

Quantile Models for the Threshold Range for k -connectivity

Henri Koskinen

Helsinki University of Technology (HUT), Networking Laboratory
P.O. Box 3000, FIN-02015 HUT, Finland

henri.koskinen@hut.fi

ABSTRACT

This study addresses the problem of k -connectivity of a wireless multihop network consisting of randomly placed nodes with a common transmission range, by utilizing empirical regression models for the threshold range for k -connectivity when the nodes are uniformly distributed in a square region. The cases $k = 1, 2, 3$ are considered: with $k = 1$, the models are based on known asymptotic results to assure correct limiting behavior; with $k = 2, 3$, an attempt is made to generalize these results and the models are built accordingly. Verification with independent simulation data shows all the models to be able to predict k -connectivity with good accuracy under this network model.

Categories and Subject Descriptors

C.2.1 [Computer-communication networks]: Network architecture and design—*wireless communication, distributed networks, network topology*; G.2.2 [Discrete mathematics]: Graph theory; G.3 [Probability and statistics]: *distribution functions, correlation and regression analysis*; I.6.6 [Simulation and modeling]: Simulation output analysis

General Terms

Reliability, Performance, Design, Theory

Keywords

Ad hoc networks, Boolean model, connectivity, empirical models, geometric random graphs, sensor networks

1. INTRODUCTION

The problem of connectivity in wireless multihop networks has been widely studied and can be formulated as follows. Suppose that two network nodes can establish a direct link between them if and only if they are each within the other's transmission range (this is commonly referred to as the Boolean network model). Suppose further that all

the nodes have a common transmission range r (or, more generally, limit for the range). Assuming that there are n nodes randomly located in some region, how are the number of nodes and the transmission range (relative to the dimensions of the region) related so that the resulting network topology is connected with high probability?

Equivalently, we may try to find the probability that the transmission range r exceeds the *threshold range* for connectivity R for a random set of nodes. The threshold range is defined as the smallest value of r for which a given set of nodes is connected, and it is equal to the greatest edge length in the minimum spanning tree of the nodes when the length of an edge is defined as the Euclidean distance between its endpoints (see e.g. [8]; hereafter, we will refer to graphs with such edge length definition as Euclidean graphs). The connectivity problem therefore reduces to knowing the distribution of the greatest edge length of the random Euclidean minimum spanning tree: the probability of connectivity of a random network with transmission range r equals the value of the corresponding cumulative distribution function at r .

The distribution in question is known asymptotically (as the number of nodes n tends to infinity), for uniformly distributed points in a square domain. The asymptotic distribution has been derived by Penrose in [4] and is presented in the next section. In contrast, no exact analytical results exist for finite n . There is hence a gap in the knowledge of the distribution of R that prevents accurately predicting the connectivity of this network model with a finite number of nodes. The importance of knowing this distribution in the physically realistic non-asymptotic regime has been recognized by e.g. D'Souza et al. in [2]: studying the distribution by simulation when the nodes are uniformly distributed in a square region, their aim was to see whether determining the distribution with various n had predictive power, by modelling the behavior of the mean of the distribution as a function of n .

The estimated parameter characterizing the model for the mean in [2] was the asymptotic value of R as n tends to infinity while the node density remains fixed. On the other hand, it has been shown already by Philips et al. in [7] that with a constant node density, R^2 must grow logarithmically with the network area - and hence the number of nodes - and therefore does not have a finite limit. Thus, the model used in [2] is in contradiction with known asymptotic results. As it happens, the purely empirical models in our earlier work [3] have the same shortcoming, as noted therein.

In this study, we have the same objective of predicting the distribution of R by extrapolation from simulation data

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MSWiM'04, October 4-6, 2004, Venezia, Italy.

Copyright 2004 ACM 1-58113-953-5/04/0010 ...\$5.00.

as in [2] and [3], but we take the known asymptotic results in [4] as the starting point. Keeping to the case of uniformly distributed nodes in a square region, we focus on the convergence of individual quantiles of R to the asymptotic distribution. Furthermore, we attempt to generalize this treatment to k -connectivity: a k -connected network remains connected after the removal of any $k - 1$ nodes. The notion of the threshold range can be readily generalized to k -connectivity: in general, the threshold range can be defined as the minimized greatest edge length of a spanning k -connected Euclidean graph, i.e., a k -connected Euclidean graph containing all the nodes in a given set. Accordingly, we denote the threshold range for k -connectivity with R_k , making R_1 the threshold range R discussed above. Also, for later reference, we denote by M_k the greatest edge length of the k -nearest neighbor graph, which is the Euclidean graph where each node is connected by an edge to the k th nearest other node.

One supplementary remark is in order. While the goal in the analytical treatment of k -connectivity in this paper is to derive an approximation to the asymptotic distribution of R_k , $k > 1$, the distribution has recently been derived exactly by Wan and Yi in [9]. Although it is straightforward to take this exact distribution as the basis of our regression models, we became aware of these results regrettably late with respect to the publication schedule of this paper to make the required modifications.

This document is organized as follows. The next section deals with modelling R_1 . The attempt to generalize the treatment to R_k , $k > 1$, is made in Section 3. Section 4 demonstrates the ability of the models to predict independent simulation results and at the same time presents one possible application scenario for the models. Section 5 contains discussion of our models and of the generality of applicability of the used approach. Finally, Section 6 presents some concluding remarks.

2. QUANTILE MODELS FOR R_1

2.1 Preliminaries

Let us use the following results given by Penrose in [4] and [6], respectively, as the starting point.

THEOREM 2.1. *For n points placed uniformly at random on the unit square, let M_1 denote the longest edge-length of the nearest neighbor graph on these points. Then*

$$\lim_{n \rightarrow \infty} P[n\pi M_1^2 - \log n \leq \alpha] = \exp(-e^{-\alpha}), \quad \alpha \in \mathbb{R}. \quad (1)$$

In other words, asymptotically $n\pi M_1^2 - \log n$ is Gumbel-distributed. As remarked by Penrose, the qualitative meaning of this theorem is that the asymptotics of M_1 are as if the nearest-neighbor distances of the points were independent. The validity of this statement will be demonstrated in detail in Section 3.1.

THEOREM 2.2. *For n points uniformly randomly distributed on the unit cube in d dimensions, with $d > 1$, let R_k (respectively M_k) denote the minimum r at which the graph, obtained by adding an edge between each pair of points distant at most r apart, is k -connected (respectively, has minimum degree k). Then $P[R_k = M_k] \rightarrow 1$ as $n \rightarrow \infty$.*

In the context of our model, this theorem means that when the number of nodes is large enough, then with high probability, if one starts to increase the common transmission range of the nodes from zero, the network becomes k -connected as soon as the last node with only $k - 1$ direct links vanishes. Thus, R_k and M_k have the same asymptotic distribution, and therefore R_1 can be substituted for M_1 in Theorem 2.1.

Fixing the right hand side of (1) to a desired probability q (which implies that $\alpha = \alpha_q = -\log(-\log q)$) then allows us to write

$$\lim_{n \rightarrow \infty} P \left[R_1 \leq \sqrt{\frac{\log n + \alpha_q}{n\pi}} \right] = q. \quad (2)$$

This means that as n tends to infinity, the q -quantile of R_1 - denote this by $r_1(q, n)$ - tends to the square-root expression in (2), which contains the q -quantile α_q of the Gumbel distribution present in (1). This gives us important prior information for finding a model for $r_1(q, n)$ with the correct asymptotic behavior: interpreting α_q as the limit of some function $\alpha_q(n)$, we assume that

$$r_1(q, n) = \sqrt{\frac{\log n + \alpha_q(n)}{n\pi}} \quad (3)$$

for all n . It then remains to find an empirical model for $\alpha_q(n) = n\pi r_1^2(q, n) - \log n$ using quantile estimates $\hat{r}_1(q, n)$ obtained by simulation.

2.2 Building the model

The sample size in the simulations was 5000 for each n , with n ranging from 5 to 350. Because the quantile estimates are obtained from simulation data with limited sample sizes, the model is initially built for the 50% quantile so that the most accurate estimates are used. Figure 1 shows the quantile estimates and the corresponding estimates $\hat{\alpha}_{0.5}(n)$ plotted against n . We know now that $\lim_{n \rightarrow \infty} \alpha_{0.5}(n) = \alpha_{0.5} = -\log(\log 2)$. We first concentrate on the tail of $\alpha_{0.5}(n)$ to find the rate at which it decays towards this limit; therefore, we observe instead the deviation from the limit, $\tilde{\alpha}_{0.5}(n) = \alpha_{0.5}(n) - \alpha_{0.5}$, and use only the data points forming the tail, i.e. those from $n = 25$ onwards.

When plotted on a log-log scale, the tail seems to decrease linearly, which implies a power-law decay. This is verified by the residuals (the differences between the data points and the fitted model) of all the data points with respect to such a tail model, which are plotted in Figure 2(a). It then remains to find a supplement to the model for the low-end data points. Figure 2(b) shows the logarithms of the first residuals after their signs have been changed. The linear trend implies that the low-end supplement could be of the

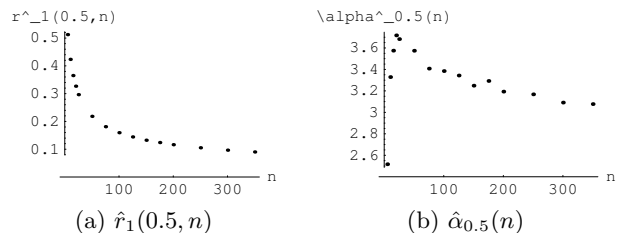


Figure 1: Values obtained from the 50% quantile estimates of R_1

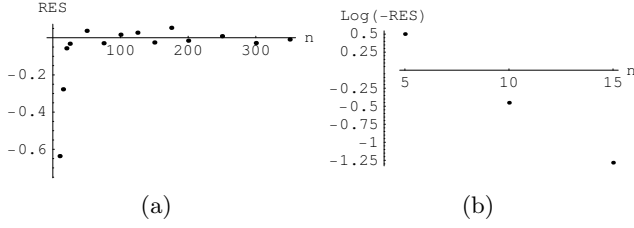


Figure 2: Residuals of the model fitted for the tail of $\tilde{\alpha}_{0.5}(n)$

form $-c \exp(-dn)$.

This proved to be a fair assumption: Figure 3(a) shows the final model of the form

$$\alpha_q(n) = an^{-b} - c \exp(-dn) + \alpha_q, \quad a, b, c, d > 0 \quad (4)$$

(where, for now, $q = 0.5$), and Figure 3(b) its residuals, obtained after all the parameters have been re-estimated using nonlinear regression to eliminate the effect of considering only part of the data points in fitting the piecewise models. In particular, the residuals seem to be evenly scattered around zero level, showing no trend as an indication of an invalid form for the model. The same applies to the overall model thus obtained for $r_1(0.5, n)$, the absolute and relative residuals of which are shown in Figures 3(c) and 3(d), respectively. Figures 3(b) and 3(d) are almost identical in pattern which implies that fitting the parameters by minimizing the sum of squared errors of $\hat{\alpha}_q(n)$ is nearly equivalent to minimizing the sum of squared relative residuals of $\hat{r}_1(q, n)$. This is a sensible choice because, as shown by Figure 3(c), the variance of R_1 (and hence that of $\hat{r}_1(q, n)$) decreases with n , and the data points with smaller variance yield more accurate information and therefore deserve more emphasis when fitting the model.

To sum up, we have found a model for the 50% quantile of R_1 given by equations (3) and (4) where $\alpha_{0.5} = -\log(\log 2)$; the parameter estimates are given in Table 1. The confi-

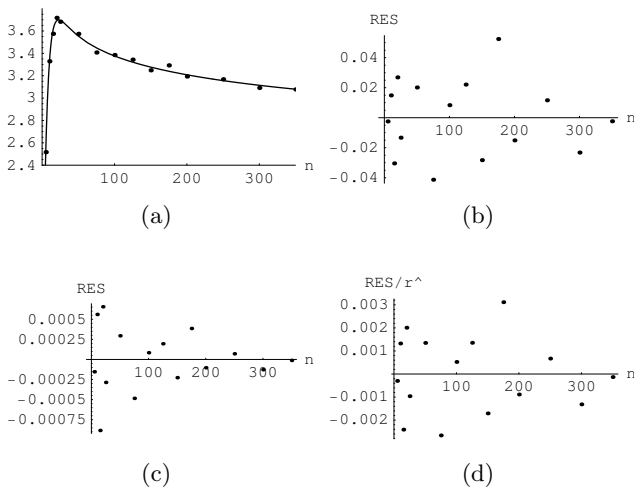


Figure 3: The model (a) and its residuals (b) obtained for $\alpha_{0.5}(n)$ and the overall absolute (c) and relative (d) residuals for $r_1(0.5, n)$.

Table 1: Parameter estimates and their confidence intervals for the model (4) for $\alpha_{0.5}(n)$

Parameter	Estimate	Confidence interval
a	4.41	[4.23,4.60]
b	0.0829	[0.0741,0.0917]
c	4.18	[3.80,4.60]
d	0.179	[0.157,0.202]

dence intervals of the parameters a and c can be seen to overlap, implying that these two parameters could be combined.

Encouraged by the well-performing model obtained for the 50% quantile, the same model was fitted to a selection of other quantiles. As shown by Figure 4, the behavior of the relative residuals of all these models resemble that of the first model, except for the variance that increases with the quantile due to the inaccuracy inherent in estimating extreme quantiles from a limited sample size. The parameter estimates are given in Table 2; the confidence intervals were no longer valid for statistical inference. It can be seen that whereas parameter b in the power law does not seem to change significantly with the quantile, the effect of the exponential part lasts longer as the rate of decay (determined by parameter d) becomes slower. The overall effect of these trends is shown by Figure 5. The "amplitudes" a and c seem to remain close to each other as observed with

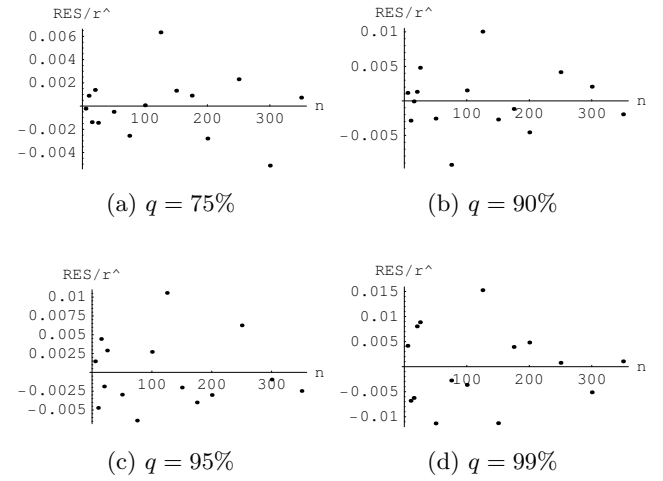


Figure 4: Relative residuals for various quantiles of R_1

Table 2: Parameter estimates for the model (4) for $\alpha_q(n)$, for various quantiles of R_1

Quantile	a	b	c	d
75%	5.49	0.0787	4.92	0.183
90%	7.18	0.0789	5.83	0.159
95%	8.03	0.0644	6.05	0.142
99%	12.2	0.0742	11.0	0.122

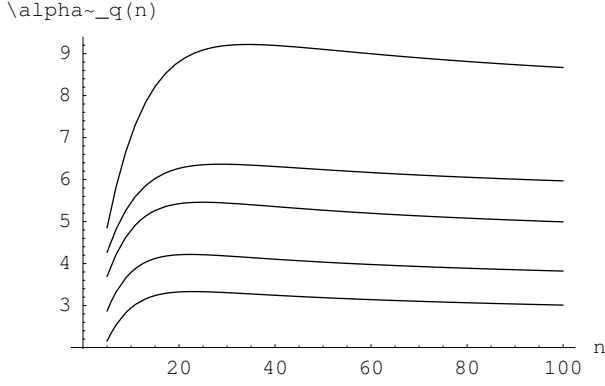


Figure 5: The models for $\tilde{\alpha}_q(n) = \alpha_q(n) - \alpha_q$, with q ranging from 50% (lowest curve) to 99% (highest curve)

the first model. However, the attempt to combine them resulted in the deterioration of both the representation of low-end data and the predictions of independent simulation data.

3. GENERALIZATION TO R_k

3.1 Preliminaries

As presented in [4], the generalization of Theorem 2.1 to the k -nearest neighbor graph is

$$\lim_{n \rightarrow \infty} P[n\pi M_k^2 - \log n - (k-1)\log(\log n) + \log(k-1)! \leq \alpha] = \exp(-e^{-\alpha}), \quad \alpha \in \mathbb{R} \quad (5)$$

which, combined again with Theorem 2.2, also applies to R_k . The problem here is that this was proven to hold for $k > 1$ only in the toroidal model where the opposite borders of the unit square are assumed to coincide, which eliminates boundary effects and in effect makes the domain a torus.

However, let us now demonstrate the qualitative meaning of Theorem 2.1. Consider the approximation derived by Bettstetter in [1] for the probability that a random network, with $n \gg 1$ nodes generated from a homogeneous Poisson process with intensity λ and employing a transmission range r , has minimum degree at least one (i.e., every node is connected to at least one other node). In the spirit of Theorem 2.1, let us focus on the unit square so that we may take $\lambda = n$. Using our notation, the approximation then states

$$P[M_1 \leq r] = (1 - \exp(-n\pi r^2))^n. \quad (6)$$

Here, $1 - \exp(-n\pi r^2)$, the probability that a single random node is not isolated, can be interpreted as the cumulative distribution function of a single nearest-neighbor distance, and the distribution of M_1 is calculated as that of the maximum of n such independent and identically distributed nearest-neighbor distances. Equivalently, we have

$$\log P[M_1 \leq r] = n \log(1 - \exp(-n\pi r^2)).$$

With the assumption $n \gg 1$, we may say that the expected degree of a node (neglecting boundary effects) $n\pi r^2 \gg 1$, for such r that the probability in question differs significantly from zero. (In fact, Theorem 2.1 states that it is very likely not much less than $\log n$, but this is the essence of

the weaker result derived already in [7] and mentioned in Section 1.) This makes $\exp(-n\pi r^2)$ small, so we may use the approximation

$$\log(1+x) \approx x, \quad |x| \ll 1 \quad (7)$$

to obtain

$$\begin{aligned} \log P[M_1 \leq r] &\approx -n \exp(-n\pi r^2) \\ \Leftrightarrow -\log(-\log P[M_1 \leq r]) &\approx n\pi r^2 - \log n \stackrel{\text{def}}{=} \alpha. \end{aligned}$$

Expressing the event $M_1 \leq r$ equivalently using the right hand side, we get

$$P[n\pi M_1^2 - \log n \leq \alpha] \approx \exp(-e^{-\alpha}),$$

which is precisely the limit in Theorem 2.1.

Having thus found that the very simple expression (6) is in fact asymptotically correct, it then seems tempting to conjecture that its generalization to $k > 1$ is as well, i.e., to assume that the asymptotics are as if also the k -nearest-neighbor distances were independent. Thus, take the generalized form of (6),

$$P[M_k \leq r] = \left(1 - \exp(-n\pi r^2) \sum_{i=0}^{k-1} \frac{(n\pi r^2)^i}{i!}\right)^n, \quad (8)$$

where we merely preclude the Poisson point probabilities up to $k-1$ independently for each node. Regarding only $k \ll n\pi r^2$, we again utilize (7) and arrive at

$$-\log(-\log P[M_k \leq r]) \approx n\pi r^2 - \log n - \log \left(\sum_{i=0}^{k-1} \frac{(n\pi r^2)^i}{i!} \right).$$

Writing the sum in the following nested form

$$\sum_{i=0}^{k-1} \frac{(n\pi r^2)^i}{i!} = \left(1 + n\pi r^2 \left(1 + \frac{n\pi r^2}{2} \left(1 + \frac{n\pi r^2}{3} (\dots)\right)\right)\right)$$

and keeping in mind that $n\pi r^2 \gg 1$, we make the approximation $\log(1 + n\pi r^2/i) \approx \log(n\pi r^2/i)$, the error of which diminishes as $n\pi r^2 \rightarrow \infty$, obtaining

$$\begin{aligned} -\log(-\log P[M_k \leq r]) &\approx \\ n\pi r^2 - \log n - (k-1)\log(n\pi r^2) + \log(k-1)! &\stackrel{\text{def}}{=} \alpha \\ \Leftrightarrow \end{aligned}$$

$$\begin{aligned} P[n\pi M_k^2 - \log n - (k-1)\log(n\pi M_k^2) + \log(k-1)! \leq \alpha] \\ \approx \exp(-e^{-\alpha}), \end{aligned} \quad (9)$$

because α is an increasing function of r when $n\pi r^2 > k-1$. Comparing this with (5), we see that the two are the same if we assume $\log(n\pi M_k^2) \approx \log(\log n)$. Although Theorem 2.1 shows this assumption to be valid when $k=1$, this is not necessarily the case with $k > 1$. In fact, equation (9) implies that $n\pi M_k^2 \approx \log n + (k-1)\log(n\pi M_k^2) - \log(k-1)!$ which, when substituted recursively into $\log(n\pi M_k^2)$ in (9), suggests that this approximation would not be very far off. As this would nevertheless result in a less conservative requirement for M_k - which might account for the relaxed conditions for network connectivity induced by the toroidal distance metric - we will not make this approximation and will therefore use equation (9) as our approximated generalization of Theorem 2.1.

Substituting again M_k with R_k according to Theorem 2.2, we may then write accordingly

$$P \left[R_k \leq \sqrt{\frac{\log n + (k-1) \log(n\pi R_k^2) - \log(k-1)! + \alpha_q}{n\pi}} \right] \xrightarrow{n \rightarrow \infty} q. \quad (10)$$

If the right-hand side of the inequality in (10) increased everywhere slower as a function of R_k than R_k itself, we could say that the q -quantile of R_k is the unique fixed point of the square-root expression with each n ; this is in general not the case. However, taking the equivalent squared inequality

$$R_k^2 \leq \frac{\log n + (k-1) \log(n\pi R_k^2) - \log(k-1)! + \alpha_q}{n\pi}$$

and requiring that the derivative with respect to R_k of the left hand side is greater than that of the right hand side yields

$$2R_k \geq \frac{2(k-1)}{n\pi R_k} \quad \Leftrightarrow \quad n\pi R_k^2 \geq (k-1),$$

i.e. the right-hand side increases slower whenever $n\pi R_k^2$, the expected degree of a node with range R_k , is at least $k-1$, which must hold almost surely. Based on the approximate equation (9), we may therefore after all say that as n tends to infinity, the q -quantile of R_k denoted by $r_k(q, n)$ satisfies

$$r_k(q, n) = \sqrt{\frac{\log n + (k-1) \log(n\pi r_k(q, n)^2) - \log(k-1)! + \alpha_q}{n\pi}}. \quad (11)$$

Just as with R_1 , we may now assume that this holds for all n with $\alpha_q = \lim_{n \rightarrow \infty} \alpha_q(n)$ and find an empirical model for $\alpha_q(n) = n\pi r_k(q, n)^2 - \log n - (k-1) \log(n\pi r_k(q, n)^2) + \log(k-1)!$.

3.2 Models for R_2 and R_3

The simulations of R_2 and R_3 were carried out with the same sample sizes and the same n as with R_1 , using the algorithms presented in [3] for determining the threshold

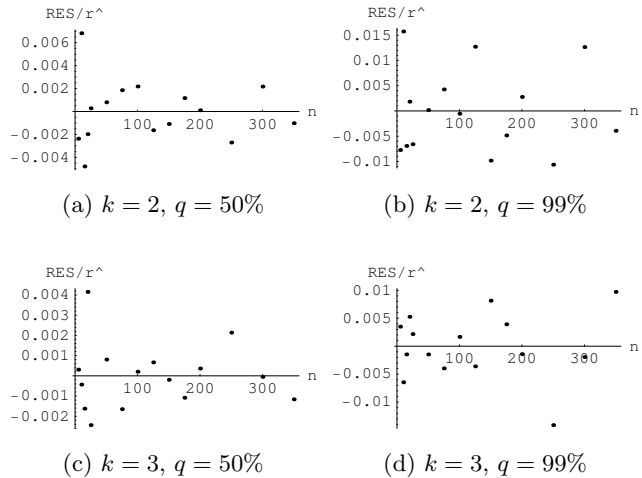


Figure 6: Relative residuals for quantiles of R_k

Table 3: Parameter estimates for the model (4) for $\alpha_q(n)$, for various quantiles of R_k

		Quantile				
		50%	75%	90%	95%	99%
$k = 2$	a	5.73	7.21	9.48	11.3	16.1
	b	0.0754	0.0650	0.0620	0.0616	0.0588
	c	4.32	5.29	7.00	9.51	13.5
	d	0.140	0.143	0.129	0.124	0.0875
$k = 3$	a	7.01	9.27	11.4	13.0	15.6
	b	0.0639	0.0621	0.0463	0.0380	0.0125
	c	5.25	6.52	9.24	10.8	14.0
	d	0.137	0.123	0.126	0.115	0.107

ranges. It turns out that the model (4) remains sufficient to describe $\alpha_q(n)$ when $k > 1$. The relative residuals of the most extreme quantiles of R_k (in terms of quantile estimate variance), $k = 2, 3$, are shown in Figure 6 and the parameter estimates for the different quantiles in Table 3.

4. MODEL VALIDATION USING INDEPENDENT SIMULATION DATA

The most important argument for the models presented here is their ability to predict the independent simulation results presented in [1]. This is demonstrated here with the aid of the following example scenario used therein:

”(Design of a large-scale wireless sensor network): A wireless sensor network should cover an area of size $A = 500 \times 500 \text{ m}^2$. Since all sensors exchange information, e.g. for environmental monitoring, the network should be connected. The sensors are equipped with transceivers that transmit a range of $r_0 = 20 \text{ m}$ in free space and do not perform power control. How many sensors do we need to distribute over the area?”

With permission of the author, Figure 7 is excerpted from [1] and shows the related simulation results (the analytical curve represents the asymptotic relation (6)); the predic-

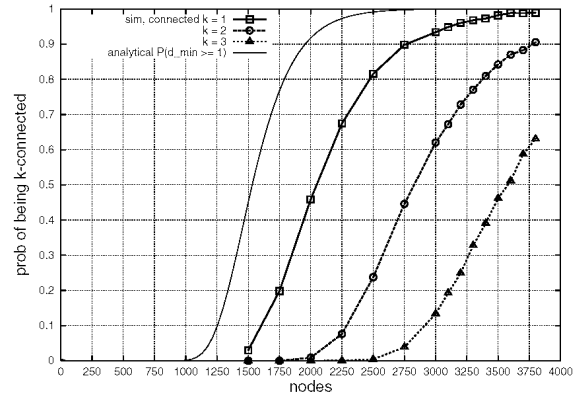


Figure 7: ”Simulation results for n nodes with $r_0 = 20 \text{ m}$ uniformly distributed on $A = 500 \times 500 \text{ m}^2$ {...}, 3000 random topologies” [1]

Table 4: Predictions of the quantile models for R_1, R_2, R_3 for the required number of nodes n

	Probability				
	50%	75%	90%	95%	99%
$k = 1$	2057	2387	2790	3144	3871
$k = 2$	2805	3262	3807	4208	
$k = 3$	3533	4065			

tions of our quantile models to this example scenario are given in Table 4. Comparison of the two shows that although the models were fitted to simulation data involving no more than $n = 350$ nodes, their predictions turn out to be quite accurate up to $n = 2000$; there are some visible deviations when $n > 2500$. Furthermore, the models for R_2 and R_3 seem to perform as well as those for R_1 , implying that equation (9) serves as a reasonable approximation to the real asymptotic distribution of R_k when $k > 1$.

5. DISCUSSION

Let us first remark that although the total of four parameters can be regarded as excessive to define the above models, it is easy to see that the effect of the exponential portion of the model is only relevant with up to 50 nodes and can be neglected in scenarios involving much more nodes. The two parameters characterizing the power-law portion are therefore sufficient to describe larger-scale networks.

The generality of applicability of the approach used in this study, i.e. modelling the convergence of the distribution of R_k to the asymptotic one by regression, is dictated by the various definitions and assumptions that it relies on. First of all, the underlying Boolean model of a wireless multihop network does not take the effects of interference of concurrent transmissions on connectivity into account. Instead, it can be used to represent such a network in the presence of a constant-level background noise only and is therefore applicable in studying ultimate limits for connectivity or connectivity in networks with low transmission activity. Furthermore, the definition of the threshold range for connectivity is based on the assumption that all network nodes have the same transmission range. As for the spatial distribution of the nodes, one can of course study the distribution of R_k with any spatial distribution (note that this also applies to the stationary spatial distribution of a mobility model), but the asymptotic distribution of R_k is not known for arbitrary spatial distributions (for R_1 it has been derived for a symmetrical normal distribution by Penrose in [5]).

Moreover, the asymptotic distribution can also depend on the shape of the network domain; in this study, we have only focused on the square-shaped region. However, the fact that the asymptotic distribution in Theorem 2.1 was reached by starting from the approximation (6) that in no way takes the domain shape into account implies that the asymptotic distribution of R_1 is to some extent independent of the shape of the domain. This is in fact verified in [9]: it turns out that the asymptotic distribution of R_1 is the same in a unit-area square and a unit-area disk. As, nevertheless, the shape does affect the distribution with finite n by determining the strength of the boundary effect, the methods presented here could be used in studying the convergence to the asymptotic

distribution in domains of different shapes. On the other hand, as shown in [9], the asymptotic distribution of R_k , $k > 1$ is different in the square and in the disk, and therefore presumably in every individual shape of domain. It thus appears that under rather general conditions for the shape of the domain, the border effect has a negligible contribution to the asymptotic distribution when $k = 1$, whereas this is no longer the case when $k > 1$. This implies that the determination of the asymptotic distribution in the latter case requires a separate analytical treatment for every shape of domain.

6. CONCLUSION

This study presented empirical regression models for individual quantiles of the threshold range for k -connectivity when the network nodes are uniformly distributed in a square region. The aim of the work was to fill some of the gap in the knowledge of the distribution of the threshold ranges which has prevented accurately predicting k -connectivity of random networks with a finite number of nodes under this network model. The models took into account known asymptotic results and were obtained by fitting to simulation data that accurately characterizes connectivity, and therefore they do not rely on approximating the connectivity of a network with its minimum degree in other than the asymptotic regime as justified by Theorem 2.2. Consequently, they were able to predict results from independent simulations with good accuracy.

7. ACKNOWLEDGMENTS

I thank C. Bettstetter for allowing the use of his simulation results in this paper, and Jorma Virtamo and the anonymous reviewers for their helpful comments. This work has been financially supported by the Academy of Finland (grant n:o 202204) and the Finnish Defence Forces Technical Research Centre, and in part by a grant from the Nokia Foundation.

8. REFERENCES

- [1] C. Bettstetter. On the minimum node degree and connectivity of a wireless multihop network. In *Proc. 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'02)*, pages 80–91, June 2002.
- [2] R. D'Souza, S. Ramanathan, and D. Lang. Measuring performance of ad hoc networks using timescales for information flow. In *Proc. The 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003)*, Apr. 2003.
- [3] H. Koskinen. A simulation-based method for predicting connectivity in wireless multihop networks. *Telecommunication Systems*, 26(2-4):321–338, June 2004.
- [4] M. D. Penrose. The longest edge of the random minimal spanning tree. *Annals of Applied Probability*, 7(2):340–361, 1997.
- [5] M. D. Penrose. Extremes for the minimal spanning tree on normally distributed points. *Advances in Applied Probability*, 30(3):628–639, 1998.

- [6] M. D. Penrose. On k -connectivity for a geometric random graph. *Random Structures and Algorithms*, 15(2):145–164, 1999.
- [7] T. K. Philips, S. S. Panwar, and A. N. Tantawi. Connectivity properties of a packet radio network model. *IEEE Transactions on Information Theory*, 35(5):1044–1047, Sept. 1989.
- [8] M. Sánchez, P. Manzoni, and Z. J. Haas. Determination of critical transmission range in Ad-Hoc Networks. In *Proceedings of Multiaccess Mobility and Teletraffic for Wireless Communications 1999 Workshop (MMT'99)*, Oct. 1999.
- [9] P.-J. Wan and C.-W. Yi. Asymptotic critical transmission radius and critical neighbor number for k -connectivity in wireless ad hoc networks. In *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, pages 1–8. ACM Press, 2004.