Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Linear Algebra and its Applications xx (2006) xxx–xxx
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Greedy pathlengths and small world graphs

Desmond J. Higham

Department of Mathematics, University of Strathclyde, Glasgow G1 1XH, UK

Received 21 April 2005; accepted 3 January 2006

Submitted by M. Neumann

Abstract

We use matrix analysis to study a cycle plus random, uniform shortcuts—the classic small world model. For such graphs, in addition to the usual edge and vertex information there is an underlying metric that determines distance between vertices. The metric induces a natural greedy algorithm for navigating between vertices and we use this to define a pathlength. This pathlength definition, which is implicit in [J. Kleinberg, The small-world phenomenon: an algorithmic perspective, in: Proceedings of the 32nd ACM Symposium on Theory of Computing, 2000] is entirely appropriate in many message passing contexts. Using a Markov chain formulation, we set up a linear system to determine the expected greedy pathlengths and then use techniques from numerical analysis to find a continuum limit. This gives an asymptotically correct expression for the expected greedy pathlength in the limit of large network size: both the leading term and a sharp estimate of the remainder are produced. The results quantify how the greedy pathlength drops as the number of shortcuts is increased. Further, they allow us to measure the amount by which the greedy pathlength, which is based on local information, exceeds the traditional pathlength, which requires knowledge of the whole network. The analysis allows for either O(1) shortcuts per node or O(1) shortcuts per network. In both cases we find that the greedy algorithm fails to exploit fully the existence of short paths.

AMS classification: 68R10; 60J10; 65L05

Keywords: Continuum limit; Differential equation; Finite difference; Greedy algorithm; Mean hitting time; Markov chain; Small world phenomenon

E-mail address: djh@maths.strath.ac.uk

0024-3795/$ - see front matter © 2006 Published by Elsevier Inc.
doi:10.1016/j.laa.2006.01.001

1. Introduction

We consider a class of graphs where there is an underlying connectivity pattern on top of which extra links have been added at random. Such *partially random* graphs, which lie between the two classical areas of deterministic and random graph theory, have attracted interest in a number of application areas within computer science [1,4,5,10–12], the social sciences [19], bioinformatics [7,14,18] and chemistry [6]. From a theoretical perspective, fascinating results have been proved, or conjectured through simulations, concerning the decrease of the expected pathlength as a function of the amount of disorder added, and the ability of navigational algorithms to find the short paths [2,12,15,16,21].

Much of the work in this area makes reference to the small world experiments of the psychologist Stanley Milgram. In [13] Milgram used the US postal service to compute short paths in a large social acquaintance graph: here the nodes in the graph are people and two nodes are connected if the two people know each other on a first name basis. In the experiment, a source person in Nebraska was given basic information about a target person in Boston. The source was asked to get the letter to the target as efficiently as possible. In the (likely) event that the source did not know the target, he/she was to send the letter to a suitable first name basis acquaintance, with the process continuing iteratively until the target was reached. Hence, using the information about the target, each person in the chain chose the next recipient with the aim of minimizing the overall number of steps. Milgram found that successful chains had a typical length of around six. Given that the social acquaintance graph is large, sparse, and highly clustered (if A knows B and B knows C, then C is very likely to know A), it seems counter-intuitive that the typical pathlength is so short. This raises a key question; what kind of highly clustered networks permit a short typical pathlength? Watts and Strogatz [21] showed via simulation that such small world networks can be constructed by randomly re-wiring a small percentage of links in a deterministic lattice. Newman, Moore and Watts [16] looked at a variation of the Watts–Strogatz model that is more amenable to analysis. Here, random links are superimposed rather than existing links being rewired. They gave a semi-heuristic mean-field derivation of an expression for the expected pathlength of a large network in the limit of either a large or small number of shortcuts. Barbour and Reinert [2] subsequently gave a fully rigorous treatment.

Kleinberg [12], see also [11], recognized that in addition to the existence of short paths in Milgram's experiment there is a second surprising discovery: the participants, using only local knowledge about their own acquaintances, were able to *construct* short paths. This leads to a second question: what combination of small world network and navigation algorithm leads to the computation of short paths? By analogy with Milgram's experiment, Kleinberg looked at the task of transmitting a message between a typical pair of nodes using a decentralized algorithm, that is, an algorithm where the current message holder knows

(i) the underlying lattice connectivity structure,
(ii) the location of the target,
(iii) its own random edges,
(iv) the random edges of each node that has previously come into contact with the message.

The measure of success of an algorithm was the expected delivery time to a randomly chosen target. Item (iv) above was used only in the derivation of negative results. A positive result was proved for the following greedy algorithm that satisfies (i)–(iii).

Algorithm 1.1. The current message holder passes the message on to a contact that is closest (in lattice distance) to the target.

This motivates the following definition.

Definition 1.2. Given a graph where, independently of the connectivity pattern, there is a measure of distance between nodes, the *greedy pathlength* between nodes i and j is the number of steps taken by the greedy algorithm 1.1 to pass a message from i to j.

With regard to Milgram's experiment, the lattice distance between nodes can be interpreted as an idealization of geographic distance: choose the person you know who lives closest to the target (although, of course, other factors such as the target's age and occupation may have had some influence). A similar greedy algorithm that uses "social distance" on a sophisticated social network model has also been studied in [20]. More generally, in many message-passing scenarios it is reasonable to assume that, independently of the precise connectivity structure, individual nodes may use some inherent metric to guide their choice. Indeed, the greedy pathlength can be regarded as the *free packet delay* for a simple routing algorithm; see, for example, [5].

The model that we consider in this work has an underlying cycle: N nodes, labeled $0, 1, \ldots, N-1$, are arranged in a ring structure, so i and j are connected when $i = j \pm 1 \bmod N$, and the distance between them is $\min\{|i-j|, N-|i-j|\}$. For each node an additional random link is added with probability p; that is, for each node, we flip a (biased) coin to determine whether to add an extra link. If an extra link is to be added, its endpoint is picked uniformly from the entire set of nodes. This model was analyzed in [16] (we have $k = 1$) and is a minor variation of the original small world network of [21]. Our aim is to study the greedy pathlength for this type of network.

We consider two regimes that add different amounts of disorder to the cycle:

$$p = \frac{K}{N}, \quad \text{for fixed } K > 0, \quad \text{as } N \to \infty \tag{1.1}$$

and

$$p \text{ is constant with } 0 < p \leqslant 1, \quad \text{as } N \to \infty. \tag{1.2}$$

In expectation, (1.1) is the case where $O(1)$ extra links are added to the network, whereas $O(N)$ extra links are added in (1.2). Both extremes have been discussed in the literature and are of practical interest.

We mention that some related work for a cycle plus random edges appeared in [3]. In that work, high probability upper and lower bounds on the maximum pathlength are derived for the case where a random matching is added to a cycle: letting $\lfloor \cdot \rfloor$ denote the integer part, exactly $\lfloor N/2 \rfloor$ extra links are inserted at random in such a way that every node has degree 3 (except, of course, that one node must miss out when N is odd).

In the $p =$ constant regime (1.2), our work relates closely to that in [12]. Kleinberg considers a 2-dimensional lattice where each node, u, is connected to its nearest neighbors up to fixed lattice distance and, in addition, has "random" directed edges to q other nodes, for some fixed q. These extra links are constructed from q independent trials where the probability of connecting node u to node v is inversely proportional to the rth power of the lattice distance between u and v. In the uniform case, $r = 0$, Kleinberg showed that any decentralized algorithm has an expected delivery time bounded below by a non-zero multiple of $N^{\frac{2}{3}}$, and hence exponential in the expected pathlength. He went on to show that this mismatch occurs for any $r \neq 2$, but for $r = 2$ the greedy algorithm has an expected delivery time bounded above by a multiple of $(\log N)^2$. Hence, Kleinberg's results allow a parametrized "range-dependent" distribution for $O(N)$ shortcuts and they distinguish a critical inverse-square distribution where the greedy algorithm works best. Our results, which apply to a different model and are restricted to the uniform ($r = 0$) setting, cover the cases of both $O(1)$ and $O(N)$ shortcuts and give precise asymptotic expressions for the leading terms, plus sharp remainders.

In the next section we show how the greedy pathlength can be analyzed through a Markov chain formulation. In §3 we state and interpret our results, which are proved in §4. General conclusions are given in §5.

2. Markov chain formulation

The expected value of the greedy pathlength between a pair of nodes in the ring network described above can be calculated using a Markov chain approach. Without loss of generality, we consider starting at node j and navigating towards node 0 using the greedy algorithm. This induces a Markov chain on the distance to node 0. Our state space is labeled $0, 1, 2, \dots, M-1$, where $M := \lfloor N/2 \rfloor + 1$. State 0 corresponds to node 0, state 1 corresponds to nodes $\{1, N-1\}$ and, generally, state i corresponds to the two nodes that are a distance i from 1. When N is even, state $M-1$ corresponds to the single node that is a distance $M-1$ from 0. For example, for $N = 12$, the states correspond to nodes

$$\{0\}, \{1, 11\}, \{2, 10\}, \{3, 9\}, \{4, 8\}, \{5, 7\}, \{6\}$$

and for $N = 13$ they correspond to

$$\{0\}, \{1, 12\}, \{2, 11\}, \{3, 10\}, \{4, 9\}, \{5, 8\}, \{6, 7\}.$$

Now, given node i, the probability that it has an extra link to any particular node j is given by

$$1 - \text{probablity of no extra links.}$$

This is $1 - (1 - p/N)(1 - p/N)$, which simplifies to $2p/N - p^2/N^2$. (This follows because independent coin flips take place for both node i and node j.) We will let $\hat{p} := 2p - p^2/N$, so that the appropriate probability may be written conveniently as $\hat{p}/N$. In considering a step of the greedy algorithm to progress towards state 0, note that

- if an extra link exists it is given by an edge that is equally likely to meet up with any node in the ring,
- the greedy algorithm will use an extra link only if it decreases the distance to node 0 by more than one, otherwise it will use the nearest neighbor edge to decrease the distance by one.

Putting this together we find that if the Markov chain at time level n has the value $X_n = i$ for some $i \geqslant 2$ then

$$X_{n+1} = \begin{cases} i-1, & \text{with probability } 1 - (2i-3)\hat{p}/N, \\ j, \quad \text{for } 1 \leqslant j \leqslant i-2, & \text{with probability } 2\hat{p}/N, \\ 0, & \text{with probability } \hat{p}/N. \end{cases}$$

Here, the event $X_{n+1} = j$ for $1 \leqslant j \leqslant i-2$ arises if there is a shortcut to either of the two nodes that are a distance j from the target node 0; hence the appropriate probability is $2\hat{p}/N$. To complete the specification, we note that if $X_n = 1$ or $X_n = 0$ then $X_{n+1} = 0$ with probability 1. The transition matrix $P \in \mathbb{R}^{M \times M}$, which has general entry $p_{ij} := \mathbb{P}(X_{n+1} = j, \text{ given } X_n = i)$, thus has the lower triangular form

$$P = \begin{bmatrix} 1 & & & & & & \\ 1 & 0 & & & & & \\ \frac{\hat{p}}{N} & 1-\frac{\hat{p}}{N} & 0 & & & & \\ \frac{\hat{p}}{N} & \frac{2\hat{p}}{N} & 1-\frac{3\hat{p}}{N} & 0 & & & \\ \frac{\hat{p}}{N} & \frac{2\hat{p}}{N} & \frac{2\hat{p}}{N} & 1-\frac{5\hat{p}}{N} & 0 & & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \\ \frac{\hat{p}}{N} & \frac{2\hat{p}}{N} & \frac{2\hat{p}}{N} & \dots & \frac{2\hat{p}}{N} & 1-\frac{(2M-5)\hat{p}}{N} & 0 \end{bmatrix}. \tag{2.1}$$

Now let h_j be the hitting time for state 0, starting from state j; that is $h_j(\omega) := \inf\{n \geqslant 0 : X_n(\omega) = 0, \text{ given } X_0 = j\}$, and let z_j be the corresponding mean hitting time for state j,

$$z_j := \mathbb{E}(h^j), \quad j = 0, 1, \dots, M-1. \tag{2.2}$$

In general, z_j is the expected value of the greedy pathlength between nodes $\{j, N-j\}$ and 0. We are interested in the average over all nodes $i \in \{0, 1, \dots, N-1\}$ of the greedy pathlength between node i and 0; that is

$$z_{\text{ave}} := \begin{cases} \frac{1}{N}(z_0 + 2(z_1 + z_2 + \cdots + z_{M-1})) & N \text{ odd}, \\ \frac{1}{N}(z_0 + 2(z_1 + z_2 + \cdots + z_{M-2}) + z_{M-1}) & N \text{ even}. \end{cases} \tag{2.3}$$

Note that z_{ave} is equivalent to the expected greedy pathlength between a pair of nodes chosen uniformly at random.

Clearly $z_0 = 0$. A classical result, see for example [17, Theorem 1.3.5], shows that the mean hitting times $\{z_1, z_2, \dots, z_{M-1}\}$ satisfy a linear system that involves the entries in the transition matrix. In our case, the system is

$$\begin{bmatrix} 1 & & & & & \\ \frac{\hat{p}}{N}-1 & 1 & & & & \\ -\frac{2\hat{p}}{N} & \frac{3\hat{p}}{N}-1 & 1 & & & \\ -\frac{2\hat{p}}{N} & -\frac{2\hat{p}}{N} & \frac{5\tilde{p}}{N}-1 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \\ -\frac{2\hat{p}}{N} & -\frac{2\hat{p}}{N} & \dots & -\frac{2\hat{p}}{N} & \frac{(2M-5)\hat{p}}{N}-1 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ \vdots \\ \vdots \\ z_{M-1} \end{bmatrix} = \mathbf{e}, \tag{2.4}$$

where $\mathbf{e} := [1, 1, \dots, 1]^{\mathrm{T}} \in \mathbb{R}^{M-1}$. This system may be re-written in the form

$$\left(-\frac{2\hat{p}}{N} \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & & \\ 1 & \vdots & \ddots & \ddots & \\ 1 & 1 & \dots & \dots & 1 \end{bmatrix} + \left(1 + \frac{2\hat{p}}{N}\right) \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix} \right.$$

$$+\frac{\hat{p}}{N}\begin{bmatrix} 0 & & & & \\ 5 & 0 & & & \\ & 7 & 0 & & \\ & & \ddots & \ddots & \\ & & & 2M-1 & 0 \end{bmatrix}\Bigg)\Bigg)\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ \vdots \\ z_{M-1} \end{bmatrix} = \mathbf{e}. \tag{2.5}$$

The essence of our analysis is to observe that the system (2.5) has the form of a finite difference method for computing discrete approximations $z_j \approx z(x_j)$, where $x_j = j\Delta x$ with $\Delta x := 1/(M-1)$ and $z(x)$ is some continuous function. The first matrix in (2.5) represents an integral operator, the second a first derivative operator (assuming $z(0) = 0$) and the third a linear scaling. Overall, the putative continuum limit function $z(x)$ satisfies

$$-\frac{4p}{N}\frac{1}{\Delta x}\int_0^x z(y)\,\mathrm{d}y + \left(1 + \frac{4p}{N}\right)\Delta x z'(x) + \frac{2p}{N}\left(\frac{2x}{\Delta x} + 1\right)z(x) = 1, \quad z(0) = 0 \tag{2.6}$$

and we may reasonably hope that $\int_0^1 z(y)\,\mathrm{d}y$ is a good approximation to z_{ave} in (2.3).

From this point of view obtaining asymptotically valid expressions for z_j and z_{ave} reduces to a convergence analysis for a numerical method applied to an integro-differential equation. We note that Eq. (2.6) itself depends upon Δx, in contrast to the usual situation in numerical analysis where a method is applied to a fixed problem. Further, when written as a second order initial value ordinary differential equation (ODE), in the regime (1.2) Eq. (2.6) has a Lipschitz constant that is unbounded as $\Delta x \to 0$, which rules out the traditional approach to establishing convergence. However, the equation does have a sufficiently simple structure that a customized convergence theory can be developed. This theory treats convergence in a relative, rather than absolute, sense; the solution grows with N, but the finite difference error remains $\mathrm{O}(1)$. In the next section we quote the final results. Proofs are given in §4.

3. Results

Theorem 3.1. *In the regime* (1.1) *the mean hitting time* z_j *in* (2.2) *satisfies*

$$z_j = N\frac{\sqrt{\pi}}{2\sqrt{2K}}\operatorname{erf}\left(\frac{j\sqrt{2K}}{N}\right) + \mathrm{O}(1) \tag{3.1}$$

and the average mean hitting time z_{ave} *in* (2.3) *satisfies*

$$z_{\mathrm{ave}} = \frac{N}{2K}\left(\frac{\sqrt{2K\pi}}{2}\operatorname{erf}\left(\frac{\sqrt{2K}}{2}\right) + \mathrm{e}^{-\frac{1}{2}K} - 1\right) + \mathrm{O}(1). \tag{3.2}$$

Here, $\operatorname{erf}(y) := \frac{2}{\sqrt{\pi}}\int_0^y \mathrm{e}^{-t^2}\,\mathrm{d}t$ *is the error function.*

Proof. See §4. □

Theorem 3.2. *In the regime* (1.2) *the mean hitting time* z_j *in* (2.2) *satisfies*

$$z_j = N^{\frac{1}{2}}\frac{\sqrt{\pi}}{2\sqrt{2p}}\operatorname{erf}\left(N^{-\frac{1}{2}}j\sqrt{2p}\right) + \mathrm{O}((\log N)^2) \tag{3.3}$$

and the average mean hitting time z_{ave} *in* (2.3) *satisfies*

$$z_{\text{ave}} = N^{\frac{1}{2}} \frac{\sqrt{\pi}}{2\sqrt{2p}} + \mathrm{O}((\log N)^2). \tag{3.4}$$

Proof. See §4. □

We focus first on Theorem 3.1. Here we are adding a fixed number, K, of shortcuts to the cycle, on average. This does not alter the power of N that governs the asymptotic behavior of the expected pathlength—it remains $\mathrm{O}(N)$—but it does affect the constant factor in the leading term. From (3.2), the factor by which the shortcuts reduce the greedy pathlength is given by

$$\frac{z_{\text{ave}} \text{ with } p = K/N}{z_{\text{ave}} \text{ with } p = 0} = \frac{2}{K}\left(\frac{\sqrt{2K\pi}}{2}\operatorname{erf}\left(\frac{\sqrt{2K}}{2}\right) + \mathrm{e}^{-\frac{1}{2}K} - 1\right) + \mathrm{O}(N^{-1}). \tag{3.5}$$

The solid line in the upper plot of Fig. 3.1 shows the leading term in this expression as a function of K. On the same picture, the circles denote the corresponding expected value for the traditional pathlength (that is, the length of the shortest path between a randomly selected a pair of nodes). Because an analytical formula for this quantity is not known, data was computed via simulation. We fixed $N = 3000$ and used $p = 10^{\alpha}$, with 40 equally spaced α values between -5 and 0. For each p we generated 500 instances of a random graph and averaged the pathlength from nodes j to 0 for $0 \leqslant j \leqslant N - 1$. The results agree, to visual accuracy, with those in [16]. For the purpose of comparison, we have also plotted the mean-field approximation to the expected traditional pathlength from [16]. We note that this approximation is not claimed to be accurate in this $\mathrm{O}(1)$ shortcut regime; as we can see it tends to underestimate the true value. The lower picture in Fig. 3.1 repeats the same data with a log scaling of the x-axis. This emphasizes the region $1 \leqslant K \leqslant 100$. Overall we see a striking discrepancy between the two pathlength measures for K larger than about 10. As reported in [16], it requires an average of around 3.5 shortcuts to reduce the traditional pathlength by a factor of two. (That is, the circles pass through height $\frac{1}{2}$ in Fig. 3.1 at $K = 3.5$.) In contrast, it takes around 16 shortcuts per network to reduce the greedy pathlength by a factor of two. With just 2 shortcuts, the expected greedy pathlength is around 16% bigger than the traditional pathlength, and with 10 shortcuts it is around 70% bigger. Hence, even when the shortcuts are sparse, there is a significant difference between taking a shortcut whenever the chance arises and taking a shortcut only when it is globally optimal to do so.

In the regime (1.2), with $\mathrm{O}(N)$ shortcuts added, Theorem 3.2 shows that the greedy pathlength behaves like a nonzero multiple of $N^{\frac{1}{2}}$. In this case, the analysis in [2] shows that the expected traditional pathlength between a pair of randomly chosen nodes behaves like a polynomial in $\log N$. Hence, on average, the greedy algorithm is exponentially worse than a global breadth first search in this regime.

4. Proofs

The following subsections give proofs of Theorems 3.2 and 3.1.

We remark that numerical tests indicate that the $\mathrm{O}(1)$ second term in (3.2) is sharp, in the sense that a non-zero, constant remainder was observed. It is, of course, difficult to distinguish numerically between $(\log N)^2$ and a constant, but we suspect that (3.4) remains true with the $\mathrm{O}((\log N)^2)$ second term replaced by $\mathrm{O}(1)$.

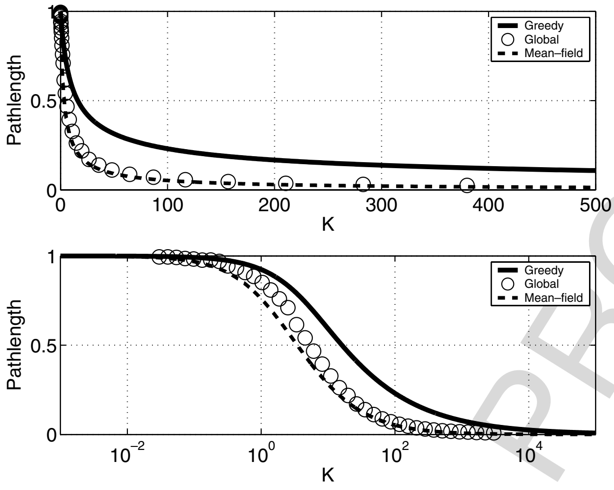

Fig. 3.1. Pathlength reduction ratios for $p = K/N$ regime (1.1). Solid curve: expected greedy pathlength (3.5). Circles: traditional pathlength. Dashed curve: mean-field approximation to traditional pathlength from [16]. Lower figure uses log scaling on x-axis.

4.1. *Proof of Theorem 3.2*

It is convenient to let $\tilde{p} = 2p$ and consider the system (2.4) with $\hat{p}$ replaced by $\tilde{p}$; that is

$$\begin{bmatrix} 1 & & & & & \\ \frac{\tilde{p}}{N} - 1 & 1 & & & & \\ -\frac{2\tilde{p}}{N} & \frac{3\tilde{p}}{N} - 1 & 1 & & & \\ -\frac{2\tilde{p}}{N} & -\frac{2\tilde{p}}{N} & \frac{5\tilde{p}}{N} - 1 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \\ -\frac{2\tilde{p}}{N} & -\frac{2\tilde{p}}{N} & \cdots & -\frac{2\tilde{p}}{N} & \frac{(2M-5)\tilde{p}}{N} - 1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \vdots \\ \vdots \\ \vdots \\ \tilde{z}_{M-1} \end{bmatrix} = \mathbf{e}. \tag{4.1}$$

We show later that, since $\hat{p} = \tilde{p} - p^2/N$, (4.1) and (2.4) have solutions that are sufficiently close.

We let $z(x)$ denote the solution to the ODE

$$z''(x) + \frac{1}{2}\tilde{p}(Nx + 1)z'(x) = 0, \quad z(0) = 0, \quad z'(0) = \frac{N}{2}. \tag{4.2}$$

Note that this equation can be derived by differentiating (2.6), setting $N\Delta x = 2$ and neglecting small terms. For convenience our notation does not reflect the dependence of $z(x)$ upon N, but in the subsequent analysis it is crucial to take account of the fact that $z(x)$, and its derivatives, grow with N.

Eq. (4.2) can be solved via an integrating factor to yield the following expressions:

$$z'(x) = \frac{N}{2}\mathrm{e}^{-\frac{N\tilde{p}x^2}{4} - \frac{\tilde{p}x}{2}}, \tag{4.3}$$

$$z(x) = \sqrt{\frac{N}{\tilde{p}}}\mathrm{e}^{\frac{\tilde{p}}{4N}}\frac{\sqrt{\pi}}{2}\mathrm{erf}\left(\frac{\sqrt{N\tilde{p}}}{2}x + \frac{1}{2}\sqrt{\frac{\tilde{p}}{N}}\right) - \sqrt{\frac{N}{\tilde{p}}}\mathrm{e}^{\frac{\tilde{p}}{4N}}\frac{\sqrt{\pi}}{2}\mathrm{erf}\left(\frac{1}{2}\sqrt{\frac{\tilde{p}}{N}}\right), \tag{4.4}$$

$$\int_0^1 z(x)\,\mathrm{d}x = \frac{1}{\tilde{p}}\mathrm{e}^{-\frac{N\tilde{p}}{4}-\frac{\tilde{p}}{2}} - \frac{1}{\tilde{p}} + \frac{(N+1)}{\sqrt{N\tilde{p}}}\frac{\sqrt{\pi}}{2} \times \mathrm{e}^{\frac{\tilde{p}}{4N}}\left[\operatorname{erf}\left(\frac{\sqrt{N\tilde{p}}}{2}+\frac{1}{2}\sqrt{\frac{\tilde{p}}{N}}\right) - \operatorname{erf}\left(\frac{1}{2}\sqrt{\frac{\tilde{p}}{N}}\right)\right]. \tag{4.5}$$

We also note for later use that

$$\max_{[0,1]} |z''(x)| = \mathrm{O}(N^{\frac{3}{2}}), \tag{4.6}$$

$$\max_{[0,1]} |z'''(x)| = \mathrm{O}(N^{2}), \tag{4.7}$$

$$\max_{[0,1]} |z^{\mathrm{iv}}(x)| = \mathrm{O}(N^{\frac{5}{2}}). \tag{4.8}$$

Now we make a precise connection between the linear system (2.4) and a numerical method applied to (4.2).

Lemma 4.1. *Let the sequences* $\{y_j^{[1]}\}_{j=0}^{M-1}$ *and* $\{y_j^{[2]}\}_{j=0}^{M-1}$ *be defined by* $\Delta x = 2/N$ *and*

$$y_j^{[1]} = y_{j-1}^{[1]} + \Delta x y_{j-1}^{[2]}, \tag{4.9}$$

$$y_j^{[2]} = y_{j-1}^{[2]} - \Delta x \frac{1}{2}\tilde{p}(N(j-1)\Delta x + 1) y_{j-1}^{[2]} \tag{4.10}$$

for $j \geqslant 1$, *with* $y_0^{[1]} = 0$ *and* $y_0^{[2]} = N/2$. *Then, for* $\tilde{z}_j$ *in* (4.1), *we have*

$$y_j^{[1]} = \tilde{z}_j \quad \text{for } 0 \leqslant j \leqslant M-1.$$

[*Note that* (4.9) *and* (4.10) *represents Euler's method, see, for example,* [8], *with stepsize* Δx *applied to the ODE* (4.2) *written as a first order system.*]

Proof. By construction, $y_0^{[1]} = 0 = \tilde{z}_0$ and $y_1^{[1]} = 1 = \tilde{z}_1$. Generally, substituting $y_{j-1}^{[2]} = \left(y_j^{[1]} - y_{j-1}^{[1]}\right)/\Delta x$ from (4.9) into (4.10) gives

$$y_{j+1}^{[1]} - 2y_j^{[1]} + y_{j-1}^{[1]} + \frac{\tilde{p}}{N}\left(y_j^{[1]} - y_{j-1}^{[1]}\right)(2j-1) = 0$$

and subtracting row j of (4.1) from row $j+1$ gives the same recurrence for $\tilde{z}_j$. □

Now, for $z(x)$ in (4.2), we let

$$e_j^{[1]} := z(x_j) - y_j^{[1]} \quad \text{and} \quad e_j^{[2]} := z'(x_j) - y_j^{[2]},$$

where $x_j := j\Delta x$ with $\Delta x = 2/N$. Here, $e_j^{[1]}$ and $e_j^{[2]}$ represent the errors in the Euler approximations to $z(x_j)$ and $z\prime(x_j)$, respectively.

Lemma 4.2. *The errors satisfy*

$$e_j^{[1]} = e_{j-1}^{[1]} + \Delta x e_{j-1}^{[2]} + \frac{1}{2}\Delta x^2 z''(\beta_j), \tag{4.11}$$

$$e_j^{[2]} = R_{j-1} e_{j-1}^{[2]} + \frac{1}{2}\Delta x^2 z'''(\gamma_j) \tag{4.12}$$

LAA 8797
23/1/2006 DISK used
ARTICLE IN PRESS
No. of pages: 14, Model 1+

for some $\beta_j, \gamma_j \in [x_{j-1}, x_j]$, *where*

$$R_{j-1} := 1 - \frac{\tilde{p}}{N}(2j-1).$$

Proof. Taking Taylor series expansions, using (4.2), we have

$$z(x_j) = z(x_{j-1}) + \Delta x z'(x_{j-1}) + \frac{1}{2}\Delta x^2 z''(\beta_j), \tag{4.13}$$

$$z'(x_j) = z'(x_{j-1}) - \Delta x \frac{1}{2}\tilde{p}(N x_{j-1} + 1) z'(x_{j-1}) + \frac{1}{2}\Delta x^2 z'''(\gamma_j). \tag{4.14}$$

Subtracting (4.9) and (4.10) from (4.13) and (4.14), respectively, gives the result. □

Our task is to show that $\max_{0 \leqslant j \leqslant M-1} |e_j^{[1]}| = O(1)$. Lemma 4.2 gives recurrences satisfied by the errors, and we see that (4.12) involves only $e_j^{[2]}$. Our approach is therefore to use (4.12) to get an estimate for $\sum_{j=0}^{M-1} e_j^{[2]}$ that can be inserted into (4.11). The result relies on subtle cancellation and we found it necessary to retain asymptotic estimates, rather than bounds, as far as possible.

To motivate the subsequent analysis, note from (4.3) that $z'(x)$, $z''(x)$ and $z'''(x)$ have a factor $e^{-\frac{N\tilde{p}x^2}{4}}$. It follows that these derivatives become negligible as x increases beyond $O\left(N^{-\frac{1}{2}}\right)$. To exploit this effect we define

$$x^{\star} := \frac{4}{\sqrt{\tilde{p}}} N^{-\frac{1}{2}} \sqrt{\log N} \quad \text{and} \quad n^{\star} := \left\lfloor \frac{x^{\star}}{\Delta x} \right\rfloor. \tag{4.15}$$

It follows directly that

$$\max_{[x^{\star},1]} \{|z''(x)|, |z'''(x)|\} = O(N^{-1}). \tag{4.16}$$

Lemma 4.3. *For* $0 \leqslant j \leqslant n^{\star}$,

$$e_j^{[2]} = e^{-\frac{N x_j^2 \tilde{p}}{4}} \sum_{k=1}^{j} e^{\frac{N x_k^2 \tilde{p}}{4}} \frac{1}{2}\Delta x^2 z'''(x_k) + O((\log N)^2).$$

Proof. Since $e_0^{[2]} = 0$, it follows from (4.12) that

$$e_j^{[2]} = \sum_{k=1}^{j} \widehat{R}_k^{(j)} \frac{1}{2}\Delta x^2 z'''(\gamma_k), \tag{4.17}$$

where

$$\widehat{R}_k^{(j)} := \prod_{i=k}^{j-1} R_i,$$

with the empty product regarded as unity. Now, for $0 \leqslant i \leqslant n^{\star}$, we have

$$\log R_i = -\frac{2i\tilde{p}}{N} + O(N^{-1}\log N)$$

and hence

$$\log \widehat{R}_k^{(j)} = \sum_{i=k}^{j-1} \log R_i = -\frac{2\tilde{p}}{N} \sum_{i=k}^{j-1} i + \mathrm{O}(n^{\star} N^{-1} \log N),$$

which simplifies to

$$\log \widehat{R}_k^{(j)} = -\frac{\tilde{p}}{N}(j^2 - k^2) + \mathrm{O}\left(N^{-\frac{1}{2}} (\log N)^{\frac{3}{2}}\right).$$

Hence,

$$\widehat{R}_k^{(j)} = \mathrm{e}^{-\frac{\tilde{p}}{N}(j^2-k^2)} \left(1 + \mathrm{O}\left(N^{-\frac{1}{2}} (\log N)^{\frac{3}{2}}\right)\right). \tag{4.18}$$

Now, from (4.8), we have $z'''(\gamma_k) = z'''(x_k) + \mathrm{O}\big(N^{\frac{3}{2}}\big)$. Using this, along with (4.7) and (4.18), in (4.17) leads to the required result. □

Lemma 4.4. *For all* $0 \leqslant j \leqslant M-1$,

$$\Delta x \sum_{k=0}^{j} e_k^{[2]} = \mathrm{O}((\log N)^2).$$

Proof. Using (4.3) in (4.2) gives an expression for $z''(x)$. Differentiating this and inserting the expression for $z'''(x)$ into the expansion for $e_j^{[2]}$ in Lemma 4.3, we find that for $0 \leqslant j \leqslant n^{\star}$

$$e_j^{[2]} = \mathrm{e}^{-\frac{N x_j^2 \tilde{p}}{4}} \frac{1}{N} \sum_{k=1}^{j} \mathrm{e}^{-\frac{\tilde{p} x_k}{2}} \left(\frac{\tilde{p}^2}{4}(N x_k + 1)^2 - \frac{N\tilde{p}}{2}\right) + \mathrm{O}((\log N)^2).$$

After some asymptotic expansion and manipulation, this simplifies to

$$e_j^{[2]} = \frac{N\tilde{p}}{4} \psi(x_j) + \mathrm{O}((\log N)^2), \tag{4.19}$$

where

$$\psi(s) := \mathrm{e}^{-\frac{N s^2 \tilde{p}}{4}} s \left(\frac{N\tilde{p} s^2}{6} - 1\right).$$

Since $\max_{[0,1]} |\psi'(s)| = \mathrm{O}(1)$, we have for $0 \leqslant x_j \leqslant x^{\star}$,

$$\int_0^{x_j} \psi(s)\, \mathrm{d}s = \Delta x \sum_{k=1}^{j} \psi(x_k) + \mathrm{O}\left(N^{-\frac{3}{2}} \sqrt{\log N}\right). \tag{4.20}$$

Using

$$\int_0^{x_j} \psi(s)\, \mathrm{d}s = \mathrm{e}^{-\frac{N x_j^2 \tilde{p}}{4}} \left(-\frac{x_j^2}{3} + \frac{2}{3N\tilde{p}}\right) - \frac{2}{3N\tilde{p}}$$

in (4.20) gives

$$\Delta x \sum_{k=1}^{j} \psi(x_k) = -\frac{x_j^2}{3} \mathrm{e}^{-\frac{N x_j^2 \tilde{p}}{4}} + \mathrm{O}(N^{-1}).$$

Hence, from (4.19),

$$\Delta x \sum_{k=1}^{j} e_j^{[2]} = -\frac{N\tilde{p}}{12} x_j^2 \mathrm{e}^{-\frac{N x_j^2 \tilde{p}}{4}} + \mathrm{O}(1) = \mathrm{O}(1)$$

for $0 \leqslant j \leqslant n^\star$.

Now, from (4.19), $e_{n^\star}^{[2]} = \mathrm{O}((\log N)^2)$. So, from (4.12) and (4.16), for $n^\star < k \leqslant M-1$

$$|e_k^{[2]}| \leqslant |e_{k-1}^{[2]}| + \frac{1}{2}\Delta x^2 |z'''(\gamma_k)| \leqslant \cdots \leqslant \left|e_{n^\star}^{[2]}\right| + \frac{1}{2}\Delta x^2 \sum_{r=n^\star+1}^{k} |z'''(\gamma_k)| = \mathrm{O}((\log N)^2)$$

and hence $\Delta x \sum_{k=n^\star}^{j} |e_k^{[2]}| = \mathrm{O}((\log N)^2)$ for $n^\star < j \leqslant M-1$, which completes the result. □

Lemma 4.5

$$\max_{0 \leqslant j \leqslant M-1} |e_j^{[1]}| = \mathrm{O}((\log N)^2).$$

Proof. From (4.11) and Lemma 4.4 we have, using $e_0^{[1]} = 0$,

$$e_j^{[1]} = \Delta x \sum_{k=1}^{j-1} e_k^{[2]} + \frac{1}{2}\Delta x^2 \sum_{k=1}^{j} z''(\beta_k) = \mathrm{O}((\log N)^2) + \frac{1}{2}\Delta x^2 \sum_{k=1}^{j} z''(\beta_k).$$

Further, (4.7) implies that

$$\Delta x \sum_{k=1}^{j} z''(\beta_k) = \int_0^1 z''(x)\,\mathrm{d}x + \mathrm{O}(N) = z'(1) - z'(0) + \mathrm{O}(N) = \mathrm{O}(N),$$

and the result follows. □

Lemmas 4.1 and 4.5 show that $\tilde{z}_j = z(x_j) + \mathrm{O}((\log N)^2)$ for all $0 \leqslant j \leqslant M-1$. Inserting the expression (4.4) for $z(x_j)$ and simplifying gives

$$\tilde{z}_j = N^{\frac{1}{2}} \frac{\sqrt{\pi}}{2\sqrt{2\tilde{p}}} \operatorname{erf}\left(N^{-\frac{1}{2}} j \sqrt{2\tilde{p}}\right) + \mathrm{O}((\log N)^2). \tag{4.21}$$

Now, we may write (4.1) and (2.4) as $M\tilde{\mathbf{z}} = \mathbf{e}$ and $(M+E)\mathbf{z} = \mathbf{e}$, respectively, where

- $\|E\|_\infty = \mathrm{O}(N^{-1})$,
- and, since $\tilde{\mathbf{z}} = M^{-1}\mathbf{e}$ and M is an M-matrix, $\|M^{-1}\|_\infty = \|\tilde{\mathbf{z}}\|_\infty = \mathrm{O}\left(N^{\frac{1}{2}}\right)$.

So,

$$M(\tilde{\mathbf{z}} - \mathbf{z}) = -E\mathbf{z} = E(\tilde{\mathbf{z}} - \mathbf{z}) - E\tilde{\mathbf{z}}$$

and hence

$$\|\tilde{\mathbf{z}} - \mathbf{z}\|_\infty \leqslant \|M^{-1}\|_\infty \|E\|_\infty \|\tilde{\mathbf{z}} - \mathbf{z}\|_\infty + \|E\|_\infty \|\tilde{\mathbf{z}}\|_\infty,$$

so that

$$\|\tilde{\mathbf{z}} - \mathbf{z}\|_\infty \leqslant \frac{\|E\|_\infty \|\tilde{\mathbf{z}}\|_\infty}{1 - \|M^{-1}\|_\infty \|E\|_\infty} = \mathrm{O}\left(N^{-\frac{1}{2}}\right).$$

Hence, (4.21) also holds for z_j, establishing (3.3) in Theorem 3.2.

Since each $z_j = \mathrm{O}\left(N^{\frac{1}{2}}\right)$, in (2.3) we have

$$\begin{aligned} z_{\text{ave}} &= \frac{1}{M}\sum_{j=0}^{M-1} z_j + \mathrm{O}(1) \\ &= \frac{1}{M}\sum_{j=0}^{M-1} z(x_j) + \mathrm{O}((\log N)^2) \\ &= \int_0^1 z(x)\,\mathrm{d}x + \mathrm{O}\left(N^{-1}\max_{[0,1]}|z'(x)|\right) + \mathrm{O}((\log N)^2) \\ &= \int_0^1 z(x)\,\mathrm{d}x + \mathrm{O}((\log N)^2). \end{aligned}$$

Using the expression (4.5) for the integral leads us to the result (3.4) in Theorem 3.2.

4.2. *Proof of Theorem 3.1*

To prove Theorem 3.1 we again appeal to the connection established in Lemma 4.1. In the regime (1.1), the continuum equation (4.2), when written as a system of two first order ODEs, has a global Lipschitz constant $L := 1 + K$ in the L_2 norm. We may thus apply a standard "Taylor series plus Gronwall inequality" argument for convergence of Euler's method, see, for example, [8, Theorem 3.4], to give

$$\sup_{1\leqslant j\leqslant M-1} |\tilde{z}_j - z(j\Delta x)| \leqslant C(L)\frac{1}{N}\max_{[0,1]}\{|z''(x)| + |z'''(x)|\},$$

where $C(L)$ depends on L (but not on N). Since $z''(x)$ and $z'''(x)$ are $\mathrm{O}(N)$, we conclude that the overall error is $\mathrm{O}(1)$. Converting from $\tilde{z}_j$ in (4.1) to z_j in (2.4), as in §4.1, leads to (3.1). The result (3.2) for z_{ave} also follows as in §4.1.

5. Summary

Partially random graphs form an appealing model for capturing features in many real-life networks and yet have yielded relatively little, so far, to rigorous analysis. This work makes three main theoretical contributions.

1. To formalize the idea of the greedy pathlength as a natural measure of the separation between nodes in a graph where there is an underlying metric.
2. To show that the expected greedy pathlength for a cycle plus shortcuts can be computed as the mean hitting time for a Markov chain.
3. To show that a rigorous continuum limit for the set of mean hitting times can be established via a convergence analysis for a finite-difference method.

Regarding item 1, we emphasize that the greedy pathlength is implicit in the work of Kleinberg [12] and has a natural interpretation as the free packet delay for a simple routing algorithm, [5]. Regarding items 2 and 3, we mention that the author has used a similar Markov chain approach to study mean hitting times for a random walk on a partially random graph [9]. In that case, the underlying Brownian motion gives rise to a diffusion term and the continuum limit is a singly perturbed boundary value problem, in contrast to the initial value problem encountered here.

The key new insight from this work is encapsulated in Fig. 3.1. Even when relatively few shortcuts are present in the network, the strategy of taking any shortcut that presents itself (without looking ahead to see if a better shortcut is coming up) is, on average, significantly sub-optimal. When a large number, $O(N)$, of shortcuts are added our results mirror those of [12] for a different model, in showing that the greedy algorithm completely fails to exploit the existence of a small world.

Acknowledgments

I thank Mike Molloy and Petter Wiberg for useful discussions on this topic. This work was carried out while I visited The Fields Institute for Research in Mathematical Sciences, Toronto, during the Thematic Year on Numerical and Computational Challenges in Science and Engineering. The final manuscript has also benefited from the attention of an extremely careful referee.

References

[1] L. Adamic, The small world web, in: Proceedings of the European Conference on Digital Libraries, 1999, pp. 443–452.
[2] A.D. Barbour, G. Reinert, Small worlds, Random Structures Algorithms 19 (2001) 54–74.
[3] B. Bollobás, F.R.K. Chung, The diameter of a cycle plus a random matching, SIAM J. Disc. Math. 1 (1988) 328–333.
[4] H. Fuks, A.T. Lawniczak, Performance of data networks with random links, Math. Comput. Simulation 51 (1999) 103–119.
[5] H. Fuks, A.T. Lawniczak, S. Volkov, Packet delay in models of data networks, ACM Trans. Model. Comput. Simulation 11 (2001) 233–250.
[6] P.M. Gleiss, P.F. Stadler, A. Wagner, D.A. Fell, Relevant cycles in chemical reaction networks, Adv. Complex Syst. 4 (2001) 207–226.
[7] P. Grindrod, Range-dependent random graphs and their application to modeling large small-world proteome datasets, Phys. Rev. E 66 (2002) 066702-1–066702-7.
[8] E. Hairer, S.P. Nørsett, G. Wanner, Solving Ordinary Differential Equations I: Nonstiff problems, second ed., Springer, Berlin, 1993.
[9] D.J. Higham, A matrix perturbation view of the small world phenomenon, SIAM J. Matrix Anal. Appl. 25 (2003) 429–444.
[10] S. Jin, A. Bestavros, Small-world characteristics of internet topologies and multicast scaling, Comput. Networks, in press.
[11] J. Kleinberg, Navigation in a small world, Nature 406 (2000) 845.
[12] J. Kleinberg, The small-world phenomenon: an algorithmic perspective, in: Proceedings of the 32nd ACM Symposium on Theory of Computing, 2000.
[13] S. Milgram, The small world problem, Psychology Today 2 (1967) 60–67.
[14] J.M. Montoya, R.V. Solé, Small world patterns in food webs, J. Theor. Biol. 214 (2002) 405–412.
[15] M.E.J. Newman, The structure and function of complex networks, SIAM Rev. 45 (2003) 167–256.
[16] M.E.J. Newman, C. Moore, D.J. Watts, Mean-field solution of the small-world network model, Phys. Rev. Lett. 84 (2000) 3201–3204.
[17] J.R. Norris, Markov Chains, Cambridge University Press, 1997.
[18] S.H. Strogatz, Exploring complex networks, Nature 410 (2001) 268–276.
[19] D.J. Watts, A simple model of global cascades on random networks, Proc. Natl. Acad. Sci. USA 99 (2002) 5751–6450.
[20] D.J. Watts, P.S. Dodds, M.E.J. Newman, Identity and search in social networks, Science 296 (2002) 1302–1305.
[21] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440–442.