Could any graph be turned into a small-world?

Philippe Duchon[1], Nicolas Hanusse[1], Emmanuelle Lebhar[2] and Nicolas Schabanel[2]

[1] *LABRI, Domaine universitaire 351, cours de la libération, 33405 Talence Cedex, France*
[2] *LIP - ENS Lyon, 46 allée d'Italie, 69007 Lyon, France*

In addition to statistical graph properties (diameter, degree, clustering, ...), Kleinberg [Kle00] showed that a small-world can also be seen as a graph in which the routing task can be efficiently and easily done in spite of a lack of global knowledge. More precisely, in a lattice network augmented by extra random edges (but not chosen uniformly), a short path of polylogarithmic expected length can be found using a greedy algorithm with a local knowledge of the nodes. We call such a graph a *navigable small-world* since short paths exist and can be followed with partial knowledge of the network. In this paper, we show that a wide class of graphs can be augmented with one extra edge per node into navigable small-worlds.

Keywords: small-world, random graph model, routing algorithm

1 Introduction

In the last decade, effective measurements of real interaction networks have revealed specific unexpected properties. Among these, most of these networks present a very small diameter and a high clustering. Furthermore, very short paths can be efficiently found between any pair of nodes without global knowledge of the network, which is known as the small-world phenomenon [Mil67]. Several models have been proposed to explain this phenomenon ([WS98], [NW99]). However, Kleinberg showed in 2000 [Kle00] that these models lack the essential *navigability* property : in spite of a polylogarithmic diameter, none of the short paths can be computed efficiently without global knowledge of the network ; i.e., routing requires the visit of a polynomial number of nodes (in the size of the network). He introduced an augmented graph model consisting of a grid where each node is given a constant number of random additional directed long range links distributed according to the harmonic distribution, i.e., the probability that a node $\mathbf{v}$ is the i-th long range contact of a node $\mathbf{u}$ is proportional to $1/|\mathbf{u}-\mathbf{v}|^s$, where $|\mathbf{u}-\mathbf{v}|$ denotes their distance in the grid and $s>0$ is a parameter of the model. In this model, the local knowledge at each node is the underlying metric of the grid and the positions on the grid of its long range neighbors. Kleinberg considered the greedy routing (decentralized) where each node forwards the message to its neighbor that is the closest to the destination. He proved that this algorithm computes between any pair of nodes a path of polylogarithmic length in the size of the network after visiting a polylogarithmic number of nodes, if and only if the exponent s is equal to the dimension of the grid. Later on, Barrière *et al.* [BFKK01] generalized this result to a regular $n\times\cdots\times n$ torus. Moreover, they showed that the expected number of steps of the greedy algorithm is $\Theta(\log^2 n)$, and that, noticeably, the number of steps is independent of the dimension. This reveals a strong correlation between the underlying grid metric and the additional long range links distribution that turns the grid into a small-world. This statement raises an essential question to capture the small-world phenomenon : are there only specific graph metrics that can be turned into small-worlds by the addition of shortcuts ?

This question on the metric structure can be reinforced by the fact that whenever the exponent s is different from the dimension of the grid, the greedy algorithm follows a path of polynomial length even when the diameter is polylogarithmic (which is the case for $d<s<2d$ [MN04, NM05]). The reader might believe that the navigability property is very specific to the grid topology, but we will show that a wide family of graphs can be turned into navigable small-worlds. In [Kle02], Kleinberg generalized his lattice-based model and showed how to turn into smallworlds tree-based or group-based structures by adding a polylogarithmic number of long range links per node. [NM05, Fra05] are other recent articles which tackle these questions. In this paper, we attempt to find a class of graph metrics as wide as possible for which the addition of a constant number of random long range links per node gives rise to the small-world phenomenon. Roughly speaking, as soon as the original graph H is homogeneous in terms of ball expansion and as soon as balls centered on each node grow up to slightly more than polynomialy with their radius, H can be augmented to become a navigable small-world. It follows that a wide class of graphs can be turned into navigable small-worlds, including in particular any Cayley graphs known up to now. In a second step, we try to catch the dimensional phenomenon by studying cartesian products of our graphs. We show that if two independent graphs can be augmented into two navigable small-worlds then their cartesian product can also be augmented into a navigable small-world. For instance, as a consequence, any unbalanced torus $C_{n_1} \times C_{n_2} \times \ldots \times C_{n_l}$ can be turned into a small-world in which the greedy algorithm computes paths of length $O(\log^{2+\varepsilon}(\max_i n_i))$, for any $\varepsilon > 0$.

In all the following, we will consider infinite graphs, but our definitions and results apply as well to infinite families of finite graphs (see [DHLS05]).

2 Small-worlds and graph metrics

For a given graph $G = (V, E)$, we write $\mathcal{B}_{G,\mathbf{u}}(r)$ for the ball centered on a node $\mathbf{u}$ with radius r, and $b_{G,\mathbf{u}}(r)$ for its cardinality. Let $b_G(r) = \max_{\mathbf{u}\in V} b_{G,\mathbf{u}}(r)$. The G subscript will be omitted in case the concerned graph is obvious. We only consider graphs with maximum degree Δ, a fixed constant.

In the following, an *underlying metric* δ_H of a graph G is the metric given by a spanning connected subgraph H (i.e., $\delta_H(\mathbf{u}, \mathbf{v})$ is the distance between $\mathbf{u}$ and $\mathbf{v}$ in H). Definitions 1 and 2 are inspired from the work of Kleinberg [Kle00].

Definition 1 *A* decentralized algorithm using an underlying metric δ_H *in a graph G is an algorithm that computes a path between any pair of nodes by navigating through the network from the source to the target, using only the knowledge 1) of* δ_H *2) of the nodes it has previously visited as well as their neighbors. But, crucially, 3) it can only visit nodes that are neighbors of previously visited nodes.*

Our definition of a navigable small-world is essentially probabilistic. We consider random graph models in which a fixed “base” graph H is *randomly augmented* by adding random links (called *long range links* below), according to some probability distribution. Routing will then be performed by a decentralized algorithm, using the base metric δ_H ; our goal is to identify such augmented graph models for which this procedure results in uniformly “fast” routing. Since the augmented graph will have a finite degree, at least some of the $b_{H,\mathbf{u}}(r)$ nodes at distance at most r of $\mathbf{u}$ will remain at distance $\Omega(\log b_{H,\mathbf{u}}(r))$ in the augmented graph This motivates the following definition.

Definition 2 *An infinite randomly augmented graph G, on a (infinite) base graph H, is a* navigable small-world *if there exists a decentralized algorithm using the underlying base metric* δ_H *that, for any two nodes* $\mathbf{u}$ *and* $\mathbf{v}$*, computes a path from* $\mathbf{u}$ *to* $\mathbf{v}$ *in G by visiting an expected number of nodes that is polylogarithmic in* $b_{H,\mathbf{v}}(\delta_H(\mathbf{u}, \mathbf{v}))$.

3 Turning graphs into small-worlds

In this section, we describe a wide class of infinite graphs, or of infinite families of finite graphs, for which we are able to define random augmentation models that result in navigable small-worlds. In all cases, our routing algorithm will be the greedy algorithm, thus the set of visited nodes will coincide with the path computed. Furthermore, even if some algorithms can compute significantly shorter paths [FGP04, MN04, LS04], it has been shown in [MNW04] that no decentralized algorithm can compute a polylogarithmic path between two nodes while visiting asymptotically fewer nodes than the greedy algorithm. All models we will consider add exactly one directed edge† leaving each node $\mathbf{u}$. For each node $\mathbf{u}$, there is a function $f_{\mathbf{u}}$ such that each other node $\mathbf{v}$ has probability proportional to $f_{\mathbf{u}}(\delta(\mathbf{u}, \mathbf{v}))$ of being the destination $L_{\mathbf{u}}$; the normalizing factor is $Z_{\mathbf{u}} = \sum_{\mathbf{v}\in V} f_{\mathbf{u}}(\delta(\mathbf{u}, \mathbf{v}))$.

† Adding a constant number k of edges instead of one would not significantly alter the results, as will be made clear by the proofs.

Definition 3 *We say that an infinite graph is* smallworldizable *if there exists, for each* $\mathbf{u}$*, a distribution* $f_{\mathbf{u}}(r)$ *such that we obtain a navigable small-world randomly augmented graph by adding one random long range link to each node* $\mathbf{u}$ *according to* $f_{\mathbf{u}}(r)$ *(each node* $\mathbf{u}$ *is the origin of one long range link whose destination is* $\mathbf{v}$ *with probability proportional to* $f_{\mathbf{u}}(\delta(\mathbf{u},\mathbf{v}))$*).*

The following class of graph is defined for the sake of readability. As shown below, it characterizes a class of smallworldizable graphs.

Definition 4 *A bounded degree infinite graph H is an* α-moderate growth graph *if there exists a constant* $\alpha > 0$*, such that the ball size of each node* $\mathbf{u}$ *of H can be written as* $b_{\mathbf{u}}(r) = r^{d_{\mathbf{u}}(r)}$*, where* $d_{\mathbf{u}}(r) : [2,\infty) \to \mathbb{R}$ *is* $\mathcal{C}^1$ *and satisfies* $\forall r \geq 2, d'_{\mathbf{u}}(r) \leq \alpha/(r \ln r)$*.*

Lemma 1 *Let G an* α*-moderate growth infinite graph. There exists a constant* $C > 0$*, s.t. for all* $\mathbf{u} \in G$*,* $r > 1$*,* $b_{\mathbf{u}}(r) - b_{\mathbf{u}}(r-1) \leq \left(\frac{C \ln\ln(r)}{r}\right) b_{\mathbf{u}}(r)$*.*

Lemma 2 *Let G a moderate growth infinite graph. Then, there exists a constant* C' *such that* $b_{\mathbf{u}}(r) \leq (\ln r)^{-\alpha \ln \beta}/(\beta^{C+C'\alpha})\, b_{\mathbf{u}}(\beta r)$*, for any* $\mathbf{u} \in G$*,* $0 < \beta < 1$ *and* $r \geq 2$*.*

Theorem 3 *Any moderate growth infinite graph is smallworldizable by the addition of one long range link per node, distributed according to* $f_{\mathbf{u}}(r) = \frac{1}{b_{\mathbf{u}}(r)\log^q r}$*, for any* $q > 1$*. Furthermore, the expected greedy path length between any pair of nodes at distance* ℓ *from each other is* $O(\ln^{1+q+\alpha \ln 5} \ell)$*.*

Proof sketch. The normalization constants $Z_{\mathbf{u}} = \sum_{\mathbf{v} \in V} f_{\mathbf{u}}(\delta_H(\mathbf{u},\mathbf{v})) = \sum_{r \geq 1} (b_{\mathbf{u}}(r) - b_{\mathbf{u}}(r-1)) f_{\mathbf{u}}(r)$ are uniformly bounded by Lemma 1. The distribution defining the randomly augmented graph is thus properly defined. Assume that $\mathbf{s}$ and $\mathbf{t}$ are respectively the source and target, we analyze the expected path length computed by the greedy algorithm. Consider some integer $r \geq 2$ and a node $\mathbf{u}$ such that $r/2 < \delta_H(\mathbf{u},\mathbf{t}) \leq r$, and denote by $L_{\mathbf{u}}$ the long range contact of $\mathbf{u}$. We give a lower bound on $\Pr[\delta_H(L_{\mathbf{u}},\mathbf{t}) \leq r/2]$, the probability that the destination node $L_{\mathbf{u}}$ belongs to $\mathcal{B}_{\mathbf{t}}(r/2)$. Since $f_{\mathbf{u}}$ is a decreasing function and $\mathcal{B}_{\mathbf{t}}(r/2) \subseteq \mathcal{B}_{\mathbf{u}}(3r/2)$, each node of $\mathcal{B}_{\mathbf{t}}(r/2)$ has probability at least $f_{\mathbf{u}}(3r/2)/Z$ of being L_u. Since, in turn, $\mathcal{B}_{\mathbf{u}}(3r/2) \subseteq \mathcal{B}_{\mathbf{t}}(5r/2)$, we can give a lower bound on $f_{\mathbf{u}}(3r/2)$ in terms of $b_{\mathbf{t}}$: $f_{\mathbf{u}}(3r/2) \geq \frac{1}{b_{\mathbf{t}}(5r/2)\ln^q(3r/2)}$. Thus, we get a lower bound, depending only on $\mathbf{t}$ and r, on the wanted probability :

$$
\begin{aligned}
\Pr[\delta_H(L_{\mathbf{u}},\mathbf{t}) \leq r/2] &\geq \frac{1}{Z \ln^q(3r/2)}\ \frac{b_{\mathbf{t}}(r/2)}{b_{\mathbf{t}}(5r/2)} \\
&\geq \left(Z 5^{C+C'\alpha} \ln^q(3r/2) \ln^{\alpha \ln 5}(5r/2)\right)^{-1} \geq \left(Z 2^{q+\alpha \ln 5} 5^{C+C'\alpha} \ln^{q+\alpha \ln 5}(r)\right)^{-1}
\end{aligned}
$$

We partition the whole graph into concentric shells centered on $\mathbf{t}$, where the k-th shell consists of all nodes whose δ_H distance to $\mathbf{t}$ is between 2^{k-1} and 2^k. The previous discussion proves that each node in the k-th shell has probability $\Omega(k^{-\gamma})$ of having its long range contact in some i-th shell with $i < k$, where $\gamma = q + \alpha \ln 5$. As a result, the expected length of the greedy path from $\mathbf{s}$ to $\mathbf{t}$ is $O(\ln^{1+q+\alpha \ln 5} \ell)$, which is polylogarithmic in $\ell = \delta_H(\mathbf{s},\mathbf{t})$, and *a fortiori* in $b_{\mathbf{t}}(\delta_H(\mathbf{s},\mathbf{t}))$. The argument used here is similar to that of Kleinberg's analysis of its original model ; our changes allow the upper bound to be expressed only in terms of the original metric (and not the total size of the graph). □

The theorem above covers graphs with ball sizes $b(r)$ growing like $r^{\alpha \log\log r}$, $\alpha > 0$, or slower. Note that we get a similar upper bound $O(\ln^{2+\varepsilon} r)$, for any $\varepsilon > 0$, on the expected length of the greedy path between any pair of nodes at distance r from each other. In particular, all known Cayley graphs[‡] are smallwordizable since groups of intermediate ball size, between polynomial and exponential, are still unknown, and it is an open question whether there exists a group with ball size $b(r)$ superpolynomial but less than $e^{\sqrt{r}}$, see for instance [Bar02] for a state of the art.

[‡] A Cayley graph is a graph defined by a group G generated by $g_1, \ldots, g_k$, whose vertices are the elements of G and such that there is an edge between x and y iff there is a generator $g_i \in G$ such that $x = g_i y$.

4 Products of small-worldizable graphs

A remarkable fact on the small-world property is its relative independence of the metric dimension in Kleinberg's model. This motivates the study of product of smallworldizable graphs.

Definition 5 *The* cartesian product $H = F \times G$ *of two undirected graphs* F *and* G *is the graph* (V_H, E_H) *where* $V_H = V_F \times V_G$ *and* $E_H = \{((f,g),(f,g')) : gg' \in E_G, f \in V_F\} \cup \{((f,g),(f',g)) : g \in V_G, ff' \in E_F\}$.

Theorem 4 *Let* F *and* G *be* α_1- *and* α_2*-moderate growth infinite graphs respectively. The cartesian product* $H = F \times G$ *is smallworldizable by the addition of one long range link per node* $\mathbf{u}$ *according to the distribution* $h_{\mathbf{u}}(r) = 1/(b_{H,\mathbf{u}}(r) \ln^{q'} r)$*, for all* $q' > q_0$*, for some constant* $q_0 > 0$*. Furthermore, the expected greedy path length between any pair of nodes at distance* ℓ *from each other is* $O(\ln^{1+q'+(\alpha_1+\alpha_2)\ln 10} \ell)$.

Note that this theorem yields another simple method to obtain a generalization of Kleinberg's graph to tori of dimension $d \geq 1$ with arbitrary side sizes, seen as cartesian products of one dimensional Kleinberg graphs of various sizes. It also gives an expected path length $O(\ln^{2+\varepsilon} \ell)$ for cartesian product of uniform growth graphs (i.e. $\alpha = 0$), close to Kleinberg model's path length.

Références

[Bar02] L. Bartholdi. Groups of intermediate growth, preprint oai :arxiv.org :math/0201293. preprint oai :arXiv.org :math/0201293 (20040622), 2002.

[BFKK01] L. Barrière, P. Fraigniaud, E. Kranakis, and D. Krizanc. Efficient Routing in Networks with Long Range Contacts. In *Proceedings of the 15th International Conference on Distributed Computing (DISC)*, pages 270–284, 2001.

[DHLS05] P. Duchon, N. Hanusse, E. Lebhar, and N. Schabanel. Could any graph be turned into a small world ? *To appear in Theoretical Computer Science special issue on Complex Networks*, 2005. Also available as Research Report LIP-RR2004-62.

[FGP04] P. Fraigniaud, C. Gavoille, and C. Paul. Eclecticism shrinks even small worlds. In *Proceedings of the 23rd ACM Symp. on Principles of Distributed Computing (PODC)*, pages 169–178, 2004.

[Fra05] P. Fraigniaud. Greedy routing in tree-decomposed graphs : a new perspective on the small-world phenomenon. preprint, 2005.

[Kle00] J. Kleinberg. The Small-World Phenomenon : An Algorithmic Perspective. In *Proceedings of the 32nd ACM Symposium on Theory of Computing (STOC)*, pages 163–170, 2000.

[Kle02] J. Kleinberg. Small-world phenomena and the dynamics of information. In T. G. Dietterich, S. Becker, and Z. Ghahramani, editors, *Advances in Neural Information Processing Systems (NIPS) 14*, Cambridge, MA, 2002. MIT Press.

[LS04] E. Lebhar and N. Schabanel. Almost optimal decentralized routing in long-range contact networks. In *LNCS proceedings of 31st International Colloquium on Automata, Languages and Programming (ICALP)*, pages 894–905, 2004.

[Mil67] S. Milgram. The small world problem. *Psychology Today*, 61(1), 1967.

[MN04] C. Martel and V. Nguyen. Analyzing kleinberg's (and other) small-world models. In *Proceedings of the Twenty-Third Annual ACM Symposium on Principles of Distributed Computing*, pages 179–188, 2004.

[MNW04] G. S. Manku, M. Naor, and U. Wieder. Know thy neighbor's neighbor : the power of lookahead in randomized p2p networks. In *Proceedings of the 36th ACM Symposium on Theory of Computing (STOC)*, 2004.

[NM05] V. Nguyen and C. Martel. Analyzing and characterizing small-world graphs. To appear in the 2005 ACM-SIAM symposium on Discrete Algorithms (SODA), 2005.

[NW99] M. E. J. Newman and D. J. Watts. Scaling and percolation in the small-world network model. *Phys. Rev.*, 60 :7332–7342, 1999.

[WS98] D. Watts and S. Strogatz. Collective dynamics of small-world networks. *Nature*, 393(440–442), 1998.