Could any graph be turned into a small-world?

Philippe Duchon, Nicolas Hanusse [a]
Emmanuelle Lebhar, Nicolas Schabanel [b]

[a] *LABRI, Domaine universitaire 351, cours de la libération, 33405 Talence Cedex - FRANCE*

[b] *LIP, École normale supérieure de Lyon, 46 allée d'Italie, 69364 Lyon Cedex 07 - FRANCE*

Abstract

In addition to statistical graph properties (diameter, degree, clustering, ...), Kleinberg [7] 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 into navigable small-worlds.

Key words: small-world, random graph model, routing algorithm

1 Introduction

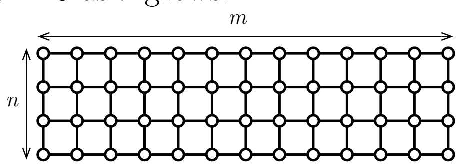
In the last decade, effective measurements of real interaction networks have revealed specific unexpected properties. Among these, most of these networks present a very small diameter and a high clustering. Furthermore, very short paths can be efficiently found between any pair of nodes without global knowledge of the network, which is known as the small-world phenomenon (exhibited by Milgram [12]). Several models have been proposed to explain this phenomenon. Among them, one approach is based upon an augmentation process: starting from a graph H and adding a relatively small set of extra edges L, we hope to obtain a new graph G sharing some graph properties with H, and exhibiting additional properties due to the design of L. For instance, the circulant graph H consisting in an n-node ring in which each node is also connected to the $k-$th closest nodes is locally clustered and is often used as a base graph for small-worlds [13,7]. Watts and Strogatz, in 1998 [15], showed that adding a controlled amount of randomness in the circulant graph gives rise to some of the small-world characteristics: rewiring an increasing fraction α of the links to random nodes *chosen uniformly* exhibits the desired characteristics (locally clustered and small diameter) for a reasonably large range of α away from 0 and 1. In fact, this model can be easily seen as an augmentation process: allowing to rewire only one link per node (defining in this way L) would not alter a lot the model and the results (cf. [13]).

However, Kleinberg showed in 2000 [7] that this model lacks 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 showed that navigability can be obtained by adjusting the amount of randomness to the underlying metric (the ring here). Precisely, 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 Manhattan distance in the grid and s is some parameter of the model. In this model, the local knowledge at each node is the underlying metric of the grid (which can be viewed as the geographic locations of the nodes) and the positions on the grid of its long range neighbors. Note that a global knowledge would be the set of positions of all the long range neighbors on the grid. Kleinberg proved that there exists a decentralized algorithm (using only local knowledge) that 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. The simplest such algorithm is *greedy routing* : each node obliviously forwards the message to its neighbor (local or long range) that is the closest (in the known metric) to the destination. Later on, Barrière *et al.* [1] generalized this result to a torus of any dimension. 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 impression 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. For Kleinberg's model, Martel and Nguyen [11] proved that the diameter is $\Theta(\log n)$ for $s = d$

and it is conjectured that for $s < 2d$, it is polylogarithmic, where d is the dimension. For a slightly different percolation model (in which degrees are unbounded), Coppersmith *et al.* [3] showed indeed a polylogarithmic upper bound on the diameter whenever $s < 2d$. 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 [8], 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. [14,4] are other recent articles which tackle these questions for other graph classes.

In section 2, we give a formal definition of navigable small-world graphs with respect to a given underlying metric. Roughly speaking, a greedy decentralized routing scheme computes a path of polylogarithmic length in the distance, between the source node and the target [1]. In this paper, we attempt to find a class of graph metrics as wide as possible for which the addition of random long range links gives rise to the small-world phenomenon. Indeed, Kleinberg's augmentation process which turns the grid into a navigable small-world fails, for instance, on unbalanced $n \times m$ grids (with $m \gg n$), since the "dimension" varies with the distance: balls of small radius grow like r^2 but larger balls grow like $r^{1+\epsilon(r)}$, where $\epsilon(r) \to 0$ as r grows.



It appears that defining the random link distribution in terms of ball growth in the original base graph, rather than in terms of distance between nodes, allows to generalize Kleinberg's process to a wide class of graphs. 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 polynomially 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 a navigable small-world, including in particular any Cayley graphs known up to now. In a second step, we try to catch the dimensional phenomenon by studying the cartesian product 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, even though it may not belong to the class itself. This reveals that the small-world phenomenon of a network may rely on multiple underlying structures. 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 $O(\log^{2+\epsilon}(\max_i n_i))$, for any $\epsilon > 0$.

[1] in Kleinberg's model, the path is of polylogarithmic length in the size of the network

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 by the work of Kleinberg [7].

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.*

The efficiency of a decentralized algorithm depends crucially on the number of nodes it visits to compute a path. Note that it upper bounds the length of the computed path.

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 links* below), according to some probability distribution. Routing will then be performed by a decentralized algorithm, using the base metric δ_H (which is obviously an upper bound on distances in the augmented graph); 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, with 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 $b_{H,\mathbf{v}}(\delta_H(\mathbf{u}, \mathbf{v}))$.*

A family of finite randomly augmented graphs $(G_i)_{i\in\mathcal{I}}$, with base graphs H_i for each G_i, is a navigable small-world family *if there exists a (uniform) polynomial p, and a decentralized algorithm using the underlying base metric δ_{H_i} on G_i that, for every $i \in \mathcal{I}$ and any pair of nodes $\mathbf{u}$ and $\mathbf{v}$ in G_i, computes a path from $\mathbf{u}$ to $\mathbf{v}$, by visiting an expected number of nodes at most $p(\log b_{H_i,\mathbf{v}}(\delta_{H_i}(\mathbf{u}, \mathbf{v})))$.*

Note that some graphs are intrinsically navigable small-worlds (and do not require any augmentation). Indeed, any graph such that each $b_{\mathbf{v}}(r)$ is bounded below by a (uniform) exponential function of r is a small-world, since the greedy algorithm, using the graph metric, computes a path of optimal length $r = \delta(\mathbf{u}, \mathbf{v})$ between two nodes at distance r, and $r = \text{polylog}(b(r))$ in this case. The simplest example of this situation is an infinite k-ary tree (with $k \geq 3$); balls of radius r have size $\Theta((k-1)^r)$. On the contrary, graphs with polynomial ball growth (i.e. $b(r) = \Theta(r^c)$ for some constant c) are not intrinsically navigable small-worlds since the length r of the optimal path computed by the greedy algorithm between two nodes at distance r is not a polylogarithmic function of $b(r)$.

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 [5,11,9], it has been shown in [10] that no decentralized algorithm can compute a polylogarithmic path between two nodes while visiting significantly fewer nodes than the greedy algorithm.

All models we will consider add exactly one directed edge [2] leaving each node $\mathbf{u}$, and the destination $L_{\mathbf{u}}$ of this outgoing edge is randomly chosen according to a random distribution that gives equal weight to any two nodes that are equally distant from $\mathbf{u}$ (in the base graph). Thus, 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 $L_{\mathbf{u}}$; the normalizing factor $Z_{\mathbf{u}}$ is

$$Z_{\mathbf{u}} = \sum_{\mathbf{v} \in V} f_{\mathbf{u}}(\delta(\mathbf{u}, \mathbf{v})) = \sum_{r > 0} (b_{\mathbf{u}}(r) - b_{\mathbf{u}}(r-1)) f_{\mathbf{u}}(r).$$

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 the randomly augmented graph obtained by the addition of one random long range link to each node* $\mathbf{u}$ *according to* $f_{\mathbf{u}}(r)$ *(any node* $\mathbf{u}$ *is the origin of one long range link whose destination is* $\mathbf{v}$ *with probability proportional to* $f_{\mathbf{u}}(\delta(\mathbf{u}, \mathbf{v}))$*), is a navigable small-world.*

Similarly, we say that an infinite family of finite graphs $(H_i)_{i \in \mathcal{I}}$ *is* smallworldizable *if there exists, for each* $\mathbf{u}$ *and* i*, a distribution* $f_{i,\mathbf{u}}(r)$*, such that the family of finite randomly augmented graphs* $(G_i)_{i \in \mathcal{I}}$*, where* G_i *is obtained*

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

by the addition of one random long range link to each node $\mathbf{u}$ *of* H_i *according to* $f_{i,\mathbf{u}}(r)$ *(any node* $\mathbf{u}$ *is the origin of one long range link whose destination is* $\mathbf{v}$ *with probability proportional to* $f_{i,\mathbf{u}}(\delta_{H_i}(\mathbf{u},\mathbf{v}))$*), is a navigable small-world family.*

The following class of graphs 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 a* 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* $\forall r \geq 2, d'_{\mathbf{u}}(r) \leq \alpha/(r \ln r)$*. Similarly, an infinite family of finite uniformly bounded degree graphs* $(H_i)_{i\in\mathcal{I}}$ *is a* moderate growth graphs family *if there exists a uniform constant* $\alpha > 0$*, such that the ball size of each node* $\mathbf{u}$ *of each* H_i *can be written as* $b_{i,\mathbf{u}}(r) = r^{d_{i,\mathbf{u}}(r)}$*, where* $d_{i,\mathbf{u}}(r) : [2,\infty) \to \mathbb{R}$ *is* $\mathcal{C}^1$ *and satisfies* $\forall r \geq 2, d'_{i,\mathbf{u}}(r) \leq \alpha/(r \ln r)$*.*

Note that the function $d_{\mathbf{u}}(r)$ is simply defined as a $\mathcal{C}^1$ interpolation of $(\ln(b_{\mathbf{u}}(r))/\ln r)_{r\in\{2,3,\ldots\}}$, then the $\mathcal{C}^1$ condition is not restrictive by itself. We can now state our main result.

Theorem 1 *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* $q > 1$*. Any infinite moderate growth graphs family* $(H_i)_{i\in\mathcal{I}}$ *is smallworldizable by the addition of one long range link per node, distributed according to* $f_{i,\mathbf{u}}(r) = \frac{1}{b_{H_i,\mathbf{u}}(r)\log^q r}$ *in each graph* H_i*, for any* $q > 1$*.*

Proof. We consider the greedy routing algorithm that, at each step, forwards the message to the closest (in the sense of the δ_H metric) neighbor (in the augmented graph) of the current node. Assume that $\mathbf{s}$ and $\mathbf{t}$ are respectively the source and target. The main argument in Kleinberg's analysis, from which our proof is inspired, is that, among a polylogarithmic number of nodes at distance between r and $r/2$ ($r \geq 2$) of the target, with constant probability, at least one node has a long range link that goes to a node at distance less than $r/2$ from the target, which gives the polylogarithmic path length. We use a similar argument, modified so that the upper bound can be expressed only in terms of the original metric (and not the total size of the graph).

Note that, in order to show that the family $(H_i)_{i\in\mathcal{I}}$ is smallworldizable, we only need to obtain uniform bounds (independent of i) on the expected path length computed by the greedy algorithm on any finite graph H_i. Since the proof is analogous for finite and infinite graphs, we will only focus on the infinite graph case.

We first need to check that the normalization constants $Z_{\mathbf{u}} = \sum_{\mathbf{v}\in V} f_{\mathbf{u}}(\delta_H(\mathbf{u},\mathbf{v}))$ are uniformly bounded (with respect to $\mathbf{u}$ and i), so that the distribution is properly defined. Let $\alpha > 0$ such that $d'_{\mathbf{u}}(r) \leq \alpha/(r \ln r)$. Since the maximum degree is bounded by Δ, integrating the upper bound on $d'_{\mathbf{u}}$ yields a uniform upper bound on $d_{\mathbf{u}}(r)$. There exists a constant $C > 0$ such that, for any $\mathbf{u}$ and $r > 1$,

$$d_{\mathbf{u}}(r) \leq C + \alpha \ln \ln r.$$

This, in turn, implies that all normalization constants $Z_{\mathbf{u}}$ are uniformly bounded. Indeed, we have

$$Z_{\mathbf{u}} = \sum_{r \geq 1} \left(b_{\mathbf{u}}(r) - b_{\mathbf{u}}(r-1)\right) f_{\mathbf{u}}(r)$$

and the upper bound on $d'_{\mathbf{u}}$ implies, for $b_{\mathbf{u}}(r) = \exp(d_{\mathbf{u}}(r) \ln r)$:

$$b_{\mathbf{u}}(r) - b_{\mathbf{u}}(r-1) \leq \left(\frac{1}{r-1} \max d_{\mathbf{u}}(t) + \ln(r) \max d'_{\mathbf{u}}(t) \right) b_{\mathbf{u}}(r),$$

where both maxima are over $t \in [r-1, r]$. Using the previous upper bounds on $d_{\mathbf{u}}$ and $d'_{\mathbf{u}}$, we get

$$\left(b_{\mathbf{u}}(r) - b_{\mathbf{u}}(r-1)\right) f_{\mathbf{u}}(r) \leq \left(\frac{C + \alpha \ln \ln(r)}{r-1} + \frac{\alpha \ln(r)}{(r-1)\ln(r-1)} \right) \frac{1}{\ln^q(r)}$$

and the sum over r of the right-hand side converges to some constant $Z < \infty$.

One can bound the ratio $b_{\mathbf{u}}(r)/b_{\mathbf{u}}(\beta r)$, for any $0 < \beta < 1$, as follows:

$$b'_{\mathbf{u}}(r) = \left(d'_{\mathbf{u}}(r) \ln r + \frac{d_{\mathbf{u}}(r)}{r} \right) b_{\mathbf{u}}(r) \leq \frac{C + \alpha + \alpha \ln \ln r}{r} b_{\mathbf{u}}(r)$$

Integrating the ratio $b'_{\mathbf{u}}(r)/b_{\mathbf{u}}(r)$ between βr and r gives:

$$\ln \left(\frac{b_{\mathbf{u}}(r)}{b_{\mathbf{u}}(\beta r)} \right) \leq -C \ln \beta + \alpha \ln r \ln \ln r - \alpha \ln(\beta r) \ln \ln(\beta r).$$

But, $\ln \ln(\beta r) = \ln \ln r + \ln(1 + \frac{\ln \beta}{\ln r}) \geq \ln \ln r + C' \frac{\ln \beta}{\ln r}$ for some constant $C' > 0$, by concavity of ln. Thus,

$$\ln \left(\frac{b_{\mathbf{u}}(r)}{b_{\mathbf{u}}(\beta r)} \right) \leq -\alpha \ln \beta \ln \ln r - (C + C'\alpha) \ln \beta.$$

We conclude that, for all $\mathbf{u}, r \geq 2$ and $0 < \beta < 1$,

$$b_{\mathbf{u}}(r) \leq \frac{(\ln r)^{-\alpha \ln \beta}}{\beta^{C + C'\alpha}} b_{\mathbf{u}}(\beta r). \tag{1}$$

We now 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 destination of the long range link from $\mathbf{u}$. We give a lower bound on $\mathbb{P}[\delta_H(L_{\mathbf{u}}, \mathbf{t}) \leq r/2]$, 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:

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

We now turn back to the initial question of the length of the greedy path from $\mathbf{s}$ to $\mathbf{t}$. 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 $\gamma = q + \alpha \ln 5$. Thus, the greedy algorithm, once it reaches the k-th shell, examines at most $O(k^\gamma)$ vertices on expectation before it finds one vertex whose long range link leads into a smaller shell. By linearity of expectation, the expected length of the greedy path from a vertex in the k-th shell is $O(k^{1+\gamma})$, with (uniform) constants that can be recovered from the above discussion. As a result, the expected length of the greedy path from $\mathbf{s}$ to $\mathbf{t}$ is polylogarithmic in $\ell = \delta_H(\mathbf{s}, \mathbf{t})$ ($O(\ln^{1+q+\alpha \ln 5} \ell)$), and *a fortiori* in $b_{\mathbf{t}}(\delta_H(\mathbf{s}, \mathbf{t}))$. Thus, the augmented graph is a randomized navigable small-world. □

This theorem 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 $O(\ln^{2+\epsilon} r)$, for any $\epsilon > 0$, on the expected length of the greedy path between any pair of nodes at distance r from each other. In a *vertex-transitive graph*, all balls grow at the same rate, since for any pair of nodes $(\mathbf{u}, \mathbf{v})$, $b_{\mathbf{u}}(r) = b_{\mathbf{v}}(r)$ for any radius r. Recall that a graph is *vertex-transitive* iff for all pair of nodes $(\mathbf{u}, \mathbf{v})$, there exists a one-to-one function σ on the vertices preserving the edges, such that $\sigma(\mathbf{u}) = \mathbf{v}$. Among these graphs, all known Cayley graphs[3] are smallwordizable: either they are covered by our theorem (polynomial expansion means $\alpha = 0$) or the diameter is polylogarithmic (exponential or almost exponential expansion, i.e. $\Omega(2^{ar^b})$ for some $a, b > 0$). Indeed, groups of intermediate ball size, between

[3] 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 $x = g_i y$.

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 [6] and [2] for a state of the art.

Products of small-worlds. A remarkable fact on the small-world property, in Kleinberg's model, is its relative independence of the metric dimension. The expected length of paths computed by the greedy algorithm is indeed unchanged whether the underlying metric is a ring or a very high dimension grid. This motivates the study of products 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* $E_H = \{((f,g),(f,g')) : gg' \in E_G, f \in V_F\} \cup \{((f,g),(f',g)) : g \in V_G, ff' \in E_F\}$.

Note that the cartesian product of two graphs of maximum degrees Δ_F and Δ_G is a graph of maximum degree $\Delta_F + \Delta_G$.

Theorem 2 *Let* F *and* G *two moderate growth infinite graphs, (*$(F_i)_{i\in\mathcal{I}}$ *and* $(G_i)_{i\in\mathcal{I}}$ *two infinite moderate growth graphs families). The cartesian product* $H = F \times G$ *(the family* $(H_i)_{i\in\mathcal{I}} = (F_i \times G_i)_{i\in\mathcal{I}}$*) is smallworldizable by the addition of one long range link per node* $\mathbf{u}$ *according to the distribution* $h_{\mathbf{u}}(r) = 1/(b_{H,\mathbf{u}}(r)\ln^{q'} r)$ *(*$h_{i,\mathbf{u}}(r) = 1/(b_{H_i,\mathbf{u}}(r)\ln^{q'} r)$ *for each graph* H_i*), for all* $q' > q_0$*, for some constant* $q_0 > 0$.

Proof. As in proof of Theorem 1, we only prove the result for infinite graphs, other cases follow. Note that it is unclear whether H is a moderate growth graph (i.e., if $\exists \alpha > 0, \forall \mathbf{u}, \forall r \geq 2, d'_{H,\mathbf{u}}(r) \leq \alpha/(r \ln r)$).
By construction of the graph H, for all $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2)$ in H, $b_{F,\mathbf{u}_1}(r/2) b_{G,\mathbf{u}_2}(r/2) \leq b_{H,(\mathbf{u}_1,\mathbf{u}_2)}(r) \leq b_{F,\mathbf{u}_1}(r) b_{G,\mathbf{u}_2}(r)$. Let α_1 and α_2 such that $d'_{F,\mathbf{u}_1}(r) \leq \alpha_1/(r \ln r)$ and $d'_{G,\mathbf{u}_2}(r) \leq \alpha_2/(r \ln r)$, for all $\mathbf{u}_1$ in F and $\mathbf{u}_2$ in G. From Equation (1), there exists a constant $A > 0$ such that

$$\begin{aligned} b_{F,\mathbf{u}_1}(r) b_{G,\mathbf{u}_2}(r) &\leq A\ (\ln r)^{(\alpha_1+\alpha_2)\ln 2} b_{F,\mathbf{u}_1}(r/2) b_{G,\mathbf{u}_2}(r/2) \\ &\leq A\ (\ln r)^{(\alpha_1+\alpha_2)\ln 2} b_{H,(\mathbf{u}_1,\mathbf{u}_2)}(r) \end{aligned}$$

We first need to check that the normalization constants $Z_{\mathbf{u}} = \sum_{\mathbf{v}\in V_H} h_{\mathbf{u}}(\delta_H(\mathbf{u},\mathbf{v}))$ are uniformly bounded, so that the distribution is properly defined. For $q' > 2 + (\alpha_1+\alpha_2)\ln 2 =_{\mathrm{def}} q_0$,

$$\begin{aligned}
Z_{\mathbf{u}} &= \sum_{\mathbf{v}\in V_H} \frac{1}{b_{H,\mathbf{u}}(\delta_H(\mathbf{u},\mathbf{v}))\ln^{q'}(\delta_H(\mathbf{u},\mathbf{v}))} \\
&\leq \sum_{\mathbf{v}\in V_H} \frac{A\ (\ln(\delta_H(\mathbf{u},\mathbf{v})))^{(\alpha_1+\alpha_2)\ln 2}}{b_{F,\mathbf{u}_1}(\delta_H(\mathbf{u},\mathbf{v}))\ b_{G,\mathbf{u}_2}(\delta_H(\mathbf{u},\mathbf{v}))\ln^{q'}(\delta_H(\mathbf{u},\mathbf{v}))} \\
&\leq A \sum_{\mathbf{v}\in V_H} \frac{(\ln(\delta_F(\mathbf{u}_1,\mathbf{v}_1)))^{\frac{(\alpha_1+\alpha_2)\ln 2 - q'}{2}}}{b_{F,\mathbf{u}_1}(\delta_F(\mathbf{u}_1,\mathbf{v}_1))} \frac{(\ln(\delta_G(\mathbf{u}_2,\mathbf{v}_2)))^{\frac{(\alpha_1+\alpha_2)\ln 2 - q'}{2}}}{b_{G,\mathbf{u}_2}(\delta_G(\mathbf{u}_2,\mathbf{v}_2))} \\
&\leq A \left(\sum_{\mathbf{v}_1\in V_F} \frac{(\ln(\delta_F(\mathbf{u}_1,\mathbf{v}_1)))^{\frac{(\alpha_1+\alpha_2)\ln 2 - q'}{2}}}{b_{F,\mathbf{u}_1}(\delta_F(\mathbf{u}_1,\mathbf{v}_1))}\right) \left(\sum_{\mathbf{v}_2\in V_G} \frac{(\ln(\delta_G(\mathbf{u}_2,\mathbf{v}_2)))^{\frac{(\alpha_1+\alpha_2)\ln 2 - q'}{2}}}{b_{G,\mathbf{u}_2}(\delta_G(\mathbf{u}_2,\mathbf{v}_2))}\right) \\
&\leq A Z_{F,\mathbf{u}_1} Z_{G,\mathbf{u}_2} < A Z_F Z_G =_{\mathrm{def}} Z_H < \infty,
\end{aligned}$$

where $Z_{F,\mathbf{u}_1}$ and $Z_{G,\mathbf{u}_2}$ are respectively the normalizing constants for nodes $\mathbf{u}_1$ in F and $\mathbf{u}_2$ in G with $q = \frac{q'-(\alpha_1+\alpha_2)\ln 2}{2}$, and Z_F and Z_G are the corresponding uniform bounds given in the proof of Theorem 1.

As in the proof of Theorem 1, we lower bound $\mathbb{P}[\delta_H(L_{\mathbf{u}},\mathbf{t}) \leq r/2]$, the probability that the long range contact $L_{\mathbf{u}}$ of a given node $\mathbf{u}$, at distance r to the target $\mathbf{t}$ in H ($\mathbf{t} = (\mathbf{t}_1,\mathbf{t}_2)$ with $\mathbf{t}_1 \in F$ and $\mathbf{t}_2 \in G$), belongs to $\mathcal{B}(\mathbf{t}, r/2)$.

$$\begin{aligned}
\mathbb{P}[\delta_H(L_{\mathbf{u}},\mathbf{t}) \leq r/2] &\geq \frac{1}{Z_H \ln^{q'}(3r/2)} \frac{b_{H,\mathbf{t}}(r/2)}{b_{H,\mathbf{t}}(5r/2)} \\
&\geq \frac{1}{Z_H \ln^{q'}(3r/2)} \frac{b_{F,\mathbf{t}_1}(r/4) b_{G,\mathbf{t}_2}(r/4)}{b_{F,\mathbf{t}_1}(5r/2) b_{G,\mathbf{t}_2}(5r/2)} \\
&\geq \Big(Z_H 10^{2C+C'(\alpha_1+\alpha_2)} \ln^{q'}(3r/2) \ln^{(\alpha_1+\alpha_2)\ln 10}(5r/2)\Big)^{-1} \\
&\geq \Big(Z_H 2^{q'+(\alpha_1+\alpha_2)\ln 10} 10^{2C+C'(\alpha_1+\alpha_2)} \ln^{q'+(\alpha_1+\alpha_2)\ln 10}(r)\Big)^{-1}.
\end{aligned}$$

We conclude as above that the expected path length computed by the greedy algorithm is polylogarithmic in $\delta_H(\mathbf{s},\mathbf{t})$ ($O\Big(\ln^{1+q'+(\alpha_1+\alpha_2)\ln 10}(\delta_H(\mathbf{s},\mathbf{t}))\Big)$ expected length) between $\mathbf{s}$ and $\mathbf{t}$, and then polylogarithmic in $b_{H,\mathbf{t}}(\delta_H(\mathbf{s},\mathbf{t}))$. H is then smallworldizable. □

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 graphs of various sizes.

4 Conclusion and open problems

In this paper, we extend the scheme introduced by Kleinberg [7], that allows to generate small-worlds from a wide variety of graph topologies. Our scheme

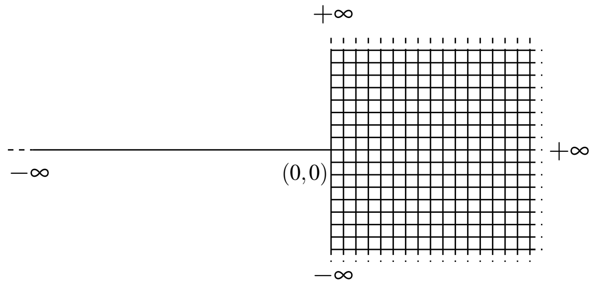


Figure 1. The *infinite fly swatter*

treats in particular all known Cayley graphs. We are also able to "smallworldize" less regular graphs whose ball growths depend on their center and their radius and can grow up to $r^{O(\log\log r)}$ for a radius r, or that are obtained as product of such graphs (e.g. unbalanced d-dimensional tori).

Theorems 1 and 2 capture wide classes of smallwordizable graphs but leave open the question of determining whether any graph is smallworldizable. The condition on the derivative of the exponent $d_{\mathbf{u}}(r)$ might be too restrictive. In proof of Theorem 2, we did not show that the rate of the cartesian product of moderate growth graphs is still moderate, but the resulting graph is however smallwordizable. If we take a closer look at graphs with polynomial ball expansion, some of them do not satisfy the derivative condition but are still smallworldizable using an augmentation with our distribution. To illustrate this idea, we give an example of an extreme graph containing two distinct parts of different density (or two types of polynomial expansion) connected only by a single node: the *infinite fly swatter*. This graph is defined by a half infinite chain connected to a half infinite square lattice. This graph has vertex set $(\mathbb{Z}^-_{<0} \times \{0\}) \bigcup (\mathbb{Z}^+_{\geq 0} \times \mathbb{Z})$ and two nodes are adjacent if the Manhattan distance is equal to one. We give some hints (but do not provide a real proof) on why the augmented infinite fly swatter is a navigable small-world. This graph is augmented according to the distribution $1/(r \log^q r)$ on the chain part, and $1/(r^2 \log^q r)$ on the lattice part, for some $q > 1$, which corresponds to our distribution in Theorem 1 on each respective subgraphs. Informally, the chain does not significantly disturb the greedy routing within the half plane. The tough part is to find a short path from $(0,0)$ to $(-n,0)$. The lattice part is indeed the attractive part of the graph. Let us consider the current node $\mathbf{u} = (-m,0)$ and the destination $\mathbf{t} = (-n,0)$ belonging to the chain with $n > m$. The size of the balls centered on $\mathbf{u}$ are $b_{\mathbf{u}}(r) = 2r+1$ for $r \leq m$ and else $b_{\mathbf{u}}(r) = 2m + (r-m) + (r-m+1)^2$. One can check that $d_{\mathbf{u}}(r) = \frac{\ln b_{\mathbf{u}}(r)}{\ln r}$ is close to 1 for $r = m$ and close to 2 for $r = 2m$. It follows that there exists $m \leq r \leq 2m$ such that $d'_{\mathbf{u}}(r) = \Omega(1/m)$ and that the infinite fly swatter does not fall into the smallwordizable graphs dealt with Theorem 1: $d'_u(t)$ cannot remain lower than $O(\frac{\alpha}{r \ln r})$ for $m \in [m, 2m]$.

However, following the same reasoning as in the proof of Theorem 1, we can first check that the partition functions Z_u are uniformly bounded. Moreover, for the destination $\mathbf{t} = (-n, 0)$, with probability $\Omega(\log^{-q} m)$, there exists a long-range link leading to a node $(-m', 0)$ with $m' \in [3m/2, 2m]$. It follows that after visiting a logarithmic number of nodes, using the greedy algorithm, the message will traverse with constant probability a long-range link and almost double its distance from the origin. This situation occurs roughly $\log n$ times before the message will be very close to $\mathbf{t}$. Of course, others cases occur but they can be reduced to the above discussion.

More generally, other classes of graphs should be smallwordizable: we suspect that smallwordizable graphs could also include subgraphs of d-dimensionnal grids or tori. The characterization of smallworldizable graphs is also open whenever the expansion of balls is locally exponential or almost exponential. A reasonable and attractive case is the one of the family of vertex-transitive graphs due to the homegeneity of balls expansion of these graphs.

References

[1] 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.

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

[3] D. Coppersmith, D. Gamarnik, and M. Sviridenko. The diameter of a long-range percolation graph. *Random Struct. Algorithms*, 21(1):1–13, 2002.

[4] P. Fraigniaud. Greedy routing in tree-decomposed graphs: a new perspective on the small-world phenomenon. Technical Report LRI-1397, Univ. Paris Sud, feb. 2005.

[5] 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.

[6] R.I. Grigorchuk. On growth in group theory. In *Proceedings of the International Congress of Mathematicians, Kyoto*, pages 325–338, 1990.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

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

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

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

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