A Doubling dimension Threshold !(log log n) for augmented graph navigability

Pierre Fraigniaud, Emmanuelle Lebhar, Zvi Lotker

Laboratoire de l'Informatique du Parallélisme

École Normale Supérieure de Lyon
Unité Mixte de Recherche CNRS-INRIA-ENS LYON-UCBL nº 5668

A Doubling Dimension Threshold $\Theta(\log\log n)$ *for Augmented Graph Navigability*

Pierre Fraigniaud
Emmanuelle Lebhar
Zvi Lotker

Février 2006

Research Report Nº 2006-09

École Normale Supérieure de Lyon
46 Allée d'Italie, 69364 Lyon Cedex 07, France
Téléphone : +33(0)4.72.72.80.37
Télécopieur : +33(0)4.72.72.80.80
Adresse électronique : `lip@ens-lyon.fr`

A Doubling Dimension Threshold $\Theta(\log \log n)$ for Augmented Graph Navigability

Pierre Fraigniaud
Emmanuelle Lebhar
Zvi Lotker

Février 2006

Abstract

In his seminal work, Kleinberg showed how to augment meshes using random edges, so that they become navigable; that is, greedy routing computes paths of polylogarithmic expected length between any pairs of nodes. This yields the crucial question of determining wether such an augmentation is possible for all graphs. In this paper, we answer negatively to this question by exhibiting a threshold on the doubling dimension, above which an infinite family of Cayley graphs cannot be augmented to become navigable whatever the distribution of random edges is. Precisely, it was known that graphs of doubling dimension at most $O(\log \log n)$ are navigable. We show that for doubling dimension $\gg \log \log n$, an infinite family of Cayley graphs cannot be augmented to become navigable. Our proof relies on a result of independent interest: we show that the analysis of greedy routing worst performances on graphs augmented by arbitrary distributions of edges can be reduced to the analysis assuming a symmetric distribution. Finally, we complete our result by studying square meshes, that we prove to always be augmentable to become navigable.

Keywords: doubling dimension, small world, greedy routing, Cayley graphs.

Résumé

Kleinberg a montré comment augmenter les grilles par des liens aléatoires de façon à ce qu'elles deviennent navigables; c'est-à-dire que le routage glouton calcule des chemins de longueur polylogarithmique en espérance entre toute paire de noeuds. Cela conduit à la question cruciale de déterminer si une telle augmentation est possible pour tout graphe. Dans cet article, nous répondons négativement à cette question en exhibant un seuil sur la dimension doublante, au-dessus duquel une famille infinie de graphes de Cayley ne peut pas être augmentée pour devenir navigable, quelle que soit la distribution de liens. Précisément, il était connu que les graphes de dimension doublante au plus $O(\log \log n)$ sont navigable. Nous montrons que pour une dimension doublante $\gg \log \log n$, une famille infinie de graphes de Cayley ne peut être augmentée pour devenir navigable. Notre preuve repose sur un résultat d'intérêt indépendant: nous montrons que l'analyse des pires performances du routage glouton sur les graphes augmentés par des distributions arbitraires peut être réduite à son analyse sous l'hypothèse de distributions symétriques. Enfin, nous complétons notre résultat en étudiant les grilles que nous démontrons pouvoir toujours être augmentées pour devenir navigables.

Mots-clés: dimension doublante, petit monde, routage glouton, graphes de Cayley.

1 Introduction

The *doubling dimension* [5, 15, 17] appeared recently as a key parameter for measuring the ability of networks to support efficient algorithms [9, 16] or to realize specific tasks efficiently [1, 2, 8, 25]. Roughly speaking, the doubling dimension of a graph G is the smallest d such that, for any integer $r \geq 1$, and for any node $u \in V(G)$, the ball $B(u, 2r)$ centered at u and of radius $2r$ can be covered by at most 2^d balls $B(u_i, r)$ centered at nodes $u_i \in V(G)$. (This definition can be extended to any metric, and, for instance, $\mathbf{Z}^d$ with the ℓ_1 norm is of doubling dimension d). In particular, the doubling dimension has an impact on the analysis of the small world phenomenon [23], precisely on the expected performances of greedy routing in *augmented* graphs [19].

An augmented graph is a pair (G, φ) where G is an n-node graph, and φ is a collection of probability distributions $\{\varphi_u, u \in V(G)\}$. Every node $u \in V(G)$ is given an extra link pointing to some node v, called the *long range contact* of u. The link from a node to its long range contact is called a *long range link*. The original links of the graph are called *local links*. The long range contact of u is chosen at random according to φ_u as follows : $\Pr\{u \to v\} = \varphi_u(v)$. Greedy routing in (G, φ) is the oblivious routing protocol where the routing decision taken at the current node u for a message of destination t consists in (1) selecting a neighbor v of u that is the closest to t according to the distance in G (this choice is performed among all neighbors of u in G and the long range contact of u), and (2) forwarding the message to v. This process assumes that every node has a knowledge of the distances in G, or at least a good approximation of them. On the other hand, every node is unaware of the long range links added to G, but its own long range link. Hence the nodes have no notion of the distances in the augmented graph. Note that the knowledge of the distances in the underlying graph G is a reasonable assumption when, for instance, G is a network in which distances can be computed from the coordinates of the nodes (e.g., in meshes, as in [19]).

An infinite family of graphs $\mathcal{G} = \{G^{(i)}, i \in I\}$ is *navigable* if there exists a family $\Phi = \{\varphi^{(i)}, i \in I\}$ of collections of probability distributions, and a function $f(n) \in O(\text{polylog}(n))$ such that, for any $i \in I$, greedy routing in $(G^{(i)}, \varphi^{(i)})$ performs in at most $f(n^{(i)})$ expected number of steps where $n^{(i)}$ is the order of the graph $G^{(i)}$. More precisely, for any pair of nodes (s, t) of $G^{(i)}$, the expected number of steps $\mathbb{E}(\varphi^{(i)}, s, t)$ for traveling from s to t using greedy routing in $(G^{(i)}, \varphi^{(i)})$ is at most $f(n^{(i)})$.

In his seminal paper, Kleinberg [19] proved that, for any fixed integer $d \geq 1$, the family of d-dimensional meshes is navigable. Duchon et al [10] generalized this result by proving that any infinite family of graphs with bounded growth is navigable. Fraigniaud [13] proved that any infinite family of graphs with bounded treewidth is navigable. Finally, Slivkins [25] recently related navigability to doubling dimension by proving that any infinite family of graphs with doubling dimension at most $O(\log \log n)$ is navigable. All these results naturally lead to the question of whether all graphs are navigable.

Let $\delta : \mathbf{N} \mapsto \mathbf{N}$, let $\mathcal{G}_{n,\delta(n)}$ be the class of n-node graphs with doubling dimension at most $\delta(n)$, and let $\mathcal{G}_\delta = \cup_{n \geq 1} \mathcal{G}_{n,\delta(n)}$. By rephrasing Slivkins result [25], we get that $\mathcal{G}_\delta$ is navigable for any function δ bounded from above by $c \log \log n$ for some constant $c > 0$. This however lets open the case of graphs of larger doubling dimensions, namely the cases of all families $\mathcal{G}_\delta$ where δ is satisfying $\delta(n) \gg \log \log n$. This paper analyzes the navigability of graphs in $\mathcal{G}_\delta$ for the entire spectrum $1 \leq \delta(n) \leq \log n$.

1.1 Our results

We prove a threshold of $\delta(n) = \Theta(\log \log n)$ for the navigability of $\mathcal{G}_\delta$. More precisely, we focus our analysis on the Cayley graphs in $\mathcal{G}_\delta$, i.e., the graphs whose nodes are the elements of a group and the edges are determined by the action of a generating set of this group on the nodes. Cayley graphs form a large class of graphs including graphs as diverse as rings, cliques, toruses, hypercubes, wrapped butterflies, as well as all graphs defined from permutation subgroups (e.g., star-graphs[1], pankace-graphs, etc.), cf [6]. Cayley graphs are highly symmetric, and can be used to build explicit expanders [24].

Let $\text{Cay}_{n,\delta(n)} \subseteq \mathcal{G}_{n,\delta}$ be the class of n-node Cayley graphs with doubling dimension at most $\delta(n)$, and let $\text{Cay}_\delta = \cup_{n \geq 1} \text{Cay}_{n,\delta(n)}$. Note that $\text{Cay}_\delta \subseteq \mathcal{G}_\delta$. We prove a threshold phenomenon for Cay_δ : bellow a certain function δ, Cay_δ is navigable, while above it Cay_δ is not navigable. More precisely, we prove that, for any function δ satisfying $\lim_{n \to \infty} (\log \log n)/\delta(n) = 0$, Cay_δ is not navigable. This implies, in particular, that $\mathcal{G}_\delta$ is not navigable for $\log \log n \ll \delta(n)$. Hence, the result in [25] is essentially the best that can be achieved by considering only the doubling dimension of graphs.

[1]Not to be mistaken for $K_{1,n}$, see [3].

Hence, we get the following view of the situation related to the doubling dimension :

$$\begin{array}{ccccc} & \mathcal{G}_\delta \text{ navigable [25]} & & \mathcal{G}_\delta \text{ not navigable [This paper]} & \\ 1 & -\,-\,-\,-\,-\,-\,-\,- & \log\log n & -\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,- & \log n \end{array}$$

Our negative result requires to prove that, for *any* distribution of the long range links, greedy routing has expected performances strictly greater than polylogarithmic. For this purpose, we prove a reduction result that has interest on its own. Namely, we show that the analysis of greedy routing in Cayley graphs can be restricted to *symmetric* probability distributions. A distribution of choice of the long range contacts is symmetric if, for any node γ, $\Pr\{u \to u\gamma\}$ is independent from u (since the nodes of a Cayley graph are elements of a group, the node $u\gamma$ is just the node corresponding to the combination of the two elements u and γ by the operation of the group — in this paper we use the multiplicative notation). We prove that if G is a Cayley graph then, for any distribution φ, there exists a symmetric distribution $\widehat{\varphi}$ such that the expected length of a path computed by the greedy routing in $(G, \widehat{\varphi})$ is at most the maximum expected length of a path computed in (G, φ), that is for any source s and any target t, $\mathbb{E}(\widehat{\varphi}, s, t) \leq \max_{s',t' \in V(G)} \mathbb{E}(\varphi, s', t')$.

Finally, we prove a somehow counter intuitive result by showing that a subgraph of a non navigable graph is not necessarily non navigable. In particular, we show that all square meshes are navigable, for all dimensions. Specificaly, we prove that although the family of non navigable Cayley graphs that we use to disprove the navigability of all graphs contains the standard square meshes of dimension δ as subgraphs, these latter family of graphs is navigable.

1.2 Related works

Kleinberg showed that greedy routing performs in $O(\log^2 n)$ steps on d-dimensional meshes augmented by the d-harmonic distribution. Since then, several results have been developed to tighten the analysis of greedy routing on randomly augmented networks. Precisely, Barrière et al. [7] showed that greedy routing expected number of steps is $\Theta(\log^2 n)$ in the d-dimensional meshes augmented by the d-harmonic distribution. In the special case of rings, Aspnes et al. [4] proved a lower bound of $\Omega(\log^2 n/\log\log n)$ expected number of steps for *any* edge distribution. For paths, Flammini et al. [11] recently showed a lower bound of $\Omega(\log^2 n)$ expected number of steps in the case of symmetric and distance monotonic edge distributions. Martel and Nguyen [22] showed however that the diameter of these networks augmented by the harmonic distribution is $\Theta(\log n)$. In another perspective, several authors developed new decentralized algorithms for the d-dimensional mesh augmented by the d-harmonic distribution. Lebhar and Schabanel [20] presented a decentralized algorithm which performs in $O(\log n(\log\log n)^2)$ expected number of steps in this graph. The algorithm Neighbor-Of-Neighbor presented by Manku et al. [21] performs in $O(\frac{1}{k\log k}(\log n)^2)$ expected number of steps, where k is the number of long range links per node in the mesh. Assuming some extra knowledge on the long range links, Fraigniaud et al. [14] described an oblivious routing which performs in $O((\log n)^{1+1/d})$ expected number of steps. Finally, Martel and Nguyen [22] presented a non oblivious routing protocol achieving the same performances under the same assumption as in [14].

1.3 Organization of the paper

The paper is organized as follows : in Section 2, we prove that the analysis of greedy routing in augmented Cayley graphs can be restricted to symmetric probability distributions. Section 3 presents the main result of the paper by exhibiting the non navigability of Cayley graphs of doubling dimension $\gg \log\log n$. Finally, in Section 4, we study the special case of square meshes, and prove that they are all navigable.

2 Symmetric augmentations

In this section, we prove a lemma that will be used later to prove our main result. This lemma is of independent interest as it proves that, without any loss of generality, one can restrict the analysis of greedy routing in Cayley graphs to a specific class of probability distributions. Recall that a Cayley graph G is defined by a pair (Γ, S) where Γ is a group, and S a generating set of Γ. We have $V(G) = \Gamma$, and $(u, v) \in E(G)$ if and only if $u^{-1}v \in S$. If $S^{-1} = S$ then G is non directed, otherwise it is directed. In this paper we use the multiplicative notation, and hence the neutral element of Γ is denoted by $\mathbf{1}$.

Definition 1 *An augmentation (G, φ) of a Cayley graph G is symmetric iff for any $\gamma \in \Gamma$, $\varphi_u(u\gamma) = \varphi_{u'}(u'\gamma)$ for any pair of nodes u and u'.*

In other words, an augmentation (G, φ) is symmetric iff for any fixed γ, $\Pr\{u \to u\gamma\}$ is independent from u. For instance, if the long range contact of a node is chosen uniformly at random among the node in $V(G)$, this augmentation is symmetric. Symmetric augmentations play an important role as far as Cayley graphs are concerned. Indeed, we have the following.

Lemma 1 *For any augmentation (G, φ) of a Cayley graph $G = (\Gamma, S)$, there exists a symmetric augmentation $(G, \widehat{\varphi})$ for which greedy routing has expected performance upper bounded by the maximum expected performance of greedy routing in (G, φ), taken over all pairs of source and target.*

Proof. Let (G, φ) be an augmentation of a Cayley graph $G = (\Gamma, S)$ of order n. For any $x \in V(G)$, let $\sigma_x : V(G) \mapsto V(G)$ defined as $\sigma_x(u) = xu$. It is folklore that σ_x is an automorphism of G, i.e., a one-to-one and onto mapping that preserves adjacency. In particular, for any path $P = (u_0, u_1, \dots, u_k)$ from a node s to a node t, $\sigma_x(P) = xP = (xu_0, xu_1, \dots, xu_k)$ is a path from xs to xt. Given $\varphi = \{\varphi_u, u \in V(G)\}$, we define, for any $x \in V(G)$, the set $\varphi^{(x)} = \{\varphi_u^{(x)}, u \in V(G)\}$ where, $\varphi_u^{(x)} = \varphi_{xu} \circ \sigma_x$. I.e. for all $v \in V$,

$$\varphi_u^{(x)}(v) = \varphi_{xu}(xv).$$

Since V is finite, we can label the set of nodes V as $V = \{u_1, \dots, u_n\}$. From φ, we define a symmetric distribution $\widehat{\varphi}$ as follows :

$$\begin{aligned} &\widehat{\varphi} = \{\varphi_u^{(x)}, u \in V(G)\}, \\ &\text{where} \quad x = u_X, \end{aligned}$$

and X is the uniform random variable that takes value i with probability $\frac{1}{n}$ for all $i \in \{1, \dots, n\}$.

Intuitively, $\widehat{\varphi}$ corresponds to picking a random origin $x = u_X$ and reconstructing distribution φ by translating it by σ_x.

This augmentation is symmetric. Indeed we show that for all $u', u, v \in V$, $\widehat{\varphi}_u(uv) = \widehat{\varphi}_{u'}(u'v)$. It is enough to show that for all $u, v \in V$, $\widehat{\varphi}_{\mathbf{1}}(v) = \widehat{\varphi}_u(uv)$, where $\mathbf{1}$ is the unit element of G.

Since X is a uniform random variable it follows that for all $u, v \in V$,

$$\Pr\{u_X = v\} = \Pr\{u_X = uv\}.$$

The proof then follows from the definition of $\widehat{\varphi}$:

$$\begin{aligned} \widehat{\varphi}_{\mathbf{1}}(v) &= \varphi_{\mathbf{1}}^{(u_X)}(v) \\ &= \sum_{i=1}^{n} \varphi_{\mathbf{1}}^{(u_i)}(v) \Pr\{X = i\} \\ &= \sum_{i=1}^{n} \varphi_{u_i}(u_i v) \Pr\{X = i\} \\ &= \sum_{x \in V} \varphi_x(xv) \cdot \Pr\{u_X = x\} \\ &= \sum_{y \in V} \varphi_{yu}(yuv) \cdot \Pr\{u_X = yu\} \\ &= \sum_{y \in V} \varphi_{yu}(yuv) \cdot \Pr\{u_X = y\} \\ &= \sum_{y \in V} \varphi_u^{(y)}(uv) \cdot \Pr\{u_X = y\} \\ &= \widehat{\varphi}_u(uv). \end{aligned}$$

The equality of the fifth line follows from the facts that σ_y is an automorphism of G.

To complete the proof, we show that greedy routing expected performances on G augmented by $\widehat{\varphi}$ are at most the maximum expected performance of greedy routing in G augmented by φ. For any source s and target t in V, recall that $\mathbb{E}(\phi, s, t)$ denotes the expected length of the path computed by greedy routing from s to t in G augmented by distribution ϕ. We have :

$$\mathbb{E}(\widehat{\varphi}, s, t) = \sum_{x \in V} \Pr\{u_X = x\} \cdot \mathbb{E}(\varphi^{(x)}, s, t),$$

since, once $u_X = x$, the distribution is $\varphi^{(x)}$ for all the nodes. Hence,

$$\begin{aligned}\mathbb{E}(\widehat{\varphi}, s, t) &= \sum_{x \in V} \Pr\{u_X = x\} \cdot \mathbb{E}(\varphi, xs, xt) \\ &\leq \left(\max_{x \in V} \mathbb{E}(\varphi, xs, xt)\right) \left(\sum_{x \in G} \Pr\{u_X = x\}\right) \\ &\leq \max_{s', t' \in V} \mathbb{E}(\varphi, s', t').\end{aligned}$$

□

3 Non navigable graphs

In this section, we prove that the result in [25] is essentially the best that can be achieved as far as doubling dimension is concerned.

Theorem 1 *Let $\delta : \mathbf{N} \mapsto \mathbf{N}$ be such that $\lim_{n \to \infty} \frac{\log \log n}{\delta(n)} = 0$. Then $\mathcal{G}_\delta$ is not navigable.*

Informally, the argument of the proof is that a doubling dimension $\gg \log \log n$ implies that the number of possible "directions" where a random link can go is greater than any polylogarithm of n. Therefore, for any source, there always exists a target such that the expected path length of greedy routing is greater than any polylogarithm of n, because of the low probability of a random link to go in the right direction. **Proof.** We show that there exists an infinite family of graphs $\{G^{(n)}, n \geq 1\}$ indexed by their number of vertices, such that $G^{(n)} \in \mathcal{G}_{n,\delta(n)}$ and for any family $\Phi = \{\varphi^{(n)}, n \geq 1\}$ of collections of probability distributions, greedy routing in $(G^{(n)}, \varphi^{(n)})$ performs in an expected number of steps $t(n) \notin O(\text{polylog}(n))$ for some couples of source and target.

Let $d : \mathbf{N} \mapsto \mathbf{N}$ be such that $d \leq \delta$, $\lim_{n \to \infty} \frac{\log \log n}{d(n)} = 0$, and $\lim_{n \to \infty} \frac{d(n)}{\sqrt{\log n}} = 0$. For the sake of simplicity, assume that $p = n^{1/d(n)}$ is integer. $G^{(n)}$ is the graph of n nodes consisting of $p^{d(n)}$ nodes labeled $(x_1, \ldots, x_{d(n)})$, $x_i \in \mathbf{Z}_p$. Node $(x_1, \ldots, x_{d(n)})$ is connected to all nodes $(x_1 + a_1, \ldots, x_{d(n)} + a_{d(n)})$ where $a_i \in \{-1, 0, 1\}$, $i = 1, \ldots, d(n)$, and all operations are taken modulo p (cf. Figure 1). Note that the diameter of this graph is $\lfloor p/2 \rfloor$. Note also that, by construction of $G^{(n)}$, the distance between two nodes $y = (y_1, \ldots, y_{d(n)})$ and $z = (z_1, \ldots, z_{d(n)})$ is $\max_{1 \leq i \leq d(n)}(|y_i - z_i|, p - |y_i - z_i|)$.

Claim 1 $G^{(n)} \in \mathcal{G}_{n,\delta(n)}$.

Clearly $G^{(n)}$ has n nodes. We prove that $G^{(n)}$ has doubling dimension $d(n)$, therefore at most $\delta(n)$. Let $\mathbf{0} = (0, \ldots, 0)$. The ball $B(\mathbf{0}, 2r)$ can be covered by $2^{d(n)}$ balls of radius r, centered at the $2^{d(n)}$ nodes $(x_1, \ldots, x_{d(n)})$, $x_i \in \{-r, +r\}$ for any $i = 1, \ldots, d(n)$. Hence the doubling dimension of $G^{(n)}$ is at most $d(n)$. On the other hand, $|B(\mathbf{0}, 2r)| = (4r+1)^{d(n)}$ and $|B(\mathbf{0}, r)| = (2r+1)^{d(n)}$. Thus at least $(4r+1)^{d(n)}/(2r+1)^{d(n)}$ balls are required to cover $B(\mathbf{0}, 2r)$, since in $G^{(n)}$, for any node u and radius r, $|B(u, r)| = |B(\mathbf{0}, r)|$. This ratio can be rewritten as $2^{d(n)}(1 - \frac{1}{2(2r+1)})^{d(n)}$. For $2r = \frac{n^{1/d(n)}}{5}$, since

$d(n) \le \sqrt{\log n}$, we get that $(2r+1) > \frac{2^{\sqrt{\log n}}}{5} > d(n)$ for $n \ge n_0$, $n_0 \ge 1$. Then, for $n \ge n_0$,

$$\begin{aligned}\left(1 - \frac{1}{2(2r+1)}\right)^{d(n)} &> \left(1 - \frac{1}{2d(n)}\right)^{d(n)} \\ &= 2^{d(n)\log\left(1-\frac{1}{2d(n)}\right)} \\ &\ge 2^{d(n)\left(-\frac{1}{2d(n)}-\frac{4}{4(d(n))^2}\right)} = 2^{-\frac{1}{2}-\frac{1}{d(n)}}\end{aligned}$$

There exists $n_1 \ge n_0$, such that $2^{-\frac{1}{2}-\frac{1}{d(n)}} > \frac{1}{2}$ for $n \ge n_1$. Then, for $n \ge n_1$, $|B(\mathbf{0}, 2r)|/|B(\mathbf{0}, r)| > 2^{d(n)-1}$. Thus the doubling dimension of $G^{(n)}$ is at least $d(n)$, which proves the claim.

Consider now a distribution $\varphi^{(n)}$ that belongs to some given collection of probability distributions $\Phi = \{\varphi^{(n)}, n \ge 1\}$. From Lemma 1, we can assume that $\varphi^{(n)}$ is symmetric since, if $\varphi^{(n)}$ is not symmetric, then one can construct a symmetric distribution $\widehat{\varphi}^{(n)}$ achieving better maximum expected performances for greedy routing in $G^{(n)}$. Thus, in order to prove a lower bound on the maximum expected performances of greedy routing, we can restrict our analysis to symmetric distributions.

Definition 2 *For any node $u = (u_1, \ldots, u_{d(n)})$, and for any $D = (\nu_1, \ldots, \nu_{d(n)}) \in \{-1, 0, +1\}^{d(n)}$, we call* direction *the set of nodes*

$$dir_u(D) = \{v = (v_1, \ldots, v_{d(n)}) : v_i = (u_i + \nu_i \cdot x) \bmod p, \ 1 \le x \le \lfloor p/2 \rfloor\}.$$

Note that, for any u, the directions $\mathrm{dir}_u(D)$ for $D \in \{-1, 0, +1\}^{d(n)}$ partition the nodes of $G^{(n)}$ (see Figure 1). There are obviously $3^{d(n)}$ directions, and the $2^{d(n)}$ directions defined on $\{-1, +1\}^{d(n)}$ have all the same cardinality.

At each node u, and for each of its direction $\mathrm{dir}_u(D)$ for $D \in \{-1, 0, +1\}^{d(n)}$, one can associate a probability $p_{\mathrm{dir}_u(D)}$ to have a long-range link in its direction $\mathrm{dir}_u(D)$. Formally, $p_{\mathrm{dir}_u(D)} = \sum_{v \in \mathrm{dir}_u(D)} \varphi_u^{(n)}(v)$. Since $\varphi^{(n)}$ is symmetric, $p_{\mathrm{dir}_u(D)}$ is independent of its origin, that is : $p_{\mathrm{dir}_u(D)} = p_{\mathrm{dir}_v(D)}$ for any two nodes u, v. Hence, the index u or v can be omitted while considering these probabilities.

Among directions $\mathrm{dir}(D)$ for $D \in \{-1, +1\}^{d(n)}$, we say that $\varphi^{(n)}$ *disadvantages* the direction $\mathrm{dir}(D)$ if $p_{\mathrm{dir}(D)} = \min_{D' \in \{-1,+1\}^{d(n)}} p_{\mathrm{dir}(D')}$. Without loss of generality, we can assume that $\varphi^{(n)}$ disadvantage direction $\mathrm{dir}(+1, +1, \ldots, +1)$ at every node. Since $\sum_{D \in \{-1,+1\}^{d(n)}} p_{\mathrm{dir}(D)} \le 1$, $p_{\mathrm{dir}(+1,+1,\ldots,+1)}$ is at most $1/2^{d(n)}$.

We analyze the performance of greedy routing from source $s = (0, \ldots, 0) = \mathbf{0}$ to target $t = (\lfloor p/2 \rfloor, \ldots, \lfloor p/2 \rfloor)$, at mutual distance $\lfloor p/2 \rfloor$.

We show that distribution $\varphi^{(n)}$ leaves the expected number of steps of greedy routing from s to t above any polylogarithm of n. For this purpose, we prove the following claim.

Claim 2 *If the current node reached by greedy routing on its way from s to t is of the form $x = (k, k, \ldots, k)$, where $0 \le k < \lfloor p/2 \rfloor$, then with probability at least $1 - 1/2^{d(n)}$, greedy routing follows the local link to node $(k+1, k+1, \ldots, k+1)$.*

Recall that greedy routing follows the link of x which leads the closest to the target among its local and long range links. Among its local contacts, the closest node to the target is $(k+1, k+1, \ldots, k+1)$, which reduces the current distance to the target by 1. Therefore, for the algorithm to follow the long range link of x, the long range contact of x has to have all its coordinates strictly greater than k in order to reduce the current distance to the target by at least 1. This means that the long range contact has to be in direction $\mathrm{dir}_x(+1, +1, \ldots, +1)$, which, from the discussion above, happens with probability at most $1/2^{d(n)}$. This proves the claim.

The following claim gives a lower bound on the expected number of steps of greedy routing from s to t.

Claim 3 *There exists $n_0 \ge 1$ such that for any $n \ge n_0$, $\mathbb{E}(\varphi^{(n)}, s, t) \ge 2^{d(n)}$.*

Let $\mathcal{E}_k$ be the event : "greedy routing follows a long range link at the kth step". From Claim 2, $\Pr \mathcal{E}_1 \le \frac{1}{2^{d(n)}}$. All nodes visited by greedy routing are distinct, and thus, their random links are independent.

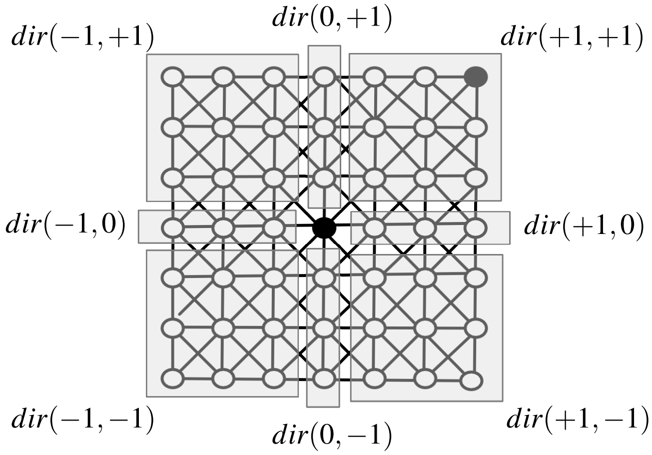

FIG. 1 – Example of graph $G^{(n)}$ defined in proof of Theorem 1 with $d(n) = 2$. Source and target are black nodes. Grey areas represent the various directions.

Therefore, still from Claim 2, $\Pr\{\mathcal{E}_k \,|\, \forall j < k,\, \overline{\mathcal{E}_j}\} \leq \frac{1}{2^{d(n)}}$. Let X be the number of steps taken by greedy routing before it follows a long range link on its way from s to t. We get :

$$\mathbb{E}(X) = \sum_{k\geq 1} \Pr\{X \geq k\} \geq \sum_{k\geq 1} \Pr\{\bigcap_{i=1}^{k-1} \overline{\mathcal{E}_j}\} \geq \sum_{k\geq 1} \left(1 - \frac{1}{2^{d(n)}}\right)^{k-1} = 2^{d(n)}.$$

On the other hand, since $d(n) \ll \sqrt{\log n}$, this expected number of steps is less than the distance $\lfloor p/2 \rfloor$ from s to t. Indeed, there exists $n_0 \geq 1$ such that, for any $n \geq n_0$, $d(n) < \sqrt{\frac{\log n}{2}}$. Then for $n \geq n_0$, $(d(n))^2 < \frac{\log n}{2}$ and $2d(n) < \frac{\log n}{d(n)}$. We get :

$$\frac{2^{d(n)}}{n^{1/d(n)}/2} = 2^{1+d(n)-\frac{\log n}{d(n)}} \leq 2^{1-d(n)} < 1.$$

Then, for $n \geq n_0$, after the $2^{d(n)} < \lfloor p/2 \rfloor$ first steps of greedy routing from s to t in $G^{(n)}$, no long range link is used. Therefore greedy routing has not yet reached the target. Thus $\mathbb{E}(\varphi^{(n)}, s, t) \geq 2^{d(n)}$.

We complete the proof of the theorem by proving the following claim.

Claim 4 $2^{d(n)} \notin O(\text{polylog } n)$.

Let $\alpha \geq 1$, we have :

$$\frac{\log^\alpha n}{2^{d(n)}} = 2^{\alpha \log\log n - d(n)} = 2^{\alpha\, d(n)\left(\frac{\log\log n}{d(n)} - \frac{1}{\alpha}\right)}.$$

Since $\lim_{n\to\infty} \frac{\log\log n}{d(n)} = 0$, there exists $n_1 \geq 1$ such that for any $n \geq n_1$, $\left(\frac{\log\log n}{d(n)} - \frac{1}{\alpha}\right) \leq -\frac{1}{2\alpha}$, and thus $\frac{\log^\alpha n}{2^{d(n)}} \leq 2^{-d(n)/2}$. Moreover, $d(n) \geq \log\log n$, then, for $n \geq n_1$,

$$\frac{\log^\alpha n}{2^{d(n)}} \leq 2^{-\frac{\log\log n}{2}} = o(1).$$

In other words, $2^{d(n)}$ is not a polylogarithm of n, which proves the claim.

Finally, for $n \geq \max\{n_0, n_1\}$, $\max_{s',t' \in V(G^{(n)})} \mathbb{E}(\varphi^{(n)}, s', t') \geq \mathbb{E}(\varphi^{(n)}, s, t) \geq 2^{d(n)} \notin O(\text{polylog } n)$, which yields the result. $\square$

4 Navigability of meshes

The family of non navigable graphs defined in the proof of Theorem 1 contains the standard square meshes of dimension $d(n)$ as subgraphs, where $d(n) \gg \log\log n$. Nevertheless, and somehow counter intuitively, a subgraph of a non navigable graph is not necessarily non navigable. In this section, we illustrate this phenomenon by focusing on the special case of d-dimensional meshes, the first graphs that were considered for the analysis of navigable networks [19]. Precisely, we show that any d-dimensional torus $C_{n^{1/d}} \times \ldots C_{n^{1/d}}$ is navigable : either it has a polylogarithmic diameter, or it admits a distribution of links such that greedy routing computes paths of polylogarithmic length. This result has partially been proven in [11] for the case of constant dimensions. We give here a complete proof that holds for any dimension.

Theorem 2 *For any function positive function $d(n)$, the n nodes $d(n)$-dimensional torus is navigable.*

Proof. We construct a random link distribution φ as follows. Let $u = (u_1, \ldots, u_{d(n)})$ and $v = (v_1, \ldots, v_{d(n)})$ be two nodes. If they differ in more than one coordinate, then $\varphi_u(v) = 0$; otherwise, i.e. they differ in only one coordinate, say the ith, then :

$$\varphi_u(v) = \frac{1}{d(n)} \cdot \frac{1}{2H_k} \cdot \frac{1}{|u_i - v_i|},$$

where $k = \frac{n^{1/d}}{2}$ and $H_k = \sum_{j=1}^{k} \frac{1}{j}$ is the harmonic sum. Note that this distribution corresponds to :

- picking a dimension uniformly at random (probability $\frac{1}{d(n)}$ to pick dimension i)
- and to draw a long-range link on this axis according to the 1-harmonic distribution over distances ($\frac{1}{2H_k}$ is the normalizing coefficient for this distribution), which is the distribution given by Kleinberg to make the 1-dimensional torus navigable.

Let now $s = (s_1, \ldots, s_{d(n)})$ and $t = (t_1, \ldots, t_{d(n)})$ be a pair of source and target in the mesh. Assume that the current message holder during an execution of greedy routing is $x = (x_1, \ldots, x_{d(n)})$, at distance X from t. The probability that x has a long range link to some node $w = (x_1, \ldots, x_{i-1}, w_i, x_{i+1}, \ldots, x_{d(n)})$, $1 \leq i \leq d(n)$ such that $|t_i - w_i| \leq |t_i - x_i|/2$ is greater than $\sum_{1 \leq i \leq d(n)} \frac{1}{3d(n)H_k} = \frac{1}{3H_k}$ along the same analysis as the analysis of Kleinberg one dimensional model, summing over the dimensions. If such a link is found, it is always preferred to the local contact of x that only reduces one of the coordinate by 1. Thus, after at most $3H_k$ steps on expectation, one of the coordinates has been divided by two. Note that since long range links only get to nodes that differs in a single coordinate from their origin, further steps cannot increase $|x_i - t_i|$ for any $1 \leq i \leq d(n)$, x being the current message holder. Repeating the analysis for all coordinates, we thus get that after $3d(n)H_k$ steps on expectation, all the coordinates have been divided by at least two, and so the current distance to the target is at most $X/2$. Finally, the algorithm reaches t after at most $3d(n)H_k \log(dist(s,t))$ steps on expectation, which is $O(\log^2 n)$. $\square$

Remark. Note that our example of non-navigability in Section 3 may appear somehow counter intuitive in contrast to our latter construction of long range links on meshes. Indeed, why not simply repeating such a construction on the graph $G^{(n)}$ defined in the proof Theorem 1 ? That is, why not selecting long range contacts on each "diagonal" using the 1-harmonic distribution, in which case greedy routing would perform efficiently between pairs (s,t) on the diagonals ? This cannot be done however because, to cover all possible pairs (s,t) on the diagonals, $2^{d(n)}$ long range links per node would be required, which is larger than any polylogarithm of n when $d(n) \gg \log\log n$.

5 Conclusion

The increasing interest in graphs and metrics of bounded doubling dimension arises partially from the hypothesis that large real graphs do present a low doubling dimension (see, e.g., [12, 18] for the Internet). Under such an hypothesis, efficient compact routing schemes and efficient distance labeling schemes designed for bounded doubling dimension graphs would have promising applications. On the other hand, the navigability of a network is actually closely related to the existence of efficient compact routing and distance labeling schemes on the network. Indeed, long range links can be turned into small labels, e.g. via the technique of rings of neighbors [25]. Interestingly enough, our paper emphasizes that

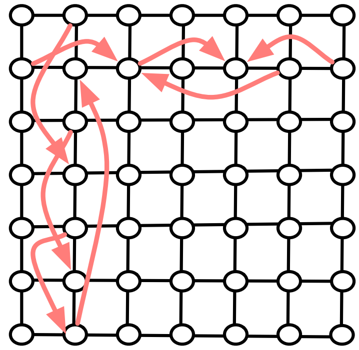

Fig. 2 – Example of 2-dimensional mesh augmented as in proof of Theorem 2 : 1-harmonic distribution of links on each axis. Bold links are long range links, they are not all represented.

the small doubling dimension hypothesis of real networks is crucial. Indeed, for doubling dimension above $\log \log n$, networks may become not navigable. It would therefore be important to study precisely to which extent real networks do present a low doubling dimension.

In a more general framework, our result of non navigability shows that the small world phenomenon, in its algorithmic definition of navigability, is not only due to the good spread of additional links over distances in a network, but is also highly dependent of the base metric itself, in particular in terms of dimensionality.

Références

[1] I. Abraham, C. Gavoille, A. V. Goldberg, and D. Malkhi. Routing in networks with low doubling dimension. Technical Report TR-2005-175, Microsoft Research, December 2005.

[2] I. Abraham and D. Malkhi. Name independent routing for growth bounded networks. In 17^{th} *Annual ACM Symposium on Parallel Algorithms and Architecture (SPAA)*, pages 49–55, 2005.

[3] S.B. Akers and B. Krishnamurthy. A group-theoretic model for symmetric interconnection networks. *IEEE Trans. Comput.*, 38 :555–566, 1989.

[4] J. Aspnes, Z. Diamadi, and G. Shah. Fault-tolerant routing in peer-to-peer systems. In *In 21st ACM Symp. on Principles of Distributed Computing (PODC)*, pages 223–232, 2002.

[5] P. Assouad. Plongements lipshitzien dans $\mathbf{R}^n$. *Bull. Soc. Math.*, 111(4) :429–448, 1983.

[6] L. Babai, W. M. Kantor, and A. Lubotzky. Small diameter cayley graphs for finite simple groups. *European Journal of Combinatorics*, (10) :507–522, 1989.

[7] L. Barrière, P. Fraigniaud, E. Kranakis, and D. Krizanc. Efficient routing in networks with long range contacts. In *LNCS Proceedings of 15th International Symposium on Distributed Computing (DISC)*, volume 2180, pages 270–284, 2001.

[8] H.T-H. Chan, A. Gupta, B. M. Maggs, and S. Zhou. On hierarchical routing in doubling metrics. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms (SODA)*, pages 762–771, 2005.

[9] K. L. Clarkson. *Nearest-Neighbor Methods for Learning and Vision : Theory and Practice*, chapter Nearest-neighbor searching and metric space dimensions. (Survey). MIT Press, T. Darrell, P. Indyk, G. Shakhnarovich, and P. Viola, editors, 2006.

[10] P. Duchon, N. Hanusse, E. Lebhar, and N. Schabanel. Could any graph be turned into a small world ? *Theoretical Computer Science*, Special issue on Complex Networks, 2005. To appear.

[11] M. Flammini, L. Moscardelli, A. Navarra, and S. Perennes. Asymptotically optimal solutions for small world graphs. In *LNCS Proceedings of the 19th International Symposium on Distributed Computing (DISC)*, volume 3724, pages 414–428, 2005.

[12] M. Fomenkov, K. Claffy, B. Huffaker, and D. Moore. Macroscopic internet topology and performance measurements from the dns root name servers. In *USENIX LISA*, 2001.

[13] P. Fraigniaud. Greedy routing in tree-decomposed graphs : a new perspective on the small-world phenomenon. In *Proceedings of the 13th Annual European Symposium on Algorithms (ESA)*, pages 791–802, 2005.

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

[15] A. Gupta, R. Krauthgamer, and J.R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 534–543, 2003.

[16] S. Har-Peled and M. Mendel. Fast construction of nets in low dimensional metrics and their applications. In *Proceedings of the* 21*th ACM Symposium on Computational Geometry (SoCG)*, pages 150–158, 2005.

[17] J. Heinonen. *Lectures on analysis on metric spaces.* Springer-Verlag, New York, 2001.

[18] D. R. Karger and M. Ruhl. Finding nearest-neighbors in growth-restricted metrics. In *Proceedings of the 34th annual ACM symposium on Theory of computing (STOC)*, pages 741–750, 2002.

[19] 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.

[20] E. Lebhar and N. Schabanel. Close to optimal decentralized routing in long-range contact networks. *Theoretical Computer Science special issue on ICALP'04*, 348(2-3) :294–310, 2005.

[21] 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)*, pages 54–63, 2004.

[22] C. Martel and V. Nguyen. Analyzing kleinberg's (and other) small-world models. In *In* 23*rd ACM Symp. on Principles of Distributed Computing (PODC)*, pages 178–187, 2004.

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

[24] E. Rozenman, A. Shalev, and A. Wigderson. A new family of cayley expanders. In *36th ACM Symposium on Theory of Computing (STOC)*, pages 445–454, 2004.

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