On the Minimum Number of Neighbours for Good Routing Performance in MANETs

Oday D. Jerew
School of Engineering, CECS
The Australian National University
Canberra ACT 0200, Australia
Email: Oday.Jerew@anu.edu.au

Haley M. Jones
School of Engineering, CECS
The Australian National University
Canberra ACT 0200, Australia
Email: Haley.Jones@anu.edu.au

Kim L. Blackmore
School of Engineering, CECS
The Australian National University
Canberra ACT 0200, Australia
Email: Kim.Blackmore@anu.edu.au

***Abstract*—**
In a mobile ad hoc network, where nodes are deployed without any wired infrastructure and communicate via multihop wireless links, the network topology is based on the nodes' locations and transmission ranges. The nodes communicate through wireless links, with each node acting as a relay when necessary to allow multihop communications. The network topology can have a major impact on network performance. We consider the impact of number and placement of neighbours on mobile network performance. Specifically, we consider how neighbour node placement affects the network overhead and routing delay. We develop an analytical model, verified by simulations, which shows widely varying performance depending on source node speed and, to a lesser extent, number of neighbour nodes.

***Index Terms*—MANET performance, neighbour node location, mobile routing, network mobility**

I. Introduction

A mobile ad hoc network (MANET) is an autonomous system of mobile nodes that does not need any fixed infrastructure. The mobile nodes use wireless transceivers to communicate with each other, and communication between distant nodes is achieved using a sequence of intermediate nodes, called relays, to forward the packets from sender to receiver. Much MANET research focuses on the routing protocol, which determines the sequences of intermediate nodes [18].

Most MANET routing protocols rely on the network topology, which is determined by node location and the nodes' communication transmission ranges. In many protocols, full knowledge of the network topology is not available, so each node collects neighbourhood information through periodic, asynchronous messages. This localized information is usually based on 1-hop information (i.e., location information of all immediate neighbours).

Routing protocols generally assume that the network is fully connected. Various approaches, such as continuum percolation [3], throughput maximization, and random graph theory [19], [13], are proposed for the evaluation of the minimum number of neighbours needed for full connectivity in a wireless network [8], [17], [4], [10], [22] . It was first proposed by Kleinrock and Silvester in [8] that six was the "magic number", i.e., on average every node should connect itself to its six nearest neighbours, and various papers since then have argued for magic numbers between five and eight [15], [5], [23], [14]. Doci et al. [2] introduced maximum node degree as a mobility metric, which represents the maximum number of neighbours for each node and an algorithm is designed to compute that metric.

Topology control in MANETs generally refers to selecting an appropriate transmission power for each node in order to reduce energy consumption and signal interference without impeding performance. Typically, each node selects a few logical neighbours from its 1-hop neighbours within the normal transmission range, and the (smaller) actual transmission range of each node is set to be the distance to its farthest logical neighbour. The schemes are designed to satisfy global constraints, such as network connectivity, reduced channel contention and other reliability and throughput related measures [9], [20], [11]. Blough et al. [1] showed that connectivity is preserved with high probability (95 percent) if every node keeps nine neighbours.

The majority of these approaches assume a static network without mobility (even though they refer to Mobile Ad Hoc networks). One exception is Wu and Dai [21], where the logical neighbour set and transmission range are first computed from the neighbourhood information of each individual node, and then adjusted to compensate for node mobility. Again, this work focuses on achieving and maintaining network connectivity. Tian et al. [16] proposed a topology based model to describe the mobility of networks by means of link duration and connectivity. Yanmaz [24] also proposed a topology based mobility model, where static nodes in the network are assigned an importance metric. Mobile nodes move according to a type of random walk, such that the probability distribution for the direction depends on the locations of the static nodes. Therefore, the mobile nodes are likely to move towards or around more important nodes, based on their individual connectivities.

We contend that, even with a fully connected network, some topologies are superior to others, and that this is best evaluated via the standard routing protocol performance metrics, such as routing overhead and routing delay. Within this context, we aim to study the performance of dynamic network topologies.

In this paper we propose an analytical model to evaluate the effect of neighbour nodes on mobile network performance. Specifically, we explore how the number of neighbour nodes affects the network overhead and routing delay. To investigate

978-1-4244-5113-5/09/$25.00

this topic, the Destination Source Routing (DSR) protocol [7] is used for route establishment. While it is likely that different routing protocols will have different saturation levels and route characteristics, the results obtained with DSR can be generalised to most on-demand ad hoc routing protocols.

Our results show that for most speeds both overhead and delay experience diminishing returns for having increasing numbers of neighbours or, correspondingly, caching a greater number of routes. This becomes more pronounced with increasing mobility. We also show that greater benefit is derived from having more neighbours when the connectivity of the network is lower.

The remainder of this paper is organized as follows. In section II we introduce the geometry of the source node and its neighbours, developing a statistical framework for distances, locations and timing. In Section III we set up the mathematical framework for later developing analytical expressions for overhead and delay. In section IV we develop expressions for expected overhead and delay, for single and multiple routes stored in the source node cache, based on the frameworks set up in the previous sections. In Section V we present a comparison of theoretical and simulation results, with conclusions in Section VI.

II. Topology Scenario

In this paper we assume the "transmission range" model of signal transmission. That is, it is assumed that each node is equipped with an omnidirectional antenna and that signal attenuation is due only to path loss related to distance transmitted. We assume that the transmission ranges of all nodes are identical and equal to r. Consider, then, a source node, n_s, with N neighbours distributed on a circle of radius $r/2$, centred on n_s. We choose an initial separation distance of $r/2$ so that the nodes are close enough to be communicating and without their links breaking too soon, but not so close that neighbour location diversity is compromised.

We assume that the velocities of the neighbour nodes are slow enough to be considered stationary and that n_s moves with a constant speed v in a random direction θ_s. The scenario could also be extended so that the speeds of the neighbouring nodes are combined with the speed of n_s, such that v is the relative speed. This scenario fits well with the random waypoint mobility model or for scenarios where packet arrival times or the transmission range are small compared with node speeds. All node-to-node communications are assumed to be bi-directional.

In this section we derive or present expressions for the respective probability density functions (PDFs) and cumulative distribution functions (CDFs) of the link breakage distance, link residual time neighbour node distance to destination node and packet arrival time. We use these expressions later in the paper to evaluate the network routing performance as the number of neighbour nodes of the source node is varied.

A. Travelling Distance to Link Breakage

The source node/neighbour nodes scenario described above, is illustrated in Fig. 1. To determine when the link between the source node and a given neighbour node, n_i breaks, we let n_i be at angle θ_i on the radius $r/2$ neighbour circle. Then the distance that n_s has to travel from the centre of the circle, in direction θ_s, before the link with n_i breaks is given by

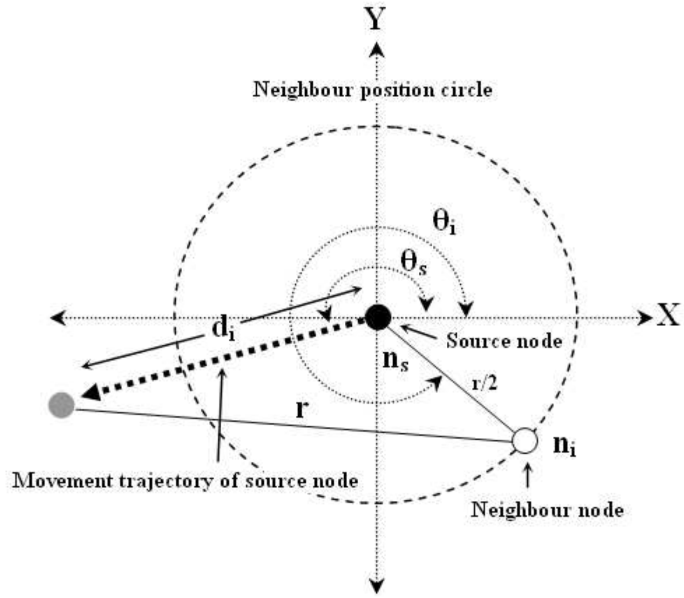

Fig. 1. Distance relationship between moving source node n_s and active neighbour node n_i, when n_s moves in a straight line in direction θ_s and n_i is on a circle of radius $r/2$, initially centred on n_s, at angle θ_i.

$$d_i = 0.5r\left(\cos(\theta_s - \theta_i) + \sqrt{\cos^2(\theta_s - \theta_i) + 3}\right). \quad (1)$$

Because θ_s and θ_i are random variables, d_i is also a random variable. Moreover, $r/2 \leq d_i \leq 3r/2$, where the limits are achieved when $\theta_s = \theta_i + \pi$ and $\theta_s = \theta_i$, respectively.

We assume that θ_s and θ_i are both uniformly distributed between 0 and 2π. Then it can be shown that the PDF of $d_{r,i} = d_i/r$ is given by

$$f_D(d_{r,i}) = \frac{1 + \frac{3}{4d_{r,i}^2}}{\pi\sqrt{1 - \left(d_{r,i} - \frac{3}{4d_{r,i}}\right)^2}} \quad (2)$$

and the CDF of $d_{r,i}$ is given by

$$F_D(d_{r,i}) = 1 - \frac{1}{\pi}\arccos\left(d_{r,i} - \frac{3}{4d_{r,i}}\right). \quad (3)$$

Finally, the expected value of $d_{r,i}$ is

$$E\{d_{r,i}\} = 1. \quad (4)$$

That is, the average distance that n_s has to travel from the centre of the circle of neighbours before the link with the active neighbour breaks is equal to the transmission range r, such that $d_{r,i} = d_i/r = 1$.

B. Link Residual Time

The time from when n_s moves from the centre of the circle to the point where the link with n_i breaks at a distance of d_i

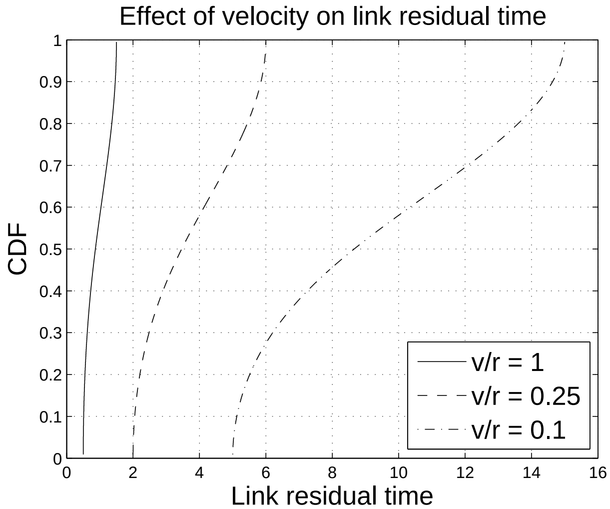

Fig. 2. The effect of velocity on the CDF of the link residual time, from (7)

is called the *link residual time*. The time required for n_s to move distance d_i at velocity v is given by

$$R_i = \frac{d_i}{v} = \frac{d_{r,i}}{v/r}. \quad (5)$$

Because $d_{r,i}$ is a random variable, the link residual time, R_i, is also a random variable. The PDF of R_i can be shown to be given by

$$f_{R_i}(r_i) = \frac{\frac{v}{r} + \frac{3r}{4vr_i^2}}{\pi\sqrt{1 - \left(\frac{vr_i}{r} - \frac{3r}{4vr_i}\right)^2}} \quad (6)$$

with the CDF of R_i given by

$$F_{R_i}(r_i) = 1 - \frac{1}{\pi}\arccos\left(\frac{vr_i}{r} - \frac{3r}{4vr_i}\right). \quad (7)$$

Finally, the expected value of R_i is

$$E\{R_i\} = \frac{r}{v}. \quad (8)$$

Using (7), we can study the effect of velocity on the link residual time as shown in Fig. 2 where CDF of R_i is plotted for different n_s speed to transmission radius ratios. It can be seen that for high speeds the link residual time is very small with a small variance. As the speed decreases, the link residual time increases with a corresponding increase in variance. The observations noted here will be used later to explain some seeming inconsistencies in the delay results.

C. Distance to Destination

We now consider the location of the destination node, n_d. Let the angle to the destination node be θ_d and the distance from the source node, n_s, to n_d be L_d, as shown in Fig. 3. Let the distance from neighbour n_i to the destination n_d be L_i. Then

$$L_i = \sqrt{\left(\frac{r}{2}\right)^2 + L_d^2 - L_d r\cos(\theta_d - \theta_i)} \quad (9)$$

Because θ_d and θ_i are random variables, L_i is also a random variable. As above, we assume that θ_d and θ_i are uniformly distributed between 0 and 2π. Then it can be shown that the PDF of L_i is given by

$$f_{L_i}(\ell_i) = \frac{2\ell_i}{\pi\sqrt{r^2L_d^2 - (\left(\frac{r}{2}\right)^2 + L_d^2 - \ell_i^2)^2}} \quad (10)$$

and the CDF of L_i is given by

$$F_{L_i}(\ell_i) = \frac{1}{\pi}\arccos\left(\frac{\left(\frac{r}{2}\right)^2 + L_d^2 - \ell_i^2}{rL_d}\right). \quad (11)$$

The minimum distance between the neighbour node and the destination is $L_{i,\min} = L_d - r/2$ when $|\theta_i - \theta_d| = 0$ and the maximum distance is $L_{i,\max} = L_d + r/2$ when $|\theta_i - \theta_d| = \pi$. Using (10), the expected value of the distance between the neighbour nodes and the destination is, then

$$E\{L_i\} = L_d. \quad (12)$$

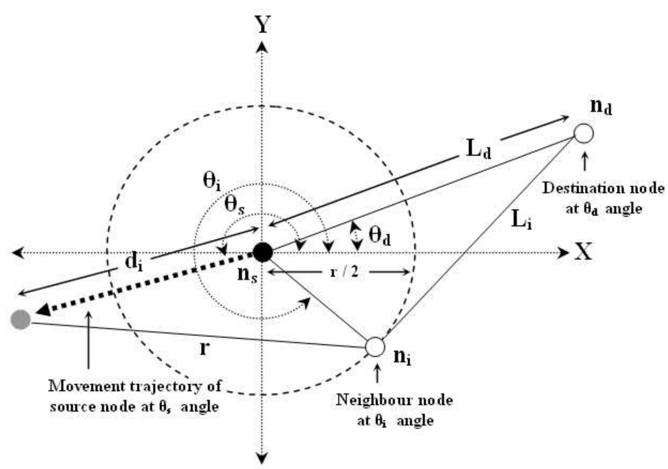

Fig. 3. Distance relationship between moving source node, n_s, active neighbour node, n_i, and destination node, n_d, when n_s moves in a straight line in direction θ_s, n_i is on a circle of radius $r/2$, initially centred on n_s, at an angle θ_i and n_d is outside the radius $r/2$ circle at a distance L_d from the circle centre, at an angle θ_d.

D. Packet Arrival Time

Packet arrival times, t_a, are generally modelled as having an exponential distribution, with parameter λ_a appropriate to the given network. We use this model here. The PDF of t_a is, therefore, given by

$$f_a(t) = \lambda_a e^{-\lambda_a t} \quad (13)$$

and the CDF of t_a is given by

$$F_a(t) = 1 - e^{-\lambda_a t}. \quad (14)$$

In this section we have presented a statistical model of the topology scenario we use for investigating the effects of neighbours on MANET performance. We will use this model to develop an analytical model of selected performance criteria in the coming sections.

III. ON-DEMAND ROUTING

In on-demand routing protocols, a source node attempts to discover a route to a destination only when it is presented with a packet for forwarding to that destination. Usually a route cache is also employed to avoid carrying out a route discovery for every new packet to the same destination. However, cached routes may become outdated if, for example, two nodes in the route have moved out of each others' transmission ranges. This stale data may degrade performance, so a process for removing old routes is employed.

Overhead models for on-demand routing protocols are generally based on the following principles [12].

- When a source node, n_s , initiates a new route discovery, the consequent route request (RREQ) packet is broadcast throughout the network. Any node receiving a duplicate RREQ discards the duplicate, so that each node is considered to have only dealt with each RREQ once.
- When a destination node, n_d, receives a RREQ from n_s, it returns to n_s a route reply (RREP) packet back along the route via which the RREQ arrived.
- If a node involved in forwarding a data packet along an established route determines that the link of which it is at the head is no longer valid, it returns an error packet (RERR) to n_s back along the route. Monitoring the correct operation of a route in use is referred to as route maintenance [7].

Note that in the following analyses, all network traffic is ignored except for the route discovery itself. This is to isolate the effects of numbers and positions of neighbours.

A. Routing Overhead

When the details of an established route are saved in the cache, a route expiry time, T, is determined and saved for that route. After this time it is assumed that the route is no longer valid. If n_s wishes to transmit to n_d after this time it must carry out a new route discovery process. That is, if we let t_a be the time of arrival of a new packet destined for n_d, then, if $t_a < T$ and the cached route is not broken, no routing overhead is incurred in sending the new data packet. However, if $t_a > T$ a route discovery process is automatically undertaken, using flooding of RREQs, as discussed above. If there are n nodes in the network, $n-1$ RREQ packets are transmitted during a flood (all nodes in the network receive the RREQ except for n_s), plus h RREPs where h is the number of hops between n_d and n_s in the route chosen by n_d as decided by the route of the first RREQ to reach n_d. This scenario assumes that only one route is cached in any route discovery process.

In summary, the number of overhead packets generated by an on-demand routing protocol in response to a packet arrival at time t_a is given by,

$$\mathrm{OH} = \begin{cases} 0 & t_a < T \\ n+h-1 & t_a > T. \end{cases} \tag{15}$$

Usually a third case, where $t_a < T$ but some later link in the cached route is broken, should be considered. However we discount this case because we wish to isolate the effect of neighbour nodes on routing performance. We assume that all links in the active route remain intact except for the first link, that between n_s and its active neighbour, n_i.

B. Routing Delay

The routing delay is the time taken to transmit any control packets necessary to establish a route, before sending a data packet between n_s and n_d. Since RREQ packets must propagate through the network to the destination, followed by the return of RREP packets, the routing delay is given by

$$D = \begin{cases} 0 & t_a < T \\ 2t_h h & t_a > T, \end{cases} \tag{16}$$

where h is the number of hops in the discovered path, and t_h is the time taken for transmission over a single hop. We assume t_h is a constant and the same for all hops in the network, and includes processing time at the nodes at either end of the hop as well as propagation time.

As each wireless transmission has a maximum range, r, there is a strong relationship between the distance, L_d, from n_s to n_d and the minimum hop count. All routes must obey [6]

$$\text{Number of hops in shortest path} \geq \frac{L_d}{r}. \tag{17}$$

Most ad hoc routing protocols utilise hop count in their route selection criteria. This approach minimises the total number of transmissions required to send a packet on the selected path. If, in addition, the network is dense, then the number of hops in the minimum hop path approaches L/r.

IV. ROUTING OVERHEAD AND DELAY

In this section we consider the expected overhead generated and the expected delay incurred with the occurrence of a new route request. In each case, the expected value of the overhead is equal to the cost of an individual route discovery process, multiplied by the probability that the route discovery is necessary. The probability of route discovery is determined by the topology scenario and the caching strategy.

We consider two different caching strategies at the source node. First, we assume that a single route to each destination is stored in the source node cache. Next, we assume that the source node cache holds multiple paths to each destination, each one commencing with one of its neighbour nodes.

A. One Route in Cache

1) Overhead: In this case we assume that the cache at node n_s stores a single route for each destination n_d, and that the route commences with neighbour i with probability $1/N$. That is, we assume that all of the neighbour nodes are equally likely to be the first hop on the active route.

In order to incur routing overhead, the route must have broken prior to the arrival of the next packet to send. That is, n_s must have reached the breaking distance with respect to n_i, such that $R_i < t_a$. Then the expected value of the overhead, given link residual times $\mathbf{R} = (R_1, R_2, \cdots, R_N)$ (i.e., given that we know the position of each of the neighbour nodes and the direction of travel of n_s) is given by

$$\begin{aligned}&E\{\mathrm{OH}|\mathbf{R}=\mathbf{r}\}\\&=\frac{(n+h-1)}{N}\sum_{i=1}^{N}[1-P(t_a<r_i)]\\&=\frac{(n+h-1)}{N}\sum_{i=1}^{N}[1-F_a(r_i)]\\&=\frac{(n+h-1)}{N}\sum_{i=1}^{N}e^{-\lambda_a r_i}.\end{aligned}\tag{18}$$

We have used the fact that neighbour nodes are independently placed, so link residual times are independent of each other.

The minimum distance travelled until a link breaks is $d_{i,\min}=r/2$ when $|\theta_s-\theta_i|=\pi$, such that $d_{r,i,\min}=1/2$ and the maximum travelled distance is $d_{i,\max}=3r/2$ when $\theta_s=\theta_i$, such that $d_{r,i,\max}=3/2$. The corresponding values of R_i are $R_{i,\min}=r/2v$ and $R_{i,\max}=3r/2v$. Using (6), and letting $v_r=v/r$, the expected value of the routing overhead is, then

$$\begin{aligned}&E\{\mathrm{OH}|\text{one cached route}\}\\&=\int_{R_{1,\min}}^{R_{1,\max}}\cdots\int_{R_{N,\min}}^{R_{N,\max}}E\{\mathrm{OH}|\mathbf{R}=\mathbf{r}\}f_{\mathbf{R}}(\mathbf{r})\,\mathbf{dr}\\&=\frac{(n+h-1)}{\pi}\int_{1/2v_r}^{3/2v_r}\frac{e^{-\lambda_a r_i}\left(v_r+\frac{3}{4v_r r_i^2}\right)}{\sqrt{1-\left(v_r r_i-\frac{3}{4v_r r_i}\right)^2}}\,dr_i\end{aligned}\tag{19}$$

Unfortunately, there is no closed form solution to (19), so it must be calculated numerically. Note that the expected overhead is *independent of the number of neighbours* when only one route is cached.

2) Delay: When considering delay incurred with only one path cached, we assume the routing protocol is configured to aggressively seek a minimum hop length path, which gives us a lower bound on delay. In this case the node n_s caches only the route via the neighbour node with minimum distance (hops) to the destination. Once this route is broken, a new route discovery process is necessary.

In order to calculate the expected route delay we need first to find the distribution of the minimum distance $L_{\min}=\min_i L_i$ between the neighbour nodes and the destination. Since neighbour node positions are independent and identically distributed in angle and distance from n_s, we find that the CDF of the minimum distance is

$$F_{L_{\min}}(\ell)=1-(1-F_L(\ell))^N\tag{20}$$

and the PDF of the minimum distance is

$$f_{L_{\min}}(\ell)=N(1-F_L(\ell))^{N-1}f_L(\ell).\tag{21}$$

The expected value of the minimum distance between the neighbour nodes and the destination is, then

$$E\{L_{\min}\}=\int_{L_{i,\min}}^{L_{i,\max}}\ell N(1-F_L(\ell))^{N-1}f_L(\ell)d\ell.\tag{22}$$

The hop between this closest neighbour to the destination and the source node will be broken when the source node moves too far away. Since the direction of movement of the source node is independently distributed from the destination node, the link residual time is unaffected by the fact that we have chosen the closest neighbour to the destination. Hence the probability that the link is broken when the route is needed is given by $f_{R_i}(t_a)$. We have the expected delay

$$\begin{aligned}&E\{D|\text{one cached route}\}\\&=2t_h(E\{L_{\min}\}/r+1)\int_{1/2v_r}^{3/2v_r}e^{-\lambda_a t}f_{R_i}(t)dt\end{aligned}\tag{23}$$

Comparing (23) with (19) we can see that the overhead and delay for one cached route are identical in form, differing only by a constant factor.

B. Multiple Routes in Cache

In this case we assume that the routing protocol is configured to avoid route discovery as long as possible. If there are multiple routes in the cache, the moving source node will progressively lose connection with the first hop in each route. If there are routes originating with each neighbour node, the last route to be broken in this way will correspond to the closest neighbour the direction of movement of the source node.

1) Overhead: The cache at node n_s stores multiple routes to n_d, each commencing with a different neighbour node n_i. A route to n_d via neighbour node n_i is stored with probability p. The probability that there are k routes cached is equal to

$$\Pr(K=k)=\frac{N!}{k!(N-k)!}p^k(1-p)^{N-k}.\tag{24}$$

In particular, if $p=1$, then there are N cached routes with probability 1.

As mentioned above, in this case the source node can use the route through the neighbour closest to its direction of movement, θ_s, and no route discovery process is incurred until the link to that closest neighbour breaks. Let $r_{\max k}$ be the maximum R_i value, $i\in\{1,\cdots,k\}$, for some $k\leq N$.

Thus the expected value of the overhead, given $R_i, i=\{1,2,\cdots,N\}$ is given by

$$\begin{aligned}&E\{OH|k=K,\mathbf{R}=\mathbf{r}\}\\&=E\{OH|k=K,R_{\max K}=r_{\max K}\}\\&=(n+h-1)P\{t_a>r_{\max K}\}\\&=(n+h-1)(1-F_a(r_{\max K}))\\&=(n+h-1)e^{-\lambda_a r_{\max K}}\end{aligned}\tag{25}$$

Since $k>0$, the PDF of the maximum $R_{\max k}$ is determined by finding the CDF first, as follows.

$$\begin{aligned}F_{R_{\max k}}(t)&=P\{R_{\max k}<t\}\\&=P\{(R_1<t)\cup(R_2<t)\cup\cdots\cup(R_k<t)\}\\&=\prod_{i=1}^{k}P\{R_i<t\}\\&=F_{R_i}^k(t).\end{aligned}\tag{26}$$

Taking the derivative of (26) it can be seen that the PDF of the maximum R_i value is given by

$$f_{R_{\max k}}(t) = kF_{R_i}^{k-1}(t)f_{R_i}(t), \quad (27)$$

where $F_{R_i}(r_i)$ is from (7) and $f_{R_i}(r_i)$ is from (6).

Combining (25) and (27), we can determine an expression for expected overhead when multiple paths are cached.

$$\begin{aligned} &E\{OH|\text{many cached routes}\} \\ &= \sum_{K=0}^{N} \int_{R_{i,\min}}^{R_{i,\max}} \Pr(k=K) \\ &\qquad \cdot E\{OH|k=K, R_{\max k}=t\} \\ &\qquad \cdot f_{R_{\max k}}(t)dt \\ &= (n+h-1)\Big[(1-p)^N + \sum_{K=1}^{N} \frac{N!}{K!(N-K)!}p^K(1-p)^{N-K} \\ &\qquad K\int_{1/2v_r}^{3/2v_r} e^{-\lambda_a t}F_R^{K-1}(t)f_R(t)dt\Big] \quad (28) \end{aligned}$$

In the special case where $p = 1$, then $k = N$. That is, the source node stores one route to the destination through *each* neighbour node. We find that the expected overhead when all N neighbours have paths in the cache is

$$\begin{aligned} &E\{OH|N \text{ paths}\} \\ &= \int_{R_{i,\min}}^{R_{i,\max}} E\{OH|R_1 = r_1, \cdots, R_N = r_N\} f_{R_{\max}}(t)\, dt \\ &= (n+h-1)N\int_{R_{i,\min}}^{R_{i,\max}} e^{-\lambda_a t}F_{R_i}^{N-1}(t)f_{R_i}(t)dt \quad (29) \end{aligned}$$

There is no closed form solution to either (28) or (29), so they must be calculated numerically. Note that in both cases the expected overhead is now dependent on the number of neighbours.

2) Delay: Following similar reasoning to above, we find that the expected value of the delay is equal to the expected time taken for an individual route discovery process, multiplied by the probability that the route discovery process is necessary. This probability is the same as in (29).

$$\begin{aligned} &E\{D|\text{many routes}\} \\ &= 2t_h E\{h\}N\int_{R_{i,\min}}^{R_{i,\max}} e^{-\lambda_a t}F_{R_i}^{N-1}(t)f_{R_i}(t)dt \quad (30) \end{aligned}$$

As discussed in section III-B, the expected number of hops from the source to the destination is approximately $E\{L_i\}/r$, where L_i is the distance from the neighbour node i to the destination node.

While the position of the neighbour node that is closest to the direction of movement of n_s will tend towards the direction of movement of the source node if N is large, there is no dependence of the distance of this neighbour node to the destination on N because the direction of the destination node from the source node, which also determines length L_i, is uniformly distributed independently of θ_s. Thus $E\{L_i\}$ is given in (21), and the expected route delay is

$$\begin{aligned} &E\{D|\text{many routes}\} \\ &= 2t_h(L_d/r)\int_{R_{i,\min}}^{R_{i,\max}} Ne^{-\lambda_a t}F_{R_i}^{N-1}(t)f_{R_i}(t)dt. \quad (31) \end{aligned}$$

Again, comparing (29) with (31) we can see that the overhead and delay for multiple cached routes are identical in form, differing only by a constant factor.

V. Results

In this section we present and discuss results of comparisons of theoretical calculations and Monte-Carlo simulations, conducted in Matlab, for the expected overhead and delay, discussed above. Note that the purpose of these simulations is to validate the accuracy of the theoretical results, under the given network assumptions, rather than provide a realistic network simulation. More realistic simulation scenarios will be the subject of future work.

The simulations were constructed based on the scenario in Figure 3 such that the relationship between the initial distance of the source node to its neighbours and the transmission range is $1/2$. The neighbour nodes are distributed around the source node at uniformly distributed random angles. Note that no units are specified as it is the relative and not absolute distance that is important.

In order to have statistically accurate results, 10000 simulation repetitions were carried out for each combination of number of neighbour nodes and source node velocity. The reported results are the average of those simulations. For each simulation repetition, the source node direction $\theta_s, \sim u[0, 2\pi)$, and the neighbour node locations $\theta_i, \sim u[0, 2\pi)$ are generated. These are used in equation (5) to calculate the link residual times. The packet arrival times, t_a, at the source node are then generated according to an exponential distribution with parameter $\lambda_a = 0.1$. The packet arrival times are compared to the link residual times, and the overhead and delay are incremented if all routes in the cache have expired.

A. One Cached Route

The integral for the theoretical expression for expected overhead with one route saved in the cache, given in (19), was calculated using a sum of the values of the integrand calculated at incremental intervals. It was found that 10^5 increments were required for an accurate representation for one cached path. The overhead has been normalized by not including the $(n + h - 1)$ factor.

Figure 4 compares the theoretical expression in (19) to simulation results. It can be seen that the theoretical and simulation results match very well. The overhead for one cached path increases with the ratio of source node velocity, v, to transmission range, r, reaching 90% of the maximum by the time the ratio is equal to 1. The increase is quite steep initially, levelling off as v/r increases past 1. This makes sense as, as the velocity increases, the difference between the link residual time of the shortest possible link and the longest possible link

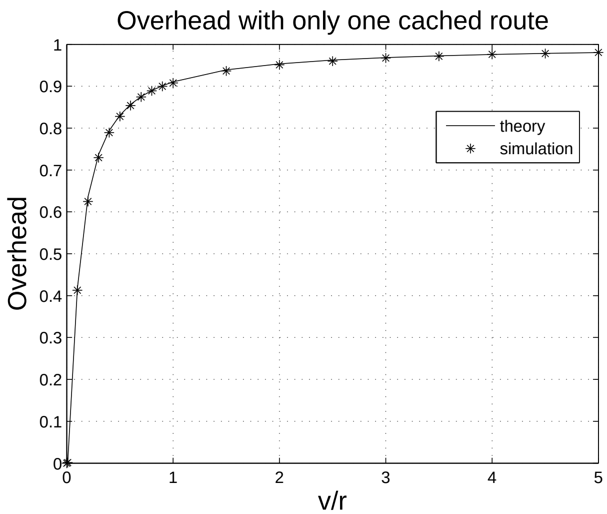

Fig. 4. Normalized overhead as it varies with ratio of n_s velocity to transmission range for a cached route from only one neighbour, from (19).

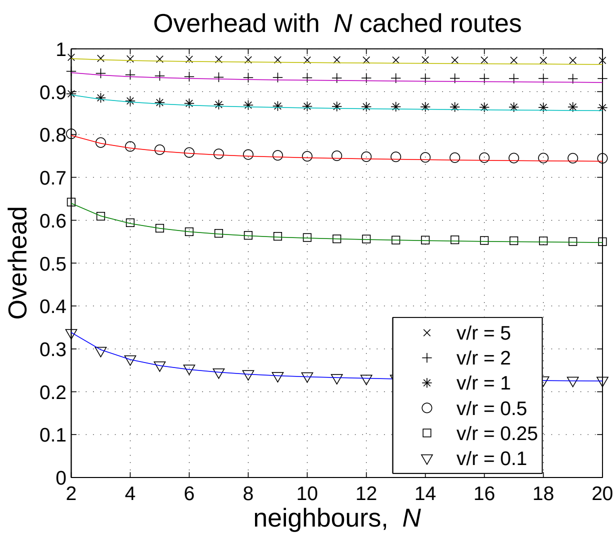

Fig. 5. Normalized overhead as it varies with number of cached paths from different neighbour nodes, for various ratios of n_s velocity to transmission range, from (29). Simulation results are depicted by markers while theoretical results are depicted by lines.

becomes more and more negligible. However, as mentioned in Section IV-A, there is no overhead benefit from having extra neighbours if only one path is cached.

The results for delay for only one cached route are shown in Figures 6 and 7, along with the delay for the N cached route case and are discussed in the next section.

B. N Cached Routes

1) Overhead: Similarly to the one cached route case, the integral for the theoretical expression for the expected overhead for N cached routes, given in (29), was calculated using a sum of the values of the integrand calculated at incremental intervals. It was found that 10^6 increments were required for an accurate representation for N cached paths. Again, the

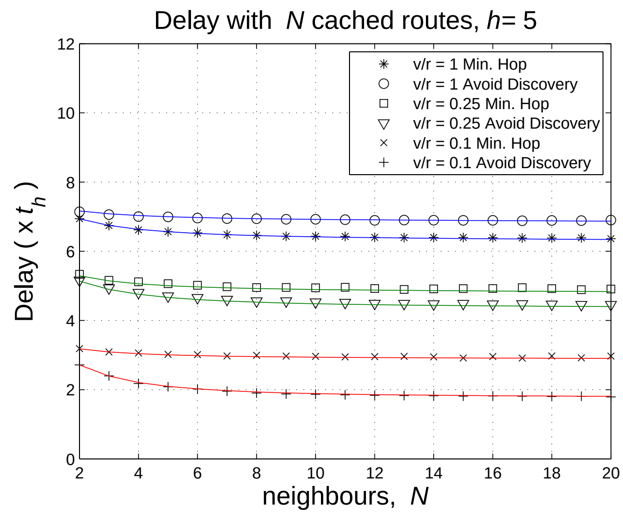

Fig. 6. Delay as it varies with the number of cached paths from different neighbour nodes, for $h = L_d/r = 5$ and various ratios of n_s velocity to transmission range, from (23) and (31). Simulation results are depicted by markers while theoretical results are depicted by lines.

overhead has been normalized by not including the $(n+h-1)$ factor. A comparison of theoretical and simulation results for overhead for N cached routes is shown in Figure 5. Again the theoretical and simulation results match very well. It can be seen that as the number of neighbouring nodes and, therefore, cached paths, increases, the overhead decreases. There is a levelling off of the overhead after an initial pronounced decrease. The overhead decrease with increase in neighbours, N, is more pronounced with small v/r ratios. As v/r increases the difference in amount of overhead for different numbers of cached paths becomes less pronounced to the point of being almost negligible when $v/r = 5$. Again, all of these trends are as would be expected. Except for the very slowest v/r ratio tested, the overhead is within 10% of its asymptotic value with 4 neighbour nodes. For the slowest ratio, 6 neighbours are required to reach this mark.

2) Delay: Figures 6 and 7 show comparisons of the theoretical expressions for delay with one and N cached paths, respectively, given in (23) and (31), with simulation results. Note that the delay is shown in terms of t_h, the one-hop packet transmission time. In the simulations, t_h was arbitrarily set to 1 time step. The "Avoid Discovery" case corresponds to the N cached route case, described by (31) and the "Min. Hop" case corresponds to the one cached route case, described by (31). Note that in Figure 6, the number of hops to the destination was kept small, at 5, and in Figure 7, the number of hops to the destination was made larger, at 16. The results show that as the number of cached paths from neighbouring nodes increases, the delay decreases, with the decrease more pronounced for smaller v/r ratios. Again the delay decreases with small v/r ratios because of a decrease in the number of route discovery processes required.

In both Figure 6 and Figure 7 it can be seen that the route delay is almost always smaller when N routes were cached

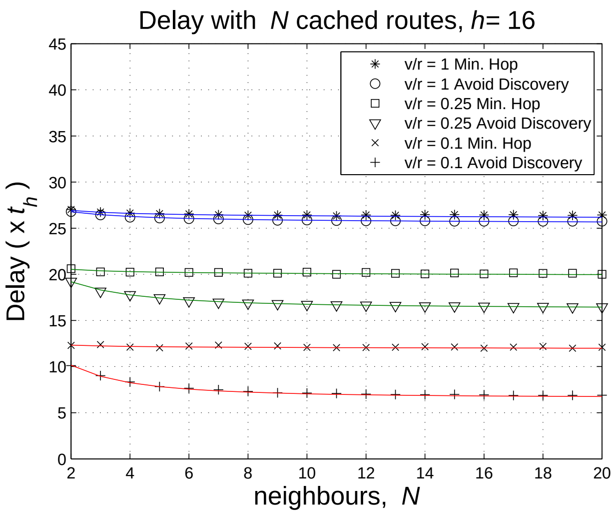

Fig. 7. Delay as it varies with the number of cached paths from different neighbour nodes, for $h = L_d/r = 16$ and various ratios of n_s velocity to transmission range, from (23) and (31). Simulation results are depicted by markers while theoretical results are depicted by lines.

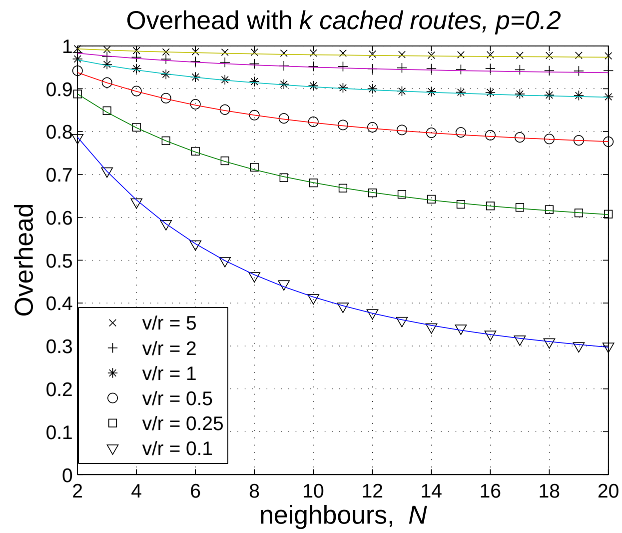

Fig. 8. Normalized overhead as it varies with the k cached paths from different neighbour nodes with probability of a path, $p = 0.2$, for various ratios of n_s velocity to transmission range, from (28). Simulation results are depicted by markers while theoretical results are depicted by lines.

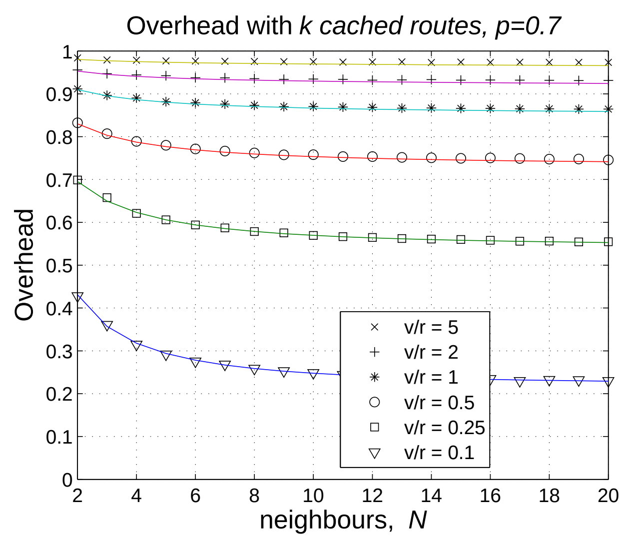

Fig. 9. Normalized overhead as it varies with k cached paths from different neighbour nodes with probability of a path, $p = 0.7$, for various ratios of n_s velocity to transmission range, from (28). Simulation results are depicted by markers while theoretical results are depicted by lines.

("Avoid Discovery") than when only the minimum hop route (one route) was cached. The only exception is for high v/r ratio, for the smaller number of hops. However, both delays are very close at high speeds and the switch is caused by a trade-off between the number of hops between neighbour nodes and the destination node and the number of times a new route discovery process is required. Recall from Figure 2 that for high velocities the link residual time is very small with only a small variation over possible positions of the neighbour nodes, meaning there will be little difference in how often a new route discovery process is required at such speeds. Also, the difference in number of hops between the routes to the destination of the neighbour with the minimum number of hops and that closest to the source (last route to fail when N routes are cached) is likely to be, at most, 2. So, in general, there is little delay benefit in caching only the minimum hop route if the speed is high.

For the small number of hops case, except for the very slowest speed v/r ratio tested, and only for the avoid discovery case, the delay is within 10% of its asymptotic value with 4 neighbour nodes. For the slowest ratio, for the avoid discovery case, 6 neighbours are required to reach this mark.

For the large number of hops case, except for the very slowest speed v/r ratio tested, and only for the avoid discovery case, the delay is within 10% of its asymptotic value with 4 neighbour nodes. For the slowest ratio, for the avoid discovery case, 8 neighbours are required to reach this mark.

C. k Cached Routes

Figures 8 and 9 illustrate the overhead incurred when k routes to the destination node are cached, from (28), with probability of any particular neighbour node's route being chosen, of 0.2 and 0.7, respectively. These two cases could, for example, correspond to a poorly connected network and a well-connected network, respectively. As the number of neighbour nodes increases, the number of cached paths from neighbour nodes correspondingly increases, and it can be seen that the overhead decreases. The overhead decrease is more pronounced with small v/r ratios, becoming almost negligible with larger speeds. In fact, with low speeds and low probability of any particular neighbour node having a cached route to the destination, it is best for the source to have as many neighbours as possible. Further, the overhead is greater for all speeds and numbers of neighbours when the probability of a route being cached for any given neighbour is smaller. All of these trends are as would be expected.

For the small probability of a cached route case, there is, again, quite a variation in the number of neighbour nodes required to achieve within 10% of the asymptotic performance. In fact, even for $N = 20$ neighbour nodes, it is not clear what the asymptotic overhead values are for any of the speeds tested, except, possibly for the very fastest, because the overhead is practically constant across all N.

VI. Conclusion

In this paper we have considered the effects of number of neighbour nodes on MANET performance, specifically for overhead and delay caused by frequency of route discovery process. We have developed analytical models to describe the network behaviour, validated by simulation results. We have shown that performance can vary quite significantly depending on the mobility (speed) of the source node, with respect to the transmission range, as well as with respect to number of neighbour nodes. Further, the required number of neighbours for a given performance level depends on the performance measure being considered.

The work in this paper is the first step in an attempt to analytically describe the topological features which are most effective in achieving good performance outcomes, for example, by considering the performance effects of different neighbour node location distributions. In this case, it is almost certain that a few spatially well-distributed neighbour nodes will serve better than a lot of closely clustered neighbour nodes. As a part of the investigation in to effect of neighbour node distributions on network performance, we will include an analysis of the effect on performance measures such as network lifetime.

Appendix

We derive the CDF and PDF of $d_{r,i}$ from (1), (3) and (2). Recall that $d_{r,i}$ is a function of $\cos(\theta_s - \theta_i)$ where $\theta_s - \theta_i$ is a uniformly distributed variable between 0 and 2π. Note that we get the same CDF and PDF when considering a uniformly distributed variable between 0 and π. As this makes the derivation slightly simpler in terms of cases to be considered, we make this assumption. Let $Y = \cos\theta$ where $\theta = (\theta_s - \theta_i)$. From (1)

$$\begin{aligned} d_{r,i} &= 0.1\left(y_i + \sqrt{y_i^2 + 3}\right) \\ \Rightarrow y_i &= d_{r,i} - \frac{3}{4d_{r,i}} \end{aligned} \tag{32}$$

Because of the non-positive slope of $\cos\theta$ for $0 \le \theta \le \pi$, we can write

$$\begin{aligned} F_y(y) &= P(\theta \ge \arccos y) \\ &= 1 - F_\theta(\arccos y) \\ &= 1 - \frac{1}{\pi}\arccos y. \end{aligned} \tag{33}$$

Now, using (32) and (33), we can derive the CDF of $d_{r,i}$, as follows.

$$\begin{aligned} F_D(d_{r,i}) &= P(Y \ge y_i) \\ &= P\left(Y \ge d_{r,i} - \frac{3}{4d_{r,i}}\right) \\ &= 1 - F_Y\left(d_{r,i} - \frac{3}{4d_{r,i}}\right) \\ &= 1 - \frac{1}{\pi}\arccos\left(d_{r,i} - \frac{3}{4d_{r,i}}\right). \end{aligned} \tag{34}$$

The PDF of $d_{r,i}$, from (2), is found by taking the derivative of (34) with respect to $d_{r,i}$. The CDF and PDF of the link residual time, R_i, from (7) and (6), can be derived in a similar way.

References

[1] D. M. Blough, M. Leoncini, G. Resta, and P. Santi. The k-neigh protocol for symmetric topology control in ad hoc networks. In *MobiHoc '03: Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pages 141–152, New York, NY, USA, 2003. ACM.

[2] A. Doci, W. Springer, and F. Xhafa. Maximum node degree mobility metric for wireless ad hoc networks. In *The Second International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, pages 463–468, 2008.

[3] O. Dousse, P. Thiran, and M. Hasler. Connectivity in ad-hoc and hybrid networks. *proceedings of IEEE INFOCOM*, 2:1079– 1088, 2002.

[4] G. Ferrari and O. K. Tonguz. Minimum number of neighbors for fully connected uniform ad hoc wireless networks. In *IEEE Communication Society*, volume 7, pages 4331– 4335, June 2004.

[5] T.-C. Hou and V. Li. Transmission range control in multihop packet radio networks. *IEEE Transactions on Communications*, 34:38– 44, January 1986.

[6] B. Ishibashi and R. Boutaba. Topology and mobility considerations in mobile ad hoc networks. *Ad Hoc Networks*, 3(6):762–776, 2005.

[7] D. B. Johnson and D. A. Maltz. Dynamic source routing in ad hoc wireless networks. In Imielinski and Korth, editors, *Mobile Computing*, volume 353. Kluwer Academic Publishers, 1996.

[8] L. Kleinrock and J. A. Silvester. Optimum transmission radii in packet radio networks or why six is a magic number. In *Proceedings of IEEE National Telecommunications Conference*, pages 04.3.1–04.3.5, December 1978.

[9] N. Li, J. C. Hou, and L. Sha. Design and analysis of an mst-based topology control algorithm. *IEEE Transactions on Wireless Communications*, 4(3):1195– 1206, May 2005.

[10] Q. Ling and Z. Tian. Minimum node degree and k-connectivity of a wireless multihop network in bounded area. In *GLOBECOM '07: Global Telecommunications Conference*, pages 1296–1301. IEEE, November 2007.

[11] J. Liu and B. Li. Mobilegrid: capacity-aware topology control in mobile ad hoc networks. In *International Conference on Computer Communications and Networks*, pages 570– 574, October 2002.

[12] M. Naserian, K. E. Tepe, and M. Tarique. Routing overhead analysis for reactive routing protocols in wireless ad hoc networks. In *IEEE International Conference on Wireless And Mobile Computing, Networking And Communications*, volume 3, pages 22–24, August 2005.

[13] P. Santi. The critical transmitting range for connectivity in mobile ad hoc networks. *IEEE Transactions on Mobile Computing*, 4(5):310 – 317, May-June 2005.

[14] S. Song, D. Goeckel, and D. F. Towsley. An improved lower bound to the number of neighbors required for the asymptotic connectivity of ad hoc networks. *IEEE Transactions on Information Theory*, 2005.

[15] H. Takagi and L. Kleinrock. Optimal transmission ranges for randomly distributed packet radio terminals. *IEEE Transactions on Communications*, 32:246–257, March 1984.

[16] G. Tian, W. Cai, and W. Wang. Topology variety model for mobile ad hoc networks. In *MOBILWARE '08: Proceedings of the 1st international conference on MOBILe Wireless MiddleWARE, Operating Systems, and Applications*, pages 1–6, 2007.

[17] O. K. Tonguz and G. Ferrari. Is the number of neighbors in ad hoc wireless networks a good indicator of connectivity? *IEEE International Zurich Seminar on Communications : Access-Transmission-Networking*, 18-20 Feb. 2004.

[18] J. Tsumochi, K. Masayama, H. Uehara, and M. Yokoyama. Impact of mobility metric on routing protocols for mobile ad hoc networks. In *Proc. of IEEE PacRim '03*, pages 322–325, Augest 2003.

[19] P.-J. Wan and C.-W. Yi. Asymptotic critical transmission radius and critical neighbor number for k-connectivity in wireless ad hoc networks. *International Symposium on Mobile Ad Hoc Networking and Computing*, pages 1–8, May 2004.

[20] Y. Wang and X.-Y. Li. Distributed spanner with bounded degree for wireless ad hoc networks. *Foundations of Computer Science*, 14(2):183–200, 2003.

[21] J. Wu and F. Dai. Mobility-sensitive topology control in mobile ad hoc networks. *IEEE Transactions on Parallel and Distributed Systems*, 17(6):522–535, June 2006.

[22] J. Wu and D. R. Stinson. Minimum node degree and k-connectivity for key predistribution schemes and distributed sensor networks. In *WiSec '08: Proceedings of the first ACM conference on Wireless network security*, pages 119–124. ACM, 2008.

[23] F. Xue and P. Kumar. The number of neighbors needed for connectivity of wireless networks. *Wireless Networks*, 10(2):169–181, March 2004.

[24] E. Yanmaz. Impact of topology-dependent and independent mobility models on the connectivity of wireless networks. In *Sarnoff Symposium, IEEE*, pages 1–5, April 2008.