0 & \text{if } G_X \text{ is not connected.} \end{cases} \tag{6}$$

The coverage of $X$ is denoted by $Cov(X)$, where

$$Cov(X) = \frac{\text{Area covered by } \bigcup_{j=1}^{k} (\Omega(x_j) \cap \Omega(x))}{\text{Area covered by } \Omega(x)}. \tag{7}$$

Therefore, the probability that $(d − k)$ random nodes in $\Omega(x)$ can be covered by the other $k$ random nodes in $\Omega(x)$ will be given by

$$P(k, d) = \int_{X \in \Omega(x)^k} I(X) Cov(X)^{d-k} \, dX. \tag{8}$$

The probability that a node of degree $d$ will be unmarked will thus be given by

$$P(\mathcal{M}(d) = F \mid d) = \sum_{k=1}^{d} P(k, d) \binom{d}{k} \left(\frac{1}{2}\right)^d \tag{9}$$

and the expected size of CDS is given by

$$E\left[|V_{cds}|\right] = N\left(1 - \sum_d \sum_{k=1}^{d} P(k,d)\frac{\exp(-\lambda)(\lambda/2)^d}{k!(d-k)!}\right). \tag{10}$$

Unfortunately, there is no simple closed-form solution for the probability $P(k, d)$, Eq. (8). We obtained the values empirically by a *random sampling* procedure, as depicted in Fig. 4.

The idea of the procedure can be sketched as follows. In the first step, we generate $Z$ random nodes, $x_1, x_2, \ldots, x_Z$, within a unit disk centered at the origin (step 1). In the second step, for each value of node degree, say $k$, we generate another $k$ random nodes, $y_1$, $y_2, \ldots, y_k$, within the same unit disk (step 2.1.1). Then, we count the *fractional number* of $x_i$s being covered by the $y_1, y_2, \ldots, y_k$ and store the value in the array $OL$ (steps 2.1.2 and 2.1.3). Next, the connectivity of the graph induced by $y_1, y_2, \ldots, y_k$ is checked (step 2.1.4). Finally the fractional number counted in the step 2.1.3 will be stored in an array $PC$ if the graph is connected. Otherwise, the stored value will be set to zero. The second step is repeated $M$ times. In our simulation, $Z$ is set to 10,000. The value of $k$ ranges from 1 to 25, and $M$ is set to 10,000. So, before the simulation, we initialize three 2D arrays ($OL$, $CN$ and $PC$) of dimension $25 \times 10,000$. Their $kj$th elements, where $k = 1, 2, \ldots, 25$ and $j = 1, 2, \ldots, 10,000$, correspond to the intermediate results obtained in the $j$th simulation for node degree $k$.

---

**Step 0:** Initialize $OL$, $CN$, $PC \in R^{25 \times 10,000}$.
**Step 1:** Let $\Omega_0$ be the unit disk centered at $(0, 0)$ and then uniformly randomly generate $x_1, x_2, \ldots, x_{10,000}$ inside $\Omega_0$.
**Step 2:** For $k = 1, 2, \ldots, 25$,
    **2.1** For $j = 1, 2, \ldots, 10,000$
    **2.1.1** Uniformly randomly generate $y_1, y_2, \ldots, y_k$ inside $\Omega_0$,
    **2.1.2** Set $NI$ equals the number of $x_i$s that are located inside $\bigcup_{\kappa=1}^{k}\Omega(y_\kappa)\bigcap\Omega_0$.
    **2.1.3** Set $OL_{kj}$ equals $NI/10,000$.
    **2.1.4** Set $CN_{kj}$ equals 1 if $y_1, \ldots, y_k$ form a connected graph.
    **2.1.5** Set $PC_{kj}$ equals $OL_{kj} \times CN_{kj}$.

---

Fig. 4. Random sampling procedure for obtaining the probability $P(k, d)$, Eq. (8).

Since the $kj$th element in the array $PC$ is the value $I(X)Cov(X)$ of the $j$th set of random $k$ nodes, the value $P(k, d)$ can then be obtained empirically by

$$P(k,d) = \frac{1}{M}\sum_{j=1}^{M} PC_{kj}^{d-k} \tag{11}$$

for all $k \leq d$. The unmarked probability of a node of degree $d$ can be obtained. Fig. 5 shows the unmarked probability $P(\mathcal{M}(d) = F \mid d)$ against node degree $d$.

It is clear that the minimum unmarked probability is attained when $d$ equals 5, where the minimum unmarked probability is equal to 0.3722. (This is due to the fact that there is a small chance for a 5-node induced graph to form a connected induced graph.)
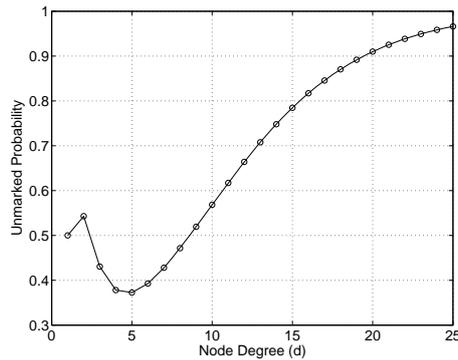
Fig. 5. Unmarked probability.

The unmarked probability reaches 0.9661 when $d = 25$. It can be further noted from the figure that the value of $P(\mathcal{M}(d) = F \,|\, d)$ increases as $d$ increases, and then approaches 1 when $d$ is large.

### 3.3 Expected Size of CDS

In accordance with the formulae derived earlier for the average number of marked nodes (Eq. (3)) and the theorem about the nature of node degree distribution (Theorem 1), the expected size of a CDS-derived restricted pruning rule can be expressed as follows:

$$\frac{E\left[\left|V_{cds}\right|\right]}{N} = \left(1 - \exp(-\lambda)\left\{\sum_d P(\mathcal{M}(d) = F \,|\, d)\frac{\lambda^d}{d!}\right\}\right), \tag{12}$$

where $\lambda$ corresponds to the average node degree. Then, the expected size of a CDS derived is evaluated by putting the values of $P(\mathcal{M}(d) = F \,|\, d)$ as shown in Fig. 5 and different values of $\lambda$ into the Eq. (12). Fig. 6 shows the expected size of CDS against $\lambda$. (For convenience, we simply let $P(\mathcal{M}(d) = F \,|\, d) = 0.9661$ for $d > 25$.) The solid line with squares corresponds to the lower bound $(\lambda + 1)^{-1}$, while the solid line with circles corresponds to our results. The solid line with squares corresponds to the lower bound $(\lambda + 1)^{-1}$. (Please refer to Appendix A for the derivation of this lower bound.) It is observed that the size is about 0.55% of the original network size when $\lambda = 6$. The factor matches the result obtained in [2] for the same $\lambda$ and $N = 200$. (Please refer to Appendix B for the reason why the comparison is only made for $\lambda = 6$, not for other values of $\lambda$ in their paper.) In accordance with Fig. 6, one can also see that the size of a CDS drops as the $\lambda$ increases. Eventually, it drops to its lower bound when $\lambda$ is close to 30.

### 3.4 A Tighter Lower Bound

A tighter lower bound for Poissonian $P(d)$ can indeed be derived from the Eq. (12). Consider a marked node of degree $d$. One condition that a marked node will be staying marked, after the pruning process, is when its ID is larger than all its neighbors. This probability is given by $(d + 1)^{-1}$ for a marked node with $d$ neighbors. Hence,
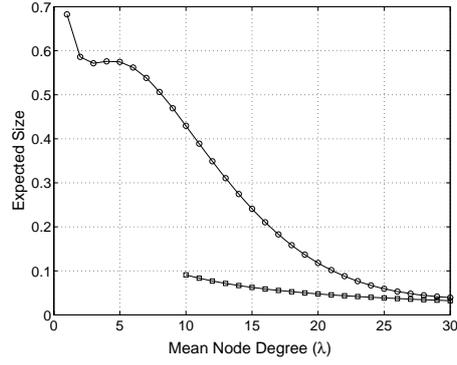
Fig. 6. Expected size of a CDS derived by restricted pruning rule. The solid line with squares cor-
responds to the lower bound $(\lambda + 1)^{-1}$, while the solid line with circles corresponds to our
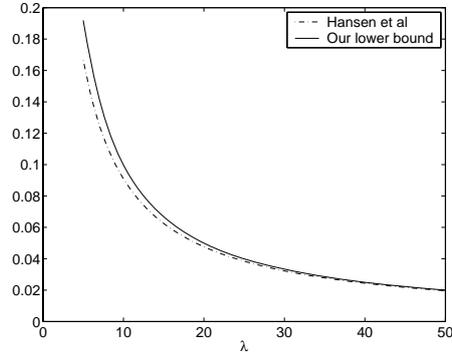results.



Fig. 7. Comaprison between Hansen *et al.* and our lower bounds. The y-axis corresponds to the
value $E[|V_{cds}|]/N$, while the x-axis corresponds to node density of $V$.

$$\frac{E\left[|V_{cds}|\right]}{N} \geq \exp(-\lambda) \sum_{d \geq 1} \frac{\lambda^d}{(d+1)!}.$$

Since

$$\sum_{d \geq 1} \frac{\lambda^d}{(d+1)!} = \frac{\exp(\lambda) - 1 - \lambda}{\lambda},$$

$$E\left[|V_{cds}|\right] \geq \left\{ \frac{1}{\lambda} - \frac{\lambda+1}{\lambda} \exp(-\lambda) \right\} N.$$

In terms of node density $\hat{n}$,

$$E\left[|V_{cds}|\right] \geq \left\{ \frac{1}{\hat{n}-1} - \frac{\hat{n}}{\hat{n}-1} \exp(-\hat{n}+1) \right\} N. \tag{13}$$

It is equal to $N/(\hat{n}-1)$ when $\hat{n}$ is large and this bound is tighter than $N/\hat{n}$ for all

$\hat{n} \geq 4$ (*i.e.* $\lambda \geq 3$). Fig. 7 compares the difference between the Hansen *et al.* lower bound and our lower bound. It is clear that there is no significant difference when $\lambda$ is large.

## 4. CONCLUSION

In this paper, we have provided an analysis on the size of a CDS derived by the restricted pruning rule-*k* algorithm. For a network of *N* nodes that are uniformly and randomly generated in a square of size $L_N \times L_N$, we have shown that the node degree distribution follows a Poisson distribution when $L_N$, $N \to \infty$. To argue that the node degree distribution of the network does not change much after the marking process has been performed; we have discussed the probability of a node being marked in the marking process and shown that such a probability tends to vanish when the node density is high. After that, we have derived an equation to evaluate the expected size of a CDS, in terms of the network node degree distribution and the unmarked probabilities. As there is no closed-form solution for the connected probability and the coverage of a graph induced by random nodes within a circle, a computer simulated procedure based on the idea of random sampling has been developed to obtain such probabilities. The probability that a node of degree *d* will be unmarked is obtained and hence the expected size of a CDS can be obtained. Finally, the expected size of a CDS derived by the restricted pruning rule-*k* is analyzed with respect to different node densities. It is found that the size is almost a decreasing function with respect to the node density. The size reaches its lower bound when the node density is equal to or greater than 30. That is to say, the CDS derived by the restricted pruning rule-*k* algorithm in a high node density situation is a minimal CDS. The results are consistent with the existing results previously obtained in [2] and [5]. More important, our results have filled in the gap, $6 \leq \lambda \leq 30$, that has not been investigated before. By applying a similar technique, analysis on other distributed methods, such as the extended works presented in [10, 14, 15], for constructing CDS might also be possible.

## REFERENCES

1. B. N. Clark, C. J. Colbourn, and D. S. Johnson, "Unit disk graphs," *Discrete Mathematics*, Vol. 86, 1990, pp. 165-177.
2. F. Dai and J. Wu, "An extended localized algorithm for connected dominating set formation in ad hoc wireless networks," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 15, 2004, pp. 1-13.
3. B. Das, R. Sivakumar, and V. Bhargavan, "Routing in ad hoc networks using a spine," in *Proceedings of the 6th IEEE International Conference on Computers Communication and Networks*, 1997, pp. 34-39.
4. S. Guha and S. Khuller, "Approximation algorithms for connected dominating sets," *Algorithmica*, Vol. 20, 1998, pp. 374-387.
5. F. Hansen, E. Schmutz, and L. Sheng, "The expected size of the rule *k* dominating set," *Algorithmica*, Vol. 46, 2006, pp. 409-418.
6. F. Ingelrest, D. Simplot-Ryl, and I. Stojmenovic, "A dominating sets and target radius based localized activity scheduling and minimum energy broadcast protocol for

ad hoc and sensor networks," in *Proceedings of the Mediterranean Ad Hoc Networking Workshop*, 2004, pp. 351-359.

7. C. R. Lin and M. Gerla, "Adaptive clustering for mobile wireless networks," *IEEE Journal on Selected Areas in Communication*, Vol. 15, 1996, pp. 1265-1275.

8. H. Liu, Y. Pan, and J. Cao, "An improved distributed algorithm for connected dominating sets in wireless ad hoc networks," in *Proceedings of the International Symposium on Parallel and Distributed Processing and Applications*, 2004, pp. 340-351.

9. P. J. Wan, K. Alzoubi, and O. Frieder, "Distributed construction of connected dominating set in wireless ad hoc networks," in *Proceedings of IEEE INFOCOM*, Vol. 3, 2002, pp. 1597-1604.

10. J. Wu, "Extended dominating-set-based routing in ad hoc wireless networks with unidirectional links," *IEEE Transactions on Parallel & Distributed Systems*, Vol. 13, 2002, pp. 866-881.

11. J. Wu and F. Dai, "Mobility control and its applications in mobile ad hoc networks," *IEEE Network*, Vol.18, 2004, pp. 30-35.

12. J. Wu and F. Dai, "Efficient broadcasting with guaranteed coverage in mobile ad hoc networks," *IEEE Transactions on Mobile Computing*, Vol. 4, 2005, pp. 259-270.

13. J. Wu and H. Li, "On calculating connected dominating set for efficient routing in adhoc wireless networks," in *Proceedings of the 3rd International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communication*, 1999, pp. 7-14.

14. J. Wu, W. Lou, and F. Dai, "Extended multipoint relays to determine connected dominating sets in MANETs," *IEEE Transactions on Computers*, Vol. 55, 2006, pp. 334-347.

15. J. Wu, B. Wu, and I. Stojmenovic, "Power-aware broadcasting and activity scheduling in ad hoc wireless networks using connected dominating sets," *Wireless Communications and Mobile Computing*, Vol. 4, 2003, pp. 425-438.

## APPENDIX

## A. Hansen *et al.* LOWER BOUND [5]

Instead of running extensive computer simulations, Hansen *et al.* have presented a theoretical analysis on the size of a CDS derived by the restricted pruning rule in [5]. In one of their theorems (Theorem 5 in [5]), they show that the size of a CDS is lower-bounded by $l_N^2 / \pi$ for $l_N \to \infty$. Here $l_N$ is the length of the square where the mobile nodes are deployed. For an ad hoc network consisting of $N$ nodes, $l_N^2 / \pi$ is equal to the total number of nodes over the node density. As node density is equal to the average node degree plus 1 (*i.e.* $\hat{n} = \lambda + 1$), the lower bound of the expected size of a CDS derived by the restricted pruning rule depends on the average node degree of the Poissonian node degree distribution:

$$E\left[\left|V_{cds}\right|\right] \ge \frac{N}{\lambda + 1}. \tag{14}$$

## B. Dai &Wu RESULT [2]

In our analysis, we assume the network is of Poisson node degree distribution. For a network of $N$ nodes deployed in a square of size $L \times L$, and each node has transmission range $r$, the Poisson node degree distribution happens when $r \ll L$ and $\frac{r}{L} = \sqrt{\frac{\lambda + 1}{\pi N}}$. This condition is equivalent to $\lambda \ll N$ for when $L$ is finite. Therefore, the node degree distribution is close to a Poisson distribution only when $\lambda$ is small. The simulated results in [2] for the expected size of a CDS at $\lambda = 6$ is consistent the results obtained in this paper. On the contrary, the node degree distribution of a large $\lambda$ network could hardly follow a Poisson distribution. A comparison between their results and the results presented here could not be made.

**John Sum (沈培輝)** received the B.Eng. in Electronic Engineering from Hong Kong Polytechnic University in 1992. Then, he received his M.Phil. and Ph.D. in Computer Science and Engineering from Chinese University of Hong Kong in 1995 and 1998 respectively. Before joining the Chung Shan Medical University (Taiwan), John was teaching in several universities in Hong Kong, including the Hong Kong Baptist University, the Open University of Hong Kong and the Hong Kong Polytechnic University. Currently, he is an assistant professor of the Department of Information Management in the Chung Shan Medical University, Taichung, R.O.C. His current research interests include neural computation, mobile sensor networks and scale-free network. John Sum is a senior member of IEEE and an associate editor of the International Journal of Computers and Applications.

**Jie Wu (吳杰)** is a Distinguished Professor at the Department of Computer Science and Engineering, Florida Atlantic University. He has published over 350 papers in various journals and conference proceedings. His research interests are in the areas of wireless networks and mobile computing, routing protocols, fault-tolerant computing, and interconnection networks. Dr. Wu was on the editorial board of IEEE Transactions on Parallel and Distributed Systems and was a co-guest-editor of IEEE Computer and Journal of Parallel and Distributed Computing. He served as the program co-chair for MASS 2004, program vice-chair for ICDCS 2001, and program vice-chair for ICPP 2000. He was also general co-chair for MASS 2006 and is general vice-chair for IPDPS 2007. He is the author of the text "Distributed System Design" published by the CRC press. He was also the recipient of the 1996-97 and 2001-2002 Researcher of the Year Award at Florida Atlantic University. Dr. Wu has served as an IEEE Computer Society Distinguished Visitor and is the Chairman of IEEE Technical Committee on Distributed Processing (TCDP). He is a Member of ACM and a Senior Member of IEEE.

**Kevin I-J Ho (何英治)** received the B.S. in Computer Engineering from National Chiao Tung University (Taiwan) in 1983. From 1985 to 1987, he was an assistant engineer of Institute of Information Industry, Taiwan. Then, he received his M.S. and Ph.D. in Computer Science from University of Texas at Dallas in 1990 and 1992, respectively. Currently, he is an associate professor of Department of Computer Science and Communication Engineering, Providence University, Taiwan. His current research interests include image processing, algorithm design and analysis, scheduling theory, and computer networks.