Influence d'une distribution des degrés en loi de puissance sur la navigabilité des petits mondes†‡

Pierre Fraigniaud[1] et George Giakkoupis[2]

[1] *CNRS et Université Paris Diderot. Email : pierre.fraigniaud@liafa.jussieu.fr*
[2] *Université Paris Diderot. Email : george.giakkoupis@liafa.jussieu.fr*

Cette note résume nos travaux sur le routage décentralisé dans les petits mondes, et plus spécifiquement dans ceux qui combinent un plongement spatial avec une grande variabilité des degrés. Plus précisément, nous considérons une variante du modèle des treillis augmentés dû à J. Kleinberg (STOC 2000) pour laquelle le nombre de contacts lointains par sommet suit une loi de puissance. Ce modèle est motivé par des observations concordantes qu'un grand nombre de réseaux ont une telle distribution de degré. Pour ces réseaux, l'exposant α de la loi de puissance est généralement entre 2 et 3. Nous prouvons que, dans notre modèle, et pour cet intervalle de valeurs $2 < \alpha < 3$, l'espérance du nombre d'étapes du routage glouton de n'importe quelle source à n'importe quelle destination est au plus $O(\log^{\alpha-1} n)$. Cette borne est exacte au sens fort. En effet, nous montrons également que l'espérance du nombre d'étapes du routage glouton entre une source et une destination choisies aléatoirement uniformément est au moins $\Omega(\log^{\alpha-1} n)$. Pour $\alpha < 2$ ou $\alpha \geq 3$, nous montrons également que le routage glouton s'exécute en $\Theta(\log^2 n)$ étapes en espérance. Et, pour $\alpha = 2$, $\Theta(\log^{1+\varepsilon} n)$ étapes sont nécessaires en espérance, où $1/3 \leq \varepsilon \leq 1/2$.

A notre connaissance, ces résultats sont les premiers permettant de quantifier l'influence d'une distribution des degrés en loi de puissance sur la navigabilité des petits mondes. De surcroît, nous montrons que cette influence est significative. En particulier, lorsque l'exposant de la loi approche 2 par excès, l'espérance du nombre d'étapes du routage glouton dans un treillis augmenté dont les degrés suivent une oi de puissance approche la racine carrée du nombre d'étapes du routage glouton dans un treillis augmenté dont les degrés sont fixés, même lorsque les deux graphes possèdent un degré moyen identique.

Keywords: Routing, Social Networks

1 Navigability of small worlds

It has been observed that many "real-world" networks, such as social, information, technological, and biological networks, exhibit the *small-world* property ; i.e., they are locally clustered, and (yet) short paths exist between almost all pairs of nodes (see [New03] and the references therein). It is also well-established that many small-world networks (e.g., the network of acquaintances between individuals) are easy to *navigate*, provided that the nodes are able to estimate the distances to other nodes with respect to some underlying metric (e.g., geography, professions, etc.) [DMW03, Mil67]. *Navigability* refers to the ability of nodes to route messages efficiently in a decentralized manner, using local information only. The most prominent example of such a routing scheme is *greedy* routing : a node handling a message destined to some target node forwards the message to its neighbor that is closer to the target, according to the underlying metric. The first formal analysis of greedy routing in a plausible model of small worlds was presented in [Kle00]. The model studied there was the *augmented lattice* : Consider the n-node d-dimensional lattice that wraps

†Both authors are supported by the ANR project ALADDIN, and by the INRIA project GANG.
‡The full version of this paper will appear in *Proc. 28th ACM Symp. on Principles of Distributed Computing (PODC 2009)*.

around, where $d \geq 1$. A node has links to its $2d$ lattice-neighbors, and also to $k \geq 1$ other nodes, its *long-range contacts*. Each of the long-range contacts of a node u is chosen using an independent random trial following the d-harmonic distribution : the probability that node v is chosen in a given trial is

$$p_{u,v} \propto 1/(\mathrm{dist}(u,v))^d, \tag{1.1}$$

where $\mathrm{dist}(u,v)$ is the lattice-distance between u and v. In [Kle00] it was shown that, in this model, greedy routing requires $\mathrm{O}(\frac{1}{k}\log^2 n)$ expected number of steps, for any source–target pair. (This complexity was later shown to be tight [MN04].) It was also shown that *any* decentralized routing algorithm performs poorly if the d-dimensional lattice is augmented using the h-harmonic distribution, for any $h \neq d$. Specifically, $\Omega(n^\gamma)$ expected steps are required, for some $\gamma > 0$ that depends on h and d.

Despite its simplicity, the augmented-lattice model seems to capture successfully the small-world and navigability properties of real-world networks. Note that in the d-dimensional lattice the d-harmonic distribution is equivalent to the "natural" distribution $p_{u,v} \propto 1/|B_u(\mathrm{dist}(u,v))|$, where $B_u(r)$ is the ball centered at u of radius r ; this latter distribution was used in [DHLS06, Sli05] to extend the results of [Kle00] to graphs of bounded ball growth, and to graphs of bounded doubling dimension. Also, the d-harmonic distribution is equivalent in the lattice to the rank-based distribution $p_{u,v} \propto 1/r_u(v)$, where $r_u(v)$ is the rank of v when nodes are sorted in increasing distance from node u ; this latter distribution was used in [KLNT06] to extend the results of [Kle00] to non-uniform population densities. In fact, it was experimentally demonstrated that two-thirds of friendships are geographically distributed this way : the probability of befriending a particular person is inversely proportional to the number of people closer to you [LNNK+05]. Finally, it was recently shown that the d-harmonic distribution of the long-range links might as well be an inherent byproduct of node mobility [CFL08]. Therefore, there is now a consensus that the augmented-lattice model is an appropriate framework for analyzing small-world navigability.

2 Power-law degree distribution

The augmented-lattice model, however, fails to capture another commonly observed property of real-world networks, the *heavy-tailed degree distribution*. Such a distribution is well approximated by a *power law*

$$\mathbb{Pr}[\deg(u) = k] \propto 1/k^\alpha, \tag{2.1}$$

where α is a real, typically between 2 and 3 [New03]. Nevertheless, it is straightforward to reconcile the augmented-lattice model with a power-law distribution for the node degrees, simply by drawing the number of long-range links added to each node independently at random from a power-law distribution [Kle06]. It is reasonable to expect that this modification would reduce the lengths of shortest paths between nodes, since the (few) high-degree nodes should provide short-cuts between most nodes. This is typically the case in networks with power-law degree sequences [CL03]. However, it is unclear how decentralized routing could benefit from the existence of these high-degree nodes [Kle06].

Utilizing the heavy-tailed degree distribution in the design of decentralized routing algorithms was suggested in [ALPH01, FGL05, KYHJ02, SBR04]. In all these works, the routing algorithms only have access to information about the degrees of neighboring nodes, not to any embedding of the graph. Although some performance improvements are observed compared to routing algorithms oblivious to the node degrees, the expected number of steps remains polynomial in the network size. Also, [SJ05] proposed a heuristic decentralized algorithm for routing in a variance of the augmented lattice where nodes have widely varying degrees. This heuristic assumes that nodes have access both to the locations of theirs neighbors, and to their degrees. Simulations showed that this algorithm performs better than decentralized algorithms using only one of these two sources of information. However, no formal analysis was provided.

3 Our framework

We consider the following variance of the augmented-lattice model. As in the original model, the long-range links are drawn independently at random according the harmonic distribution with exponent equal to

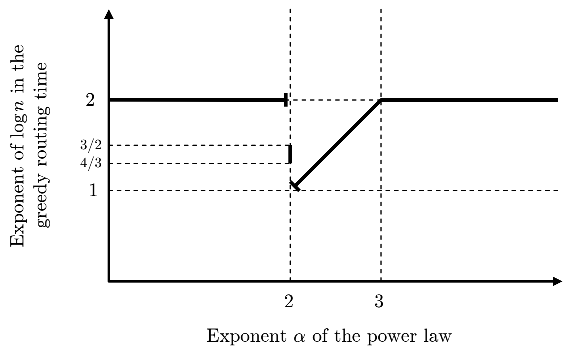

FIGURE 1: Summary of the results.

the dimensionality of the lattice (cf. Eq. 1.1). Unlike the original model, however, the number of long-range contacts each node has is not fixed, but it is drawn independently at random from the power-law distribution with exponent $\alpha \geq 0$ (cf. Eq. 2.1). This distribution is scaled so that its expectation is constant and each node has at least one long-range contact. We then remove the orientation of each of the long-range links to get a non-directed network. We study the performance of greedy routing in this network.

4 Our results

In this section, we ignore $O(\log \log n)$ multiplicative factors in the statement of the asymptotic bounds.

We prove that, for $2 < \alpha < 3$, which is the case for most real-world networks, the expected number of steps of greedy routing *from any source to any target* is $O(\log^{\alpha-1} n)$ steps. Thus, for this range of values for α, the effect of the power-law degree distribution is significant. In particular, when α approaches 2, the expected number of steps of greedy routing in the augmented lattice with *power-law degrees* approaches the square-root of the expected number of steps of greedy routing in the augmented lattice with *fixed degrees*, although both networks have the same *average degree*. For both $\alpha < 2$ and $\alpha \geq 3$, we show that the expected number of steps of greedy routing from any source to any target is $O(\log^2 n)$ steps, which is the same order of magnitude as the performance of greedy routing in the augmented lattice with fixed degrees. For the critical value $\alpha = 2$, we prove that the expected number of steps of greedy routing from any source to any target is $O(\log^{3/2} n)$ steps.

All these upper bounds are tight (but, perhaps, for $\alpha = 2$). For $\alpha > 2$, the upper bounds are even tight in a strong sense. Indeed, we prove that the expected number of steps of greedy routing for a *uniformly-random* pair of source–target nodes is $\Omega(\log^{\alpha-1} n)$ steps if $2 < \alpha < 3$, and $\Omega(\log^2 n)$ steps if $\alpha \geq 3$. For $\alpha < 2$, we prove that there exists a source–target pair for which greedy routing requires $\Omega(\log^2 n)$ expected steps. For $\alpha = 2$, we show that the expected number of steps for a uniformly-random source–target pair is $\Omega(\log^{4/3} n)$.

We formally prove the above results for the case of the 1-dimensional lattice, i.e., the ring. Nevertheless, none of the arguments we use is specifically tied to the ring, and the *exact* same results can be easily shown for d-dimensional lattices, for constant values of d. Note that unlike the results in [Kle00], where the critical value of the exponent depends on the dimensionality d of the lattice, our results do not depend on d.

To the best of our knowledge, these results are the first to formally quantify the effect of the power-law degree distribution on the navigability of small worlds.

The following picture emerges from our analysis. For $\alpha \geq 3$, almost all nodes are of small degree, and the nodes of higher degree are too few to contribute significantly. Hence greedy routing performs essentially the same as when the degrees are fixed. For $2 < \alpha < 3$, there are still very few nodes of high degree. However, nodes of degree roughly $\log n$ are relatively abundant, and there are more and more such nodes as α approaches 2. It is the contribution of these nodes that reduces the routing time from $\log^2 n$ to $\log^{\alpha-1} n$.

The case $\alpha = 2$ is special. All "degree scales" are present, and each is equally likely to contribute. On the one hand, this results in greater routing speed than in the case $2 < \alpha < 3$ when the current node is far from the target, since there are many high-degree nodes between the current node and the target in the lattice. On the other hand, the balance in the degree scales means that as we get closer to the target the number of high-degree nodes available decreases faster than when $2 < \alpha < 3$; and when we get at distance sub-polynomial from the target (essentially at distance less than $e^{\sqrt{\ln n}}$), greedy routing performs the same as when the degrees are fixed.

Finally, for $\alpha < 2$, there are very many nodes of high degree, and the role of the cut-off point $k_{\max}$ of the power law becomes critical. We assumed that $k_{\max} \sim n^{\gamma}$, for some $0 < \gamma \leq 1$. In this setting, only the contribution of nodes with degree close to $k_{\max}$ is significant. However, when the current node is at distance less than $k_{\max}$ from the target, it is very likely that greedy routing will not find a node of such degree, and from that distance it starts performing the same as when the degrees are fixed. Note that for $\alpha < 2$, nodes that are further away from the target may, in expectation, require fewer steps to reach that target than nodes closer to the target, which is not the case when $\alpha > 2$.

Références

[ALPH01] L. A. Adamic, R. M. Lukose, A. R. Puniyani, and B. A. Huberman. Search in power-law networks. *Phys. Rev. E*, 64 :46135, 2001.

[CFL08] A. Chaintreau, P. Fraigniaud, and E. Lebhar. Networks become navigable as nodes move and forget. In *Proc. 35th Int. Colloq. on Automata, Languages and Programming (ICALP)*, pages 133–144, 2008.

[CL03] F. R. K. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Mathematics*, 1(1) :91–113, 2003.

[DHLS06] P. Duchon, N. Hanusse, E. Lebhar, and N. Schabanel. Could any graph be turned into a small-world ? *Theoretical Computer Science*, 355(1) :96–103, 2006.

[DMW03] P. S. Dodds, R. Muhamad, and D. J. Watts. An experimental study of search in global social networks. *Science*, 301(5634) :827–829, 2003.

[FGL05] P. Fraigniaud, P. Gauron, and M. Latapy. Combining the use of clustering and scale-free nature of user exchanges into a simple and efficient P2P system. In *Proc. 11th European Conf. on Parallel Computing (Euro-Par)*, pages 1163–1172, 2005.

[Kle00] J. Kleinberg. The small-world phenomenon : An algorithm perspective. In *Proc. 32nd ACM Symp. on Theory of Computing (STOC)*, pages 163–170, 2000.

[Kle06] Jon Kleinberg. Complex networks and decentralized search algorithms. In *Proc. International Congress of Mathematicians (ICM)*, 2006.

[KLNT06] R. Kumar, D. Liben-Nowell, and A. Tomkins. Navigating low-dimensional and hierarchical population networks. In *Proc. 14th Ann. European Symp. on Algorithms (ESA)*, pages 480–491, 2006.

[KYHJ02] B. Kim, C. Yoon, S. Han, and H. Jeong. Path finding strategies in scale-free networks. *Phys. Rev. E*, 65 :027103, 2002.

[LNNK+05] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins. Geographic routing in social networks. *Proc. National Academy of Sciences of the USA*, 102(33) :11623–11628, 2005.

[Mil67] S. Milgram. The small world problem. *Psychology Today*, 67(1) :60–67, 1967.

[MN04] C. Martel and V. Nguyen. Analyzing Kleinberg's (and other) small-world models. In *Proc. 23rd ACM Symp. on Principles of Distributed Computing (PODC)*, pages 179–188, 2004.

[New03] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2) :167–256, 2003.

[SBR04] N. Sarshar, P. O. Boykin, and V. P. Roychowdhury. Percolation search in power law networks : Making unstructured peer-to-peer networks scalable. In *Proc. 4th Int. Conf. on Peer-to-Peer Computing*, pages 2–9, 2004.

[SJ05] Ö. Simsek and D. Jensen. Decentralized search in networks using homophily and degree disparity. In *Proc. 19th Int. Joint Conf. on Artificial Intelligence (IJCAI)*, pages 304–310, 2005.

[Sli05] A. Slivkins. Distance estimation and object location via rings of neighbors. In *Proc. 24th ACM Symp. on Principles of Distributed Computing (PODC)*, pages 41–50, 2005.