EPJ B
Condensed Matter and Complex Systems
EPJ.org
your physics journal

Eur. Phys. J. B **73**, 597–604 (2010) DOI: 10.1140/epjb/e2010-00049-x

Heterogeneous attachment strategies optimize the topology of dynamic wireless networks

P. Holme, B.J. Kim and V. Fodor

EDP SCIENCES
Società Italiana di Fisica
Springer

Eur. Phys. J. B **73**, 597–604 (2010)
DOI: 10.1140/epjb/e2010-00049-x

THE EUROPEAN PHYSICAL JOURNAL B

Regular Article

Heterogeneous attachment strategies optimize the topology of dynamic wireless networks

P. Holme[1,2,3,a], B.J. Kim[3,4], and V. Fodor[5]

[1] Department of Physics, Umeå University, 901 87 Umeå, Sweden
[2] Department of Energy Science, Sungkyunkwan University, 440–746 Suwon, Korea
[3] School of Computer Science and Communication, Royal Institute of Technology, 100 44 Stockholm, Sweden
[4] Department of Physics, BK21 Physics Research Division and Institute of Basic Science, Sungkyunkwan University, 440–746 Suwon, Korea
[5] Access Linneaus Centre, Royal Institute of Technology, 100 44 Stockholm, Sweden

Received 6 November 2009 / Received in final form 14 January 2010
Published online 9 February 2010 –

Abstract. In optimizing the topology of wireless networks built of a dynamic set of spatially embedded agents, there are many trade-offs to be dealt with. The network should preferably be as small (in the sense that the average, or maximal, pathlength is short) as possible, it should be robust to failures, not consume too much power, and so on. In this paper, we investigate simple models of how agents can choose their neighbors in such an environment. In our model of attachment, we can tune from one situation where agents prefer to attach to others in closest proximity, to a situation where agents attach to random others regardless of distance (which thus are, on average, further away than the connections to the spatial neighbors). We evaluate this scenario with several performance measures and find that the optimal topologies, for most of the quantities, is obtained for strategies resulting in a mix of most local and a few random connections.

1 Introduction

The performance of a network is a consequence of three factors – the hardware, the network protocols and the network topology. For some networked systems, like the wireless mesh networks, the topology can be controlled rather easily through the medium access layer, and it is thus practically possible to optimize th topology of the network. In this work, we investigate a scenario of static agents (sensors, individuals with Wi-Fi devices, etc.), embedded in space, joining and leaving the system (so called "churn"). Our scenario is based on the following three assumptions. First, only new agents and neighbors of an agent leaving the network create new edges. Second, the power consumption of each agent is, for fairness and practical reasons, to be restricted. Third, the agents are localized with uniform randomness over the unit square. From these starting points, we evaluate different simple attachment strategies to optimize several network-structural quantities, capturing efficiency of communication, scalability and robustness to failure.

Many of the previous studies on topological optimization of wireless networks have concerned static point patterns [1], not (as in our study) systems under churn. Nevertheless, these studies form a theoretical backdrop of ours. The simplest approach to connecting wireless agents in the plane is to let all agents within range be connected. Assuming identical agents, one arrives at the Unit Disc Graph (UDG) model of wireless networks [2]. Maintaining and routing information in a network with this architecture, however, is a waste of power and memory resources. Many theoretical efforts in this area have been focusing on energy efficient, yet scalable topologies – a seminal work is [3], where the "Yao graph" was proposed. This graph is cowinstructed by connecting each vertex to its closest neighbor in k equiangular sectors. It has several topological features, such as (almost surely) full connectivity and scalable power consumption, which are desirable for wireless networks. Subsequent studies have proposed improved proximity graphs, with localized construction algorithms more [2,4,5] or less [6–8] similar to the Yao graph, and studied general features of minimum power topologies [9,10]. Our approach is from a slightly different angle, taking more inspiration from the recent complex network literature [11–13]. These works discuss how dynamic properties of networks in general (including communication performance in technological networks) are related to network structure (how the observed network topology differs from what is expected in a unconstrained random network model). This approach lends itself naturally to analyzing topics like point-to-point communication, and data distribution scenarios in large, to some extent random, systems. On a yet more theoretical level, the "Waxman graph" model [14] has been investigated much. The Waxman graph, like the UDG and Yao graphs is typically defined on an unit square with randomly distributed

[a] e-mail: petter.holme@physics.umu.se

vertices. The vertices are connected with a probability decreasing exponentially with the distance between the vertices. While the UDG and Yao graphs are, given a set of points on the unit square, deterministic; the Waxman graph is a stochastic model. Still the Waxman graph is analytically tractable and much research has focused on finding the region in parameter space where the graph is almost surely connected. In this work we evaluate the possibility to optimize network topology in the case when connections cost power, and consequently, due to limitations of energy storage, agents maintaining long connections can keep only a limited number of neighbors.

In the next section, we give a precise description of the problem and our proposed attachment models and discuss related practical issues. We evaluate the proposed models considering a large set of performance measures in Section 3. Finally, Section 4 concludes our work.

2 Preliminaries

2.1 Problem statement

We consider a set V of N agents spread out on the unit square. In our numerical study, the spatial distribution will be uniformly random, but generalizing to other point patterns is straightforward. At every time step, one random agent leaves the system and a new one enters (at a random coordinate). The agents form a graph with bidirectional connections. For two agents i and j at coordinates $\mathbf{r}_i$ and $\mathbf{r}_j$ to be connected it costs both agents a power consumption of $p(i,j) = |\mathbf{r}_i - \mathbf{r}_j|^\delta$ $(2 \leq \delta \leq 4)$ [6]. In this work we focus on the limit $\delta = 2$. We assume the agents to be selfish and homogeneous in the sense that no agent i accept a higher power-consumption $\sum_{j\in\Gamma_i} p(i,j)$ (where Γ_i is i's neighborhood), than $P_{\max}$. This also ensures scalable, total power consumption. We also assume a lower power limit $P_{\min}$, such that while an agent consumes less power than $P_{\min}$, it will attempt to create new edges in the network (unless creating any new edge would result in a larger power consumption than $P_{\max}$). Thus $P_{\min}$ represents a minimal investment all agents agree to contribute with to the system. Based on these precepts, we try to find simple attachment strategies (rules for agents to select other agents to link to) that optimize various objective functions. To assess unperturbed network performance, we measure the following quantities (that will be discussed in greater detail in the Sect. 3).

- *Connectivity.* We measure connectivity as the fraction of time steps when the graph is completely connected. This quantity should, in any functioning network, be very close to one. Except when we explicitly study this quantity, we tune the parameters so that the graph is almost always connected.
- *Power consumption and power efficiency.* We consider the value of aggregate power consumption and the power efficiency (the ratio of power consumption across the shortest multihop path and a direct edge). Low values are desirable for efficient communication.
- *Diameter.* $d_{\max} = \max_{i,j\in V} d(i,j)$. Where $d(i,j)$ is the *distance* (the number of edges in the shortest path) between i and j.
- *Average distance* $\bar{d} = \frac{2}{N(N-1)} \sum_{i,j\in V} d(i,j)$. The overhead in communication time grows with distance; thus, a small diameter and short average distance are desirable.
- *Spectral gap*, E – the difference between the largest and second largest eigenvector of the adjacency matrix. This is an approximate measure of the link-expander property [15] that embodies several desirable properties for wireless networks (little redundancy, high robustness, efficient broadcast, etc.).

Since we do not construct an optimal network each time step, only update the network incrementally, it is important to monitor the effect of churn. Indeed, this problem is similar to another important aspect – the robustness to failure and we measure it by the two quantities:

- The change, $\Delta d_{\max}$, of the diameter as a random vertex is deleted from the network.
- The change, $\Delta\bar{d}$, of the average distance.

We aim at finding simple, local attachment strategies optimizing these quantities – local in the sense that each agent follows its own rules.

2.2 Model definition

We will investigate two models[1], both having in common that they interpolate between situations where the power budget is spent only on short-range connections, to a situation where connections are made at random, regardless of location. At every time step, a random agent i' is removed along with its connections $\Gamma_{i'}$ and a new agent added. This can cause both the new agent and $\Gamma_{i'}$ to have power consumptions below $P_{\min}$. All agents i with $P(i) < P_{\min}$ try to attach to new agents j. If adding an edge (i,j) would make either $P(i)$ or $P(j)$ larger than $P_{\max}$, or if (i,j) already is an edge in the graph, the edge is not added. These properties are the same for both models. The differences between the models are:

- *Model A.* With probability q_A, attach i to the closest possible agent, otherwise attach i to a random agent.
- *Model B.* There are two kinds of agents – a fraction q_B of them always attach to the closest possible agent, the rest attaches to a random agent.

Note that in the limits $q_A = q_B = 0$ and $q_A = q_B = 1$, the models are identical.

2.3 Practical considerations

The main assumption of the proposed models is that the power consumption of an agent i is proportional to $\sum_{j\in\Gamma_i} p(i,j)$, where Γ_i is the set of agents connected to i

[1] A Java Applet implementation of the model can be found at `http://statphys.skku.ac.kr/~bjkim/Applet/mobile.html`.

and $p(i,j) = |\mathbf{r}_i - \mathbf{r}_j|^\delta$. This requires transmission power control to be implemented at the agents. We motivate the assumption on power consumption as follows. First, once the network topology is defined, and if the traffic load in the network is balanced, both unicast and broadcast routing scenarios are likely to have a similar average usage of the connections. (In a situation with heterogeneous agents of very different capacities and network positions, this would not hold.) Second, considering dynamic power control, the required transmission power depends on the transmission channel characteristics. In systems designed for tolerable outage probability it is sufficient to consider the distance-dependent path loss to set the required transmission power [9]. Finally, we do not consider the energy spent for data reception and processing. These are distance independent additive values and consequently would not change the main conclusions of our study.

For the implementation of the attachment strategies described in Section 2.2, the problems of distance-dependent neighbor selection and efficient access control to the shared wireless medium have to be solved.

If the network, for efficient higher layer protocols (like geographic routing [16]), keeps information about the location of the vertices, then such information can also improve the proposed attachment models. If absolute vertex location information is not maintained, agents can select neighbors by discovering a given area via beaconing with increasing transmission power (see for example Ref. [10] for detailed solutions).

The medium access control is responsible for sharing the wireless medium in an efficient way. TDMA, CSMA and mixed solutions are usually proposed, depending on the expected network load [17]. Since the interference region of a connection depends on the distance of the transmitting and receiving agent, the construction of an optimal access control scheme is not trivial in the addressed scenario. An optimal access control should take the attachment model into account and schedule most local and random connections separately. The efficiency of the medium access control can potentially affect the network performance, which calls for extended future studies.

3 Numerical results

In this section we will discuss results from our numerical simulations. Unless otherwise stated, we use 10^6 time steps for equilibration and 10^6 time steps for measuring averages.

3.1 Average degree

Many of the performance measures, like connectivity and network diameter are directly related to the number of connections the agents maintain – that is, the average degree. Since the number of connections an agent can maintain in the considered case depends on the attachment strategy, we start investigating the average degree.

We can derive approximate results for the two extreme cases, $q_A = q_B = 0$ and $q_A = q_B = 1$. For $q_A = q_B = 0$,

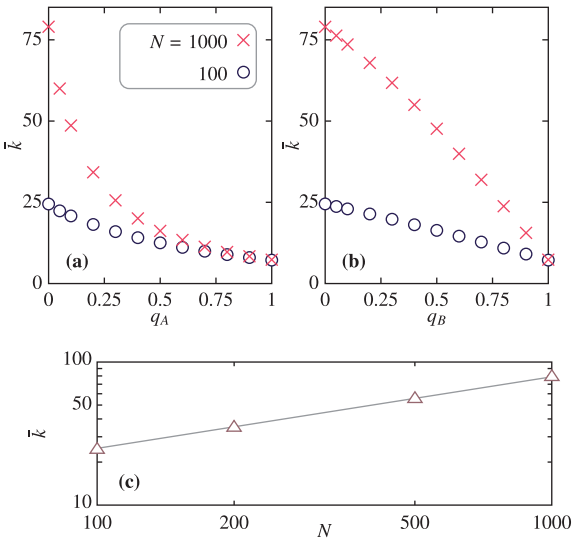

Fig. 1. (Color online) The average degree $\bar{k}$ for our models. (a) shows results for Model A, (b) for Model B. $P_{\max} = 2P_{\min} = 2$, and (c) shows the size scaling for $q_A = q_B = 0$. The line in (c) represents growth proportional to $N^{1/2}$. Standard errors are smaller than the symbol size.

each agent tries to connect to the closest available neighbor. Assuming that all connections can be made and that the agents are uniformly spread out (so that the number of agents n_r within distance r from point is $\pi r^2 N$), this means that the power consumption to attach to agents within a distance R is

$$\int_0^R r^\delta 2\pi r N \, dr = \frac{2\pi N}{2+\delta} R^{2+\delta}. \tag{1}$$

For a given minimal power $P_{\min}$ the number of agents within reach is

$$\pi N \left(\frac{2+\delta}{2\pi N} P_{\min} \right)^{2/(2+\delta)} \in O\left(N^{\delta/(2+\delta)} \right) \tag{2}$$

or in other words, for $\delta = 2$ the average degree $\bar{k} \in O(N^{1/2})$. This relationship is confirmed in Figure 1. In the opposite limit, $q_A = q_B = 1$, all agents attach to others at random (which is a limit in common to Waxman graphs [14]). Since the average distance to another, random, agent is independent of N, then so is $\bar{k}$ (which also can be seen in Fig. 1). The functional forms of $\bar{k}$, as functions of $q_{A,B}$ are strikingly different. Even if $\bar{k}$ itself is not a performance measure, it affects network performance; therefore a stable response to the parameter values is beneficial. We note that the curves for Model B has a smaller slope for small values of q_B, whereas Model A has larger magnitude of the derivative when q_A is close to zero.

The key to understanding the difference in degree between the two model is the observation that the number of long-range edges is much lower in Model B than in Model A (Fig. 1). $P_{\max}$ controls the power consumption so that it is relatively similar for both models. With the power consumption constrained, more power-expensive,

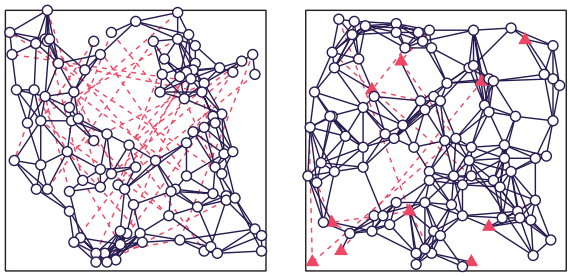

Fig. 2. (Color online) Snapshot of the models with $N = 100$, $P_{\min} = 0.1$, $P_{\max} = 1$, and $q_A = q_B = 0.1$. The dashed lines are formed by random attachment. The triangles in (b) are the agents with a random attachment strategy rather than attachment to the closest other vertex.

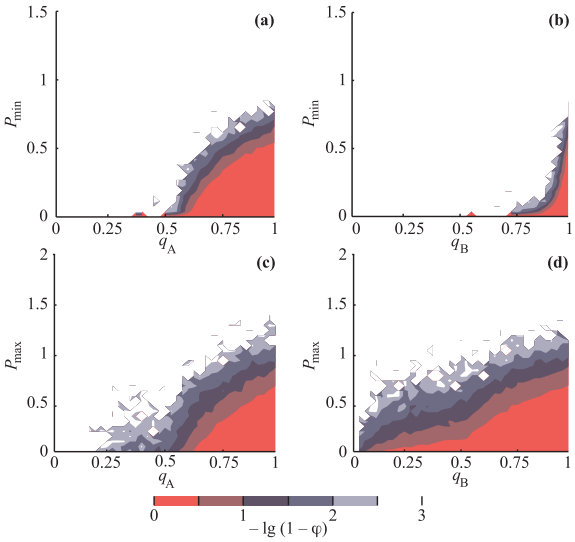

Fig. 3. (Color online) The connectivity is shown as a function of $q_{A,B}$ and $P_{\min}$ (panel (a) and (b)) or $P_{\max}$ for Model A (panel (a) and (c)) and Model B (panel (b) and (d)). The fraction φ of time steps when the network is fully connected is measured and displayed as color scales so that the numbers at the color bars correspond to $-\lg(1-\varphi)$, white means that the network is almost always connected. In all panels we set $2P_{\min} = P_{\max}$ and use $N = 1000$.

random edges leads to a sparser network. So why does Model A have more random edges? New edges are added until the power consumption reaches above $P_{\min}$. Since the steps in power consumption are larger when a random connection is added, the chance that the last added edge, for Model A, is by random attachment is larger than q_A. For model B, on the other hand, the chance the last added connection is long range is q_B, so $\bar{k}$ will be a linear combination of the $q_B = 0$ and $q_B = 1$ limits. Therefore, for $q_A = q_B$ the number of long-range edges and the power consumption will be higher for Model A than Model B. These mechanisms can also be understood from our example figure, Figure 2. To make this figure readable we use smaller values of N and $P_{\min}$ than in the rest of the paper. We also use a relatively large $P_{\max}$, making the $\bar{k}$ more similar between the models at the expense of a more different power consumption. From this figure we see the difference in number of long-range edges even more pronounced than in Figure 1.

From Figures 1a and 1b we learn that the degree can be controlled continuously by the model parameters q_A and q_B. It would maybe be convenient for the discussion if we could neglect the model parameters and investigate the models in terms of some basic network quantity like $\bar{k}$. However, such an approach is complicated by the fact that, for $q_{A,B} < 1$, the average degree also depends on N (see Fig. 1c and the above discussion). Moreover, the size scaling of $\bar{k}$ depends on q_A and q_B, going continuously (but with different functional forms) from $\bar{k} \sim N^{1/2}$ for $q_A = q_B = 0$, to size independence for $q_A = q_B = 1$. For this reason we keep the control parameters $q_{A,B}$ as our main independent parameters in this investigation and do not go into details of functional forms when discussing the size dependence of performance measures.

3.2 Connectivity

A fundamental functional requirement is that the network should be connected. There is, in principle, nothing that guarantees connectivity in our construction algorithm. We investigate the connectivity, quantified as φ – the fraction of time steps when the network is fully connected. Since φ is close to one, we rather plot $-\lg(1-\varphi)$ (Fig. 3). The larger are values of the energy limits (both $P_{\min}$ and $P_{\max}$), the better is the connectivity – by consuming more power, the network can always be made connected – a trivial but necessary fact. In the white region of the density plot in Figure 3, the network is fully connected more than 99.9% of simulation runs ($\varphi > 0.999$ and thus $-\lg(1-\varphi) > 3$). In the rest of this paper we will (unless stated otherwise) investigate a region of parameter space where the network is almost surely connected – $P_{\min} = 1$ and $P_{\max} = 2$.

As seen in Figure 3, a qualitative difference between the two models is that the connectivity of Model B is less dependent on q_B than Model A is dependent on q_A. For a large range of q_B (Fig. 3b), the network is connected almost regardless of $P_{\min}$. One explanation is the higher degrees of Model B – adding edges, increasing the degree, can connect, but not fragment, a network. The connectivity depending on the $P_{\max}$ value (Figs. 3c, 3d) shows different tendencies: model A becomes connected at lower $P_{\max}$-values (and consequently lower $P_{\min}$-values). To explain this feature we evaluate the actual power consumption of the two models in the next section.

3.3 Power consumption

The power consumption of an agent is a stochastic variable in the interval $[0, P_{\max}]$, changing as agents join and leave the network. (While in both models the agents try to allocate at least $P_{\min}$, we have to note that if an agent cannot

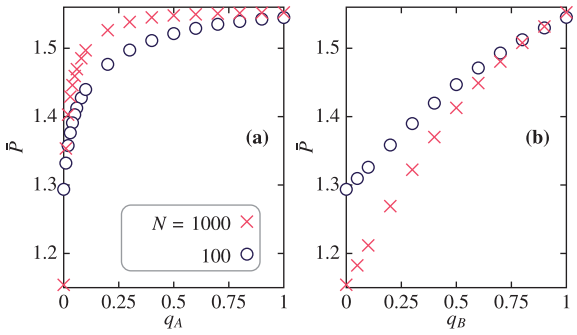

Fig. 4. (Color online) Power consumption for the two attachment models A and B as functions of their respective parameters, q_A and q_B. Other parameter values are the same as in Figure 1. Standard errors are smaller than the symbol size.

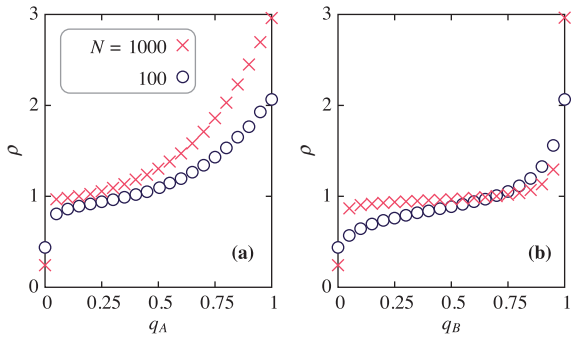

Fig. 5. (Color online) The power efficiency – the average (per vertex pair) ratio between the actual consumed in a shortest path routing power and the power of a direct connection – for the two attachment schemes as a function of their respective control parameter. Other parameter values are the same as in Figure 1. Standard errors are smaller than the symbol size.

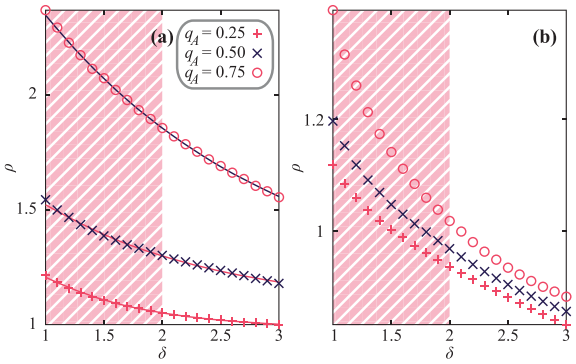

Fig. 6. (Color online) The power efficiency of the two attachment schemes as a function of δ, the exponent controlling the power consumption per distance of the connection. All curves are for system size $N = 1000$. The shaded areas represent the non-feasible range of δ. The curves in (a) are best (non-linear least-square) fits to exponential forms $\rho_\infty + (\rho_0 - \rho_\infty)\exp(-\delta/\Delta)$ (where ρ_∞, ρ_0 and δ are fitting parameters). Other parameter values are the same as in Figure 1. Standard errors are smaller than the symbol size.

connect without raising the power consumption over $P_{\max}$, its power consumption will be below $P_{\min}$.) If the average power consumption is higher, so more resources are invested in the infrastructure, one can expect the network properties to improve. Therefore, to better be able to analyze our performance measures, we investigate the level of power consumption. Figures 4a and 4b shows the average power consumption with $P_{\min} = 1$ and $P_{\max} = 2$. The power consumption maximally fluctuates $\sim 20\%$ and it is always lower for model B. This difference increases with system size for all values except close to $q_{A,B} = 1$, meaning that performance advantages for Model A in these limits, especially for larger systems, should be evaluated with this increased power consumption in mind. Or, alternatively, a specific q_A-value can be translated to a (higher) q_B-value, corresponding to the same power consumption.

Why is the power consumption higher in Model A? Again, we use Figure 2 as an illustration. Above, we explained why, if $q_A = q_B$, Model A has larger fraction of long-range edges than Model B. Since the steps in power consumption is larger when adding random edges, if a random edge is the last one added, then the power consumption overshoots $P_{\min}$ more. In sum, more randomly added edges mean more overshoot and thus higher power consumption.

Power efficiency in multihop networks is often measured by the ratio ρ of actual consumed power to the power needed for a direct connection, averaged over all vertex pairs (scaling like the distance squared):

$$\rho = \frac{1}{\binom{N}{2}} \sum_{i<j\in V} \frac{\sum_{(i',j')\in\pi(i,j)} p(i',j')}{p(i,j)} \qquad (3)$$

where $\pi(i,j)$ is the set of edges forming the shortest path (in number of hops) between i and j [2]. If the route between the source and the target is circuitous, ρ might be much larger than one, but if it proceeds in several steps relatively straight towards the target, ρ can be smaller than unity. Indeed, the infimum of ρ is zero – assume i and j, at distance r are connected by a straight path of n equidistant edges. Then $p(i',j')$ for all edges along the path is r^2/n^2, so the whole sum in the numerator equals r^2/n, while the denominator is r^2, giving a contribution $1/n$ from the vertex pair $\{i,j\}$ to ρ. This argument can be generalized to arbitrary partitions of a path by Jensen's inequality. In Figure 5 we plot this quantity for Model A and B, and two different network sizes. The ρ-curves are monotonically increasing for both sizes and attachment models. Both models have some regions of their parameter space where $\rho < 1$, which is the definition for being a *power spanner* [2]. For Model B and the larger system size, ρ plateaus around unity (which is a coincidence, other $P_{\min}$-values shifts the ρ-value of the plateau). Model B shows lower values of ρ for a large region of its parameter space, which means that shortest path routes are straighter compared to the ones in Model A, a results of the different distribution of long-range random connections.

We furthermore investigate ρ as a function of δ, the parameter governing the power consumption as a function of the distance between a sender and a receiver, see Figure 6. In this plot we do not rescale the power-limits so larger δ means that most of the possible connections, especially the

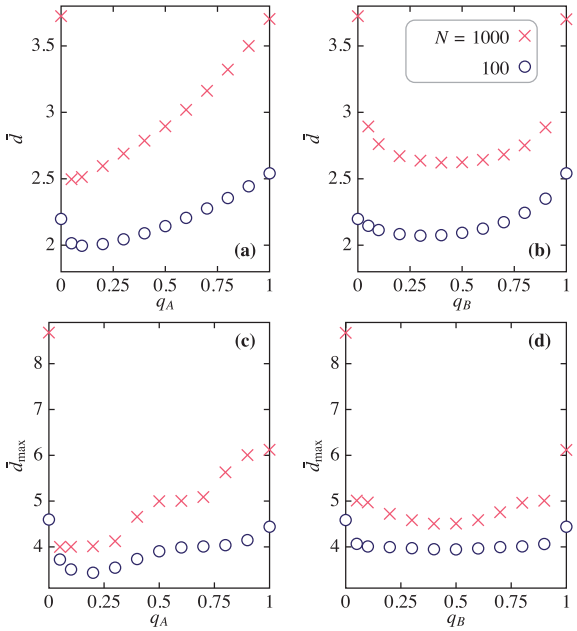

Fig. 7. Distance-related measures for different sizes and $q_{A,B}$-values. Panels (a) and (b) shows the average distance $\bar{d}$ while (c) and (d) shows the diameter $d_{\max}$. Other parameter values are the same as in Figure 1. Standard errors are smaller than the symbol size.

shorter, gets cheaper. This then means that the network gets denser so that there will be more direct connections and thus lower ρ. In the $\delta \to \infty$ the models become identical to the UDG model. In this limit, if there is any pair of vertices separated more than a unit distance, then the cost for a direct connection would be infinite and thus $\rho = 0$. In practice δ is thought to be close to but larger than 2 and in this region the behavior is far from the large-size limit. Indeed the curves for Model A fits well to an exponential decay that approaches values close to one. (Fitting the curves of Fig. 6a to $\rho_\infty + (\rho_0 - \rho_\infty)\exp(-\delta/\Delta)$ gives $\rho_\infty(q = 0.25) = 0.98 \pm 0.03$, $\rho_\infty(q = 0.5) = 1.08 \pm 0.03$ and $\rho_\infty(q = 0.75) = 1.11 \pm 0.05$.) The important observation from this figure is that the conclusion from Figure 5, that Model B has the lowest ρ-values (except the extreme points $q_{A,B} = 0, 1$), holds stronger for larger δ.

3.4 Distances

Both for unicast and multicast communication, short distances (in number of hops) are beneficial, as they lead to lower delays [18]. In Figure 7 we investigate the average distances and diameter, corresponding to mean and extreme-case connections as a function of the model parameters q_A and q_B. For both Model A and B, and both distance measures, the minima are attained for intermediate values of $q_{A,B}$. In other words, a mix between random and short-range connections optimizes network distances. Furthermore, the minimum occurs for low $q_{A,B}$-values, i.e. with larger proportion short-range connections compared with the random connections. The decreasing distances for $q_{A,B} \approx 0$ can be understood as a "small-world effect" – we only need to introduce a few random edges to a regular graph for it to go from algebraic to logarithmic distance scaling [19,20]. The increasing distances for larger $q_{A,B}$-values can be explained by the decreasing average degrees. Comparing Models A and B, we note that Model A gives the smallest distances (both average and maximal). If one increases the degree of a graph by adding edges to it, its average distance cannot increase. Similarly, most network models have distances decreasing with the average degree. With that in mind, it is a little surprising that Model A has both shortest distances and smallest average degree. This non-trivial, purely topological effect can only be related to the distribution of longer-range, random edges. In Model A, the long-range edges are more evenly distributed out between the agents (compared with Model B where, at least one side of the every long-range edge belong to one class of vertices). Model A loses this advantage as more random edges are inserted in the network, then the average and maximal distances increase together with the decreasing degree according to Figure 1. Model B has the advantage that it is not as sensitive to q_B as Model A is to q_A.

3.5 Clustering coefficient

The clustering coefficient C is a measure of the number of triangles, or local redundancy, in the network. Technically it is defined as [11]

$$C = \frac{3 \times \text{the number of triangles}}{\text{the number of connected triples}}, \tag{4}$$

where a connected triple is any subgraph of three vertices and two edges. The factor three is to normalize the measure to the $[0, 1]$ interval. A high clustering coefficient indicates that the network is robustness to edge failure on a local scale. On a global scale, one of the triangle's edges might be better of better use (to create robustness) if rewired to some distant location (we discuss a global measure of robustness, the spectral gap of the adjacency matrix, below). The clustering coefficient for Model A (Fig. 8a) follows an algebraic decay $C_\infty + C'N^{-\xi}$ (where C_∞, C' and ξ are fitting parameters) very well. By using this form we can extract the large system behavior C_∞ (plotted as a line in Fig. 8a). For Model B the situation is different (Fig. 8b) – the size-scaling curves, like the inset of Figure 8a, does not follow any universal functional form. For some q_B-values, the clustering coefficient is even growing with the system size. There are more triangles for intermediate q_A-values of Model A than for similar q_B-values of Model B.

3.6 Expander property

For multicast and broadcast communication, redundancy is an important topological factor. Imagine a triangle between agents 1, 2 and 3. A broadcast originating from 1 reaches 2 and 3 in two steps. Then, however, the edge between 2 and 3 is redundant – both 2 and 3 already have the

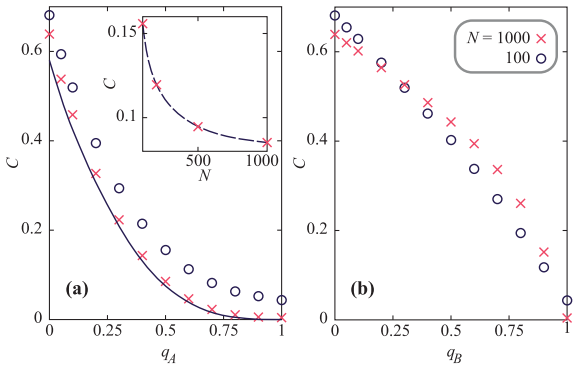

Fig. 8. (Color online) Clustering coefficient C for the two models (Model A in (a) and B in (b)) as a function of $q_{A,B}$. The inset in (a) shows the size-scaling for Model A with $q_A = 0.5$. The curve in the inset is a best non-linear least-squares fit to a power-law decay form – $C_\infty + C'N^{-\xi}$ where C_∞, C' and ξ are fitting parameters. The solid line in (a) is C_∞, the $N \to \infty$ limit of C. Other parameter values are the same as in Figure 8. Standard errors are smaller than the symbol size.

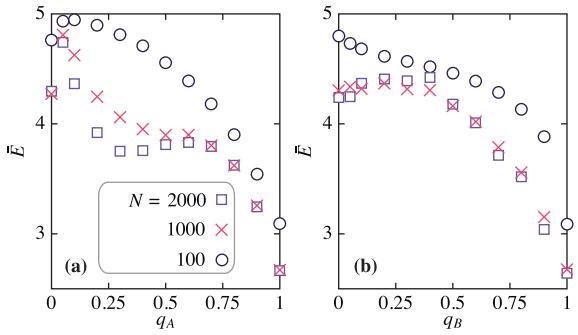

Fig. 9. (Color online) The spectral gap, the difference between the two largest eigenvalues of the adjacency matrix, for Model A (a) and B (b). Other parameter values are the same as in Figure 1. Standard errors are about the size of the symbols, but omitted for readability.

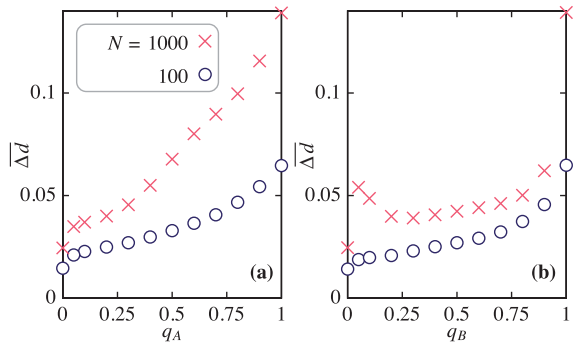

Fig. 10. (Color online) The expected difference in average distance as a random vertex is deleted from the network. Panels (a) and (b) shows curves for Model A and Model B respectively. Other parameter values are the same as in Figure 1. Standard errors are smaller than the symbol size.

packet. If a graph has few such redundancies, it is said to have good expander properties. The usual way of defining this property is via the expansion factor Φ – the minimum over all subsets S, smaller than half the entire vertex set, of the ratio between the size of the neighborhood of S and the size of S [21]. Φ is a computationally hard quantity. Instead, we measure the spectral gap $\bar{E}$ – the difference between the leading and second largest eigenvalues of the adjacency matrix, which is known to be a lower bound to Φ (apart from a factor 2) [15].

In addition to efficient broadcast communication, large spectral gap also leads to two other desirable network properties. First, it reflects robustness. It is known that if the graph has a large spectral gap one needs to cut many edges to split it into two disjoint subnetworks [15]. Consequently, the spectral gap is a measure of resilience against a worst-case scenario of an adversary deliberately trying to disconnect different parts of the network [22]. Second, networks with large spectral gaps are easier to synchronize [23]. This is especially relevant for synchronizing clocks of sensor networks [24].

In Figure 9, we plot the expander coefficient for the two models corresponding to Figure 7. The best parameter range is for small $q_{A,B}$. In contrast with the distance measures, the maximum $\bar{E}$ for Model B is network size dependent and does not always occur for an intermediate q_B-value. Model A, on the other hand, has a maximum for small q_A-values. For both models there is a small $q_{A,B}$ value where $\bar{E}$ is high and stable (indeed accentuated) as the system grows. This stable value is higher for Model A, which indicates that Model A can achieve higher performance than Model B for the same system size. For intermediate $q_{A,B}$-values, $\bar{E}$ seems to converge for Model B, whereas it decreases (roughly) logarithmically for Model A. Because of this, in implementations where the system size is unknown a priori, and the model parameter is hard to control exactly, Model B might be advantageous. Model A also shows a second incipient peak for the larger sizes, much lower than the low-q_A peak and probably uninteresting for practical applications.

3.7 Robustness

In addition to resilience to worst case attacks measured by the spectral gap, it is desirable that network graphs are robust to random failures. This is, in fact, similar to demand that the turnover of agents should not worsen our objective functions much. To investigate how random failures affect the network we measure $\Delta\bar{d}$ – the difference in average pathlength if a random vertex is disconnected from the network. From Figure 10 we can see that this quantity is also optimized at low $q_{A,B}$-values. But in contrast to the behavior of E, $q_A = 0$ is optimal for Model A and a small q_B (around $q_B = 0.2$) is optimal for Model B ($N = 1000$). Like other quantities, Model B is less q_B-dependent than Model A is q_A-dependent. Both models have similar magnitude of $\Delta\bar{d}$.

4 Summary and conclusions

We have investigated the optimization of network topology of the medium access layer of wireless networks, and

found that agents with simple but heterogeneous rules for finding neighbors are capable to form network with shorter average distances, better expander properties and relatively high robustness (compared to random or UDG approaches). The two models we test have parameters controlling the ratio of long-range interactions. We find that close to (but not at) one end of parameter space (where most connections are short) is the region where optimal topologies exist.

For most score functions – average distance, diameter and spectral gap – our Model A (where agents attach to spatially close others with a probability q_A, otherwise to random agents) gives better values than our Model B (where there are two classes of agents, one making only short connections, one making only random connections). The good topological properties of Model A are however outweighed by the effect of fast decreasing average degree as q_A increases. Model B, on the other hand, has the advantage that its performance is less dependent on the parameter value.

This work was supported by the Swedish Research Foundation through the Linneaus Centre ACCESS. PH acknowledges financial support from the Swedish Foundation for Strategic Research and the Swedish Research Council and the WCU program through NRF Korea funded by MEST (R31–2008–000–10029–0). BJK acknowledges the support from the Korea Research Foundation Grant KRF–2009–013–C00021.

References

1. X.Y. Li, Wireless Communications and Mobile Computing **3**, 119 (2003)
2. S. Rührup, C. Schindelhauer, M. Grünewald, K. Volbert, Performance of Distributed Algorithms for Topology Control in Wireless Networks, in *In 17th International Parallel and Distributed Processing Symposium* (2008), p. 2003
3. A.C.C. Yao, SIAM J. Comput. **11**, 721 (1982)
4. X.Y. Li, P.J. Wang, Y. Wang, O. Frieder, Sparse Power Efficient Topology for Wireless Networks, in *Proceedings of the 35th Annual Hawaii International Conference on System Sciences* (2002), pp. 3839–3848
5. X.Y. Li, P.J. Wan, Y. Wang, Power efficient and sparse spanner for wireless ad hoc networks, in *IEEE ICSN* (2001)
6. J. Liu, B. Li, Distributed Topology Control in Wireless Sensor Networks with Asymmetric Links, in *Global Telecommunications Conference* (2003)
7. F. Kuhn, A. Zollinger, Ad-hoc networks beyond unit disk graphs, in *DIALM-POMC '03: Proceedings of the 2003 joint workshop on Foundations of mobile computing* (ACM, New York, NY, 2003), pp. 69–78, ISBN 1-58113-765-6
8. N. Li, J.C. Hou, Topology Control in Heterogeneous Wireless Networks: Problems and Solutions, in *IEEE INFOCOM 2004* (2004), pp. 232–243
9. V. Rodoplu, T.H. Meng, IEEE Journal on Selected Areas in Communications **17**, 1333 (1999)
10. R. Wattenhofer, L. Li, P. Bahl, Y.M. Wang, *Distributed Topology Control for Wireless Multihop Ad-hoc Networks*, in *INFOCOM* (2001), pp. 1388–1397
11. M.E.J. Newman, SIAM Rev. **45**, 167 (2003)
12. S.N. Dorogovtsev, J.F.F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003)
13. P.T. Eugster, R. Guerraoui, A.M. Kermarrec, L. Massoulie, IEEE Computer **37**, 60 (2004)
14. B.M. Waxman, IEEE Journal on Selected Areas in Communications **6**, 1617 (1988)
15. S. Hoory, N. Linial, A. Wigderson, Bulletin of the American Mathematics Society **43**, 439 (2006)
16. H. Takagi, L. Kleinrock, IEEE Transactions on Communications **32**, 246 (1984)
17. V. Loscrı, J. Parallel Distrib. Comput. **68**, 387 (2008), ISSN 0743-7315
18. A. El Gamal, J. Mammen, B. Prabhakar, D. Shah, IEEE/ACM Trans. Netw. **14**, 2568 (2006)
19. D.J. Watts, S.H. Strogatz, Nature **393**, 440 (1998)
20. B. Bollobás, F.R.K. Chung, SIAM J. Discrete Math. **1**, 328 (1988)
21. E. Estrada, Phys. Rev. E **77**, 036111 (2008)
22. P. Holme, B.J. Kim, C.N. Yoon, S.K. Han, Phys. Rev. E **65**, 056109 (2002)
23. L. Donetti, P.I. Hurtado, M.A. Muñoz, Phys. Rev. Lett. **95**, 188701 (2005)
24. Q. Li, D. Rus, IEEE Transactions on Computers **55**, 214 (2006)