



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

MASSIMO FRANCESCHETTI,* *University of California, San Diego*

RONALD MEESTER,** *Vrije Universiteit Amsterdam*

Abstract

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 both in nature and in engineering systems like the World Wide Web. Work of Jon Kleinberg has shown that people, using only local information, are very effective at finding short paths in a network of social contacts. In this paper we argue that the underlying key to finding short paths is scale invariance. In order to appreciate scale invariance we suggest a continuum setting, since true scale invariance happens at all scales, something which cannot be observed in a discrete model. We introduce a random-connection model that is related to continuum percolation, and we prove the existence of a unique scale-free model, among a large class of models, that allows the construction, with high probability, of short paths between pairs of points separated by any distance.

Keywords: Small-world network; scale invariance; random-connection model; Poisson process

2000 Mathematics Subject Classification: Primary 60D05
Secondary 60K37

1. The small-world phenomenon

1.1. Introduction and background

It often happens that we 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 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 US state, with instructions to deliver them to a single target person in another state by mailing the letter to an acquaintance whom the subject deemed closer to the target. This 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 – remarkably close to Karinthy's prediction 40 years earlier. This striking result continues to be an object of fascination and

Received 14 June 2005; revision received 14 September 2006.

* Postal address: Electrical and Computer Engineering Department, University of California, San Diego, 9500 Gillman Drive, MC 0407, La Jolla, CA 92093-0407, USA.

** Postal address: Divisie Wiskunde en Informatica, Faculteit der Exacte Wetenschappen, Vrije Universiteit Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands. Email address: rmeester@cs.vu.nl

Presented at the ICMS Workshop on Spatial Stochastic Modelling with Applications to Communications Networks (Edinburgh, June 2004).

amusement today, and was popularized in the 1990s in John Guare's successful 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, 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 [9]. Those authors noted that many real-world networks (like the social contacts networks investigated by Milgram, and also biological networks) and artificial networks (like power grids and the 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 a small number of edges uniformly at random to a structured subgraph like a ring or a mesh, it is possible to drastically reduce its diameter. Similar kinds of network have been investigated in the field of probabilistic combinatorics [2]. For example, Bollóbas and Chung [1] gave bounds on the diameter of the random graph obtained by adding a random matching to the nodes of a cycle.

There is, however, 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 to *find* them. Milgram's simple instruction of forwarding a letter to the 'closest' acquaintance to the target was 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 knowledge only of their local contacts and of the final 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 considered a regular lattice and, rather than adding long-range contacts uniformly at random, added them in a biased way, such that connections are more likely to exist between lattice sites that are close together in the Euclidean space defined by the lattice. He showed that the number of sites visited before reaching the target using a routing algorithm with only local information is minimized when the probability of having a connection between two sites decays with the *square* of their distance. This is the only case in which 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 two leads to a polynomial number of steps. For this reason, Kleinberg concluded, many random networks differ from social networks and do not allow fast routing with local information: their connection probabilities scale wrongly.

1.2. Scale invariance and a continuum setting

Although it represents an interesting contribution, Kleinberg's model is a somewhat unnatural way to describe the small-world phenomenon. For one thing, it is a discrete model that assumes all nodes to be located on a lattice, which is often not the case in the real world. More importantly, the numbers of local and long-range contacts that he considered are uniformly bounded in the system size. The local contacts are deterministically formed by connecting each site to a fixed number of p -nearest neighbors on the square lattice, and the number of long-range contacts that are randomly added, q , is also fixed. For example, the case in which $p = 1$ and $q = 2$

corresponds to a social network where people live on a square grid, each one having exactly four neighbors as their local acquaintances and exactly two long-range acquaintances that live somewhere at random grid points. A uniform upper bound on the number of acquaintances (i.e. one independent of the system size) is somewhat unnatural.

In this paper, we argue that the underlying reason for fast delivery (i.e. a short chain) is *scale invariance*. Loosely speaking, we say that a (random) graph is scale invariant if the graph has no natural scale; we will be more precise in a moment. Scale invariance will only occur when the parameter, α , which governs the probabilistic connections at a given distance takes the value two, and this is the reason why fast delivery can only occur in this situation.

Now, true scale invariance can only be described in a continuum setting that accounts for *all* distance scales, including the very small and the very large. This is why our setup is in the context of continuum percolation [7]. Indeed, when the vertices are all on a discrete grid, like the square lattice, then there is no way to observe scale invariance at small distances.

When trying to model the small-world phenomenon using 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* $g(\cdot)$, according to which two people with Euclidean separation 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. Indeed, a connection function acting on a fixed density of points still does not allow proper scaling of the entire model on all scales.

In order to realize the full scaling we suggest the following variation, which describes a highly idealized situation but does reveal the underlying reason for fast delivery. We start with an individual located at the origin 0 (the *source*) who has a message which he wants to deliver to an individual at location t (the *target*). The message holder has a random number of acquaintances randomly located in the plane according to an inhomogeneous Poisson process with density function $g(x) = \lambda|x|^{-\alpha}$, where $\alpha, \lambda > 0$. Each acquaintance, for example one at position z , has a random number of further acquaintances which can be described using an inhomogeneous Poisson process with density function $h(x) = \lambda|x - z|^{-\alpha}$, independent of the other Poisson processes. For any $\varepsilon > 0$, we say that an ε -*delivery* has taken place if the message has been forwarded to an individual within a distance ε of the target.

Some remarks are in order. First, we can now explain what we mean by the scale invariance mentioned earlier. To this end, let $V \subset \mathbb{R}^2$ be a bounded set, not containing the origin, over which the (Lebesgue) integral of g can be defined. The (random) number of acquaintances of 0 in V has a Poisson distribution with mean $\int_V g(x) dx$. A simple substitution shows that

$$\int_{aV} g(x) dx = a^{2-\alpha} \int_V g(x) dx$$

for any $a > 0$, where aV is the set $\{av : v \in V\}$. From this it follows that the choice $\alpha = 2$ is special, since in that case $\int_{aV} g(x) dx$ is independent of a . The interpretation of this is that for $\alpha = 2$ the model has no natural scale; changing the unit of length does not make any difference.

Secondly, the simplifications and natural interpretations of the idealization come at the price of introducing simplifications, and possibly singular behavior at the origin, that are certainly not present in the real world. Arbitrarily small scale lengths play no role in the real world, but they are necessary to understand the true nature of scale invariance. It would certainly be possible to complicate the model to better reflect real-world scenarios, but our point is that the main reason behind efficient routing would still be, simply, scale invariance, which has its simplest interpretation in our continuum setting.

We will show that for $\alpha < 2$ any routing algorithm using only local information (i.e. such that the current message holder in a given step of the algorithm knows only the location of its acquaintances and 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 $1/\varepsilon$. 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 $1/\varepsilon$ and d .

We would like to stress that our message is primarily a philosophical one, supported by mathematics which is elementary. As already mentioned, we could make the model more realistic, at the cost of less insight into the reasons underlying efficient routing.

1.3. Comparison with Kleinberg's result

Kleinberg's result that choosing $\alpha < 2$ (giving higher probability to longer links) increases the delivery time is more an artefact of his model that uniformly bounds the number of contacts each node has rather than it is a true effect of scaling. To see this, let us consider a modified (and more natural) version of his model that does not bound the number of contacts of each node. Accordingly, let us consider a model in which long-range contact is made between each pair of nodes of a square grid with probability proportional to $d^{-\alpha}$, where d is the distance, in units of grid edges, that needs to be traversed to connect the two nodes. Note that in this case the number of long-range contacts each node has 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 contacts. Clearly, the minimum delivery time of a message routed between two randomly selected nodes must be nondecreasing in α , because there are on average fewer long-range edges departing from each node as α increases. Hence, there appears to be nothing special about a scaling distribution with exponent $\alpha = 2$. What makes Kleinberg's model behave differently is that in his formulation the number of long-range contacts each node has is uniformly bounded in the system size, and for this reason $\alpha = 2$ turns out to be the best possible scaling exponent for the random graph obtained. Our continuum formulation shows that having a scale-free distribution 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 random-connection model considered.

2. Precise results and proofs

Consider the full plane \mathbb{R}^2 as our model of the real world. (We restrict to two dimensions for obvious reasons, but it is not hard to see that we could do everything in arbitrary dimension n . In that case, choosing $\alpha = n$ would give us scale invariance.)

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 inhomogeneous Poisson point process X with density function $h(y) = \lambda g(|z - y|)$, for some $\lambda > 0$. Let d be the Euclidean 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 by being sequentially passed along a chain of intermediate acquaintances. That is, the current message holder in a given step, u , knows

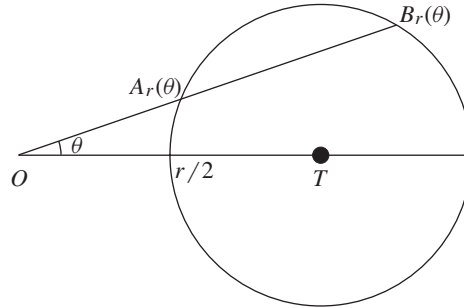


FIGURE 1: Decreasing the distance to the target by a factor of one-half.

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 $\varepsilon > 0$, 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 by at each step making a forwarding decision based on the rules of \mathcal{A} . Finally, let $\bar{\mathcal{A}}$ be the decentralized algorithm whereby at each step the message is forwarded 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 of the model, α , influences the ε -delivery time (over a distance d) of a decentralized algorithm as follows.*

Case 1: *For $\alpha = 2$ there exists a constant $c > 0$ such that, for any $\varepsilon > 0$ and $d > \varepsilon$, the expected ε -delivery time of the decentralized algorithm $\bar{\mathcal{A}}$ is at most $c(\log d + \log(1/\varepsilon))$.*

Case 2: *For $\alpha < 2$ there exists a constant $c(\alpha) > 0$ such that, for any $\varepsilon > 0$, the expected ε -delivery time of any decentralized algorithm \mathcal{A} is at least $c(\alpha)(1/\varepsilon)^{2-\alpha}$.*

Case 3: *For $\alpha > 2$ and any $\varepsilon > 0$ and $d > 1$, the expected ε -delivery time of any decentralized algorithm \mathcal{A} is at least cd^β , for any $\beta < (\alpha - 2)/(\alpha - 1)$ and some constant $c \equiv c(\alpha, \beta) > 0$.*

Essentially, the theorem states that for $\alpha = 2$ it is possible to approach the target on any distance scale in a logarithmic number of steps, steadily improving at each step. However, for $\alpha < 2$ a decentralized algorithm starts off quickly but then slows down as it approaches the target, and has trouble making the last small steps. Finally, for $\alpha > 2$ the situation is reversed, as the performance bottleneck is not near the target but is instead at large distances $d \gg \varepsilon$. Our case 3 corresponds to Kleinberg’s in [6, Theorem 3(b)]. It is interesting that he obtains the same exponent. On the other hand, our case 1, which corresponds to [6, Theorem 2], presents a bound slightly better than the squared logarithm bound obtained there.

Proof of case 1. We first compute the probability that at any step of the algorithm an intermediate node has a neighbor at a distance from the target that is less than half the distance between the target and the intermediate node, and show that this probability is positive and independent of the distance. We refer to Figure 1. Let $OT = r$ be the distance to the target. The (random) number of neighbors that are at a distance less than $r/2$ from the target T has a

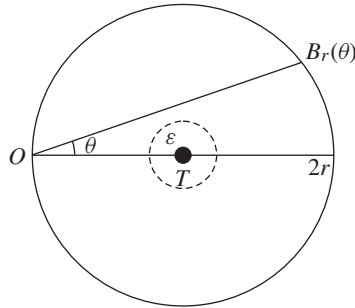


FIGURE 2: Getting closer to the target.

Poisson distribution with mean

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

which is positive and, since $\alpha = 2$, already known to be 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 circle depicted in Figure 1, i.e. closer to T than O by at least half the distance between T and O . Hence, algorithm \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 one-half 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 positive probability. We now show that a step that simply decreases the distance to the target has probability 1. The number of points that are less than a distance r from 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; hence, the number of such points is infinite with probability 1. It follows that the probability of *decreasing* the distance to the target has probability 1. Therefore, even when a step of the algorithm is not successful, it will not increase the distance to the target. It follows that at most a total number of n successful steps is needed to reach an ε -neighborhood of T , starting at a distance $d > \varepsilon$, where

$$\left(\frac{1}{2}\right)^n d < \varepsilon \iff n < \frac{\log d + \log(1/\varepsilon)}{\log 2}.$$

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

$$E(\varepsilon\text{-delivery time}) < \frac{\log d + \log(1/\varepsilon)}{\tau \log 2},$$

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 a distance $r \geq \varepsilon$ from the target, and start by computing the number of acquaintances of point O that are closer to the target. We refer to Figure 2. The number of such points has a

Poisson distribution, and since $\alpha < 2$ it has a finite mean

$$\begin{aligned} \mu(r, \alpha) &= \lambda \int_{-\pi/2}^{\pi/2} \int_0^{B_r(\theta)} g(r)r \, dr \, d\theta \\ &= \lambda \int_{-\pi/2}^{\pi/2} \int_0^{r B_1(\theta)} \frac{1}{r^{\alpha-1}} \, dr \, d\theta \\ &= \frac{\lambda}{2-\alpha} r^{2-\alpha} \int_{-\pi/2}^{\pi/2} B_1(\theta)^{2-\alpha} \, d\theta \\ &= c(\alpha)r^{2-\alpha}. \end{aligned}$$

Let an *improving* step of any decentralized algorithm be one that forwards the message to a neighbor that is closer to the target than is O . The above computation shows that, when the message is at a distance ε from the source, the probability of an improving step being made is bounded above by $c(\alpha)\varepsilon^{2-\alpha}$. When the distance to the target is greater than ε , the probability of entering the ε -neighborhood is easily seen to be smaller than this probability, since the density of the Poisson processes decreases with distance. Hence, at *any* step in the algorithm the probability of an ε -delivery is at most $c(\alpha)\varepsilon^{2-\alpha}$. It follows that the expected number of steps required to enter an ε -neighborhood of the target is at least

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

Proof of case 3. Consider the collection of acquaintances of a given individual, and denote by D the distance to the acquaintance furthest away. We find that

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

for some constant c . This quantity tends to 0 as $r \rightarrow \infty$, since $\alpha > 2$.

We next estimate the probability that, starting at a distance $d > 1$, 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 step of size at least $d^{1-\beta}$. According to the computation above, the probability that this happens is at most

$$\frac{c}{\alpha-2} d^\beta d^{(1-\beta)(2-\alpha)} = d^{2-\alpha-\beta+\alpha\beta}.$$

Writing X_d for the delivery time starting at distance d , it follows that

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

and, therefore, that

$$E(X_d) \geq d^\beta \left(1 - \frac{c}{\alpha-2} d^{2-\alpha-\beta+\alpha\beta} \right).$$

Whenever $2 - \alpha - \beta + \alpha\beta < 0$ or, equivalently,

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

this expression is, for some constant c , at least cd^β . The result now follows.

Acknowledgements

This work was partially supported by DARPA, grant number F33615-01-C-1895, at the University of California, Berkeley, and by the Ennio De Giorgi Center for Mathematical Research of the Scuola Normale Superiore of Pisa. The authors would like to thank the center for providing both this funding and an ideal setting where part of this research was done.

References

- [1] BOLLÓBAS, B. AND CHUNG, F. R. K. (1988). The diameter of a cycle plus a random matching. *SIAM J. Discrete Math.* **1**, 328–333.
- [2] BOLLÓBAS, B. AND RIORDAN, O. M. (2002). Mathematical results on scale-free random graphs. In *Handbook of Graphs and Networks*, eds S. Bornholdt and H. G. Schuster, Wiley-VCH, Berlin, pp. 1–34.
- [3] GUARE, J. (1990). *Six Degrees of Separation*. Vintage, New York.
- [4] KARINTHY, F. (1929). *Chains. Everything is Different*. Budapest.
- [5] KLEINBERG, J. M. (2000). Navigation in the small world. *Nature* **406**, 845.
- [6] KLEINBERG, J. M. (2000). The small-world phenomenon: an algorithmic perspective. In *Proc. 32nd Annual ACM Symp. Theory Comput.*, ACM, New York, pp. 163–170.
- [7] MEESTER, R. AND ROY, R. (1996). *Continuum Percolation*. Cambridge University Press.
- [8] MILGRAM, S. (1967). The small world problem. *Psychology Today* **2**, 60–67.
- [9] WATTS, D. J. AND STROGATZ, S. H. (1998). Collective dynamics of small-world networks. *Nature* **393**, 440–442.