On the Properties of One-Dimensional Infrastructure-based Wireless Multi-hop Networks

Seh Chun Ng, Student Member, IEEE, Guoqiang Mao, Senior Member, IEEE, and Brian D. O. Anderson, Life Fellow, IEEE

Abstract—Many wireless multi-hop networks are deployed with some infrastructure support. Existing results on ad-hoc networks are inadequate to fully understand the properties of those networks. In this paper, we study the properties of 1-D infrastructure-based multi-hop networks. Specifically, we consider networks with two types of nodes, i.e., ordinary nodes and powerful nodes. Ordinary nodes are i.i.d. and Poissonly distributed in a unit interval. Powerful nodes are arbitrarily distributed within the same unit interval. These powerful nodes are inter-connected via some backbone infrastructure. The network is said to be connected if each ordinary node is connected (possibly through a multi-hop path) to at least one of the powerful nodes. We obtain analytical results for the connectivity probability and the average number of clusters in the network. We also prove for the first time that the optimum powerful node distribution that minimizes the average number of clusters, and maximizes the asymptotic connectivity probability, is to deploy these powerful nodes in an equi-distant fashion. These results are important for the design and deployment of 1-D infrastructure-based networks.

Index Terms—Wireless networks, 1-D networks, random geometric graph, connectivity, clusters.

I. INTRODUCTION

Connectivity is one of the most important properties of a wireless multi-hop network and many network functions depend on the underlying network to be connected. A wireless ad-hoc network is said to be connected if there is a path between any pair of distinct nodes. The connectivity of wireless ad-hoc networks, i.e. infrastructureless multi-hop networks, has been extensively studied and the most well-known result is given by Gupta and Kumar in [1]. Specifically, they investigated the critical transmission range required for a random network with nodes independently, identically (i.i.d.) and uniformly distributed in a unit disk area in \( \mathbb{R}^2 \) to be connected as the number of nodes goes to infinity, under the unit disk communication model. Besides, there are many other studies on the connectivity of one or higher dimensional networks, either analytically or empirically (e.g., see [2], [3], [4], [5], [6], [7], [8]).

However, the aforementioned studies are only applicable to ad-hoc networks in which no infrastructure is deployed. Due to the limitations of ad-hoc networks [9], real networks may often not be modelable in this way. Several examples can be found in vehicular ad-hoc networks and wireless sensor networks (see [10], [11] for more examples). In vehicular networks, roadside infrastructure plays an important role in the reliable and timely distribution of important information to the vehicles on the road [12]. The active research projects in this field include IntelliDrive [13] in USA and a series of projects under CAR 2 CAR Communication Consortium [14] in Europe. In wireless sensor networks, data sinks gather the useful information collected by the sensors via multi-hop paths. Then the data sinks may either store the data for later retrieval, or aggregate and transfer the data immediately via a backbone network to the remote base station or the Internet. An example is the sensor network deployed on Great Duck Island for habitat monitoring [15]. From the above examples, it can be summarized that an infrastructure-based wireless multi-hop network has the following characteristics: (a) The communication between “ordinary” nodes (vehicles / sensors) and “powerful” nodes (roadside infrastructure / data sinks) is important for the core functions of the networks to be carried out properly. (b) The powerful nodes are always interconnected, either by wired or wireless links. Their locations are usually deterministic. (c) The locations of ordinary nodes are usually random. Based on the above characteristics, existing studies on the connectivity of ad-hoc networks are inadequate to understand these networks. Indeed, a novel conceptual framework is required to investigate the properties of these networks.

In this paper, we propose a new concept of connectivity for infrastructure-based wireless multi-hop networks, which we term type-II connectivity. We say that a network is type-II connected if every ordinary node in the network is connected (via one-hop or multi-hop paths) to at least one of a small
subset of powerful nodes. Type-II connectivity problem is a broad topic. In this paper we study the type-II connectivity of 1-D networks. It is expected that the results will be useful for many real world applications modelable by 1-D networks, e.g. a vehicular network built along a highway or a sensor network deployed along the border of a defined region for intrusion detection. The connectivity probability results can be used to solve some network design problems. For example, they provide a guideline of how far away two powerful nodes can be placed to meet a designated connectivity requirement. In addition to connectivity probability, we analyze the average number of clusters. A cluster is defined to be a maximal set of nodes where there is a (multi-hop) path between any two nodes in the set. Hence the average number of clusters can be considered as an alternative measure of network connectivity, which measures how fragmented a network is if it is not connected. It tells us how many additional powerful nodes are required for all ordinary nodes in a network to be connected to at least one powerful node with high probability. Assuming these additional powerful nodes are mobile, then there exist a number of ways [16] that they can connect the ordinary nodes together to achieve certain purposes such as maximizing the communication reliability between nodes [17] or balancing the traffic load among the nodes [18]. Such problems are important in the network topology control and routing. Based on the above connectivity and clustering results, we obtain the optimum powerful node distribution that minimizes the average number of clusters and maximizes the asymptotic connectivity probability of the network. Finally we verify the analytical results with the simulation results.

The rest of this paper is organized as follows: In Section II we introduce the related work. In Section III we define the network model. Then we present the analysis of type-II connectivity probability in Section IV and the average number of clusters in Section V. In Section VI, we discuss the optimum distribution of powerful nodes. In Section VII we further discuss some interesting observations on the network properties. Finally, Section VIII concludes this paper and discusses future work.

II. RELATED WORK

Connectivity of 1-D wireless ad-hoc networks has been extensively studied [2], [5], [6], [7]. However, those results do not accommodate the incorporation of infrastructure into the networks. Among the studies, Miorandi and Altman [7] assumed that there is a pre-determined node located at the origin. They investigated the probability of other nodes, which are either arbitrarily or uniformly distributed along a semi-infinite line, being connected (either directly or via multi-hop paths) to the node at the origin. A unit disk model and a Boolean model with random transmission range were considered. This scenario can be considered as a special case of type-II connectivity with only one powerful node placed at the origin. In this paper, we consider multiple powerful nodes in a network.

Dousse et al. have done a study closely related to type-II connectivity in [19], considering 1-D networks under the unit disk model. The nodes are assumed to be Poissonly distributed on a line segment of length $L$ with a known density. Two base stations are placed at both ends of the line segment. Based on the above model, they obtained analytically $p(x)$, the probability that a node at distance $x$ from the left base station is connected to at least one base station. Based on $p(x)$, the authors concluded that the existence of base stations improves the probability that two arbitrary nodes are connected. The authors considered this line segment as a “reduced” version of a more generic network with an infinite number of base stations placed every $L$ units distance on an infinite line. This work was later extended in [20] to consider 2-D networks but only simulation results were reported.

The work of Dousse et al. is different from ours in three aspects. First, Dousse et al. analyzed the probability that a node at location $x$ is connected to at least one base station, denoted by $p(x)$, whereas we analyze the probability that all nodes are connected to at least one powerful node (or base station), i.e. the network is type-II connected. It is not trivial to derive the probability that a network is type-II connected using $p(x)$. The difficulty lies in the fact that the event that one node located at $x$ is connected to a base station and the event that another node at $y$ is connected to a base station are not independent, but correlated in a complicated way. Therefore, a different technique is used in this paper to analyze the probability that a network is type-II connected. Using our technique, $p(x)$ can be readily derived but using the technique in [19] to derive the probability that a network is type-II connected appears rather complicated. Second, we consider the situation that powerful nodes and ordinary nodes may have different transmission ranges; such an assumption at least sometimes should better reflect physical reality. Last, we analyze a number of network properties beyond the mere type-II connectivity probability. These include the average number of clusters, which is an important performance indicator, and the optimal placement of powerful nodes. The results help to obtain a better understanding of these networks.

More recently, Sou [21] also studied the network connectivity of an infrastructure-based wireless multi-hop networks. Similar to [19], Sou considered a 1-D vehicular network where base stations are equally spaced on a road. However in their paper, all vehicles in a road segment bounded by two adjacent base stations must be connected to both base stations. In this paper, we consider each ordinary node is connected to at least one (but not necessarily both) powerful node(s). In addition, the powerful nodes are not necessarily equally spaced in the network. Using the technique used in this paper, their results can be readily derived. Other work can also be found in the literature which studied the $k$-connectivity problem in vehicular networks (e.g. see [22], [23], [24]). However, those work considered only inter-vehicle communications without involving the base stations. Such assumption simplifies vehicular networks from infrastructure-based networks to ad hoc networks. As shown in [25], vehicular networks will involve both vehicle-to-vehicle and vehicle-to-infrastructure communications. In this paper, we show the impact of powerful nodes to the network connectivity. The results will be significantly different if no powerful nodes are involved in the communications.

Another related problem has been studied in the context of Multihop Cellular Networks (MCNs) [26]. MCNs combine
the features of conventional cellular networks and ad-hoc networks to reduce the required number of base stations in an area while limiting path vulnerability encountered in ad-hoc networks. Wu et al. compared the performance of MCNs with the conventional cellular networks in terms of the call blocking/dropping probability, throughput and signaling overhead in their work [27], [28]. In [29], Yannaz and Tonguz investigated the impact of number of available relay channels to the performance of MCNs, such as call blocking/dropping probability. Given the number of available channels, Liu et al. [30] suggested a channel allocation and routing strategy to maximize the cell throughput. In [31], Venkataraman et al. [30] studied the impact of number of hops to the channel usage. For a 2-hop network, Venkataraman and Muntean [32] proposed a resource allocation technique to support high data rate traffic. Other discussion on the design and implementation issues of MCNs can also be found in [33], [34], [35]. Even though the aforementioned work studied MCNs from various aspects, only few of them investigated the fundamental and crucial issue of network connectivity, which is the focus of this paper. Among the studies, the work by Ojha et al. [36] and Mukherjee et al. [37] are related to network connectivity.

For a network of $n$ uniformly distributed nodes in a circular area of unit radius, Ojha et al. [36] obtain a lower bound on the transmission range required for all nodes in the network to be asymptotically connected to the base station at the center of the area as $n \to \infty$ under the unit disk model. Under a more generic assumption of having both base stations and subscriber stations Poissonly distributed in $\mathbb{R}^2$ and the log-normal shadowing model, Mukherjee et al. [37] obtain a lower bound on the probability that an arbitrary subscriber station cannot reach any base station in at most $t$ hops using the independence assumption, i.e. the event that one subscriber station can reach any base station in $k$ hops is independent of the event that another subscriber station being able to reach any base station in $k$ hops. While this assumption frequently simplifies analysis, it is typically not true [38], [39]. In this paper, we are neither restricted to considering a single base station nor to limiting the maximum hop count between two connected nodes. Our analysis also does not rely on the independence assumption.

III. Network Model

Based on the observation in Section I, we define the network model we consider in this paper as follows.

Definition 1. Denote by $G(\lambda, n_p; L; r_o, r_p)$ a wireless multi-hop network with two types of nodes: ordinary nodes and powerful nodes. Ordinary nodes are i.i.d. and Poissonly distributed with a known density $\lambda$ in the interval $[0, L]$. There are $n_p \geq 2$ powerful nodes in the network, where two of them are placed at both ends of the interval and the rest are arbitrarily distributed in the interior of the same interval. A direct connection between two ordinary nodes (respectively, between an ordinary node and a powerful node) exists if their Euclidean distance is smaller than or equal to $r_o$ (respectively, $r_p$); all powerful nodes are assumed to be inter-connected to each other by default.

An example of our model is illustrated in Fig. 1. In the model, the powerful nodes divide the interval $[0, L]$ into $n_p - 1$ sub-intervals and each sub-interval $i$ has length $w_i$ for $1 \leq i \leq n_p - 1$.

Note that in our model the direct connections between nodes follow the well-known unit disk model (with transmission ranges $r_o$ and $r_p$). In general, we assume that $r_p \geq r_o$. This assumption is justified because it is often the case that a powerful node can not only transmit at a larger transmission power than an ordinary node, it can also be equipped with more sophisticated antennas, which make it more sensitive to the transmitted signal from an ordinary node [40].

Consideration of the unit disk model increases the usefulness and applicability of our results. First, the analysis becomes tractable under the unit disk model and all equations obtained in this paper are closed form equations which offer better insight into the interactions of various performance-impacting parameters. As will be shown later, in this paper we mostly focus on the network with $L = 1$, i.e. on the unit interval. Using the space scaling technique [41], the results for the network on the unit interval can be easily applied to a network on the interval $[0, L]$ where $L \neq 1$. Second, the results obtained under the unit disk model provide bounds for networks under other connection models. For example, consider a real-life scenario (or a connection model other than the unit disk model) and let $r_o$ (respectively, $r_p$) to be the distance threshold such that any two ordinary nodes (respectively, an ordinary node and a powerful node) separated by a distance less than or equal to the threshold are directly connected with high probability. Then the connectivity probability obtained under the unit disk model with transmission ranges $r_o$ and $r_p$ will provide a lower bound for the connectivity probability in the real-life scenario (or under other connection model) [42]. Finally, the qualitative conclusions obtained under the unit disk model are normally also valid for other connection models. The examples include the phase transition behavior of network connectivity [43] and the energy saving achievable when only requiring most, but not all, nodes in the network to be connected [44]. In Section VII the simulation results are obtained under the log-normal model\(^1\) and the plots show that the results obtained under the unit disk model in this paper are qualitatively applicable to the networks under the log-normal model.

\(^1\)The log-normal model is commonly used to model the real world signal propagation where the transmit power loss increases logarithmically with the Euclidean distance between two nodes and varies log-normally due to the shadowing effect caused by the surrounding environment [40]. More details are included in Section VII.
IV. TYPE-II CONNECTIVITY PROBABILITY

In this section we investigate the type-II connectivity probability of a network $G(\lambda, n_p; 1, r_o, r_p)$, i.e. on the unit interval. Note that the network is said to be type-II connected if each ordinary node is connected (directly or via a multi-hop path) to at least one of the powerful nodes. Under the unit disk model, the connectivity probability can be derived by first examining each sub-interval bounded by two consecutive powerful nodes.

Let $A_i(w_i)$ be the event that sub-interval $i$ with length $w_i$ is type-II connected under the general assumption that $r_p \geq r_o$. It is trivial to show that $\Pr\{A_i(w_i)\}$, the probability that $A_i(w_i)$ occurs, is 1 when $w_i \leq 2r_p$. For $w_i > 2r_p$, a realization of a type-II connected sub-interval $i$ for $r_p \geq r_o$ remains type-II connected if and only if after removing an interval of length $r_p - r_o$ (and the ordinary nodes within that interval) from the left end and right end of sub-interval $i$ respectively, the resulting sub-interval with $r_p = r_o$ is still type-II connected. Hence for $w_i > 2r_p$,

$$\Pr\{A_i(w_i)\} = \Pr\{A_i^q(w_i - 2(r_p - r_o))\}$$

(1)

where $A_i^q(x)$ is the event that sub-interval $i$ with length $x$ is type-II connected under the situation that $r_p = r_o$. In next sub-section we provide the derivation of $\Pr\{A_i^q(x)\}$.

A network is type-II connected if and only if each sub-interval is type-II connected. Under the unit disk model, the event that one sub-interval is type-II connected is independent of the event that another sub-interval is type-II connected. Hence, the probability that a network with each sub-interval $i$ having length $w_i$ is type-II connected (say, event $B(w_1, \ldots, w_{n_p-1})$) is

$$\Pr\{B(w_1, \ldots, w_{n_p-1})\} = \prod_{i=1}^{n_p-1} \Pr\{A_i(w_i)\}$$

(2)

Based on Eq. (2), we can state the following theorem:

**Theorem 1.** Denote by $B$ the event that a random instance of $G(\lambda, n_p; 1, r_o, r_p)$ is type-II connected. Then the probability that event $B$ occurs is

$$\Pr\{B\} = \int_{\mathbb{D}} \left( \prod_{i=1}^{n_p-1} \Pr\{A_i(w_i)\} \right) f(w) \, dw$$

(3)

where $\mathbb{D} = \left\{(w_1, \ldots, w_{n_p-1}) : \sum_{i=1}^{n_p-1} w_i = 1\right\}$; $f(w) = f(w_1, \ldots, w_{n_p-1})$ is the joint probability density function of the distances between adjacent powerful nodes; $\Pr\{A_i(w_i)\}$ is given in Eq. (1) and $\Pr\{A_i^q(w_i)\}$ is given in Eq. (10).

Using Eq. (3), we can calculate the type-II connectivity probability of a network with any distribution of powerful nodes as long as $f(w)$ of that distribution is known. For example, if the powerful nodes are uniformly distributed, then $f(w) = (n_p - 2)!$ [45]. If the powerful nodes are placed in an equi-distant fashion, then Eq. (3) simplifies into

$$\Pr\{B\} = \left[\Pr\{A_i(w)\}\right]^{n_p-1}$$

(4)

with $w = \frac{1}{n_p-1}$. In the following sub-sections, we derive $\Pr\{A_i^q(x)\}$ and its asymptotic approximation.

A. Exact probability that a sub-interval is type-II connected for $r_p = r_o$

As mentioned earlier, the result for $r_p = r_o$ can be used to obtain the result for the general case where $r_p \geq r_o$. Let $A_i^q(m_i, w_i)$ be the event that sub-interval $i$ with length $0 < w_i \leq 1$ is type-II connected given that there are $m_i$ ordinary nodes in the sub-interval. Denote the common transmission range by $r$, i.e. $r = r_p = r_o$. The derivation of $\Pr\{A_i^q(m_i, w_i)\}$ relies on the following lemma from [46].

**Lemma 1** (Lemma 1 in [46]). Let $[x, x+y]$ be a sub-interval of length $y$ within $[0,1]$. Assume two of $k$ given vertices have been placed at the borders of this sub-interval. Define two vertices to be neighbors if and only if they are at distance $r$ or less apart, let $Z_{k,y,r}$ be the event that $k - 2$ vertices, corresponding to the remaining vertices and uniformly placed in $[0,1]$, are inside $[x, x+y]$ and “join” the borders, that is, the $k$ vertices form a connected subgraph of length $y$; and let $P(k, y, r) = \Pr(Z_{k,y,r})$. Then for $k \geq 2$,

$$P(k, y, r) = \min_{j=1}^{\min(k-1, \lfloor y/r \rfloor)} \frac{(k-1)!}{j!} \frac{(-1)^j}{y^j} r^{k-2}.$$  

(5)

A sub-interval is type-II connected if all ordinary nodes within the sub-interval are connected to at least one of the two powerful nodes located at both ends of the sub-interval. Hence, event $A_i^q(m_i, w_i)$ occurs with probability

$$\Pr\{A_i^q(m_i, w_i)\} = P(m_i + 2, 1, \hat{r}) + m_i(m_i + 1) \int_0^{1-\hat{r}} P(m_i + 1, \hat{x}, \hat{r}) \, d\hat{x}$$

(6)

where $\hat{r} = \frac{r}{w_i}$ is the normalized transmission range, and $\hat{x} = \frac{x}{w_i}$ is the normalized distance of $x$. The two terms on the right hand side of Eq. (6) represent the two possible cases of the event, as illustrated in Fig. 2. Fig. 2(a) corresponds to the first term in Eq. (6), and Fig. 2(b) corresponds to the second term. [1]

Fig. 2 shows a possible case where all $m_i$ ordinary nodes within sub-interval $i$ are connected to both powerful nodes. That is, none of the $m_i + 1$ spacings between the adjacent ordinary nodes and the two powerful nodes is larger than $r$. In this case, all ordinary nodes and the two powerful nodes in sub-interval $i$ form a connected “subgraph” of length $w_i$. From Lemma 1, the probability of this case is $P(m_i + 2, 1, \hat{r})$,
where \( m_i + 2 \) is the sum of the number of ordinary nodes and the two powerful nodes.

Fig. 2(b) shows the other possible case where the \( m_i \) ordinary nodes inside sub-interval \( i \) are connected to either one of the two powerful nodes but not both. Then among the \( m_i + 1 \) spacings between the adjacent nodes, there is exactly one spacing with length \( s > r \). Suppose temporarily that the big spacing of length \( s \) and the ordinary node attached to the left end of the big spacing are removed from sub-interval \( i \), as illustrated in Fig. 3; then the \( m_i - 1 \) remaining ordinary nodes and the two powerful nodes form a connected “subgraph” of length \( x = w_i = s - s \). [1] A special case occurs when the big spacing is the left most spacing in the sub-interval. If this is the case, then we remove the ordinary node attached to the right end of the big spacing instead. The probability that the \( m_i - 1 \) remaining ordinary nodes and the two powerful nodes, with the sub-interval having the new length of \( x \), form a connected interval is given by \( P(m_i + 1, x, \hat{r}, \hat{r}) \), where \( \hat{x} = \frac{x}{w_i} \). Following the convention of [46] that nodes are treated as distinguishable, the event that a particular node \( i \) attached to the left end of the big spacing is removed together with the big spacing and the remaining nodes form a connected interval, and the event that a particular node \( j \) attached to the left end of the big spacing is removed together with the big spacing and the remaining nodes form a connected interval, are treated as different events. Therefore, any of the \( m_i \) ordinary nodes can be attached to the left end of the big spacing (or attached to the right end for the special case), and the big spacing can be any of the \( m_i + 1 \) spacings in sub-interval \( i \). As a result, the probability that events like Fig. 2(b) occur is then \( m_i(m_i + 1)P(m_i + 1, x, \hat{r}, \hat{r}) \hat{x} \), for \( \hat{x} \) ranging from zero to \( 1 - \hat{r} \). So we obtain the second term in Eq. (6).

After applying Eq. (5) into the second term of Eq. (6), we can get rid of the integral in the second term by moving the inner sum outside the integral, with changes to the range of summation and integral we obtain

\[
m_i(m_i + 1) \int_0^{1-\hat{r}} P(m_i + 1, x, \hat{r}, \hat{r}) d\hat{x} = (m_i + 1) \sum_{j=0}^{\text{min}(m_i, 1/\hat{r}) - 1} \left( \frac{m_i}{j} \right) (-1)^j (1 - (j + 1) \hat{r})^{m_i}.
\]

Using Eq. (7), and replacing the first term in Eq. (6) by Eq. (5), we can simplify Eq. (6):

\[
\Pr \{ A_i^q(w_i) \} = \min(m_i + 1, \lfloor w_i/r \rfloor) \sum_{j=0}^{\text{min}(m_i+1, \lfloor w_i/r \rfloor) - 1} \left( \frac{m_i}{j} \right) (-1)^j (1 - j \hat{r})^{m_i}.
\]

Since all sub-intervals bounded by powerful nodes are non-overlapping segments with length \( w_i \) and note that ordinary nodes are Poissonly distributed, \( m_i \) is a Poisson random variables with mean \( w_i \lambda \), and \( m_i \) and \( m_j \) are mutually independent for \( i \neq j \). Let \( A_i^q(w_i) \) be the event that sub-interval \( i \) with length \( 0 < w_i \leq 1 \) is type-II connected. Then,

\[
\Pr \{ A_i^q(w_i) \} = \sum_{m_i=0}^{\infty} \Pr \{ A_i^q(m_i, w_i) \} \frac{(w_i \lambda)^{m_i}}{m_i!} \exp(-w_i \lambda)
\]

Fig. 4 shows that Eq. (13) serve as a good approximation for the exact result in Eq. (10) provided \( \mu \geq 6 \), and virtually all values of \( \hat{w} \geq 2 \), not just large values of \( \hat{w} \). [1] Solving
Eq. (13) for \( \hat{w} \) leads to
\[
\hat{w} = \left\{ -W_{-1}[-(1-\beta)^2 \exp(-(1-2\beta))\phi(\hat{w})] \\
-(1-2\beta) \right\} \frac{1-\beta}{\beta}
\]
(14)
where \( W_{-1}[\cdot] \) is the real-valued, non-principal branch of the LambertW function \([47]\). Given the required connectivity probability \( \phi(\hat{w}) \), and the value of \( \beta \), which is related to the ordinary node density \( \lambda \) and the transmission range \( r \), we can use Eq. (14) to obtain \( \hat{w} \), the maximum distance between two adjacent powerful nodes so that the designated connectivity probability requirement is fulfilled.

V. AVERAGE NUMBER OF CLUSTERS

Besides connectivity probability, another variable of interest in a network is the number of clusters in the network. It is an indicator of how fragmented a network is. In our network model, all ordinary nodes that are connected to at least one of the powerful nodes belong to the same cluster. Denote by “main cluster” the cluster formed by the ordinary nodes which are connected to at least one powerful node. Other ordinary nodes which are not connected to any powerful nodes, if they exist, form one or more “secondary clusters”. Therefore there is always one main cluster and zero or more secondary clusters in a network. Note that a network is type-II connected if and only if there is no secondary cluster in the network. In this section we investigate the average number of connected if and only if there is no secondary cluster in the network. Note that a network is type-II connected if and only if there is no secondary cluster in the network. In this section we investigate the average number of clusters in the network. In our network model, all ordinary nodes that are connected to at least one of the powerful nodes belong to the same cluster. Denote by “main cluster” the cluster formed by the ordinary nodes which are connected to at least one powerful node. Other ordinary nodes which are not connected to any powerful nodes, if they exist, form one or more “secondary clusters”. Therefore there is always one main cluster and zero or more secondary clusters in a network. Note that a network is type-II connected if and only if there is no secondary cluster in the network. In this section we investigate the average number of clusters in the network. In our network model, all ordinary nodes that are connected to at least one of the powerful nodes belong to the same cluster. Denote by “main cluster” the cluster formed by the ordinary nodes which are connected to at least one powerful node. Other ordinary nodes which are not connected to any powerful nodes, if they exist, form one or more “secondary clusters”. Therefore there is always one main cluster and zero or more secondary clusters in a network.

Let \( C_i(w) \) be the number of secondary clusters in the sub-interval \( i \) with length \( w \) under the general assumption that \( r_p \geq r_o \). Note that \( w = \frac{1}{r_p-1} \) for equi-distant powerful node distribution\(^2\). Also note that the ordinary nodes in sub-interval \( i \), which are at most \( r_p-r_o \) Euclidean distance away from the powerful nodes, belong to the main cluster with probability 1 and any other ordinary nodes which are directly connected to these ordinary nodes are also directly connected to the powerful nodes. As a result, the number of secondary clusters remains the same after we remove an interval of length \( r_p-r_o \) (and the ordinary nodes within that interval) from the left end and right end of sub-interval \( i \) respectively and then assume that \( r_p = r_o \). So, we have \( C_i(w) = C_i^{eq}(w-2(r_p-r_o)) \) where \( C_i^{eq}(x) \) is the number of secondary clusters in the sub-interval with length \( x \) under the special assumption that \( r_p = r_o \).

To obtain \( C_i^{eq}(x) \), let \( t_i(x) \) be the number of spacings with length greater than \( r \) in sub-interval \( i \) of length \( x \) where \( r = r_p = r_o \) as usual. Then \( C_i^{eq}(x) \) and \( t_i(x) \) have the following relationship:
\[
C_i^{eq}(x) = \begin{cases} 
1 & \text{for } t_i(x) \geq 1, \\
0 & \text{for } t_i(x) = 0.
\end{cases}
\]
(15)
Assume that there are \( m_i \) ordinary nodes in sub-interval \( i \) of length \( x \) and let \( 1_{i,j}(x) \) be an indicator function such that
\[
1_{i,j}(x) = \begin{cases} 
1 & \text{if the } j \text{-th spacing in sub-interval } i \\
0 & \text{otherwise}
\end{cases}
\]
where \( 1 \leq j \leq m_i + 1 \). Then, the expected value of \( t_i(x) \) given \( m_i \) ordinary nodes in sub-interval \( i \) with length \( x \) is
\[
E[t_i(x) \mid m_i] = E[ \sum_{1\leq j\leq m_i+1} 1_{i,j}(x) \mid m_i ] = (m_i + 1)E[1_{i,j}(x) \mid m_i] \quad \text{for any } j
\]
(16)
\[
= (m_i + 1)(1 - \frac{r}{x})^{m_i}.
\]
(17)
where \( E[1_{i,j}(x) \mid m_i] \) is equal to the probability that the \( j \)-th spacing in sub-interval \( i \) of length \( x \) has length greater than \( r \). Since this probability is equal to the probability that the \( m_i \) ordinary nodes fall into a smaller interval of length \( 1 - \frac{r}{x} \) in sub-interval \( i \), we obtain Eq. (17).

Since \( m_i \) is a Poisson random variable with mean \( x\lambda \), it follows immediately that
\[
E[t_i(x)] = \sum_{m_i=0}^{\infty} E[t_i(x) \mid m_i] \frac{(x\lambda)^{m_i}}{m_i!} \exp(-x\lambda)
\]
\[
= (x\lambda - r\lambda + 1)\exp(-r\lambda).
\]
From Eq. (15) we have
\[
E[C_i^{eq}(x)] = E[t_i(x)] - 1 + Pr \{ t_i(x) = 0 \}
\]
(18)
\[
= \sum_{j=2}^{\lfloor x/r \rfloor} (-1)^j \frac{1}{j!} (x + j\lambda - jr\lambda)
\]
\times (x\lambda - jr\lambda)^{j-1} \exp(-jr\lambda)
\]
(19)
where from Eq. (18) to Eq. (19) we apply
\[
Pr \{ t_i(x) = 0 \} = \sum_{j=0}^{\lfloor x/r \rfloor} (-1)^j \frac{1}{j!} (x + j\lambda - jr\lambda)
\]
\times (x\lambda - jr\lambda)^{j-1} \exp(-jr\lambda)
\]
(20)
which is obtained from the first term in Eq. (6) and simplified using the same procedure resulting in Eq. (10). Finally, let $D(n_p)$ be the number of clusters in a network with $n_p$ powerful nodes equally spaced and assume $r_p \geq r_o$. Then,

$$D(n_p) = \sum_{i=1}^{n_p-1} C_i(w) + 1 \tag{21}$$

$$= \sum_{i=1}^{n_p-1} C_i^{eq}(w - 2(r_p - r_o)) + 1 \tag{22}$$

where $w = \frac{1}{n_p-1}$. That is, we add up the number of secondary clusters in each sub-interval and one (the only) main cluster in the whole network. Based on Eq. (22), we can now state the following theorem.

**Theorem 2.** For $G(\lambda, n_p; 1; r_o, r_p)$ with powerful nodes placed in an equi-distant fashion, the expected number of clusters in the network is then

$$E[D(n_p)] = (n_p - 1)E[C_i^{eq}(w - 2(r_p - r_o))] + 1 \tag{23}$$

where $E[C_i^{eq}(x)]$ is given in Eq. (19); $w = \frac{1}{n_p-1}$.

VI. THE OPTIMAL DISTRIBUTION OF POWERFUL NODES

In this section, we prove that the equi-distant placement of powerful nodes will minimize the average number of clusters in a network and maximize the asymptotic type-II connectivity probability.

A. Minimizing the average number of clusters

From Eq. (21), we have the average number of clusters in a network $G(\lambda, n_p; 1; r_o, r_p)$ given each sub-interval $i$ has length $w_i$, is

$$E[D(w_1, w_2, \cdots, w_{n_p-1})] = \sum_{i=1}^{n_p-1} E[C_i(w_i)] + 1 \tag{24}$$

where $E[C_i(w_i)]$ is the average number of secondary clusters in sub-interval $i$. Finding the optimal powerful node placement to minimize the average number of clusters can be treated as a constrained optimization problem:

minimize $E[D(w_1, w_2, \cdots, w_{n_p-1})]$

subject to $\sum_{i=1}^{n_p-1} w_i = 1$.

In the following we prove that $E[C_i(w_i)]$ is a convex function of $w_i$.

Proof: Recall that for $w_i > 2r_p$, we have $E[C_i(w_i)] = E[C_i^{eq}(w_i - 2(r_p - r_o))]$. And from Eq. (18) we have,

$$E[C_i^{eq}(x)] = (x\lambda - r_o\lambda + 1)e^{-r_o\lambda} - 1 + \Pr\{t_i(x) = 0\}$$

where $\Pr\{t_i(x) = 0\}$ is given by Eq. (20). With some arithmetic steps we can derive the second derivative of $E[C_i^{eq}(x)]$ and obtain

$$\frac{d^2}{dx^2}E[C_i^{eq}(x)] = \frac{d^2}{dx^2} \Pr\{t_i(x) = 0\} \geq 0.$$  

Hence, the second derivative of $E[C_i(w_i)]$ is also greater or equal to zero for $w_i > 2r_p$. It is trivial to show that the second derivative is zero for $w_i \leq 2r_p$ as $E[C_i(w_i)] = 0$ in that range.

Since $E[C_i(w_i)]$ is a convex function, so by Eq. (24), $E[D(w_1, w_2, \cdots, w_{n_p-1})]$ is also a convex function. Hence the optimization problem is a convex optimization problem. It is then straightforward to prove, e.g. using the method of Lagrange multipliers, that the minimum of the average number of clusters is achieved when $w_1 = \cdots = w_{n_p-1} = \frac{1}{n_p-1}$ and by convexity it is a global minimum.

B. Maximizing the asymptotic type-II connectivity probability

Using Eq. (13) we approximate the type-II connectivity probability

$$\Pr\{B(w_1, \cdots, w_{n_p-1})\} = \prod_{i=1}^{n_p-1} \Pr\{A_i(w_i)\} \approx \prod_{i=1}^{n_p-1} \phi(x_i)$$

where $x_i = w_i - 2(r_p - r_o)$, and

$$\phi(x_i) = \begin{cases} \frac{1 - 2^\beta}{(1-\beta)^\beta} + \frac{3x_i/r_o}{(1-\beta)^3} & \exp(-\frac{3x_i/r_o}{1-\beta}) \text{ if } x_i > 2r_o, \\ \frac{1}{1-\beta} & \text{otherwise.} \end{cases}$$

Since both expressions $\frac{1 - 2^\beta}{(1-\beta)^\beta} + \frac{3x_i/r_o}{(1-\beta)^3}$ and $\exp(-\frac{3x_i/r_o}{1-\beta})$ are log-concave on $x_i \geq 0$, and the product of log-concave functions is a log-concave function [48], we have $\phi(x_i)$ is log-concave on $x_i$ and $\Pr\{B(w_1, \cdots, w_{n_p-1})\}$ is also a log-concave function of the lengths of sub-intervals. Using this property, it can be readily shown that the maximum of the probability that the network $G(\lambda, n_p; 1; r_o, r_p)$ is type-II connected is also achieved when powerful nodes are distributed in an equi-distant fashion.

VII. DISCUSSION

In this section, we investigate the impact of different parameters to the performance of a network $G(\lambda, n_p; 1; r_o, r_p)$. Note that all figures are plotted under the condition that powerful nodes are placed in an equi-distant fashion.

[1] First, Fig. 5 shows the probability that a network is connected given different values of $\lambda$, $n_p$, and $r_p = r_o = 0.05$.
under the unit disk model. The analytical results are verified by simulation results obtained from 40000 randomly generated network topologies\(^3\). The number of powerful nodes has been varied from 2 to 10. With \( r_p = 0.05 \), the network will be fully covered by the powerful nodes for \( n_p > 10 \). It is shown that an increase in \( n_p \) significantly improves network connectivity probability. The impact of \( \lambda \) on connectivity is rather interesting. When \( \lambda \) is small, the network connectivity probability drops as \( \lambda \) increases. That is because when the number of ordinary nodes is small, the probability that an ordinary node is connected to a powerful node via a multi-hop connection is small and can be almost neglected. Therefore, one ordinary node has to be close to a powerful node in order to be connected. Thus, when the number of ordinary nodes is small, an increase in the number of ordinary nodes causes a drop in the probability that all ordinary nodes are connected to at least one nearby powerful node. As the number of ordinary nodes further increases, the probability that an ordinary node far away from a powerful node can establish a multi-hop connection to the powerful node increases, which consequently causes an increase in the probability of having a type-II connected network\(^5\).

\[^3\] The simulations are conducted using a program written in Java. The random numbers required in the simulations, e.g., the locations of the ordinary nodes, are generated using the default pseudorandom number generator provided by Java.

\[^4\] As the numbers of instances of random networks used in the simulations are very large, the confidence interval is too small to be distinguishable and hence ignored in this plot and the later plots.

\[^5\] Note that the properties observed in Fig. 5, i.e., under the unit disk model, are also observed when the log-normal shadowing model is considered. In the log-normal shadowing model, two nodes separated by a Euclidean distance \( x \) are directly connected with probability

\[
g(x) = Q\left(\frac{10\lambda}{\sigma \log_{10}\frac{x}{r}}\right)
\]

where \( Q(y) = \frac{1}{\sqrt{2\pi}} \int_y^\infty \exp(-\frac{z^2}{2})dz \) is the tail probability of the standard normal distribution, \( \alpha \) is the path loss exponent, \( \sigma^2 \) is the shadowing variance, \( r \) is the transmission range ignoring shadowing effect. Refer to [25] for more details of the log-normal model. In order to further accurately model the direct connection between nodes, we consider channel correlation in our simulation. That is, we follow the approach in [49], [50], [51] and use an exponential model to model the fading correlations between wireless links. In the model, the received signals at two nearby nodes from the same transmitting node are correlated with correlation coefficient

\[
\rho(x) = \exp\left(-\frac{x}{d_{cor}}\log_{10} 2\right)
\]

where \( x \) is the Euclidean distance between two receiving nodes, \( d_{cor} \) is the de-correlation distance whose typical value is 20 meters for the urban environment and is 5 meters for the indoor environment [50]\(^6\). Fig. 6 is plotted using simulation results obtained from 40000 randomly generated network topologies, and following the correlated log-normal model. To have a fairer comparison between different shadowing effect assumptions, we adjust the density \( \lambda \) of ordinary nodes in each simulation so that the average node degree \( \mu \) of an arbitrary ordinary node is preserved under different path loss exponent \( \alpha \) and shadowing variance \( \sigma^2 \) settings. Ignoring the border effect, we have \( \mu = 2\lambda r_o \exp\left(\frac{\sigma^2}{2(\log_{10} 10)^2}\right) \). The steps to derive the equation are omitted here but a similar calculation can be found for two-dimensional plane in [52]. When \( \sigma = 0 \), the log-normal model reduces to the unit disk model and we have \( \mu = 2\lambda r_o \). As a result, Fig. 5 can be directly compared with Fig. 6 as the former is also plotted with the average node degree ranges from 0 to 8. Fig. 5 and 6 together show that the impact of the powerful nodes on the type-II connectivity probability under the correlated log-normal model has quantitatively little difference to the impact of the powerful nodes under the unit disk model. They are effectively the same from the qualitative point of view. Further, an increase in shadowing variance \( \sigma^2 \) will improve the connectivity probability even if the node density has been reduced to preserve the same average node degree. The better type-II connectivity probability observed under the log-normal model is consistent with the results in ad-hoc networks without infrastructure support (e.g. see [3], [53]).

Next we investigate the impact of \( \lambda \) and \( n_p \) on the average number of clusters under the unit disk model. The analytical formula Eq. (23) are verified by simulations obtained from 40000 randomly generated network topologies. [t] Fig. 7 shows virtually an exact match of Eq. (23) with the simulation results. In addition, the curves in the figure also agree with the curves in Fig. 5 and show that the connectivity probability

\[^6\] This is the well-known Gudmundson model. It works well for 2-D networks but it may not be able to accurately model some situations in 1-D networks. Particularly, consider a big obstacle located between a powerful node and two ordinary nodes in proximity where the radio signals between the powerful node and the ordinary nodes cannot propagate across. One ordinary node cannot receive the signals from the powerful node implies that the other ordinary node also cannot receive the signals either. That is, two nodes "hiding" behind a big obstacle in a 1-D network are highly correlated compared to the 2-D case. Gudmundson model is less suitable in modeling such situation. Nevertheless, Gudmundson model is still used in this paper due to its popularity in the literature.
reaches its minimum when the average number of clusters is maximized; conversely, the connectivity probability approaches one when the average number of clusters approaches one.

VIII. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a concept termed type-II connectivity to investigate the connectivity problem of infrastructure-based wireless multi-hop networks. Analytical results for the connectivity probability and the average number of clusters were obtained for 1-D networks with arbitrarily distributed powerful nodes and Poissonly distributed ordinary nodes which communicate following the unit disk model. First, the result proves that an increase in the number of powerful nodes in a network always has a positive impact on the connectivity probability and the average number of clusters. On the other hand, an increase in the number of ordinary nodes first degrade the connectivity probability. However, once the number of ordinary nodes increases beyond a certain value, the connectivity probability improves as the number of ordinary nodes further increases. Second, we proved that equidistant placement of powerful nodes will minimize the average number of clusters and maximize the type-II connectivity probability of a network. Finally, simulation results showed that the qualitative conclusions obtained under the unit disk model are also valid for other connection models, taking the log-normal model as a specific example.

Several issues remain for future work. Extending the results beyond one dimension is clearly desirable. Apart from that, in this paper we consider the unit disk model. In future, we may replace the unit disk model with a more realistic/generic random connection model. Furthermore, we assume that the ordinary nodes are Poissonly distributed. Other types of node distribution can be considered in the future. In addition, the impact of node mobility and work/sleep cycle on type-II connectivity may be studied.

REFERENCES


