NAVIGATION IN SMALL WORLD NETWORKS, A SCALE-FREE CONTINUUM MODEL

by

Massimo Franceschetti amd Ronald Meester

Memorandum No. UCB/ERL M03/33

25 August 2003

Cover

NAVIGATION IN SMALL WORLD NETWORKS, A SCALE-FREE CONTINUUM MODEL

by

Massimo Franceschetti and Ronald Meester

Memorandum No. UCB/ERL M03/33

25 August 2003

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Navigation in small world networks, a scale-free continuum model *

Massimo Franceschetti
University of California at Berkeley

Ronald Meester
Vrije Universiteit Amsterdam

Abstract

"Everybody on this planet is separated by only six other people. Six degrees of separation. Between us and everybody else on this planet. The president of the United States. A gondolier in Venice ... It's a profound thought ... How every person is a new door, opening up into other worlds."

This powerful observation is made by one of the characters in John Guare's play "six degrees of separation". Long a matter of anecdotal evidence, the small world phenomenon, the principle that we are all linked by a short chain of intermediate acquaintances, has been investigated in mathematics and social sciences. It has been shown to be pervasive in both nature and engineering systems, like the World Wide Web. Recent work of Jon Kleinberg has pointed out that a peculiar feature of the phenomenon is that people, using only local information, are very effective at finding short paths in a network of social contacts.

In this paper, we depart from the common practice to use probabilistic combinatorial models to explain the small world phenomenon, and we address the problem raised by Kleinberg in a more natural continuum setting. We introduce a random connection model that is related to continuum percolation, and we show the existence of a unique *scale-free* model, among a large class of models, that allows construction, with high probability, of short paths between pair of points separated by any distance scale. Our model supports the idea that the real world of social contacts is scale-free, and provides guidelines to construct networks on which it is possible to perform efficient routing using only local knowledge of the network at each node.

*This work was partially supported by DARPA grant number F33615-01-C-1895 at Berkeley, and by the center for mathematical research Ennio De Giorgi of the Scuola Normale Superiore of Pisa, where the authors were hosted in July 2003.

1 Introduction

1.1 The small world phenomenon

It often happens to meet a stranger and discover that we are linked through a short chain of intermediate acquaintances. As early as 1929, the Hungarian writer Frigyes Karinthy [4] speculated that anyone in the world could be connected to anyone else through a chain consisting of no more than five intermediaries. This was long a matter of anecdotal evidence, until the famous experiment of Milgram [8]. In this experiment letters were given to subjects in one of the United States, with instructions to deliver them to a single target person in another state, by mailing the letter to an acquaintance who the subject deemed closer to the target. The acquaintance then got the same set of instructions, thus setting up a chain of intermediaries. Milgram found that the average length of the chains that completed was about six—quite remarkably close to Karinthys prediction 40 years earlier. This striking result continues today to be an object of fascination and amusement, and has been popularized in the nineties by John Guare's succesful play "six degrees of separation" [3].

Naturally, the experimental discovery quickly led to analytical work aimed at explaining the phenomenon. For many years the typical explanation has been that random graphs have low diameter. When pairs of vertices are joined uniformly at random, with some probability, then any two vertices are connected by a short chain with high probability. This simple model, however, fails to capture the local structure of a social network. A refinement has been proposed in a paper by Watts and Strogatz [11]. These authors noted that many real world networks, like the social contacts networks investigated by Milgram, but also biological networks, and artificial networks (power grid, world wide web), tend to be highly clustered, like lattices, but have small diameters, like random graphs. In these networks it is possible to find short chains connecting any two vertices, but many of the neighbors of a node are also neighbors of each other. To capture both of these properties, Watts and Strogatz proposed a model that is a superposition of a structured subgraph of "local contacts" and a random subgraph of "long range contacts". They noted that by adding uniformly at random few edges to a structured subgraph like a ring or a mesh, it is possible to drastically reduce its diameter. Similar kind of networks have also been investigated in the field of probabilistic combinatorics [1]. For example, Bollóbas and Chung [2] gave bounds on the diameter of the random graph obtained by adding a random matching to the nodes of a cycle.

1.2 Navigation in the small world

There is another, more surprising conclusion to be drawn from Milgram's experiment. As pointed out by Kleinberg [5] [6], Milgram's result demonstrates not only the *existence* of short paths in the network, but also the ability of people at *finding* them. Milgram's simple instructions of forwarding a letter to the "closest" acquaintance to the target were sufficient to identify such paths. Note that there is a fundamental difference between the *existential* discovery and the *algorithmic* discovery. It is quite possible that short paths exist, but that these cannot be found by any algorithm using only local knowledge of the network. In

Milgram's experiment, the subjects had only knowledge of the their local contacts and of the final target. Nevertheless, they were able to find a short path to the target.

Motivated by this observation, Kleinberg [5] [6] proposed a model that is a variant of the small-world model of Watts and Strogatz. He considers a regular lattice, and rather than adding long range contacts uniformly at random, he adds them in a biased way, having connections more likely to exist between lattice sites that are close together in the Euclidean space defined by the lattice. He shows that the fewest number of sites visited to reach the target using a routing algorithm with only local information, is achieved when the probability of having a connection between two sites decays with the *square* of their distance. This is the only case when it is possible for an algorithm using only local information to reach the target in a logarithmic number of steps. Any power law exponent other than 2 leads to a polynomial number of steps. For this reason, Kleinberg concludes, many random networks differ from social networks and do not allow fast routing with local information: their connection probability scales wrongly.

1.3 Proposed continuum model

Although it represents a seminal contribution, we claim that Kleinberg's model is slightly unnatural to describe the small world phenomenon. For one thing, it is a discrete model that assumes all nodes to be located on a lattice, and this is often not the case in the real world. More importantly, the number of local and long range contacts that he considers are uniformly bounded in the system size. Namely, the local contacts are deterministically formed by connecting each site to a constant number p of nearest neighbors on the square lattice, and the long range contacts that are randomly added are also a constant number q. For example, the case $p = 1$, $q = 2$, corresponds to a social network where people live on a square grid, each one having exactly 4 neighbors as their local acquaintances, and exactly 2 long range acquaintances that live somewhere at random grid points, see left-hand side of Figure 1.3. A uniform (independent of the system size) upper bound on the number of acquaintances is somewhat unnatural.

In this paper, we present a random connection model that is related to continuum percolation [7], and that more naturally describes the phenomenon. When one tries to model the small world phenomenon with a continuum percolation approach, probably the first idea that comes to mind is to represent people by points of a Poisson point process on the plane, and connect them according to a so called *connection function* $g(\cdot)$, that is, two people at Euclidean distance x are connected to each other (i.e., are acquainted) with probability $g(x)$. Although this can certainly be done, a model along these lines fails to capture problems in delivering messages that may arise at all scales, including very small scales. A connection function $g(x)$, acting on a fixed density of points, does not allow proper scaling of the entire model. Accordingly, we suggest a slight variation in the following.

We start with an individual located at s, the *source*, which has a message which he wants to delver to an individual at location t, the *target*. The message holder has a random number of acquaintances, randomly located in the plane according to a non-homogeneous Poisson process with density function $g(x) = \lambda|x|^{-\alpha}$, where $\alpha, \lambda > 0$; see the right hand side of Figure 1.3. Note that there is no uniform upper bound on the number of acquaintances of each individual. Moreover, it is the *density* of acquaintances of each node that scales

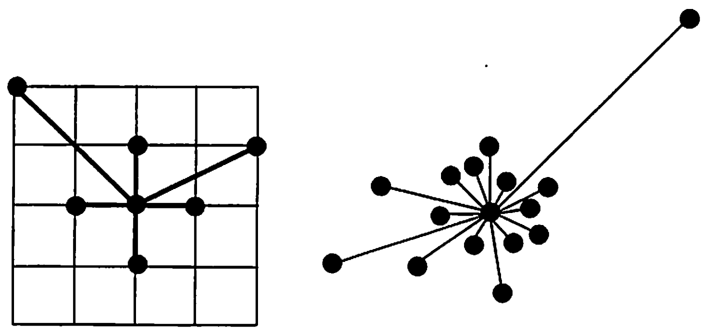

Figure 1: **Discrete vs. continuum model.** *Schematic representation of the connections of one node in two different models. Left-hand side: Kleinberg's model with parameters $p = 1$, $q = 2$. Right-hand side: continuum model where a node has a random number of connections, randomly located in the plane.*

with a power of the distance to that node. Any acquaintance at position z, say, has itself a random number of further acquaintances, which can be described by a non-homogeneous Poisson process with density function $h(x) = \lambda|x - z|^{-\alpha}$, independent of previous Poisson processes, etcetera. For any $\epsilon > 0$, we say that an *ϵ-delivery* has taken place if the message has been forwarded to an individual within distance ϵ of the target.

We will show that for $\alpha < 2$, any routing algorithm using only local information (that is, the current message holder in a given step of the algorithm knows only the location of its acquaintances and of the destination) does not perform well *near* the target, as the number of steps required to deliver a message in a small ϵ-neighborhood of the target grows polynomially in $1/\epsilon$. On the other hand, for $\alpha > 2$, the performance bottleneck is at *large* distances, as the number of steps in this case grows polynomially in the distance d between the source and the target. Finally, for $\alpha = 2$, the simple greedy algorithm that forwards the message to the acquaintance that is *closest* to the target, performs well at all scales, leading to a bound on the expectation of the ϵ-delivery time that is logarithmic in *both* $1/\epsilon$ and d.

Our results show that when $\alpha < 2$, a decentralized algorithm can perform well in the far range, exploiting long links, but then slows down in the close range, where there are not enough short links. Similarly, when $\alpha > 2$, the algorithm can perform well in the close range, exploiting short links, but it slows down in the far range, where there are only few long links. The only case when a decentralized algorithm can keep a constant pace and deliver a message in a logarithmic number of steps at all distance scales (very near, as well as very far) is when $\alpha = 2$. In this case the underlying random connection model is *scale-free*, as it allows fast delivery at all scales.

1.4 Advantages of a continuum model

The main advantage of the proposed continuum model is that it allows analysis at all scales. It reveals the effect of the scaling exponent α on the delivery time of an algorithm when the routed message is either far, or close to the target. A discrete model, as the continuum percolation model that first came to mind, fails to capture what happens in the small scale, because once a message gets sufficiently close to the target, it can always be routed through the (deterministic) network of local contacts. Hence, adding more long range contacts can only improve the expected delivery time. Kleinberg's result that $\alpha < 2$, (giving higher probability to longer links) increases the delivery time, is more an artefact of his model, rather than a true effect of the scaling of the distribution. To see this, let us consider a modified (and more natural) version of his model that does not bound the number of contacts enjoyed by each node. Accordingly, let us consider a model where long range contacts are added between each pair of nodes of a square grid, with probability proportional to $d^{-\alpha}$, where d is the distance in terms of grid edges that need to be traversed to connect the two nodes. Note that in this case the number of long range contacts departing from each node is not bounded, because each pair has some probability, independent of all other pairs, of being joined by an edge. The scaling of the probability distribution simply makes it less likely that two nodes that are far away are joined by an edge. This is more natural than assuming each node to have a fixed number of long distance acquaintances. Clearly, the minimum delivery time of a message routed between two randomly selected nodes must be non-decreasing in α, because there are on average fewer long range edges departing from each node as α increases. Hence, it appears as there is nothing special about a scaling distribution with exponent $\alpha = 2$. What makes Kleinberg's model behave differently is that in his formulation the number q of long range contacts enjoyed by each node is uniformly bounded in the system size, and for this reason $\alpha = 2$ turns out to be the best possible scaling exponent for the obtained random graph. Our continuum formulation shows that the role of the scaling exponent is important even if we do not bound the number of long range connections per node. We show that different scaling laws affect the delivery time at different distance scales, and that there is only one scale free distribution that allows fast delivery across all distances in the considered random connection model.

1.5 Applications

There are several reasons to investigate small world networks, besides trying to explain the small world phenomenon in social science. By identifying the core properties of small world networks one can try to construct artificial networks that exhibit a similar structure, and on which fast routing using only local information is possible. This applies to a variety of situations. In large communication networks one basic goal is to keep the size of the routing tables small, while guaranteeing fast delivery service. Hence, a fast algorithm that needs to store at each node only the location of its neighbors is highly desirable. Moreover, in several types of wireless networks the topology of the network is unknown and can change over time, due to the mobility of the users and to channel fluctuations, and routing must be performed using local information. Recently, there has been a growing interest in large networks of wireless sensors [10]. In this case, nodes are often deployed randomly and each one connects to a random number of neighbors, that on average decreases with the distance to the node. Given the limited memory and processing capabilities of each sensor node, routing using only local information becomes a key issue. A basic design guideline that can be extracted from this paper is to try to build sensor networks for which the number of neighbors of each node scales with the square of the distance from the node. Other important applications are robot navigation in unknown environments, data mining, web crawling, etc. Finally, small world networks that allow efficient, decentralized navigation, appear in nature in many biological systems [11].

2 Main results

We start with a few definitions. Consider the full plane $\mathbb{R}^2$ as our model of the real world. Let $g(x) = 1/x^\alpha$, for some scaling exponent $\alpha > 0$ and $x \in \mathbb{R}$. For any given point located at position $z \in \mathbb{R}^2$, let its acquaintances be given by an non-homogeneous Poisson point process X with density function $h(y) = \lambda g(|z-y|)$, for some $\lambda > 0$. Let d be the Euclidian distance between a source point $s \in \mathbb{R}^2$ and a target point $t \in \mathbb{R}^2$. We define a *decentralized algorithm* as a mechanism whereby a message is sent from s to t, being sequentially passed along a chain of intermediate acquaintances. That is, the current message holder u in a given step knows only the location of its acquaintances in $\mathbb{R}^2$ and the location t of the target. Based on this information, u forwards the message to one of its acquaintances. For some $\epsilon > 0$, define the ϵ-delivery time of a decentralized algorithm $\mathcal{A}$ as the number of steps required for the message originating at s to reach an ϵ-neighborhood of t, making at each step the forwarding decision based on the rules of $\mathcal{A}$. Finally, let $\overline{\mathcal{A}}$ be the decentralized algorithm that at each steps forwards the message to the local acquaintance that is closest in Euclidian distance to the target. We state our results in the following theorem.

Theorem 2.1 *The scaling exponent α of the model influences the ϵ-delivery time of a decentralized algorithm as follows:*

- *CASE 1. For $\alpha = 2$ and any $\epsilon > 0$ and $d > \epsilon$, there is a constant $c > 0$, such that the expected ϵ-delivery time of the decentralized algorithm $\overline{\mathcal{A}}$ is at most $c(\log d + \log 1/\epsilon)$.*
- *CASE 2. For $\alpha < 2$ and any $\epsilon > 0$, there exists a constant $c(\alpha) > 0$, such that the expected ϵ-delivery time of any decentralized algorithm $\mathcal{A}$ is at least $c(\alpha)(1/\epsilon)^{2-\alpha}$.*
- *CASE 3. For $\alpha > 2$ and any $\epsilon > 0$ and $d > 1$, the expected ϵ-delivery time of any decentralized algorithm $\mathcal{A}$ is at least cd^β, for any $\beta < \frac{\alpha-2}{\alpha-1}$ and some constant $c = c(\alpha, \beta) > 0$.*

Essentially, the theorem above says that for $\alpha = 2$ it is possible to approach the target at any distance scale in a logarithmic number of steps, steadily improving at each step. On the other hand, when $\alpha < 2$ a decentralized algorithm starts off quickly, but then slows down as it approaches the target, having trouble to make the last small steps. Finally, for $\alpha > 2$, the situation is reversed, as the performance bottleneck is not near the target, but is at large distances $d \gg \epsilon$. Our Case 3 corresponds to Kleinberg's [6] Theorem 3(b). It is interesting that he obtains the same exponent. On the other hand, our Case 1, that corresponds to his Theorem 2, presents a slightly better bound compared to his square of logarithm bound.

Our proof is based on the following properties of the scaling exponent α. For $\alpha < 2$ the probability that at each step a decentralized algorithm can make progress, tends to zero as the distance to the target gets smaller. For this reason, the algorithm slows down as

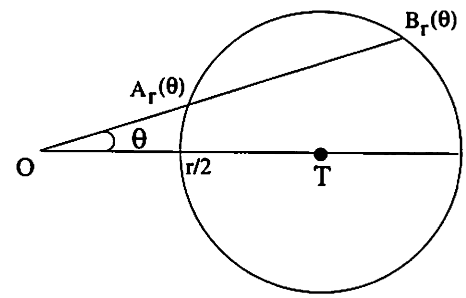

Figure 2: **Decreasing the distance to the target by a factor 1/2.**

it approaches the target. For $\alpha > 2$ a decentralized algorithm can make progress at each step with probability one, but the amount of progress is limited and cannot be of the same order of the distance to the target. For this reason, a logarithmic bound on the running time cannot be achieved. Finally, for $\alpha = 2$, progress towards the target is guaranteed with probability one, and with uniform positive probability this is of the same order as the distance to the target. In this case, by possibly reducing the distance to the target by a constant factor at each step, and with the certainty of not drifting away from the target, a logarithmic bound is achieved.

3 Proof of Theorem 2.1

Proof of CASE 1. We first compute the probability that at any step of the algorithm an intermediate node has a neighbor that is at less than half of the distance to the target and show that this is positive and independent of distance. We refer to Figure 2. Let $\overline{OT} = r$ be the distance to the target. The (random) number of neighbors that are closer than $r/2$ to the target T has a Poisson distribution with mean

$$\mu = \lambda \int_{-\pi/6}^{\pi/6} \int_{A_r(\theta)}^{B_r(\theta)} g(x) x dx d\theta. \tag{1}$$

By scaling we have that $B_r(\theta) = rB_1(\theta)$, $A_r(\theta) = rA_1(\theta)$; and by substituting $g(r) = 1/r^2$ into (1) we have,

$$\mu = \lambda \int_{-\pi/6}^{\pi/6} \int_{rA_1(\theta)}^{rB_1(\theta)} \frac{1}{x^2} x dx d\theta = \lambda \int_{-\pi/6}^{\pi/6} \log \frac{B_1(\theta)}{A_1(\theta)} d\theta, \tag{2}$$

which is independent of r. It follows that there is always a positive probability $\tau = 1 - e^{-\mu}$, independent of r, that point O has a neighbor inside the line disc depicted in Figure 2, i.e., at least half times nearer to the target T. Hence, algorithm $\overline{\mathcal{A}}$, forwarding the message to the node closest to the target, can reduce the distance to the target by a factor of at least 1/2 with uniform positive probability at each step. Whenever this occurs we say that the algorithm has taken a *successful* step. We have seen that a successful step has uniform

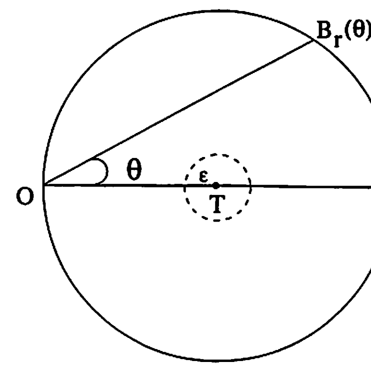

Figure 3: **Getting closer to the target.**

positive probability, we now show that a step that simply decreases the distance to the target has probability one. The number of points that are closer than r to the target is again Poisson distributed, with mean given by the integral of λg over the disc of radius r centered at T. It is easy to see that this integral diverges, and hence this number is infinite with probability one. It follows that the probability of *decreasing* the distance to the target has probability one. Hence, even when a step of the algorithm is not successful, it won't increase the distance to the target. It follows that at most a total number of n successful steps are needed to reach an ϵ-neighborhood of T, starting at a distance $d > \epsilon$, where

$$\left(\frac{1}{2}\right)^n d \leq \epsilon \Leftrightarrow n \leq \frac{\log d + \log 1/\epsilon}{\log 2}. \tag{3}$$

The expected waiting time for the n-th successful step is n/τ, and therefore our bound on the expected ϵ-delivery time is

$$E(\epsilon\text{-delivery time}) \leq \frac{\log d + \log 1/\epsilon}{\tau \log 2}, \tag{4}$$

which concludes the proof in this case. □

Proof of CASE 2. We consider a generic step of an algorithm, where the message is at point O, at distance $r \geq \epsilon$ from the target. We refer to Figure 3 and start by computing the number of acquaintances of point O that are closer to the target. This follows a Poisson distribution, and since $\alpha < 2$ it has a finite mean

$$\mu(r, \alpha) = \lambda \int_{-\pi/2}^{\pi/2} \int_0^{B_r(\theta)} g(r) r dr d\theta$$

$$= \lambda \int_{-\pi/2}^{\pi/2} \int_0^{rB_1(\theta)} \frac{1}{r^{\alpha-1}} dr d\theta = \frac{\lambda}{2-\alpha} r^{2-\alpha} \int_{-\pi/2}^{\pi/2} B_1(\theta)^{2-\alpha} d\theta = c(\alpha) r^{2-\alpha}. \tag{5}$$

Let an *improving* step of any decentralized algorithm be one that forwards the message to a neighbor that is closer to the target. Since $r \geq \epsilon$, entering an ϵ-neighborhood of the target also requires getting closer to the target, and we have

$$\begin{aligned} P(\text{enter } \epsilon\text{-neighborhood of T}) &\leq P(\text{improving step}) \\ &\leq 1 - e^{-c(\alpha)r^{2-\alpha}} \leq c(\alpha)r^{2-\alpha} \\ &\leq c(\alpha)\epsilon^{2-\alpha}, \end{aligned}$$

for $r \geq \epsilon$. It follows that the expected number of steps required to enter an ϵ-neighborhood of the target is at least

$$E(\epsilon\text{-delivery time}) \geq \frac{1}{c(\alpha)\epsilon^{2-\alpha}}. \tag{6}$$

□

Proof of CASE 3. Consider the collection of acquaintances of a given individual, and denote by D the distance to the acquaintance farthest away. We compute

$$\begin{aligned} P(D > r) &= 2\pi\lambda \int_r^\infty x^{-\alpha} x dx \\ &= \frac{c}{\alpha - 2} r^{2-\alpha}, \end{aligned}$$

for some constant c. This quantity tends to zero as $r \longrightarrow \infty$, since $\alpha > 2$.

We next estimate the probability that starting at distance d, an ϵ-delivery can take place in at most d^β steps, for some $\beta > 0$. Delivery in at most d^β steps implies that in one of the first d^β steps of the algorithm, there must be at least one stepsize of size at least $d^{1-\beta}$. According to the computation above, the probability that this happens is at most $d^\beta d^{(1-\beta)(2-\alpha)} = d^{2-\alpha-\beta+\alpha\beta}$. Writing X_d for the delivery time starting at distance d, we have shown that

$$P(X_d \geq d^\beta) \geq 1 - d^{2-\alpha-\beta+\alpha\beta} \tag{7}$$

and therefore

$$E(X_d) \geq d^\beta(1 - d^{2-\alpha-\beta+\alpha\beta}). \tag{8}$$

Whenever $2 - \alpha - \beta + \alpha\beta < 0$, that is, whenever

$$\beta < \frac{\alpha - 2}{\alpha - 1},$$

this expression is at least cd^β (recall that $d > 1$). The result now follows. □

4 Conclusion

Routing with only local information is an important problem, that has been investigated in different settings, including the natural sciences and engineering. We have presented a natural continuum model that identifies the scaling properties of a network that make routing with local information possible. We have identified a unique model, among a large class of continuum models, that is scale-free, allowing fast delivery of messages at all distance scales. Beside being useful to explain natural phenomena, if one can construct networks that resemble this model, a simple routing algorithm can achieve fast delivery using only local knowledge of the network at each node. In particular, this can be useful in the context of large networks of wireless sensors.

We believe that our continuum approach, based on a random connection model, is quite general and can be applied to explain other properties of real world networks.

5 Acknowledgment

The authors would like to thank the center for mathematical research Ennio De Giorgi of the Scuola Normale Superiore of Pisa, for providing funding and an ideal setting where part of this research was developed.

References

[1] B. Bollóbas (2003). Mathematical results on scale-free random graphs. *Preprint.*

[2] B. Bollóbas, and F. R. K. Chung (1988). The diameter of a cycle plus a random matching. *SIAM Journal on discrete mathematics 1, 328-333.*

[3] J. Guare (1990). Six degrees of separation: a play. *Vintage, New York.*

[4] F. Karinthy (1929). Chains. Everything is different. *Budapest.*

[5] J. M. Kleinberg (2000). Navigation in the small world. *Nature 406, p. 845.*

[6] J. M. Kleinberg (2000). The small world phenomenon: an algorithmic perspective. *Proceedings of the 32nd annual ACM symposium on the theory of computing, pp. 163-170.*

[7] R. Meester and R. Roy (1996). Continuum percolation. *Cambridge University Press.*

[8] S. Milgram (1967). The small world problem.' *Psychology today, 2, pp. 60-67.*

[9] M. E. J. Newman (2003). The structure and function of complex networks. *SIAM Review 45(2), pp. 167-256.*

[10] B. Sinopoli, C. Scharp, S. Schaffert, L. Schenato, and S. S. Sastry (2003). Distributed control applications within sensor networks. *Proceedings of the IEEE. Special issue on distributed sensor networks, to appear.*

[11] D. J. Watts, and S. H. Strogatz (1998). Collective dynamics of small-world networks. *Nature 393, pp.440-442.*