

# The $K$ -Neigh Protocol for Symmetric Topology Control in Ad Hoc Networks

Douglas M. Blough  
Georgia Inst. of Technology  
Atlanta GA – USA  
doug.blough@ece.gatech.edu

Giovanni Resta  
IIT – CNR  
Pisa – ITALY  
giovanni.resta@iit.cnr.it

Mauro Leoncini  
Università di Modena e  
Reggio Emilia - ITALY  
leoncini@unimore.it

Paolo Santi  
IIT – CNR  
Pisa – ITALY  
paolo.santi@iit.cnr.it

## ABSTRACT

We propose an approach to topology control based on the principle of maintaining the number of neighbors of every node equal to or slightly below a specific value  $k$ . The approach enforces symmetry on the resulting communication graph, thereby easing the operation of higher layer protocols. To evaluate the performance of our approach, we estimate the value of  $k$  that guarantees connectivity of the communication graph with high probability. We then define  $k$ -NEIGH, a fully distributed, asynchronous, and localized protocol that follows the above approach and uses distance estimation. We prove that  $k$ -NEIGH terminates at every node after a total of  $2n$  messages have been exchanged (with  $n$  nodes in the network) and within strictly bounded time. Finally, we present simulation results which show that our approach is about 20% more energy-efficient than a widely-studied existing protocol.

## Categories and Subject Descriptors

C.2.1 [Computer Systems Organization]: Computer-Communication Networks—*Network Architecture and Design: Wireless Communication*

## General Terms

Algorithms, Performance

## Keywords

Wireless ad hoc networks, topology control, energy consumption, symmetric range assignment, connectivity.

## 1. INTRODUCTION

It is a widely accepted fact that the limited energy available at the nodes of a wireless ad hoc network must be used

as efficiently as possible. If energy conservation techniques are used at different levels, the functional lifetime of both individual nodes and the network can be extended considerably. For this reason, energy conserving protocols at the MAC, routing, and upper layers have been proposed [7, 12, 15]. Further energy can be saved if the network topology itself is energy-efficient, i.e., if the nodes' transmitting ranges are set in such a way that a target property (e.g., connectivity) of the resulting network topology is guaranteed, while the global energy consumption is reduced. A protocol that attempts to achieve this is called a *topology control* protocol. Several examples of topology control mechanisms have been recently introduced [1, 5, 11, 16, 17, 23, 25, 30].

In order to be easily implementable in a realistic scenario, a topology control protocol should be *fully distributed*, *asynchronous*, and *localized* (i.e., the computation at every node should be based on information concerning neighbor nodes only). These features in general ensure that the protocol is fast and requires little message exchange; thus, it can be easily adapted to deal with dynamic and/or mobile networks. Another desirable property of a topology control protocol is that it does not rely on information that can be provided only by expensive devices, such as location information generated by a GPS receiver. In these conditions, the task of ensuring a global network property (e.g., connectivity) while reducing energy consumption is challenging.

In this paper, we evaluate the effectiveness of a topology control approach based on the generation of a symmetric subgraph of the  $k$ -neighbors graph. We evaluate (both theoretically and experimentally) the value of  $k$  that ensures connectivity with high probability (w.h.p.). We also present a specific protocol based on this approach, called  $k$ -NEIGH, that generates the desired topology in a fully distributed, asynchronous, and localized way. Our  $k$ -NEIGH protocol relies on distance estimation, a technique which can be implemented at a reasonable cost in many realistic scenarios. However, other implementations are possible (e.g., with GPS or using the approach of [30], in which a suitable set of neighbor nodes is computed by stepping up the transmission power). In case of  $k$ -NEIGH, we are able to prove that the overall number of messages exchanged is exactly  $2n$ , where  $n$  is the number of nodes in the network. We report the results of simulation experiments which show

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.

MobiHoc'03, June 1–3, 2003, Annapolis, Maryland, USA.  
Copyright 2003 ACM 1-58113-684-6/03/0006 ...\$5.00.

that our approach reduces energy consumption considerably with respect to the case where no topology control is used, and that it compares favorably with the CBTC protocol of [16, 30].

## 2. RELATED WORK

In [25], Rodoplu and Meng presented a distributed topology control algorithm that leverages on position information (provided by low-power GPS receivers) to build a topology that is proved to minimize the energy required to communicate with a given master node. Unfortunately, the protocol relies on global knowledge and specialized hardware (the GPS receiver), which makes it infeasible in many application scenarios. Further, the topology generated by the Rodoplu and Meng protocol (which is optimal for communications directed towards a single master node) can be significantly different from the energy optimal topology for the all-to-all communication scheme.

In [30], Wattenhofer et al. introduced a distributed topology control protocol based on directional information, called CBTC (Cone Based Topology Control). The basic idea is that a node  $i$  transmits with the minimum power  $p_{i,\rho}$  such that there is at least one neighbor in every cone of angle  $\rho$  centered at  $i$ . The obtained communication graph is made symmetric by adding the reverse edge to every asymmetric link. The authors show that setting  $\rho \leq 2\pi/3$  is a sufficient condition to ensure connectivity. A set of optimizations aimed at pruning energy-inefficient edges without impairing connectivity (and symmetry) is also presented. Further, the authors prove that if  $\rho \leq \pi/2$ , every node in the final communication graph has degree at most 6. A more detailed analysis of CBTC, along with an improved set of optimizations (which, however, rely on distance estimation), can be found in [16]. The CBTC protocol has been extended to the three-dimensional case in [1]. The authors of [1] also presented a modification of the protocol aimed at ensuring  $k$ -connectivity. In [11], the CBTC protocol is implemented using directional antennas.

In [5], Borbash and Jennings introduced a protocol which is also based on directional information. The goal of the protocol is to build the Relative Neighbor Graph of the network in a distributed fashion. The choice of the RNG as the target graph of the protocol is due to the fact that it guarantees connectivity and it shows good performance in terms of average transmitting range, node degree and hop diameter.

Contrary to the protocols described above, the goal of COMPOW [19] is to determine, using a fully distributed approach, the minimum *common* value of the transmitting range needed to ensure network connectivity. Recently, this approach has been extended to the case of non-uniform node distribution [13], a scenario in which COMPOW would perform poorly.

The protocols that are most closely related to our work are the MobileGrid protocol of [17] and the LINT protocol of [23]. Both protocols try to keep the number of neighbors of a node within a low and high threshold centered around an optimal value. When the actual number of neighbors is below (above) the threshold, the transmitting range is increased (decreased), until the number of neighbors is in the proper range. However, for both protocols no characterization of the optimal value of the number of neighbors is given, and, consequently, no guarantee on the connectivity of the result-

ing communication graph is provided. Another problem of the MobileGrid and LINT protocols is that they estimate the number of neighbors by simply overhearing control and data messages at different layers. This approach has the advantage of requiring no overhead, but the accuracy of the resulting neighbor number estimate heavily depends on the traffic present in the network. In the extreme case, a node which remains silent is not detected by any of its actual neighbors.

Similarly to MobileGrid and LINT, the goal of our approach is to keep the number of neighbors of a node equal to, or slightly below, a given value  $k$  that guarantees connectivity w.h.p. However, differently from the cited works, we are able to formally characterize the optimal (i.e., minimum) value of  $k$ .

Contrary to MobileGrid and LINT, our approach enforces symmetry on the resulting communication graph. As the support of unidirectional links is in general technically difficult and expensive (in terms of number of messages exchanged), we believe that the explicit requirement for a connected backbone of bidirectional links is vital in the design of a topology control mechanism. For further motivations of our symmetry requirement see Section 4.

Compared with CBTC, our  $k$ -NEIGH protocol (which implements the  $k$  neighbors approach outlined above) relies on a weaker assumption, i.e., distance estimation vs. directional information. Furthermore, CBTC has no bound on the number of messages nor on the energy expended in determining the proper transmit power, whereas in our algorithm each node transmits only two messages at a predefined power (the maximum transmit power). Finally, the simulation results reported in Section 6.3 show that the topologies generated by our protocol are (on the average) 20% more energy efficient than those generated by CBTC.

## 3. PRELIMINARIES

Let  $N$  be a set of  $n$  nodes placed in  $[0, 1]^2$  according to some distribution. A *range assignment* for  $N$  is a positive real valued function  $RA: N \rightarrow (0, r_{max}]$  that assigns to every element of  $N$  a value in  $(0, r_{max}]$ , representing its transmitting range. Parameter  $r_{max}$  is called the *maximum transmitting range* of the nodes in the network and depends on the features of the radio transceivers equipping the nodes. We assume that all the nodes are equipped with transceivers having the same features; hence, we have a single value of  $r_{max}$  for all the nodes in the network.

Given  $N$  and a range assignment  $RA$ , the *communication graph* induced by  $RA$  on  $N$  is defined as the directed graph  $G = (N, E)$ , where the directed edge  $[i, j]$  exists if and only if  $RA(i) \geq \delta(i, j)$ , and  $\delta(i, j)$  denotes the distance between nodes  $i$  and  $j$ . In this paper, we are concerned with two variants of this graph, defined as follows:

**DEFINITION 1.** *The symmetric super-graph of  $G$  is defined as the undirected graph  $G^+$  obtained from  $G$  by adding the undirected edge  $(i, j)$  whenever edge  $[i, j]$  or  $[j, i]$  is in  $G$ . Formally,  $G^+ = (N, E^+)$ , where  $E^+ = \{(i, j) | ([i, j] \in E) \text{ or } ([j, i] \in E)\}$ .*

**DEFINITION 2.** *The symmetric sub-graph of  $G$  is defined as the undirected graph  $G^-$  obtained from  $G$  by removing all the non-symmetric edges. Formally,  $G^- = (N, E^-)$ , where  $E^- = \{(i, j) | ([i, j] \in E) \text{ and } ([j, i] \in E)\}$ .*

The set of neighbors of a node  $i$ , denoted  $N(i)$ , is defined as the set of nodes to which  $i$  is directly connected, i.e.  $N(i) = \{j \mid [i, j] \in E\}$ . Neighbor sets are defined similarly in graphs  $G^+$  and  $G^-$ . Note that for these graphs  $i \in N(j)$  if and only if  $j \in N(i)$ .

Given a parameter  $k$ , with  $0 < k < n$ , the  $k$ -neighbors graph is the communication graph  $G_k$  in which every node is directly connected to its  $k$  nearest nodes. Formally,  $G_k$  is the communication graph induced by the range assignment  $RA_k$ , where  $RA_k(i) = \delta(i, j)$  and  $j$  is the  $k$ -th nearest node to node  $i$ .

Several connectivity problems on the communication graph have been studied in the literature. Before formally defining these problems, which are related to some extent to the problem considered in this paper, we need some further definitions.

A range assignment  $RA$  is said to be *connecting* if it induces a strongly connected communication graph, while it is said to be  *$r$ -homogeneous* if all the nodes have the same transmitting range  $r$ , with  $0 < r \leq r_{max}$ .

It is known [20] that the power  $p_i$  required by node  $i$  to correctly transmit data to node  $j$  must satisfy inequality  $\frac{p_i}{\delta_{ij}^\alpha} \geq \beta$ , where  $\alpha \geq 2$  is the *distance-power gradient* and  $\beta \geq 1$  is the *transmission quality* parameter. In ideal conditions we have  $\alpha = 2$ ; however, in general it is  $2 \leq \alpha \leq 6$  depending on environmental conditions. Setting  $\beta = 1$ , we can define the *energy cost* of a range assignment  $RA$  as  $c(RA) = \sum_{i \in N} (RA(i))^\alpha$ .

We are now ready to formally define the range assignment problems:

**DEFINITION 3.** *Let  $N$  be a set of points in  $[0, 1]^2$ , and assume that the  $r_{max}$ -homogeneous range assignment is connecting.*

**RA:** *Determine a connecting range assignment  $RA$  such that  $c(RA)$  is minimum.*

**WSRA:** *Determine a range assignment  $RA$  such that the symmetric sub-graph of the communication graph is connected and  $c(RA)$  is minimum.*

**HRA:** *Determine the minimum value of  $r$  such that the  $r$ -homogeneous range assignment is connecting.*

Note that HRA can be equivalently restated as the problem of finding a connecting homogeneous range assignment of minimum energy cost.

RA and WSRA have been shown to be NP-hard in the two and three-dimensional cases [3, 8, 14], while HRA can be easily solved if node positions are known. HRA has been studied also in the case of nodes distributed according to some probability distribution [10, 27].

In this paper, we are concerned with the following connectivity problem on the symmetric sub-graph of the  $k$ -neighbors graph. Motivations for our interest in  $G_k^-$  can be found in Sections 4 and 5.

**DEFINITION 4 (KNRA).** *(Same assumptions as in Definition 3). Determine the minimum value of  $k$  such that  $G_k^-$  is connected.*

As in the case of HRA, the problem can be equivalently restated in terms of minimum energy cost; furthermore, the optimal solution can be easily found if node positions are known. In the next section, we analyze KNRA in the hypothesis that nodes are distributed uniformly at random in  $[0, 1]^2$ . Our analysis will be used to provide a (probabilistic)

guarantee on the connectivity of the topology generated by our  $k$ -NEIGH protocol.

## 4. THE MINIMUM NUMBER OF NEIGHBORS FOR CONNECTIVITY

Our approach to topology control consists of setting the nodes' transmitting ranges in such a way that the resulting symmetric sub-graph  $G_k^-$  is connected w.h.p., using local information only. The choice of limiting our consideration to  $G_k^-$  is motivated by the following reasons:

- although implementing wireless unidirectional links is technically feasible (see [2, 21, 22, 24] for unidirectional link support at different layers), the actual advantage of using unidirectional links is questionable. For example, in [18] Marina and Das have shown that the high overhead needed to handle unidirectional links in routing protocols outweighs the benefits that they provide, and better performance can be achieved by simply avoiding unidirectional links;
- a recent theoretical result [3] has shown that the optimal solution to RA and WSRA have the same energy cost (asymptotically). In other words, starting from a strongly connected graph, obtaining a connected backbone of symmetric edges incurs in no additional (asymptotic) energy cost.

While a formal analysis of the conditions on  $k$  under which  $G_k$  is strongly connected w.h.p. is not straightforward, a recent result by Xue and Kumar [31] gives us the necessary technical machinery to work on its symmetric variants.

**THEOREM 1.** *Assume that  $n$  nodes are placed uniformly at random in  $[0, 1]^2$ , and let  $G_k^+$  be the symmetric super-graph of the  $k$ -neighbors graph. There exist two constants  $c_1, c_2$ , with  $0 < c_1 < c_2$ , such that:*

$$\lim_{n \rightarrow \infty} \text{Prob}\{G_{c_1 \log n}^+ \text{ is disconnected}\} = 1, \text{ and}$$

$$\lim_{n \rightarrow \infty} \text{Prob}\{G_{c_2 \log n}^+ \text{ is connected}\} = 1.$$

The authors also provide explicit values for  $c_1$  and  $c_2$ , which are  $c_1 = 0.074$  and  $c_2 > 5.1774$ .

Although the difference between the number of neighbors necessary and sufficient for connectivity is quite large, Theorem 1 is very important, since it states that  $\Theta(\log n)$  neighbors are necessary and sufficient for connectivity w.h.p.

Theorem 1 refers to the symmetric super-graph of  $G_k$ , in which a link that is physically unidirectional is considered as bidirectional. In other words, the connectivity of  $G_k^+$  is in general higher than that of  $G_k$ , since in  $G_k^+$  there are links that do not exist in the actual communication graph. As a consequence, the number of neighbors stated as sufficient to obtain connectivity w.h.p. in Theorem 1 may not be so in the actual communication graph. The following Theorem extends Xue and Kumar' result to the symmetric sub-graph  $G_k$ .

**THEOREM 2.** *The same result of Theorem 1, with  $G_k^+$  replaced by  $G_k^-$ .*

**PROOF.** The necessity part follows immediately by Theorem 1, since  $G_{c_1 \log n}^-$  is a sub-graph of  $G_{c_1 \log n}^+$ . To prove the sufficiency part, we have to show that the construction used in the proof of Theorem 1 holds for  $G_{c_2 \log n}^-$  also. The proof of Theorem 1 is based on the fact (proved in

[31]) that any node in  $G_{c_2 \log n}^+$  is directly connected w.h.p. to every node that is within distance of  $(1 - \epsilon)r_n$ , where  $r_n = \sqrt{\frac{\eta \log n}{\pi n}}$ ,  $\epsilon$  is an arbitrary constant in  $(0, 1)$ , and  $\eta$  is a constant that depends on  $\epsilon$ . In words, this means that the communication graph  $G_{(1-\epsilon)r_n}$  generated by the  $(1 - \epsilon)r_n$ -homogeneous range assignment is a sub-graph of  $G_{c_2 \log n}^+$  (asymptotically, for  $n \rightarrow \infty$ ). Since  $G_{(1-\epsilon)r_n}$  is connected w.h.p. (for  $n \rightarrow \infty$ ) by Theorem 3.2 in [10], then  $G_{c_2 \log n}^+$  is also connected w.h.p.. The proof of our Theorem follows immediately by observing that, since any node is directly connected w.h.p. to every node that is within distance of  $(1 - \epsilon)r_n$ , and distance is obviously symmetric,  $G_{(1-\epsilon)r_n}$  is a sub-graph of  $G_{c_2 \log n}^-$  too.  $\square$

Having a connected backbone of symmetric edges, as provided by the  $G_k^-$  graph, allows us to use standard bidirectional link-based protocols in the upper layers, avoiding the expensive and technically difficult implementation of unidirectional links. Given the theoretical result of [3] and Theorem 2, this additional requirement on the communication graph will come with a limited additional energy cost<sup>1</sup>.

## 5. THE $k$ -NEIGH PROTOCOL

In this section, we describe the  $k$ -NEIGH topology control protocol - an implementation of the computation of  $G_k^-$  - and prove its correctness and complexity.

The protocol is based on the following assumptions:

1. nodes are stationary;
2. the maximum transmission power  $P$  is the same for all the nodes;
3. given  $n$ ,  $P$  is chosen in such a way that the communication graph that results when all the nodes transmit at power  $P$  is connected w.h.p.;
4. a distance estimation mechanism, possibly error prone, is available to every node;
5. the nodes initiate the  $k$ -NEIGH protocol at different times. However, the difference between node wake up times is upper bounded by a known constant  $\Delta$ .

Assumption 4 is clearly the most critical and deserves some comments. The distance estimation techniques proposed in the literature so far are based on:

- *Radio Signal Strength*: distance is estimated comparing the transmitted power at the sender (which is piggybacked in the message) and the received power at the receiver of the message. This technique can be implemented at virtually no cost (RSSI registers are a standard feature in many wireless network cards [28]), but provides poor accuracy. In [28], it is shown that RSSI-based distance estimation is feasible only in a quite idealized setting (football field with all the nodes positioned at the ground level).

- *Time of Arrival*: distance is estimated comparing the time of arrival of different kinds of signals. Typically, the radio signal is used in combination with acoustic, ultrasound or infrared signals. ToA-based techniques provide a much better accuracy than RSSI-based mechanisms, and can be implemented at a reasonable cost. For example, the technique proposed in [9] uses a standard PC sound card to generate an acoustic signal, which is received by a cheap microphone. The authors show that this technique provides good accuracy (below 3%) in realistic conditions. However, accuracy

<sup>1</sup>In Section 6.1 we will validate this statement through extensive simulation.

drops to only 23% when the line of sight between the nodes is obstructed by heavy obstacles. In order to overcome this problem, several signals of different kind can be combined together.

Other well-known topology control protocols are based on assumptions that appear at least as much strong. For example, the protocol of [25] is based on location information, which is provided by a GPS receiver. Although the cost of such devices has decreased in recent years, and their form factor reduced, they are still expensive and cumbersome devices. Furthermore, the GPS signal can be received only in open air environments. The protocol of [30] and all of its variants [1, 11, 16], and the protocol of [5], are based on directional information, which can be provided using directional antennas (which are also very expensive).

At least in outdoor environments, such as in case of, e.g., sensor networks, our solution is then an acceptable implementation. Moreover, we did not overlook the potential effects of errors in the distance estimation mechanism. Hence, we have included realistic distance estimation error models in our simulator (see Section 6.2 for details).

In the protocol specification below, we assume without loss of generality that the first node wakes up at time 0. The protocol is as follows:

THE  $k$ -NEIGH PROTOCOL (FOR A GENERIC NODE  $i$ ):

1. Node  $i$  wakes up at time  $t_i$ , with  $t_i \in [0, \Delta]$ . At random time  $t_i^1$  chosen in the interval  $[t_i + \Delta, t_i + \Delta + d]$  (the value of the parameter  $d$  is set in Lemma 1), node  $i$  announces its ID at maximum power;
2. For every message received from other nodes,  $i$  stores the identity and the estimated distance of the sender;
3. At time  $t_i + 2\Delta + d$ ,  $i$  orders the list of its neighbors (i.e., of the nodes from which it has received the announcement message) based on the estimated distance; let  $L_i$  be the list of the  $k$  nearest neighbors of node  $i$  (if  $i$  has less than  $k$  neighbors,  $L_i$  is the list of all its neighbors).
4. At random time  $t_i^2$  chosen in the interval  $[t_i + 2\Delta + d + \tau, t_i + 2\Delta + 2d + \tau]$  ( $\tau$  is an upper bound on the duration of step 3), node  $i$  announces its ID and the list  $L_i$  at maximum power.
5. At time  $t_i + 3\Delta + 2d + \tau$  node  $i$ , based on the lists  $L_j$  received from its neighbors, calculates the set of *symmetric neighbors*<sup>2</sup> in  $L_i$ . Let  $L_i^S$  be the list of symmetric neighbors of node  $i$ , and let  $j$  be the farthest node in  $L_i^S$ .
6. Node  $i$  sets its transmitting power  $P_i$  to the power needed to transmit at distance  $\delta_{ij}^e$ , where  $\delta_{ij}^e$  is the estimated distance between nodes  $i$  and  $j$ .
7. (OPTIONAL PRUNING STAGE) Apply an optimization procedure to reduce the number of edges in the graph obtained so far (see below).

At the end of the protocol execution, node  $i$  considers as neighbors (e.g., for the purpose of routing) only the nodes in the list  $L_i^S$ . Note that these are *logical* neighbors, and the set of *physical* neighbors in general is larger than  $L_i^S$ : when  $i$  transmits at power  $P_i$ , it is possible that some node  $j \notin L_i^S$  receives the message. However, these are asymmetric neighbors, which are not considered. Also, the following pruning stage can be executed to further reduce the number of logical neighbors and (possibly) the actual transmission power required at some node.

<sup>2</sup>Nodes  $i$  and  $j$  are said to be *symmetric neighbors* if and only if  $i \in L_j$  and  $j \in L_i$ .

PRUNING STAGE OF THE  $k$ -NEIGH PROTOCOL (FOR A GENERIC NODE  $i$ ):

Let  $G_k^- = (N, E)$  be the undirected graph obtained as the results of steps 1–6 of the  $k$ -NEIGH protocol, and, for any  $(i, j) \in E$ , let  $P(i, j)$  denote the transmission power sufficient for  $i$  to reach node  $j$ . This information is included in the message sent by node  $i$  during step 4.

1. Node  $i$  sorts the list  $L_i^S$  according to increasing values of  $P(i, j)$  (initially, this is equivalent to the order given by the increasing distances from  $i$ ). Let  $j_1, \dots, j_k$  be the sorted list (without loss of generality, we assume that  $L_i^S$  contains  $k$  elements; otherwise, the sorted list will be composed by  $k_1 < k$  elements).

2. For  $l = 2, \dots, k$ , do the following.

- a. Check whether  $j_l$  can be reached using a transmission power lower than  $P(i, j_l)$  by routing through some  $j_q$ ,  $q < l$ . Clearly, given the information available to node  $i$ , this is possible only if  $(j_q, j_l) \in E$ , a circumstance that is known to  $i$  from step 5 of the  $k$ -NEIGH protocol.
- b. If  $P(i, j_q) + P(j_q, j_l) \leq P(i, j_l)$ , logically delete the (outgoing) edge  $(i, j_l)$  and set  $P(i, j_l) = P(i, j_q) + P(j_q, j_l)$ . If more than one node satisfies this requirement, choose the node  $q$  such that  $P(i, j_q) + P(j_q, j_l)$  is minimum.

3. Set the transmitting power to the power needed to reach the farthest node in  $L_i^S$  which is still an immediate neighbor of node  $i$ .

The following results prove that the  $k$ -NEIGH protocol is correct.

LEMMA 1. *Let  $\bar{t}$  be the time necessary to transmit a message. For  $d = m\bar{t}$ , the probability that no contention will occur in the wireless channel during step 1 of the  $k$ -NEIGH protocol is strictly greater than  $e^{-\frac{3h(h-1)}{2m}}$ , where  $h$  is the number of nodes that are contending for the channel when transmission is done at maximum power.*

PROOF. See the full version of the paper [4].  $\square$

Lemma 1 can be used to lower bound the probability of no contention when accessing the wireless channel. For example, if  $n=100$  nodes are distributed uniformly at random in a square region and  $P$  is chosen in accordance with Assumption 3, the expected number of nodes within the maximum transmitting range is about 33 (see Section 6 for details). Given these settings,  $d$  must be around  $16000\bar{t}$  to obtain a probabilistic guarantee of no contention of at least 0.9. With  $\bar{t}$  in the order of, say, milliseconds,  $d$  will be in the order of tenth of seconds, which is reasonable for most topology control scenarios. Clearly, Lemma 1 provides only a crude lower bound on  $Pr\{\text{no contention}\}$ , and smaller values of  $d$  can be used in practice.

LEMMA 2. *Let  $G_k^- = (N, E)$  be the undirected graph computed by steps 1–6 of the  $k$ -NEIGH protocol, and suppose  $G_k^-$  is connected. Let  $G' = (N, E')$  be the directed graph obtained as the result of the pruning stage of  $k$ -NEIGH. Then,  $G'$  is strongly connected and symmetric.*

PROOF. See the full version of the paper [4].  $\square$

THEOREM 3. *Assume that  $k$  is chosen in accordance with Theorem 2. Then the  $k$ -NEIGH protocol:*

- a. *terminates at time at most  $4\Delta + 2d + \tau$  (where  $d$  is set in Lemma 1), i.e., by this time all the nodes have set their transmitting power correctly and terminated the protocol execution;*
- b. *generates a symmetric communication graph which is connected w.h.p. under the hypothesis that nodes are distributed uniformly at random in  $[0, 1]^2$ ;*
- c. *has communication complexity  $\Theta(n)$ .*

PROOF. A generic node  $i$  wakes up at an arbitrary time  $t_i$  in  $[0, \Delta]$ . Before announcing its ID, node  $i$  has to wait at least time  $\Delta$  to avoid that its message is not received by nodes that are not yet awake. The additional random time (in the interval  $[0, d]$ ) is needed to avoid (with high probability) contention in accessing the wireless channel. Once the node has announced its ID at step 1, it has to wait for messages coming from other nodes. The waiting time is  $\Delta + d$ , accounting for the difference in the initial wake up times and for the maximum possible difference between random time choices. Thus, at time  $t_i + 2\Delta + d$ , node  $i$  can safely order neighboring nodes based on the distance estimated when the announcement messages are received. We recall that messages are sent at maximum power, which is the same for all the nodes by assumption. This implies that at time  $t_i + 2\Delta + d$  node  $i$  has received the announcement messages of all the nodes within its maximum transmitting range. The ordering phase lasts at most time  $\tau$ , and at time  $t_i + 2\Delta + d + \tau$  node  $i$  is ready to send the message containing its  $k$ -neighbors list. Once more, the node waits for an additional random time chosen in the interval  $[0, d]$  to avoid contention. Before ending the protocol, node  $i$  must be sure to have received the  $k$ -neighbors lists of all its neighbors, so that asymmetric neighbors can be removed. Thus, starting at time  $t_i + 2\Delta + d + \tau$  node  $i$  waits for further  $\Delta + d$  units of time. At time  $t_i + 3\Delta + 2d + \tau$ , node  $i$  is then ready to set its transmitting power correctly and the protocol execution in node  $i$  terminates. The proof of the part a. of the theorem follows by observing that the maximum possible value for the wake up time  $t_i$  is  $\Delta$ .

The proof of the part b. of the theorem follows by Theorem 2. The part c. of the Theorem is immediate, since every node sends exactly two messages. By definition of the pruning stage and by Lemma 2, it is immediate that pruning occurs with no further message exchange, and produces a graph which is connected w.h.p. and symmetric.  $\square$

## 6. SIMULATION RESULTS

To evaluate our  $k$ -NEIGH protocol we have designed an ad hoc simulator and performed a considerable body of experiments. The goals of our simulations include evaluation of:

- *preferred value of  $k$* : the result stated in Theorem 2 is mainly of theoretical interest. In the first set of experiments, we have evaluated which values of  $k$  should be used in practice to achieve a target probability (e.g., 0.95) of connectivity. We call this value the *preferred value of  $k$* ;
- *effect of errors*: as discussed in Section 5, distance estimation techniques are error-prone. We have evaluated the effect of errors in distance estimation on the preferred value of  $k$ ;
- *energy cost*: in the third set of experiments, we have compared the performance of our algorithm (in terms of energy

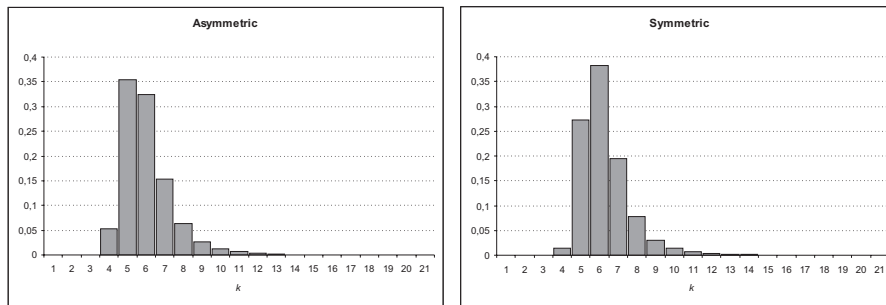


Figure 1: Empirical distribution of the minimum  $k$  for connectivity in the asymmetric (left) and symmetric (right) case for  $n=100$ . Data are shown as frequencies.

cost, as defined in Section 3) to that of other algorithms.

The results of each set of simulations are presented in separate subsections.

### 6.1 Preferred value of $k$

The preferred value of  $k$  is defined as the minimum value of the node degree  $k$  which guarantees that  $Pr(G_k^-$  is connected) is above a certain target probability. Since a theoretical characterization of this value is very difficult, we have evaluated it through extensive simulations.

The setting used for our experiments is the following. The  $n$  nodes, all with the same maximum transmitting range  $R_n$ , are distributed uniformly at random in  $[0, 1]^2$ . According to Assumption 3 of Section 5,  $R_n$  should be chosen so that the communication graph that results when all nodes transmit at maximum power is connected with high probability. An easy choice would be to make  $R_n$  independent of  $n$ , and sufficiently high to ensure connectivity w.h.p. even with very few nodes (e.g.,  $R_n = 1$ ). However, the choice of  $R_n$  has a strong influence on the energy cost of the graph generated by topology control algorithms in general. For example, in the CBTC protocol of [30], boundary nodes are very likely to transmit at full power (after the first phase of the protocol). As a consequence, larger values of  $R_n$  produce higher energy costs for CBTC.

For this reason, we have decided to choose  $R_n$  according to the following procedure: for every value of  $n$  considered in the simulations, we have generated 10000 random placements and, for every placement, we have evaluated the longest edge of the Euclidean MST.<sup>3</sup> Using these values, we have built the empirical distribution of the critical transmitting range, and taken the 0.95 quantile.<sup>4</sup> This value, further increased by 50% for safety, gives  $R_n$ . For all practical purposes, the transmitting range  $R_n$  calculated in this way accomplishes Assumption 3, and gives a uniform parameter that can be used in the implementation of  $k$ -NEIGH and other topology control protocols. The values of  $n$  used in our simulations and the corresponding values of  $R_n$  are shown in Table 1.

We have investigated the preferred value of  $k$  for different values of  $n$ . In the first experiment,  $n$  ranged from 10 to 100 in steps of 10. The reason for the small steps of  $n$

<sup>3</sup>It is known that this value corresponds to the critical transmitting range, in case the range assignment is homogeneous (see [26]).

<sup>4</sup>We recall that the  $q$  quantile of a series of data gives the point such that 100 $q$  percent of the data lie before.

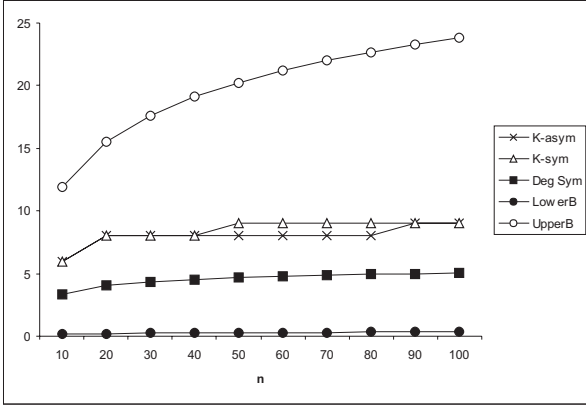
$n$	$R_n$	$n$	$R_n$
10	0.86622	75	0.37041
20	0.66420	80	0.36291
25	0.60431	90	0.34787
30	0.55589	100	0.33326
40	0.48635	250	0.23634
50	0.44526	500	0.19691
60	0.41456	750	0.17885
70	0.38336	1000	0.17274

Table 1: Values of the maximum transmitting range  $R_n$  used in our simulations.

is that in most ad hoc network applications the number of nodes is expected to be in this range. For every value of  $n$ , and for every random node placement, we have calculated the minimum value of  $k$  such that  $G_k$  is strongly connected (denoted  $k_{asym}$ ), and the minimum value of  $k$  such that  $G_k^-$  is connected (denoted  $k_{sym}$ ), subject to the constraint that every node has maximum transmitting range  $R_n$ . Given our choice for  $R_n$ , such minimum values for  $k$  always exist in practice. For each setting of  $n$ , we generated 100000 random node placements, and recorded  $k_{asym}$  and  $k_{sym}$  for each of them. These data gave us the empirical distribution of  $k_{asym}$  and  $k_{sym}$ , which can be used to evaluate the preferred value of  $k$ . The two distributions for the case of  $n = 100$  are shown in Figure 1. From the figure, it is evident that the requirement for symmetry has little influence on the minimum value of  $k$  for connectivity. This is made clearer by Figure 2, which reports the preferred value of  $k$  in the asymmetric and symmetric cases when the target probability of connectivity is set to 0.95. These values can be easily obtained by the cumulative distribution of  $k_{asym}$  and  $k_{sym}$ : the preferred value is the minimum value of  $k$  such that the cumulative frequency is above 0.95.

The plots reported in Figure 2 show that the preferred value of  $k$  in the symmetric case is at most 1 greater than the value in the asymmetric case. To a certain extent, this confirms the theoretical results of Theorem 2 and of [3]. Figure 2 also reports the average node degree in the symmetric case. We recall that  $k$  is the number of asymmetric neighbors, while only symmetric neighbors contribute to the node degree of  $G_k^-$ . The plot seems to confirm the logarithmic behavior predicted by Theorem 2.

In the second experiment, we have evaluated how the preferred value of  $k$  varies for larger values of  $n$ . We have used the following settings for  $n$ : 10, 25, 50, 75, 100, 250, 500,



**Figure 2:** Preferred values of  $k$  in the asymmetric and symmetric cases ( $y$ -axis), with target probability 0.95, for different values of  $n$  ( $x$ -axis). The graphic also reports the average node degree in the symmetric case, and the lower and upper bounds on  $k$  derived from Th. 2.

750, 1000. For every value of  $n$ , we have calculated the preferred value of  $k$  in the asymmetric and symmetric cases (with target probability 0.95), proceeding as in the previous experiment. The results of this experiment are shown in Figure 3, along with the average node degree in the symmetric case. Again, the difference between the preferred value of  $k$  in the asymmetric and symmetric cases is at most 1, and the two values are the same for many settings of  $n$ . Concerning the average node degree in the symmetric case, the logarithmic scaling with  $n$  is confirmed.

Interestingly, setting  $k = 9$  produces a symmetric graph which is connected with probability at least 0.95 for values of  $n$  in the range 50–500. In [17], it is shown that when all the nodes have the same transmitting range, a number of neighbors in the range 3–9 is optimal from the network capacity point of view, and it is also close to the optimal value for power efficiency. In this respect, our result can be seen as an improvement of [17], since we achieve connectivity with adaptive transmitting ranges.

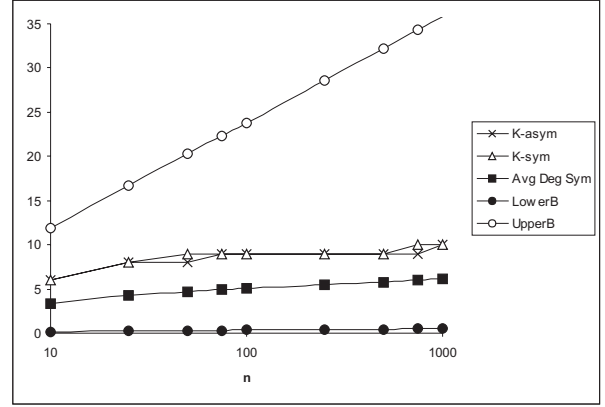
A final investigation concerned the number of asymmetric neighbors when  $k = k_{sym}$ , i.e., in the minimal scenario for achieving connectivity in  $G_k^-$ . We recall that asymmetric neighbors (and the corresponding asymmetric links) will be removed by the  $k$ -NEIGH protocol. From our experiment, whose results are not reported for lack of space, we observed that the average number of asymmetric links removed per node is slightly above 1.2, independently of  $n$ .

Overall, the results of this first set of simulations have shown that the requirement for symmetry has little influence on the preferred value of  $k$ , and that setting  $k = 9$  in the  $k$ -NEIGH protocol provides connectivity w.h.p. for a wide range of network sizes (from 50 to 500 nodes).

## 6.2 Errors in distance estimation

In this Section, we investigate how the preferred value of  $k$  is influenced by errors in distance estimation. To this purpose, we have implemented two models, which account for errors in RSSI- and ToA-based techniques.

In case of RSSI, error is due to the fact that the propa-



**Figure 3:** Preferred values of  $k$  in the asymmetric and symmetric cases ( $y$ -axis), with target probability 0.95, for different values of  $n$  ( $x$ -axis). The graphic also reports the average node degree in the symmetric case, and the lower and upper bounds on  $k$  derived from Th. 2. Values on the  $x$ -axis are reported in logarithmic scale.

gation of the radio signal in the air is influenced by many factors (weather changes, obstacles, and so on), and, consequently, an accurate model of the signal attenuation with distance is very difficult to obtain. Thus, the transformation of the difference between the transmitted and received power into a distance estimation induces a considerable error, which can be unacceptable in many situations. In [28], it is shown that the accuracy of RSSI-based distance estimation is reasonable only in quite idealized settings, such as all the nodes placed in a flat open environment.

We have modeled the error in RSSI-based distance estimation using the scheme proposed in [29], which is defined as follows:

$$RSSI(\delta) = \delta(1 - 10^{\frac{X_\sigma}{10\alpha}}),$$

where  $\delta$  is the actual distance,  $\alpha$  is the distance-power gradient, and  $X_\sigma$  is a random variable with normal distribution of parameters  $(0, \sigma)$ . According to the measurements reported in [28], in our simulations we set  $\sigma = 0.84$  and  $\alpha = 2$ . With these settings, 70% of the estimations are within 10% of the actual distance  $\delta$ .

To model errors in ToA-based distance estimation, we have simplified the scheme of [29], which is based on the acoustic ranging technique of [9]. In this case, the error can be seen as the sum of three independent components:

- *speed of sound error*: changes in the atmospheric conditions can generate both a positive and a negative error in the distance reading. We denote this error with *SSE*.
- *Non-Line-Of-Sight error*: this error, which is always positive, occurs when obstacles obstruct the line of sight between nodes. We denote this error with *NLOS*.
- *orientation error*: this error, which is always positive, occurs when the emitter and the receiver of the acoustic signal have different orientations. We denote this error with *OE*.

In our simulations, we have used the following settings for *SSE*, *NLOS* and *OE*, which are based on the measurements reported in [9]:

- *SSE* is modeled as a uniform error centered at  $\delta$ . More

precisely,  $SSE(\delta) = U[-0.005\delta, 0.005\delta]$ , where  $U[-0.005\delta, 0.005\delta]$  is a random variable with uniform distribution in the interval  $[-0.005\delta, 0.005\delta]$ .

– the experiments reported in [9] have shown that, while “light” obstacles (e.g., a stack of small cardboard boxes) have little influence on the accuracy of distance estimation, “heavy” obstacles (e.g., a large mattress) cause a relevant error. In our model, we have considered three types of obstructions: no obstruction, light obstruction, and heavy obstruction. In case of no obstruction,  $NLOS(\delta) = 0$ ; with light obstacles, we have  $NLOS(\delta) = U[0.006\delta, 0.01\delta]$ , and with heavy obstacles we set  $NLOS(\delta) = U[0.18\delta, 0.22\delta]$ . For every pair of nodes within each other maximum transmitting range, we perform an independent random experiment, and choose “no obstruction” with probability  $p_1$ , “light obstruction” with probability  $p_2$ , and “heavy obstruction” with probability  $1 - (p_1 + p_2)$ . In our experiments, we have set  $p_1$  and  $p_2$  to the values 0.5 and 0.25 respectively, which describe an open air environment with relatively few heavy obstacles. Admittedly, modeling actual NLOS errors (that *are not* independent) is a complicated task, and more investigations are needed on this subject.

– for every pair of nodes within maximum transmitting range, we perform an independent random experiment with four possible equiprobable outcomes, namely 0, 90, 180, 270. These values correspond to an orientation error of 0 degrees, 90 degrees, and so on. We set  $OE(\delta) = 0$  when the outcome is 0,  $OE(\delta) = U[0.004\delta, 0.006\delta]$  when the outcome is 90 or 270, and  $OE(\delta) = U[0.014\delta, 0.016\delta]$  when the outcome is 180. As in the case of NLOS error, our independence assumption introduces a slight approximation, but simplifies the model considerably.

In summary, the ToA-based distance estimation error is defined as follows:

$$ToA(\delta) = SSE(\delta) + NLOS(\delta) + OE(\delta) .$$

We have incorporated the two distance estimation error models in the simulator, and performed a set of experiments to evaluate the impact of errors on the preferred value of  $k$ . To account for possible errors in distance estimation, the simulator has been modified as follows. For every node, we store two neighbor lists: the list  $L$  with the actual distances, and the list  $L_e$  with the estimated distances. Both lists are ordered for increasing values of distance. The estimated distances are generated during a preprocessing phase in which, for every pair of nodes within maximum transmitting range  $R_n$ , we calculate the estimated distance according to the chosen error model. We assume that errors in distance estimation are *symmetric*: if node  $i$  estimates that node  $j$  is at distance  $\delta_{ij}^e$ , also node  $j$  performs the same estimation  $\delta_{ji}^e$ . Since in the  $k$ -NEIGH protocol nodes estimate distances to their neighbors in a very narrow time interval, this assumption is coherent with our error models.

Based on the list  $L_e$ , node  $i$  sets its transmitting power to the value needed to reach the  $k$ -th node in the list, say node  $j$ , which is at estimated distance  $\delta_{ij}^e$ . Since  $\delta_{ij}^e$  is only an estimate of the actual distance, there could exist one or more nodes  $h$  such that  $h$  precedes  $j$  in the list  $L_e$ , but  $\delta_{ih} > \delta_{ij}^e$ . Similarly, there could exist some node  $v$  that follows  $j$  in  $L_e$ , but such that  $\delta_{iv} \leq \delta_{ij}^e$ . In words, the  $k$ -th node in  $L_e$  may not be the actual  $k$ -th nearest neighbor of  $i$ . For this reason, once we have set the transmitting range to  $\delta_{ij}^e$ , for every node  $h$  that precedes  $j$  in  $L_e$  we check (using the list  $L$  of the

$n$	ToA	RSSI	$n$	ToA	RSSI
10	6	7	10	6	7
20	8	9	25	8	9
30	8	9	50	9	9
40	8	9	75	9	9
50	9	9	100	9	10
60	9	9	250	9	10
70	9	9	500	9	10
80	9	9	750	10	10
90	9	10	1000	10	11
100	9	10			

**Table 2: Preferred values of  $k$  (with target probability 0.95) with ToA and RSSI error. The values of  $k$  with ToA error always coincide with those of the exact case.**

actual distances) whether the link to  $h$  actually exists. Note that, since the  $k$ -NEIGH protocol will only consider the first  $k$  nodes in  $L_e$ , possible links to nodes that follows  $j$  in  $L_e$  are not included in the generated graph. Once we have generated the (asymmetric) communication graph according to the procedure described above, we consider only symmetric links and check for connectivity, as in the previous set of experiments.

As in Section 6.1, we have simulated networks of sizes in the range 10–100 in steps of 10, and in the range 10–1000, and evaluated the preferred value of  $k$  (with target probability 0.95) in case of RSSI and ToA errors. The results of our simulations are reported in Table 2. As expected, ToA distance estimation performs much better than the simpler RSSI technique: for all the values of  $n$  considered, the preferred value of  $k$  with ToA error was always the same as in the exact case (see Figures 2 and 3). With RSSI error, the preferred value of  $k$  is at most 1 greater than the value in the exact scenario, and it is the same value for many settings of  $n$ . The relatively little influence of error in distance estimation on the minimum value of  $k$  for connectivity is further evidenced in Figure 4, which shows the empirical distribution of  $k$  in the exact, ToA error, and RSSI error cases, for  $n = 100$ . The better performance of ToA with respect to RSSI distance estimation is due to the fact that ToA error, although occasionally large when heavy obstacles obstruct the line of sight, is essentially on the positive side. Thus, the situation described above in which a “close” neighbor cannot be actually reached is less likely to occur.

Overall, the results of this second set of experiments have shown that the  $k$ -NEIGH protocol is resilient to errors in distance estimation, also in the scenario in which obstacles obstruct the line of sight of a considerable fraction of node pairs.

### 6.3 Energy cost

In the last set of experiments, we have compared the performance of  $k$ -NEIGH with that of other topology control algorithms. The performance is measured in terms of energy cost, which, we recall, is defined as  $c(RA) = \sum_{i \in N} (RA(i))^\alpha$ , where  $RA$  is the range assignment as defined at the end of the protocol execution. The energy cost gives a measure of the “energy efficiency” of the topology generated by a topology control algorithm.

Another important parameter used in the literature to evaluate the protocols is the average node degree. We recall that, besides reducing energy consumption, topology con-



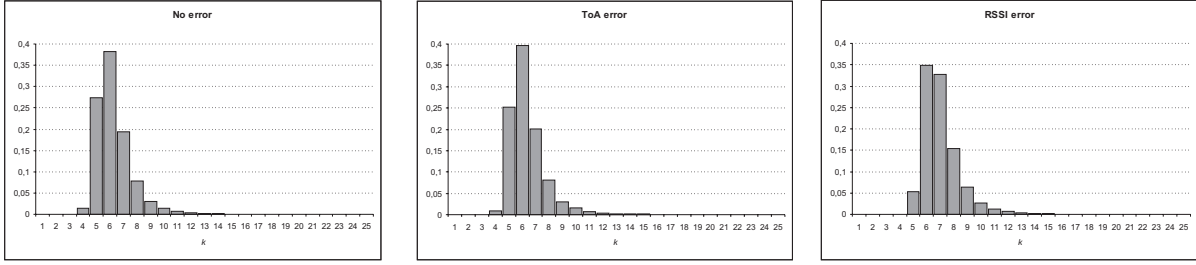


Figure 4: Empirical distribution of the minimum  $k$  for connectivity in the exact (left), ToA error (center), and RSSI error (right) case for  $n=100$ . Data are shown as frequencies.

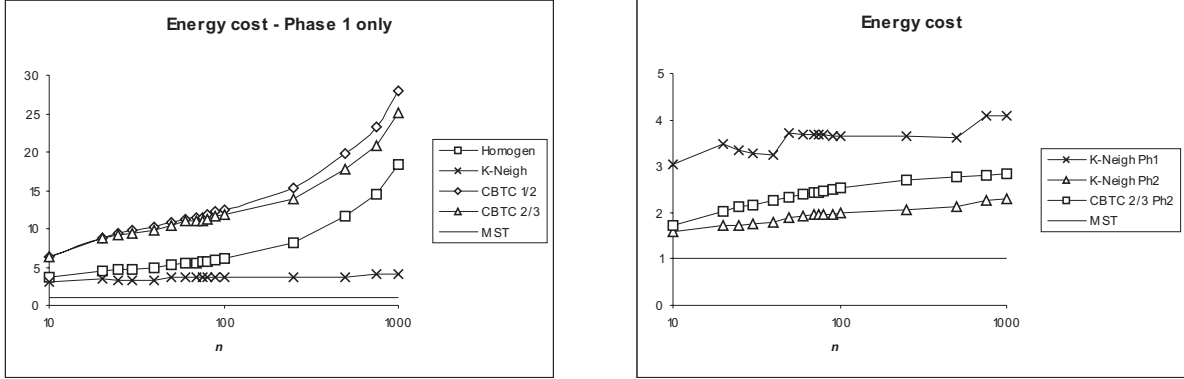


Figure 5: Energy cost of different topology control protocols. For  $k$ -NEIGH and CBTC, we have considered Phase 1 only (left), and Phases 1 and 2 implemented (right). The energy cost is normalized with respect to the cost of the MST. Values on the  $x$ -axis are reported in logarithmic scale.

control mechanisms have the positive effect of increasing spatial reuse, which means that fewer nodes are expected to experience contention in accessing the wireless link. Hence, a reduced average node degree in general implies that contention is reduced as well. However, it is important to note that what really matters is the *physical*, rather than the *logical*, degree. In fact, many protocols (such as  $k$ -NEIGH and CBTC, for instance) generate a logical topology, in which some of the actual links are not considered, because they are either asymmetric or energy inefficient. Thus, the number of logical neighbors, which determines the logical node degree, could be significantly smaller than the actual number of neighbors, which “measures” the likelihood of contention. Given the same physical degree, a higher logical degree has a positive effect on network capacity, since fewer bottlenecks are likely to occur in the topology. This point has often been disregarded in the previous analyses of topology control protocols, and will be carefully investigated in our simulations.

In our simulations, we have considered values of  $n$  ranging from 10 to 1000, as in Sections 6.1 and 6.2. For each value of  $n$ , we have generated 10000 random node placements, and executed the following topology control algorithms:

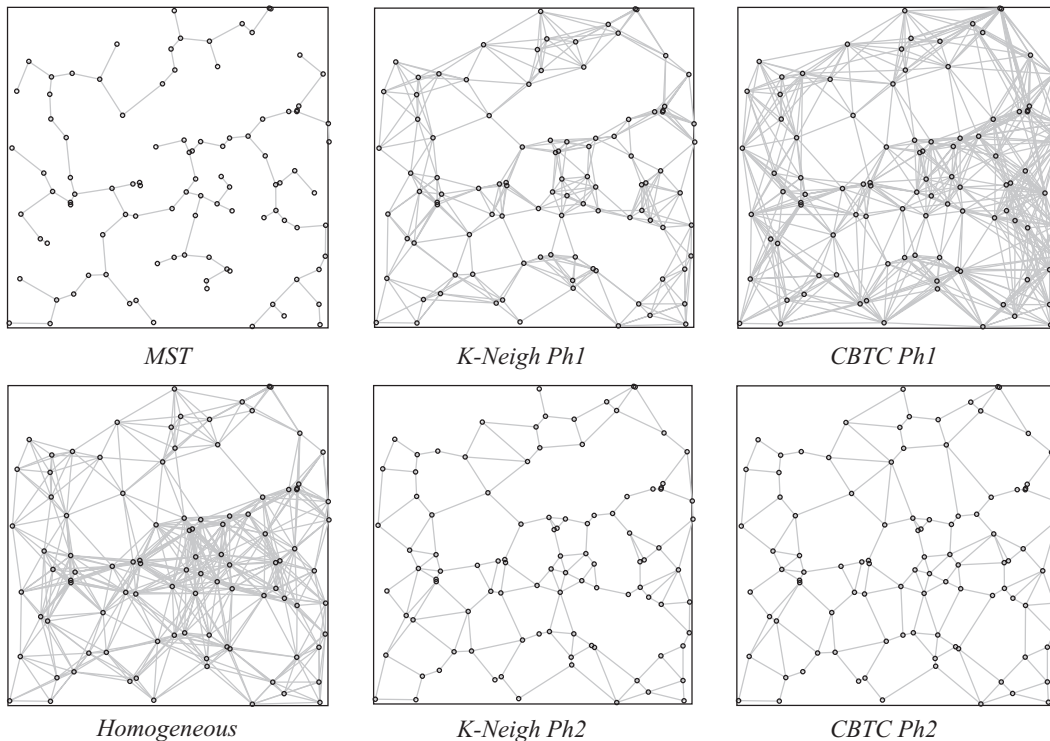
- *MST*: although impractical (its computation requires global knowledge), the Euclidean Minimum Spanning Tree produces a range assignment that is within a factor of 2 from the optimal weakly symmetric range assignment (see [3]). We have used the MST as the “optimal” topology against which the topologies generated by the other protocols will

be compared.

- *k*-NEIGH: for each setting of  $n$ , the value of  $k$  used in the protocol is the preferred value as evaluated in Section 6.1.
- *CBTC*: we have simulated CBTC using two values for  $\rho$  (the maximum angular gap required):  $\rho = \frac{2}{3}\pi$  and  $\rho = \frac{\pi}{2}$ .
- *Homogeneous*: we have also considered the situation in which no topology control is used. In this case, the value of the transmitting range is defined as the 0.95 quantile of the empirical distribution of the critical transmitting range (see Section 6.1).

First, we have evaluated the energy cost of the different protocols. For the  $k$ -NEIGH and CBTC protocols, we have considered both the result of the Phase 1 only (without pruning), and of the protocols with the pruning stage implemented. The rationale for this investigation is that in some situations (e.g., high mobility scenario), implementing the pruning step could be very difficult. We have considered two values for the distance-power gradient  $\alpha$ , i.e.,  $\alpha = 2$  and  $\alpha = 4$ . The value of the distance-power gradient has a strong influence on the pruning phase of  $k$ -NEIGH and CBTC, which are essentially based on triangular inequalities on the power function: the higher  $\alpha$ , the more edges are pruned.

In Figure 5, we show the energy cost (normalized with respect to the cost of the MST) of the different protocols when  $\alpha = 2$ , for increasing values of  $n$ . As can be seen, the Phase 1 of our  $k$ -NEIGH protocol performs much better than that of CBTC, for both values of  $\rho$ : for  $n = 1000$ , the energy cost of  $k$ -NEIGH is 83% lower than Phase 1 of



**Figure 6:** Sample topologies produced by the different topology control protocols with  $n=100$ . In CBTC,  $\rho$  is set to  $\frac{2}{3}\pi$ .

CBTC with  $\rho = \frac{2}{3}\pi$ . Compared to the case of no topology control,  $k$ -NEIGH-Phase 1 provides an improvement of 16% when  $n = 10$ , and of 77% when  $n = 1000$ . Observe that CBTC performs worse than the case of no topology control: this is due to the fact that the maximum transmitting range  $R_n$  used in CBTC is 50% larger than the 0.95 quantile of the critical transmitting range distribution used in the Homogeneous protocol. In case a lower value of  $R_n$  would be used, CBTC would perform better, at the expense of an increased probability of generating a disconnected graph. Contrary to CBTC, our  $k$ -NEIGH protocol is almost independent of the choice of  $R_n$ : using the same value of  $R_n$  of CBTC,  $k$ -NEIGH is several times better than Homogeneous. This is due to the fact that in CBTC, several nodes (those lying on the boundary of the region) are expected to transmit at maximum power, since it is very unlikely that the required angular gap  $\rho$  can actually be achieved. Conversely, in  $k$ -NEIGH we require the connection to the  $k$  nearest nodes, independently of their direction.

The implementation of the pruning stage decreases the energy cost significantly in both  $k$ -NEIGH and CBTC protocols (see Figure 5 – right). Nevertheless,  $k$ -NEIGH still performs better than CBTC: except for small network sizes ( $n=10$  to 20), the energy cost of  $k$ -NEIGH is about 20% smaller than that of CBTC. The experiments show that the topologies generated by  $k$ -NEIGH can be as much as 87% more energy efficient than in those with no topology control, while they are at most a factor of 2.28 away from the cost of the “optimal” MST topology. A sample of the topologies generated by the various protocols for  $n=100$  is shown in Figure 6.

In Figure 7 we report the average logical (left) and phys-

ical (right) node degree of the topologies generated using  $k$ -NEIGH and CBTC. As in the previous case, we have considered both protocols without and with the pruning stage implemented. From Figure 7, it is evident that  $k$ -NEIGH-Phase 1 outperforms CBTC-Phase 1 in terms of both logical and physical degree. Observe that in  $k$ -NEIGH we have the upper bound  $k$  on the number of *physical* neighbors of any node, which holds for Phase 1 also. On the contrary, the result of [30] on the maximum number of neighbors (which, we recall, is 6) regards the topology generated by CBTC *after* pruning; furthermore, the upper bound is on the number of *logical* neighbors. Finally, note that  $k$ -NEIGH performs better than CBTC also when Phase 2 is implemented.

We have performed the same simulations with  $\alpha=4$ . The results of these experiments, which are not reported for lack of space, confirmed on a larger scale that  $k$ -NEIGH performs better than CBTC in terms of energy cost, logical and physical average node degree. In terms of energy cost,  $k$ -NEIGH-Phase 1 performs as much as 97% better than CBTC-Phase 1, and as much as 94% better than the case of no topology control. With Phase 2 implemented,  $k$ -NEIGH is as much as 29% better than CBTC, and as much as 98% better than the case of no topology control.

Overall, the results of this last set of experiments have shown that:

- $k$ -NEIGH-Phase 1 performs significantly better than CBTC-Phase 1. Essentially, this is due to the fact that, contrary to the case of CBTC, after the execution of  $k$ -NEIGH-Phase 1 relatively few nodes are expected to transmit at maximum power. To some extent, this seems to indicate that  $k$ -NEIGH is well suited to be implemented in a high mobility scenario

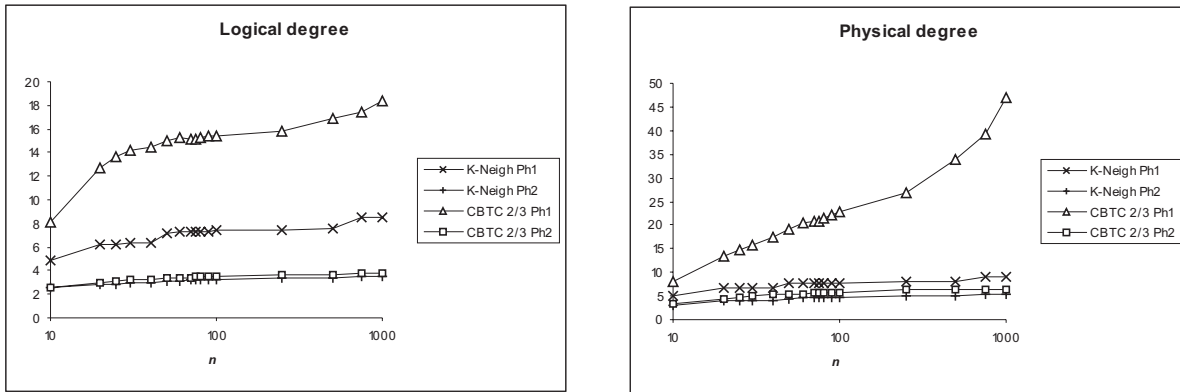


Figure 7: Average logical (left) and physical (right) degree of the topologies generated by the  $k$ -NEIGH and CBTC protocols. Values on the  $x$ -axis are reported in logarithmic scale.

(see Section 7 for further discussion on this point).

– If pruning is implemented,  $k$ -NEIGH still performs better than CBTC, in terms of energy cost, as well as logical and physical average node degree.

## 7. DISCUSSION AND FUTURE WORK

In this paper we have presented an approach to topology control, based on the computation of a symmetric subgraph of the  $k$  nearest neighbors graph, and a fully distributed and localized protocol (called  $k$ -NEIGH) that implements this approach. We have seen that  $k$ -NEIGH does not require the knowledge of the exact number  $n$  of nodes in the network to work, as  $k$  is only loosely dependent on  $n$  (e.g.,  $k = 9$  for  $n$  in the range 50-500). Also, the maximum transmitting range of nodes can be overestimated without problems, since our protocol (contrary, for instance, of CBTC) is not influenced by the choice of a specific maximum transmitting range. Whenever distance estimation is a viable choice, our protocol can then be implemented in practice.

A first possible direction for future work is studying the protocol behavior in a dynamic setting, i.e., when nodes join the network at unpredictable times. From a preliminary analysis, the number of messages exchanged to update the  $G_k^-$  graph in response to a newcomer’s “hello” message seems approximately equal to  $k'$ , where  $k'$  is the number of nodes that are within the maximum transmitting range of the newcomer, which is in a sense the minimum achievable for a single update. For the whole sequence this leads to an  $O(n^2)$  messages worst case bound, which applies when all the nodes are close together. However, we suspect that the actual average bound is much smaller.

We also plan to investigate the performance of our protocol in presence of multi-hop data traffic, and using a more sophisticated model for the radio signal propagation, such as that recently proposed in [6].

Another important topic for future work is to adapt the  $k$ -NEIGH protocol to deal with mobility. In a mobile network, the topology is continuously changing and the topology control protocol must be reexecuted periodically. A quantitative evaluation of  $k$ -NEIGH in mobile environments is beyond the scope of this paper. However, here we present a brief qualitative discussion of how the protocol can be adapted for mobile environments and how it compares to other algo-

gorithms in this case.

In the  $k$ -NEIGH protocol presented herein, the number of neighbors is set to a very precise value. If this protocol is extended to mobile networks, it would be quite expensive to control the neighbor set size so precisely: this could require re-execution of the protocol each time the neighbor set changed. Instead, we adopt the approach taken in MobileGrid and LINT, where low and high water marks are specified such that the neighbor set size falling below the low water mark or exceeding the high water mark causes the protocol to be reexecuted. Since the value of  $k$  determined in Theorem 2 is sufficient for connectivity, it is a likely candidate for the low water mark for the mobile version of the protocol. The high water mark could be determined based on the velocity of nodes and the expected transmitting range to ensure that the protocol does not need to be reexecuted too often. The initial value of the desired number of neighbors in the protocol should then be set to the average of the low and high water marks.

Since any topology control protocol needs to be executed periodically in a mobile network, the energy consumed during the protocol execution becomes even more important than the “quality” of the topology produced. Thus, we believe that the benefits of the  $k$ -NEIGH protocol will be even greater in this situation. This is because ours is the only known protocol with a proven upper bound on the number of messages exchanged during its execution. The number of messages exchanged in the  $k$ -NEIGH protocol is expected to be far lower than CBTC in practice since CBTC iteratively sends messages in a first phase and then sends even more messages during a second optimization phase. Our future work will focus on specification and evaluation of a mobile  $k$ -NEIGH protocol, with the goal of showing that the significant benefits shown herein for the stationary version of the protocol are maintained in mobile environments.

## 8. REFERENCES

- [1] M. Bahramgiri, M. Hajiaghayi, V.S. Mirrokni, “Fault-tolerant ad 3-Dimensional Distributed Topology Control Algorithms in Wireless Multi-hop Networks”, *Proc. IEEE Int. Conference on Computer Communications and Networks*, pp. 392–397, 2002.
- [2] L. Bao, J.J.Garcia-Luna-Aceves, “Channel Access

- Scheduling in Ad Hoc Networks with Unidirectional Links”, *Proc. DIALM 01*, pp. 9–18, 2001.
- [3] D.M. Blough, M. Leoncini, G. Resta, P. Santi, “On the Symmetric Range Assignment Problem in Wireless Ad Hoc Networks”, *Proc. IFIP Conf. on Theoretical Computer Science*, pp. 71–82, 2002.
- [4] D.M. Blough, M. Leoncini, G. Resta, P. Santi, “The k-Neighbors Protocol for Symmetric Topology Control in Ad Hoc Networks”, *Tech. Rep. IIT-TR-23/2002*, Istituto di Informatica e Telematica, Pisa – Italy, Dec. 2002.
- [5] S.A. Borbash, E.H. Jennings, “Distributed Topology Control Algorithm for Multihop Wireless Networks”, *Proc. IEEE Int. Joint Conference on Neural Networks*, pp. 355–360, 2002.
- [6] J. Bruck, M. Franceschetti, L. Schulman, “Microcellular Systems, Random Walks, and Wave Propagation”, *Proc. IEEE Symposium on Antennas and Propagation*, pp. 220–223, 2002.
- [7] J. Chang, L. Tassiulas, “Energy Conserving Routing in Wireless Ad Hoc Networks”, *Proc. IEEE INFOCOM 2000*, pp. 22–31, 2000.
- [8] A.E.F. Clementi, P. Penna, R. Silvestri, “Hardness Results for the Power Range Assignment Problem in Packet Radio Networks”, *Proc. 2nd International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (RANDOM/APPROX ’99)*, LNCS (1671), pp. 197–208, 1999.
- [9] L. Girod, D. Estrin, “Robust Range Estimation Using Acoustic and Multimodal Sensing”, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2001.
- [10] P. Gupta, P.R. Kumar, “Critical Power for Asymptotic Connectivity in Wireless Networks”, *Stochastic Analysis, Control, Optimization and Applications*, Birkhauser, Boston, pp. 547–566, 1998.
- [11] Z. Huang, C. Shen, C. Srisathapornphat, C. Jaikaeo, “Topology Control for Ad Hoc Networks with Directional Antennas”, *Proc. IEEE Int. Conference on Computer Communications and Networks*, pp. 16–21, 2002.
- [12] E. Jung, N.H. Vaidya, “A Power Control MAC Protocol for Ad Hoc Networks”, *Proc. ACM Mobicom 02*, pp. 36–47, 2002.
- [13] V. Kawadia, P.R. Kumar, “Power Control and Clustering in Ad Hoc Networks”, *to appear in Proc. IEEE Infocom 03*, 2003.
- [14] L.M. Kirousis, E. Kranakis, D. Krizanc, A. Pelc, “Power Consumption in Packet Radio Networks”, *Theoretical Computer Science*, Vol. 243, pp. 289–305, 2000.
- [15] R. Krashinsky, H. Balakrishnan, “Minimizing Energy for Wireless Web Access with Bounded Slowdown”, *Proc. ACM Mobicom 02*, pp. 119–130, 2002.
- [16] L. Li, J.H. Halpern, P. Bahl, Y. Wang, R. Wattenhofer, “Analysis of a Cone-Based Distributed Topology Control Algorithm for Wireless Multi-hop Networks”, *Proc. ACM PODC 2001*, pp. 264–273, 2001.
- [17] J. Liu, B. Li, “MobileGrid: Capacity-aware Topology Control in Mobile Ad Hoc Networks”, *Proc. IEEE Int. Conference on Computer Communications and Networks*, pp. 570–574, 2002.
- [18] M.K. Marina, S.R. Das, “Routing Performance in the Presence of Unidirectional Links in Multihop Wireless Networks”, *Proc. ACM MobiHoc 02*, pp. 12–23, 2002.
- [19] S. Narayanaswamy, V. Kawadia, R.S. Sreenivas, P.R. Kumar, “Power Control in Ad Hoc Networks: Theory, Architecture, Algorithm and Implementation of the COMPOW Protocol”, *Proc. European Wireless 2002*, pp. 156–162, 2002.
- [20] K. Pahlavan, A. Levesque, *Wireless Information Networks*, John Wiley and Sons, New York, 1995.
- [21] M.R. Pearlman, Z.J. Haas, B.P. Manvell, “Using Multi-Hop Acknowledgements to Discover and Reliably Communicate over Unidirectional Links in Ad Hoc Networks”, *Proc. Wireless Communications and Networking Conference (WCNC)*, pp. 532–537, 2000.
- [22] R. Prakash, “A Routing Algorithm for Wireless Ad Hoc Networks with Unidirectional Links”, *ACM/Kluwer Wireless Networks*, Vol.7, n. 6, pp. 617–625, 2001.
- [23] R. Ramanathan, R. Rosales-Hain, “Topology Control of Multihop Wireless Networks using Transmit Power Adjustment”, *Proc. IEEE Infocom 2000*, pp. 404–413, 2000.
- [24] V. Ramasubramanian, R. Chandra, D. Mosse, “Providing a Bidirectional Abstraction for Unidirectional Ad Hoc Networks”, *Proc. IEEE Infocom 02*, pp. 1258–1267, 2002.
- [25] V. Rodoplu, T.H. Meng, “Minimum Energy Mobile Wireless Networks”, *IEEE Journal Selected Areas in Comm.*, Vol. 17, n. 8, pp. 1333–1344, 1999.
- [26] M. Sanchez, P. Manzoni, Z.J. Haas, “Determination of Critical Transmitting Range in Ad Hoc Networks”, *Proc. Multiaccess, Mobility and Teletraffic for Wireless Communications Conference*, 1999.
- [27] P. Santi, D.M. Blough, “The Critical Transmitting Range for Connectivity in Sparse Wireless Ad Hoc Networks”, *IEEE Transactions on Mobile Computing*, Vol. 2, n. 1, pp. 1–15, January-March 2003.
- [28] A. Savvides, C. Han, M. Srivastava, “Dynamic Fine-Grained Localization in Ad Hoc Networks of Sensors”, *Proc. ACM Mobicom 01*, pp. 166–179, 2001.
- [29] S. Slijepcevic, S. Megerian, M. Potkonjak, “Location Errors in Wireless Embedded Sensor Networks: Sources, Models, and Effect on Applications”, *Mobile Computing and Communications Review*, Vol. 6, n. 3, pp. 67–78, 2002.
- [30] R. Wattenhofer, L. Li, P. Bahl, Y. Wang, “Distributed Topology Control for Power Efficient Operation in Multihop Wireless Ad Hoc Networks”, *Proc. IEEE Infocom 2001*, pp. 1388–1397, 2001.
- [31] F. Xue, P.R. Kumar, “The Number of Neighbors Needed for Connectivity of Wireless Networks”, *internet draft*, available at [http://decision.csl.uiuc.edu/~prkumar/postscript\\_files.html#Wireless%20Networks](http://decision.csl.uiuc.edu/~prkumar/postscript_files.html#Wireless%20Networks).