2013 International Conference on Computing, Networking and Communications (ICNC)

Utilizing Distance Distribution in Determining Topological Characteristics of Multi-hop Wireless Networks

Husnu Narman
Department of Computer Science
University of Oklahoma
husnu@ou.edu

Turgay Korkmaz
Department of Computer Science
University of Texas at San Antonio
korkmaz@cs.utsa.edu

Suleyman Tek
Department of Mathematics
University of the Incarnate Word
tek@uiwtx.edu

***Abstract*—Understanding the topological characteristics of communications networks is essential to the design and operation of such networks. Accordingly, the researchers have been analyzing the topological characteristics of various networks. In this paper, we focus on *multi-hop wireless networks* with uniformly distributed nodes in *2D* and *3D* environments. Using the geometrical relations in *2D*, researchers have provided analytical formulas for some topological characteristics. However, generalizing this approach to *3D* is difficult as it requires to carefully analyze several geometrical relations that overlap. Instead, we consider the *distance distribution* between two random points, which has been extensively studied in both *2D* and *3D*. Using these important results, we derive analytical formulas for various topological characteristics in both *2D* and *3D*. Using simulations, we verify the correctness of our analytical formulas.**

Keywords: Random Networks; Topological characteristics; Distance distribution

I. Introduction

Multi-hop wireless networks have been receiving significant attention due to their potential use in various real-life applications [1]. Accordingly, the research community has been extensively investigating these networks in different forms such as WSN [2], [3], MANET [4], Mesh [5]. In many applications (e.g., forest or border monitoring), wireless nodes are expected to be uniformly distributed through random deployment in a given environment. One of the key issues during the deployment and operation of such multi-hop wireless networks is how to determine various system parameters (e.g., transmission range r, the number of nodes n) so that the resulting topology can support the given mission while leading to efficient use of the underlying resources.

For example, an operator may want to know what would be the best value for r and n so that the average path length is less than a certain number. In another case, an operator may want to select r such that the average degree will be less than a certain number to reduce interference between neighboring nodes. To answer such questions, it is necessary to understand how various parameters impact the topological characteristics of multi-hop wireless networks. This understanding is also essential to performance evaluation studies requiring realistic topologies.

To understand topological characteristics, one can use simulations to randomly generate networks under various parameters and analyze them. However, simulations cannot consider every possible case or answer various what-if questions in a timely manner, necessitating the development of analytical formulas. Therefore, we focus on developing analytical formulas for the *topological characteristics* of multi-hop wireless networks. Since such networks can be randomly deployed in *2D* (e.g., field, forest) or *3D* (e.g., ocean, air) environments, we derive analytical formulas for both *2D* and *3D*. We use simulations to verify the correctness of the analytical formulas.

As we review in Section II, the most related work to ours considers geometrical relations in *2D* and obtains analytical formulas for *link probabilities* [6], [7]. These probabilities are then used to determine other characteristics such as average degree. However, generalizing this approach to *3D* was difficult (if not impossible) due to the excessive amount of geometrical relations that need to be considered while avoiding overlaps. Instead, we took another approach, where we compute the *link probabilities* using the *distance distribution* between any two random points. Finding the *distance distribution* between two random points is a challenging problem by itself. Fortunately, we found out that the researchers in other areas have done significant work on finding the *distance distribution* between two random points in both *2D* and *3D* [8], [9].

Accordingly, we consider these results (specifically, the ones collectively presented in [10]) and compute link probabilities in both *2D* and *3D*. We then use these link probabilities to determine analytical formulas for average node degree and average path length (hop count) of the randomly deployed wireless networks in *2D* and *3D*. We also give an approximation result for diameter for *3D* by using similar idea with [6]. Using simulations, we verify the correctness of our analytical formulas.

The rest of this paper is organized as follows. In section II we review the related work. In Section III we describe the network model and simulation environment used throughout this paper. In Section IV we derive the link probabilities by

This research is supported by DoD Infrastructure Support Program for HBCU/MI, Grant: 54477-CI-ISP (UNCLASSIFIED).

using distance distributions in *2D* and *3D*, and verify the correctness of these link probabilities using simulations.In Section V, we derive the analytical formulas for the three topological characteristics that we mentioned before, and verify their correctness using simulations. Finally, we conclude this paper in Section VI.

II. Related Works

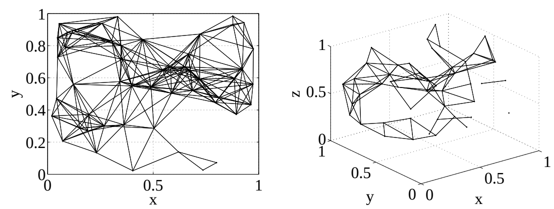
Researchers have studied the topological characteristic of wireless networks under various assumptions while mainly considering networks in *2D*. In [11], the authors simulated random, wireless sensor, regular, and small world networks to compare the network by using average path length, cluster coefficient, and average degree.They concluded that the topology of wireless network is between random and small-world network. The authors in [12], [13], [14] mainly researched on connectivity of wireless network using poisson and geometric network models. In [14] the author analyzed k-connectivity on homogeneous and heterogeneous ad hoc networks by using $\frac{\pi r^2}{A}$ as the estimate for link probability, where *A* is the area of the field where the network is deployed. In [12] the authors used $\frac{\pi r^2}{A}$ in finding upper and lower bound disjoint paths between two nodes by making their network model be similar to poisson random graph. In [13], the authors again used $\frac{\pi r^2}{A}$ for boundless area network to approximate communication link probability formula for bounded area network and analyzed k-connectivity and k-edge-connectivity.

Since the methods in [12], [13], [14] rely on approximate link probability, they cannot give exact results about topological characteristics. As discussed in the previous section, the authors in [6], [7] developed a better formula for communication link probability between two nodes in a unit square. In this paper, we made the link probability formula more accurate using the distance distribution. We also generalized our approach and developed analytical formulas for the topological characteristics of the networks in *3D* while it was difficult to generalize the methods using geometrical relations to *3D*.

III. Network Model and Simulation Environment

In *2D* we consider the same multi-hop wireless network model that is used in the literature [6], [7]. In this model, the wireless nodes having the same transmission range r are uniformly distributed in a normalized unit square of 1×1, where $0 < r \leq 1$. Similarly, we generalize this model for *3D* such that the wireless nodes having the same transmission range r are uniformly distributed in a unit cube of $1 \times 1 \times 1$, where again $0 < r \leq 1$. In both *2D* and *3D*, wireless nodes can directly communicate with each other by broadcasting radio waves if the distance between them is less than the transmission range r. Otherwise, these nodes can indirectly communicate with each other through a path if there is one. So the multi-hop network that we consider in *2D* or *3D* is a graph $G = (V, E)$, where $V = \{v_1, v_2, v_3, v_4, \ldots, v_n\}$ and $E = \{(v_i, v_j) \mid d(v_i, v_j) \leq r, \quad i \neq j, \quad i = 1, 2, \ldots n, \quad j = 1, 2, \ldots n\}$. Note that $d(v_i, v_j)$ is the distance between v_i and v_j, and r is the normalized transmission range.

In our simulations, we simply follow the above model and randomly place the nodes in a unit square and a unit cube. After positioning the nodes, we find the distance between every pair of nodes by simply using the well known Euclidian distance formula. If the distance between two nodes is smaller than the given r, we put a link between them.



(a) *2D* (b) *3D*

Fig. 1: Sample multi-hop wireless networks in *2D* and *3D* with $n = 50$ and $r = 0.30$.

For instance, Figure 1 shows a sample network with $n = 50$ and $r = 0.30$ in *2D* and *3D*. In actual simulations conducted in the following two sections, we consider several replications of the networks with varying the number of nodes n and transmission ranges r.

IV. Link Probabilities in *2D* and *3D*

To derive analytical formulas for various topological characteristics of multi-hop wireless networks, we need to first determine the *link probability*. We can formally define the link probability as $Prob[d(u, v) \leq r]$, where $d(u, v)$ is the distance between two arbitrary nodes u and v, and r is the communication range. In this section, we determine analytical formulas for the *link probabilities* in *2D* and *3D*, and verify their correctness using simulations. In the next section, we will use these probabilities to determine analytical formulas for the topological characteristics mentioned before.

A. Link Probability in 2D

Using geometric relations between the communication circle with radius r and a unit square of 1×1, the author in [6] computed the *link probability* as a function of r. The idea here is to first position an arbitrary node u and then determine what percent of the other nodes will fall within the communication circle of node u. For this, we need to compute the area of the intersection between the communication circle of the node u and the unit square. Since we are considering a unit square and a normalized transmission range, this area directly gives us the probability of having links between node u and other nodes in the network, i.e., the *link probability*.

Computing this area requires to consider several cases. For example, the simplest case is that node u is located r units away from every edge of the unit square and r is less than 0.5. In this case, the communication circle of the node u will be fully contained in the unit square. So the intersection area (i.e., the link probability) is simply computed as πr^2. In other cases (e.g., node u is located less than r units away from one of the edges of the unit square), the communication circle of node u will not be fully contained in the unit square, requiring some efforts to compute the area. The author in [6] examined several such cases and derived the following formula for *link probability* in *2D* when $0 \leq r \leq 1$:

$$F_{geo}(r) \stackrel{\text{def}}{=} \pi r^2 - \frac{8}{3}r^3 + \left(\frac{11}{3} - \pi\right) r^4 \tag{1}$$

Finding such a formula in *3D* requires to analyze several geometric relations. Also it is difficult (if not impossible) to avoid the overlaps between different cases even in *2D*. Therefore, we decided to consider the extensively studied *distance distributions* between randomly positioned nodes in a unit square and a unit cube. Actually, the distance distribution between two random points has been first studied back in 1950s [8]. In a recent study [10], the author collectively presented several useful results regarding distance distributions inside convex bodies of any dimensions. For our purposes, we will consider only the probability density function (pdf) and cumulative distribution function (cdf) of the distance between two random points in a unit square and a unit cube. Note that when determining *link probabilities*, we will consider only the parts of the pdf or cdf that are accounting for the distances less than 1 because the normalized transmission range r is less than 1. However, when considering average path length in Section V, we will consider the cases where the distance is greater than 1.

Let us first consider how to use distance distribution to find an analytical formula for link probability in *2D* and show why this new formula is more accurate than the one given in (1). In 2D, we can directly use the cdf in [10] (page 858). Fortunately, since we are interested in the distance less than the communication range r, we will take just the part accounting for the distance less than 1. Accordingly, we can give the link probability with our notation as follows when $0 \leq r \leq 1$.

$$F_{2D}(r) \stackrel{\text{def}}{=} r^2 \left(\frac{r^2}{2} - \frac{8r}{3} + \pi\right) \tag{2}$$

Clearly, $F_{geo}(r)$ and $F_{2D}(r)$ look similar. However, there is a small difference as shown below:

$$\epsilon \quad = \quad F_{geo}(r) - F_{2D}(r) = r^4 \frac{1}{42} \tag{3}$$

$F_{2D}(r)$ is more accurate than $F_{geo}(r)$ because some geometric overlaps are not accurately accounted for when determining F_{geo}(r). Another way to show the accuracy of $F_{2D}(r)$ over $F_{geo}(r)$ is to test them when $r = 1$. In that case, $F_{2D}(1)$ is approximately 0.97 while $F_{geo}(1)$ is 1. But the latter cannot be true because even when $r = 1$ there will be no edges between the nodes located around the opposite corners of the unit square as the length of the diagonal is $\sqrt{2}$. So even when $r = 1$, the link probability should be less than 1, as more accurately computed by $F_{2D}(r)$.

B. Link Probability in 3D

To obtain *link probability* in *3D*, we use the distance distribution in a unit cube, for which a pdf was first found in [9] and recently also presented in [10] (page 862). The complete form of this pdf is rather length. However, again since we are just interested in the distance less than the communication range r, we will take the part of this pdf accounting for the distance less than 1. With our notation, the pdf of the distance inside a cube will be as follows when $0 \leq r \leq 1$.

$$f_{3D}(r) = 4\pi r^2 - 6\pi r^3 + 8r^4 - r^5 \tag{4}$$

To obtain the *link probability* in *3D*, we need to find cdf by integrating equation (4) from 0 to r. With our notation, the *link probability* that the distance between two random points in *3D* is less than r can be computed as follows when $0 \leq r \leq 1$:

$$\begin{aligned} F_{3D}(r) &\stackrel{\text{def}}{=} \int_0^r f_{3D}(t)dt \\ &= \int_0^r (4\pi t^2 - 6\pi t^3 + 8t^4 - t^5)dt \end{aligned} \tag{5}$$

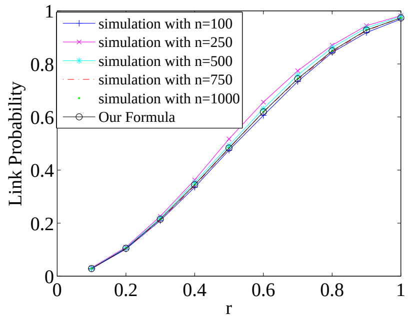
Note that we change the variable name r to t in pdf function so that we can take the integral without any confusion. After taking the integral from 0 to r, we obtain the link probability in *3D* as follows when $0 \leq r \leq 1$:

$$F_{3D}(r) \stackrel{\text{def}}{=} \frac{4\pi r^3}{3} - \frac{6\pi r^4}{4} + \frac{8r^5}{5} - \frac{r^6}{6} \tag{6}$$

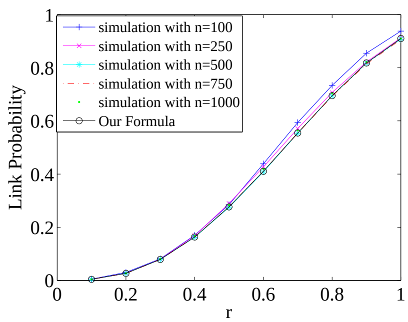
C. Verification of Analytically Obtained Link Probabilities by Using Simulations

In order to show the correctness of the analytical formulas given in (2) and (6), we conduct some simulations with different number of nodes n and different values of r. We have developed our simulation by using Matlab. First of all, we have randomly distributed nodes in unit square and unit cube. We have chosen unit cube for easiness but it can easily be modified for any size of square and cubes. Then we have counted number of links between nodes by using Euler distance calculation for a given communication link r. We have found link probability by dividing number of links to possible number of links which is $n(n-1)$ where n is number of nodes. In all cases, our analytical formula perfectly matches with the simulation results, as shown in Figure 2.

Note that all analytical formulas are actually assume that the number of nodes n goes to ∞. But through these simulations, we show that these analytical formulas are still giving accurate results even when the number of nodes are relatively small. We should also mention that each data point in the figure is actually the average of 20 replications, resulting in 95% confidence intervals to be very small; and thus we did not included them in the figure. Figure 2 also allows us to make a few other interesting observations. The *link probability* in *2D* increases faster than that in *3D* as r increases. So, to reach the same link probability, *3D* networks need to use more energy than *2D* networks. In both *2D* and *3D*, the increase in *link probability* before $r \simeq 0.6$ is much faster than that after



(a) *2D*



(b) *3D*

Fig. 2: Comparison of link probabilities from simulations and analytical formulas.

$r \simeq 0.6$. This information can help operators to determine up to what level increasing power can significantly improve link probability. After this point instead of increasing power, the operators may want to change other parameters to achieve certain topological characteristics.

For example, suppose an operator wants to increase the connectivity. He can achieve this by increasing the transmission power of each node or by increasing the number of nodes with less transmission power. If the power level is already increased such that $r \simeq 0.6$, then the operator may consider increasing the number of nodes rather than their powers.

V. Topological Characteristics and Their Verification using Simulations

In this section, we derive analytical formulas for the following topological characteristics, and verify their correctness through simulations.

- *Average Degree:* The degree of a node is the number of nodes that are within its communication range. We try to determine the average of degrees of all nodes. Degree is an important characteristic as it is used to determine transmission interference.
- *Diameter:* The length of the shortest path that has the maximum number of hops among all shortest paths between every pair of nodes. Knowing diameter of a network will enable operators to determine the worst-case data transmission latency between any two nodes.
- *Average Path Length:* The average length of the shortest paths between every pair of nodes. This can be used to determine how efficient data transmission will be on average.

A. Average Degree

The idea behind average degree is to compute the expected number of links between an arbitrary node u and the other $n-1$ nodes. Since we know the *link probabilities* from the previous section,we can easily find average (expected) degree as follows

$$E_{\aleph D} = (n-1) F_{\aleph D}(r)$$

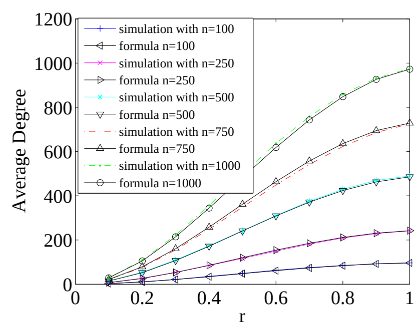
where $\aleph$ is 2 and 3. Accordingly, the analytical formula for average degree in *2D* would be

$$E_{2D} = (n-1)\left(r^2\left(\frac{r^2}{2} - \frac{8r}{3} + \pi\right)\right) \tag{7}$$

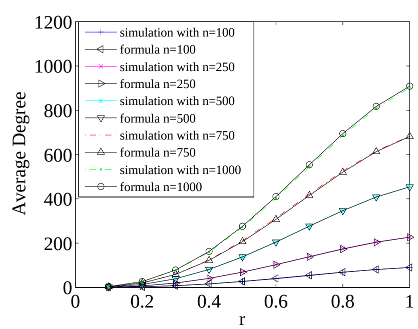
and the average degree in *3D* would be

$$E_{3D} = (n-1)\left(\frac{4\pi r^3}{3} - \frac{6\pi r^4}{4} + \frac{8r^5}{5} - \frac{r^6}{6}\right) \tag{8}$$

In order to show the correctness of the analytical formulas given in (7) and (8), we conduct some simulations with different number of nodes n and different values of r. In all cases, our analytical formula perfectly matches with the simulation results, as shown in Figure 3.



(a) *2D*



(b) *3D*

Fig. 3: Comparison of average degrees from simulations and analytical formulas.

Again each data point in the figure is actually the average of 20 replications, resulting in 95% confidence intervals to be very small; and thus we did not included them in the figure. As we observed before, the link probability in *2D* increases faster than that in *3D*. This also cause the average degree in *2D* to grow faster than that in *3D*, as shown in Figure 3.

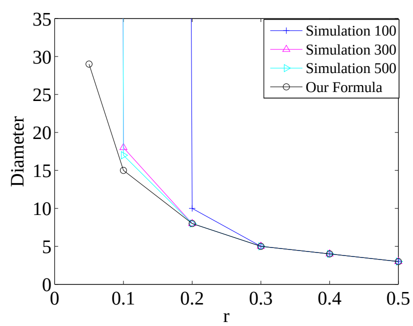
B. Diameter

The maximum number of hops are expected to appear on the shortest paths between the nodes placed in the opposite corners of the unit square or the unit cube. The distance between such nodes is approximately equal to the length of the longest diagonal, which is $\sqrt{2}$ in a unit square and $\sqrt{3}$ in a unit cube. So a lower bound on the diameter in terms of hop count can be determined by dividing the length of the longest diagonal by r as follows:

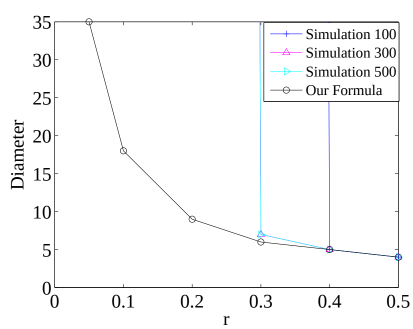
$$Diameter_{\aleph D} \geq \left\lceil \frac{\sqrt{\aleph}}{r} \right\rceil \tag{9}$$

where $\aleph$ is 2 and 3. As also shown in [6], the diameter of a randomly deployed network will be equal to the above lower bound as the number of nodes, n, goes to ∞.

Using simulations with large number of nodes, we verified the above asymptotic formula. However, instead of reporting such exact matches, we would like to see how the above asymptotic formula captures the general trend even with smaller number of nodes. Accordingly, we conducted simulations with n=100, 300, and 500 while varying r.



(a) $2D$



(b) $3D$

Fig. 4: Comparison of diameter from simulations and analytical formulas when n=100, 300, and 500.

Figure 4 presents the comparison results when n=100, 300, and 500. In both *2D* and *3D*, analytical diameter formula almost exactly matches to simulation curve for large values of r even when n is 100. As the number of nodes increases, we see better matches with smaller values of r. However, for very small values of r, the simulated networks are usually not connected because of the small number of nodes. Consequently, the diameter goes to infinity. Except these very small values of r and n, the analytical formulas provide good estimate for diameters in moderate size networks. As the number of nodes increases, the simulated networks become connected even with small values of r and their diameters approaches to the analytical formula.

C. Average Path Length

A lower bound on average path length in terms of hop count can be determined by dividing the *expected distance* between two random points by r. So we need to first find the expected distance. Since we know the distance pdf's in *2D* and *3D*, say $f_{2D}(t)$ and $f_{3D}(t)$, from [8], [9], we can determine the expected distance using

$$\int_0^{\sqrt{\aleph}} t f_{\aleph D}(t) dt, \tag{10}$$

where $\aleph$ is 2 and 3. Note that when computing the *link probabilities*, we just considered the part of the pdf accounting for the distances less than 1 because the normalized transmission range r was less than or equal to 1. But now we are interested in the average distance in the whole unit square or unit cube. So we need to consider all the parts of the pdf and take the above integral from 0 to $\sqrt{\aleph}$, which is the length of the longest diagonal. We can then compute the lower bound on the average path length (in terms of hop counts) as follows:

$$E_{hop\aleph D} \geq \frac{\int_0^{\sqrt{\aleph}} t f_{\aleph}(t) dt}{r} \tag{11}$$

where $\aleph$ is 2 and 3.

While the parts of the pdf's accounting for $0 \leq t \leq 1$ were easy to integrate, the other parts accounting for $1 < t < \sqrt{\aleph}$ cannot easily be evaluated due to some trigonometric terms in them. Fortunately, we have encountered that the average distances in a unit square and a unit cube are numerically computed in [15] as 0.52140543 and 0.661707182, respectively.Using these results along with (11),we can determine the lower bound on average path length in *2D* as

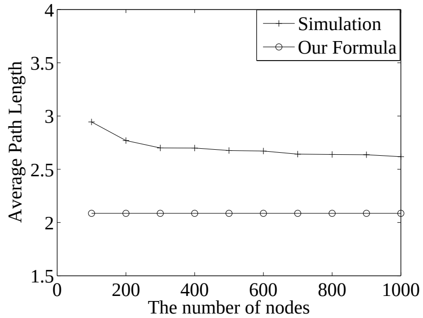
$$E_{hop2D} \geq \frac{0.52140543}{r} \tag{12}$$

and in *3D* as

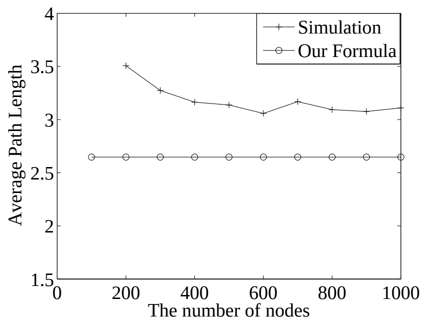
$$E_{hop3D} \geq \frac{0.661707182}{r}. \tag{13}$$

The average path lengths will be equal to the above lower bounds as n goes to ∞. Actually, we verified this using large number of nodes in simulations. Instead of these expected results, again we report the simulations showing how the above analytical formulas provide a good approximation even with relatively small number of nodes. For example, Figure 5 shows the result for average path length with $r = 0.25$ and varying number of nodes n. Simulation results are close to the analytical ones and they are getting closer as the number of nodes increases in both *2D* and *3D*.

We also conducted simulations with different values of r while fixing n to moderate sizes such as 100 and 300. As shown in Figure 6, even when $n = 100$ the path length closely follows the lower bound provided by the analytical formulas.

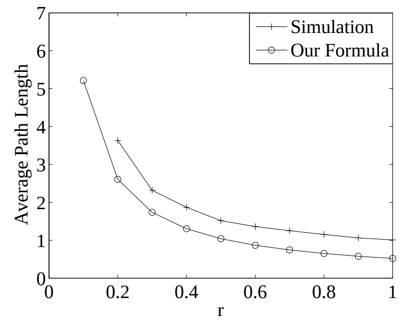


(a) *2D*

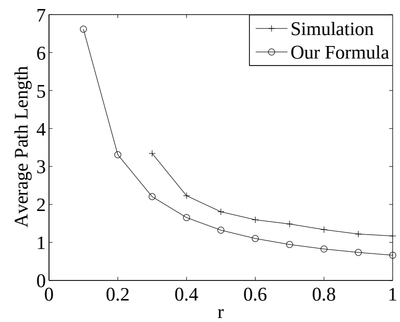


(b) *3D*

Fig. 5: Comparison of average path length (number of hops) from simulations and analytical formulas under different network sizes.



(a) *2D*



(b) *3D*

Fig. 6: Comparison of average path length (number of hops) from simulations and analytical formulas under different values of r when $n = 100$.

VI. Conclusions

We considered the topological characteristics of multi-hop wireless networks with uniformly distributed nodes in *2D* and *3D* environments. Using the extensively studied distance distributions between two random points in *2D* and *3D*, we first determined more accurate analytical formulas for the link probabilities. We then used these link probabilities and derived analytical formulas for three topological characteristics of multi-hop wireless networks. We also conducted simulations to verify our analytical formulas. Such analytical results are essential to the operators as they try to answer several what-if questions when setting various system parameters.

References

[1] L. Gavrilovska, S. Krco, V. Milutinovic, I. Stojmenovic, and R. Trobec, Eds., *Application and Multidisciplinary Aspects of Wireless Sensor Networks: Concepts, Integration, and Case Studies*. Spring-Verlag, London Limited, 2011.

[2] H. Karl and A. Willig, *Protocols and Architectures for Wireless Sensor Networks*. Wiley, 2006.

[3] I. Stojmenovic, *Handbook of Sensor Networks: Algorithms and Architectures*. Wiley-InterScience, 2005.

[4] I. Chlamtac, M. Conti, and J. J.-N. Liu, "Mobile ad hoc networking: imperatives and challenges," *Ad Hoc Networks*, vol. 1, pp. 13–64, July 2003.

[5] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer Networks*, vol. 47, no. 4, pp. 445–487, March 2005.

[6] K. Li, "Topological characteristics of random multihop wireless networks," *Cluster Computing*, vol. 8, pp. 119–126, July 2005.

[7] G.-L. Lu and H.-X. Wang, "Continue research on topological characteristics of random multihop wireless networks," in *International Conference on Communication Systems*, Nov 2008.

[8] B. Ghosh, "Random distances within a rectangle and between two rectangles," *Bull. Calcutta Math. Soc*, 1951.

[9] F. Piefke, "Beziehungen zwischen der Sehnenlängenverteilung und der Verteilung des Abstandes zweier zufälliger Punkte im Eikörper," *Probability Theory and Related Fields*, vol. 43, pp. 129–134, 1978.

[10] A. Mazzolo, "Probability density distribution of random line segments inside a convex body:application to random media," *Mathematical Physics*, vol. 44, no. 2, Feb 2003.

[11] R. Yueqing and X. Lixin, "A study on topological characteristics of wireless sensor network based on complex network," in *International Conference on Computer Application and System Modeling*, Oct 2010, pp. 486–489.

[12] F. Sun and M. Shayman, "On the average pairwise connectivity of wireless multihop networks," in *IEEE GLOBECOM*, Nov 2005.

[13] Q. Ling and Z. Tian, "Minimum node degree and k-connectivity of a wireless multihop network in bounded area," in *IEEE GLOBECOM*, Nov 2007, pp. 1296–1301.

[14] C. Bettstetter, "On the connectivity of wireless multihop networks with homogeneous and inhomogeneous range assignment," in *Vehicular Technology, IEEE Conference*, vol. 3, Dec 2002.

[15] J. Philip, "The probability distribution of the distance between two random points in a box," *Mathematics Subject Classification*, 1991.