Analyzing and Characterizing Small-World Graphs *

Van Nguyen and Chip Martel
{martel,nguyenvk}@cs.ucdavis.edu
Computer Science, UC Davis, CA 95616

Abstract

We study variants of Kleinberg's small-world model where we start with a k-dimensional grid and add a random directed edge from each node. The probability $u's$ random edge is to v is proportional to $d(u,v)^{-r}$ where $d(u,v)$ is the lattice distance and r is a parameter of the model.

For a k-dimensional grid, we show that these graphs have poly-log expected diameter when $k < r < 2k$, but have polynomial expected diameter when $r > 2k$. This shows an interesting phase-transition between small-world and "large-world" graphs.

We also present a general framework to construct classes of small-world graphs with $\Theta(\log n)$ expected diameter, which includes several existing settings such as Kleinberg's grid-based and tree-based settings [15].

We also generalize the idea of 'adding links with probability $\propto$ the inverse distance' to design small-world graphs. We use semi-metric and metric functions to abstract distance to create a class of random graphs where almost all pairs of nodes are connected by a path of length $O(\log n)$, and using only local information we can find paths of poly-log length.

1 Introduction

Small-world networks are being used and studied in many disciplines, including the social and natural sciences. These networks possess a striking property, the so called small-world phenomenon, also often spoken of as "six degrees of separation" (between any two people in the United States)[1]. Since many real networks exhibit small-world properties, a number of network models have been proposed as a framework to study this phenomenon. Watts and S. Strogatz [24] introduced a random graph setting to model certain small-world graphs. This model features two main properties, low average path length and significant clustering. We use *small-world graphs* to mean graphs with poly-log (expected) diameters, to focus on this property of small separation between nodes.

Recently, Kleinberg [16] proposed a family of small-world networks to study another compelling aspect of Milgram's findings: a greedy algorithm using only local information can construct short paths. Kleinberg adds directed long-range random links to an undirected $n \times n$ lattice network. The long-range links have a non-uniform distribution which favors arcs to close nodes over more distant ones. These graph models have generated considerable interest and recent work. Applications have been found using Kleinberg's or related small-world models to decentralized search protocols in peer-to-peer systems [21, 25], and gossip protocols for a communication network [14].

Kleinberg's model starts with a simple *base* graph and randomly adds new arcs. The base graph models local "contacts". The additional random links model long-range contacts which can connect distant components. This greatly shrinks the diameter of the graph. Thus we see a promising formula: a simple base graph plus some random links can add nice properties (such as Kleinberg's setting with expected small diameter and short greedy paths for all $s-t$ pairs). Kleinberg's setting is a very specific one, so we ask: what are the essential features, underlying the distribution of random links and the grid structure which produce these nice properties? We address this question in two ways. First, we mostly complete the picture of the diameter problem in Kleinberg's grid-based setting by identifying the critical point where the graph changes from expected poly-log to expected polynomial diameter, depending on how much we favor links to close nodes. Then we construct a framework, which starts with an arbitrary base graph and some general rules for adding random arcs. We then refine our model to identify properties which lead to small expected diameter. Further refinement allows us to find short paths using local information only.

Some of our graphs have small expected diameter, yet need not use a distance measure to describe the random link distribution[2]. Kleinberg's models (grid-based

*This work was supported by NSF grant CCR-85961

[1] Milgram discovered this in his pioneering work in the 1960's [22], and recent work by Dodds et al. suggests its still true [9].

[2] Thus, links no longer favor close nodes over distant nodes.

setting [16], tree-based and group-induced settings [15]) and several other well-known small-world graphs fit our abstract models and thus can be analyzed using our general results on diameter and routing. Moreover, we introduce or generalize several techniques used for bounding a graph's diameter.

We briefly review Kleinberg's setting then summarize our results in the next subsection. Kleinberg's basic model uses a two-dimensional grid as a base with long-range random links added between any two nodes u and v with a probability proportional to $d^{-2}(u, v)$, the inverse square of the lattice distance between u and v. In the basic model, each node has an undirected *local* link to each of its four grid neighbors and one directed *long-range* random link. A straightforward extension of this basic model is to have multiple random links from each node and use a k-dimensional grid for any $k = 1, 2, 3 \ldots$; also use an inverse r^{th} power distribution (of the random links), for any real constant r, instead of $r = 2$.

In [20], we proved a tight $\Theta(\log n)$ bound for the expected diameter of Kleinberg's extended model: for a k-dimensional grid and an inverse r^{th} power distribution when $0 \leq r \leq k$, i.e. for $0 \leq r \leq 2$ in the 2-D case. However, the diameter problem for $r > k$ was open before this paper. Note that the complexity of greedy routing in Kleinberg's grid-based setting has already been analyzed. For $r = k$ it takes $\Theta(\log^2 n)$ expected steps while for $r \neq k$, greedy routing takes expected polynomial time[16, 2, 20, 11].

1.1 Our results First, we mostly complete the analysis of the diameter of Kleinberg's grid-based setting. For a k-D grid, we show that the model still has poly-log expected diameter when $k < r < 2k$, but has polynomial expected diameter when $r > 2k$. However, interestingly enough, the case $r = 2k$ is still open, though our initial experiments suggest that the model is a large-world. In particular, for Kleinberg's 1-D model, for any $r < 2$ the expected diameter is upper-bounded by poly-log functions ($O(\log n)$ for $r \leq 1$), however, for $r > 2$, the expected diameter can be lower bounded by a (low-degree) polynomial function. This shows a phase-transition between small-world and "large-world" graphs.

We also present a framework to construct several classes of small-world graphs with $\Theta(\log n)$ expected diameter. These include several existing settings such as Kleinberg's grid-based and tree-based settings [15]. Our framework starts with a very abstract class of random graphs, then we gradually add in conditions to achieve more refined classes, which are more likely small-world candidates.

We also design graphs with poly-log greedy-like paths. Again, we start with a general class, based on an abstract semi-metric function (abstracted from the use of distance), and then add in refining criteria to construct a hierarchy of classes with interesting properties. As a result, we obtain an abstract class of random graphs such that under some easy conditions, almost all pairs of nodes are connected by a path of length $O(\log n)$, and using only local information we can find paths of expected poly-log length.

1.2 Related work There has been considerable work on the small-world phenomenon. See [17] for early surveys and [16] for a more recent account on modeling small-world networks. Before Kleinberg's model, Watts and Strogatz [24] proposed randomly rewiring the edges of a ring lattice each with a probability parameter p. Watts and Strogatz observed that for small p the model reflects many practical small-world networks with small typical path length and a non-negligible clustering coefficient. Kleinberg has generalized his basic model in several ways in [15] including a generalization that encompasses both lattice-based and tree-based ("taxonomic" or "hierarchical") small-world networks.

The diameter of random graphs is a classic problem [5, 6, 7, 10] but most results use uniformly distributed arcs. Bollobas and Chung [6], study a graph model very similar to Watts and Strogatz in [24] with the nodes of a cycle (or a "ring") randomly matched to form additional long-range links. The closest diameter work with non-uniform arc probabilities is on long-range percolation graphs (LRPGs) which have been used to study physical properties. As in Kleinberg's model, a grid with (undirected) local links is augmented by long-range random links whose probability is inversely related to their distance. Note that in contrast to Kleinberg's model, the added links are undirected, and the degree of a node is not fixed. Thus the analysis techniques for LRPGs are somewhat different than those to analyze Kleinberg's and related models. Benjamini and Berger study the diameter of 1-D LRPGs [3] and Coppersmith et al. extend this to k-D grids [8]. Both papers prove diameter results which show how the expected diameter changes as the arc probability parameters change. Biskup improves these results by proving tighter bounds [4]. These papers show there are critical points where the expected diameter changes from constant, to poly-log and then to polynomial as the probability parameter changes. We show some similar transitions occur in Kleinberg's setting.

There have also been several recent papers which analyze greedy routing in other small-world like networks [1, 2, 15, 18, 20, 11]. Though our focus is on diameter results, we show how to incorporate greedy-like routing (to find short paths) into an abstract class which already has expected $O(\log n)$ diameter.

The structure of the paper. We present new diameter results for Kleinberg's grid settings, which complement previous diameter results. In §3 we start with the most basic setting, i.e. the (one-dimensional) cycle augmented by random links.

We then generalize our approach in §4 (for analyzing Kleinberg's grid model) and introduce several abstract families of random graphs which can be constructors for small-worlds. From these abstract families, by adding some proper additional conditions, we obtain different classes of small-world graphs with polylog expected diameter. In §5 we create classes with short paths which can be found by decentralized algorithms (using local information only), and present a generalization of §3's results.

2 Preliminaries

To generalize Kleinberg's small-world models, we develop an abstract class of random graphs, which includes Kleinberg's small-world settings (in [16, 15]). We then use this abstract class as a platform to create a general framework to analyze the diameter (and other related issues) in a variety of settings.

Consider the following random assignment (or matching) operation: for a given node u in a graph G, make a random trial under a specific distribution rule τ to select another node v. We write this as $v \overset{R_\tau}{\leftarrow} u$ or $v = R_\tau(u)$. For example, in Kleinberg's basic grid setting, τ is defined as having $v \overset{R_\tau}{\leftarrow} u$ with probability proportional to the inverse square of the lattice distance between u and v, i.e. $Pr[v \overset{R_\tau}{\leftarrow} u] \propto d^{-2}(u,v)$. We can think of a random graph constructor using this operation which forms a family of random graphs. We use a given base graph H and a compatible graph constructor, where each additional (u,v) link (with $v \overset{R_\tau}{\leftarrow} u$) is called a random link. Random links are generated for a node, not for pairs of nodes as in traditional random graphs [3]. This operation is implicitly used in Kleinberg's small-world models [16, 15].

We restrict the distribution rules (τ) we use to ones which have the following property: each R_τ call performs an independent trial. Multiple R_τ calls on the same input node (u), also are independent trials. We now define an abstract class of random graphs, which includes all of Kleinberg's small-world settings.

[3]Even when we use undirected random links, we can consider that: each node u generates and, so, "owns" certain random links, while some other random links also incident to u are not owned by u but by some other nodes (which generated these links)

Definition 1. *Given a set of undirected base graphs $\mathcal{H}$, a distribution τ and a constant integer $q \geq 1$, a Family of Random Graphs $\mathcal{FRG}(\mathcal{H}, \tau, q)$ consists of graphs, each of which is a base graph $H \in \mathcal{H}$ plus q outgoing random links[4] generated under distribution τ for each node.*

All the families of random graphs we consider in this paper are $\mathcal{FRG}$ families. For example, Kleinberg's basic grid model ([16]) is a $\mathcal{FRG}(\mathcal{H}, \tau, q)$ family, where $\mathcal{H}$ consists of all $n \times n$ grids ($n = 1, 2, 3 \ldots$), $q = 1$, and τ is the inverse square distribution. Note that there is no restriction on the set of fixed edges E in the base graphs. For example, the fixed edges can be the local links in Kleinberg's grid model, a complete graph, or nothing at all as in Kleinberg's tree-based model.

We now consider some useful basic lemmas, the proofs of which are fairly simple and omitted. Consider a family $\mathcal{F} = \mathcal{FRG}(\mathcal{H}, \tau, q)$ and a graph $G \in \mathcal{F}$, which has base graph $H = (V, E)$.

Lemma 2.1. *For any graph G from a family $\mathcal{FRG}(\mathcal{H}, \tau, q)$, any two disjoint subset of vertices S and T chosen without any knowledge of the random links from S, the probability of having a random link from some node in S to at least one node in T, is $Pr[S \to T] \geq 1 - e^{-q\epsilon|T||S|}$ (where $\epsilon = \epsilon(S,T)$ denotes the minimum value of $Pr[R_\tau(u) = v]$ for all $u \in S$ and $v \in T$).*

We use lemma 2.1 where usually the sizes of S and T are large enough so that $\epsilon|T||S| = \Omega(\log n)$ and thus, for some $\theta > 0$, $Pr[S \to T] \geq 1 - O(n^{-\theta})$, which tends to 1 when n goes to the infinity. So, almost surely, T is apart from S by just one random link.

Lemma 2.2. *If each of n events $\{B_i\}_{i=1}^n$ occurs with probability at least $1 - p$, where $p < 1/n$, then the combining event $\cap_{i=1}^n B_i$ occurs with probability at least $1 - np$*

Note that lemma 2.2 applies even if the B_i are not independent.

3 Diameter transitions in Kleinberg's model

For simplicity, we first look at the 1-D setting and then extend our results to more general settings. Define $\mathcal{C}(r, n)$ as the setting where nodes are labeled $0, 1, 2, \ldots, n-1$ and each node i has 2 undirected *local* links: to $(i-1) \bmod n$ and $(i+1) \bmod n$ for $0 \leq i \leq n-1$. Each node i also has one directed random link to some node $j \neq i$. The probability its

[4]They are directed by our default assumption.

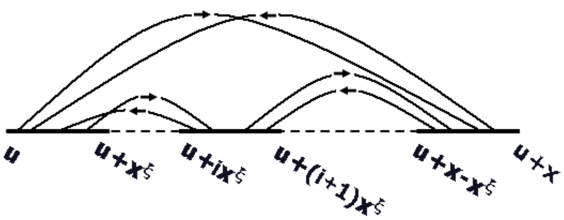

Figure 1: I_x^u is ξ-complete with directed random edges crossing between any two subsegments of length x^ξ

random link is to j, is proportional to $|i-j|^r$, where $r \geq 0$ is a parameter to be specified. For $0 \leq r \leq 1$ this cycle setting is known to have expected $\theta(\log n)$ diameter [20]. We now consider the diameter of $\mathcal{C}(r,n)$ when $r > 1$.

3.1 The $\mathcal{C}(r,n)$ setting with $1 < r < 2$.

We present our notation and basic definitions, then a sketch of our basic approach, and finally our theorems and proofs in detail.

For $r > 1$, the normalized coefficient $L = 1/(2\sum_{d=1}^{n/2} d^{-r}) = \theta(1)$; in fact, $\frac{1}{2C_r} < L < \frac{1}{C_r}$ for n large enough, where $C_r = \sum_{i=1}^{\infty} i^{-r}$ is a constant depending on r only. So, $Pr[i \to j] = L|i-j|^{-r} = \theta(|i-j|^{-r})$. Let $I_l(u)$ or I_l^u denote a 'segment' of length l, starting at node u, i.e. $I_l^u = \{u, (u+1) \mod n, \ldots, (u+l-1) \mod n\}$.

Consider segment I_x^u of length x for some arbitrary node u. Let $0 < \xi < 1$. Divide I_x^u into $x^{1-\xi}$ (disjoint) subsegments of length x^ξ. Let $\mathcal{D}_\xi(I_x^u) = \{J_1, J_2, \ldots, J_{x^{1-\xi}}\}$ be this set of subsegments, i.e. $J_k = I_{x^\xi}(u+(k-1)x^\xi)$ for $1 \leq k \leq x^\xi$. For simplicity, we assume $x^\xi, x^{1-\xi}$ and the like are integers.

Definition 2. *For each node u, I_x^u is ξ-complete if for any ordered pair of segments (J_i, J_k) from $\mathcal{D}_\xi(I_x^u)$, there is an edge from J_i to J_k* [5] *(see figure 1).*

Let $\delta(I_x^u)$ be the diameter of the subgraph induced by nodes in the segment I_x^u. Here, $\delta(I_x^u)$ is a random variable with a value for each instance of our random graph (once the random links are set). $E[\delta(I_x^u)]$ is independent of position u, so we let $\delta_x = E[\delta(I_x^u)]$.

The main idea. In order to upper bound the diameter of our random graph in this 1-D setting, we use a probabilistic recurrence approach[6]. We establish a (probabilistic) relation between the diameter of a segment and that of a smaller one. In particular, we relate $\delta(I_x)$ (the diameter of a segment of length x) to $\delta(I_y)$, where $y = x^\xi$ for some $\xi \in (0,1)$. Intuitively, with high probability, $\delta(I_x)$ is bounded by a constant multiple of $\delta(I_y)$. Thus, we use standard recurrence techniques to bound δ_n (the graph's expected diameter) based on δ_{x_0} for a small initial length x_0 (so δ_{x_0} is upper bounded by a poly-log function of n).

We use this crucial observation: I_x is almost surely ξ-complete for x and $\xi < 1$ large enough. So, $\delta(I_x)$ is almost surely not larger than twice the maximum diameter of any subsegment in $\mathcal{D}_\xi(I_x)$. We formalize the above ideas in the following lemmas and then prove our main theorem. The next two results follow directly.

Lemma 3.1. *If a segment I_x^u is ξ-complete then* $\delta(I_x^u) \leq 2 \max_{J \in \mathcal{D}_\xi(I_x)} \delta(J) + 1$.

Corollary 3.1. *If I_x^u is ξ-complete for each u = 0..n-1 then* $\max_{u=0..n-1} \delta(I_x^u) \leq 2 \max_{u=0..n-1} \delta(I_{x^\xi}^u) + 1$.

Note that for $0 < \xi < .5$, I_x^u is not ξ-complete for any u. Since x^ξ, the number of random links from nodes in a subsegment $J_i \in \mathcal{D}_\xi(I_x)$, is smaller than $x^{1-\xi} - 1$, the number of other subsegments $J_k \in \mathcal{D}_\xi(I_x)$.

Lemma 3.2. *For $r/2 < \xi < 1$ ($1 < r < 2$),*

$$Pr[I_x^u \text{ is } \xi\text{-complete}, \forall u = 0..n-1] \geq 1 - n^{-2}$$

for $x \geq \hat{c} \ln^{\frac{1}{2\xi - r}} n$, where $\hat{c} = (10C_r)^{\frac{1}{2\xi-r}}$.

Proof. We need to lower bound the probability of the event that there exists an edge connecting J_a and J_b for all possible pairs (J_a, J_b). Using lemma 2.1, $Pr[J_a \to J_b] \geq 1 - e^{-q\epsilon|J_a||J_b|}$, where $\epsilon = \epsilon(J_a, J_b)$. Note, $|J_a| = |J_b| = x^\xi$, $\epsilon(J_a, J_b) \geq Lx^{-r} > .5Lx^{-r}/C_r$ and $q = 1$, so

$$Pr[J_a \to J_b] \geq 1 - e^{-Lx^{-r} \times x^{2\xi}} \geq 1 - e^{-.5x^{2\xi-r}/C_r} \quad (3.1)$$

I_x is ξ-complete if there exists an arc between J_a and J_b for all possible pairs (J_a, J_b). The number of such pairs is $< x^{2(1-\xi)}$, hence using lemma 2.2, $P_x = Pr[I_x \text{ is } \xi\text{-complete}] \geq 1 - (e^{-.5x^{2\xi-r}/C_r} \times x^{2-2\xi})$. Let E be the event that I_x^u is ξ-complete, $\forall u = 0..n-1$. Again, using lemma 2.2:

$$Pr[E] \geq 1 - n(1-P_x) \geq 1 - (ne^{-.5x^{2\xi-r}/C_r} \times x^{2-2\xi}).$$

Now, for $x \geq (10C_r)^{\frac{1}{2\xi-r}} \times \ln^{\frac{1}{2\xi-r}} n$, clearly $ne^{-.5x^{2\xi-r}/C_r} \leq ne^{-5\ln n} = n^{-4}$, hence

$$Pr[E] \geq 1 - (n^{-4} \times x^{2-2\xi}) \geq 1 - n^{-2}$$

since $x^{2-2\xi} < n^2$.

Theorem 3.1. *For any r such that $1 < r < 2$, there exists a constant β such that the expected diameter of $\mathcal{C}(r,n)$ is $O(\log^\beta n)$.*

[5] If we think of a super-graph with the J_i's as it's nodes then these crossing links make it a complete graph

[6] Although our approach is similar to Karp's [13], his theorems necessity conditions are not met here.

Proof. Since $r < 2$ we can choose $r/2 < \xi < 1$. Let $\phi(x)$ be a random variable s.t. $\phi(x) = \max_{u=0..n-1} \delta(I_x^u)$. $\phi(x)$ is determined for each instance of our random graph. If I_x^u is ξ-complete for all $u = 0..n-1$ then from corollary 3.1, $\phi(x) \leq 2\phi(x^\xi) + 1$. Thus from lemma 3.2, for $x \geq x_0 = (10C_r)^{\frac{1}{2\xi - r}} \log^{\frac{1}{2\xi-r}} n$,

$$Pr[\phi(x) \leq 2\phi(x^\xi) + 1] \geq 1 - n^{-2} \tag{3.2}$$

We can use a standard recurrence technique to upper bound $\phi(n)$, based on $\phi(x_0)$ and n only.

Define the sequence $\{x_i\}_{i=0}^{t+1}$, where $x_{i+1} = x_i^b$ with $b = 1/\xi$, $x_0 = c \log^{\frac{1}{2\xi-r}} n$, and

$t = \lfloor \log_b(\log_{x_0} n) \rfloor = \lfloor \frac{\log(\frac{\log n}{\log x_0})}{\log b} \rfloor = \frac{\log \log n}{\log b} + 0(1)$

Thus $x_t \leq n < x_{t+1}$. Now we look closer at this sequence $\{\phi(x_i)\}_{i=0}^t$ and use (3.2) to upper bound the last term (which differs from $\phi(n)$ by a constant multiple), based on the first term and t. We claim that each of the events E_i : "$\phi(x_i) \leq 2\phi(x_{i-1}) + 1$", $i = 1, 2, \ldots, t$ and E_{t+1} : "$\phi(n) \leq 2\phi(x_t) + 1$" occurs with probability at least $1 - n^{-2}$. The first t events can be justified directly from (3.2), while we can also easily extend our proof of lemma 3.1 to justify the last event. Let E be the event that $E_1, E_2, \ldots, E_{t+1}$ all occur. Using lemma 2.2, E occurs with probability at least $1 - (t+1) \times n^{-2} \geq 1 - O(n^{-1})$.

It is easy to see that event E implies $\phi(x_i) \leq 2^i \phi(x_0) + 2^i - 1, \forall i = 1..t$ and thus,
$\phi(n) \leq 2^{t+1}\phi(x_0) + 2^{t+1} - 1 \leq O((\log n)^{\log_b 2}) \times \phi(x_0)$.
Note that $\phi(x_0) \leq x_0 = (10C_r)^{\frac{1}{2\xi-r}} \log^{\frac{1}{2\xi-r}} n$. That is,

$$\Pr[\delta(I_n) \leq c \log^\beta n)] \geq 1 - O(n^{-1})$$

where $\beta = \log_{1/\xi} 2 + \frac{1}{2\xi - r}$ and c depends on r and ξ only. Thus, $Pr[\delta(I_n) \leq O(\log^\beta n)]$ tends to 1 when n goes to infinity, and almost surely $\delta(I_n) = O(\log^\beta n)$.

Note that our bound on β grows rapidly as r approaches 2.

3.2 The $\mathcal{C}(r, n)$ setting with $2 < r$

THEOREM 3.2. *For $r > 2$, $\mathcal{C}(r, n)$ is a 'large' world with expected diameter $\Omega(n^{\frac{r-2}{r-1} - o(1)})$.*

Proof. Let $\frac{1}{r-1} < \gamma < 1$. For any node i, the probability that i's random contact is at most a distance n^γ from i, is $1 - O(\sum_{d=n^\gamma}^{n/2} d^{-r}) = 1 - O(n^{-\gamma(r-1)})$. Using lemma 2.2, the probability that all random links have length at most n^γ, is $\geq 1 - n \times O(n^{-\gamma(r-1)}) = 1 - O(n^{1-\gamma(r-1)})$. Since $\frac{1}{r-1} < \gamma$, this probability tends to 1 when n goes to infinity. Thus the diameter is at least $\frac{n}{n^\gamma} = n^{1-\gamma}$ with overwhelming probability (tending to 1 when n goes to infinity). So, the expected diameter is $\Omega(n^{\frac{r-2}{r-1} - o(1)})$

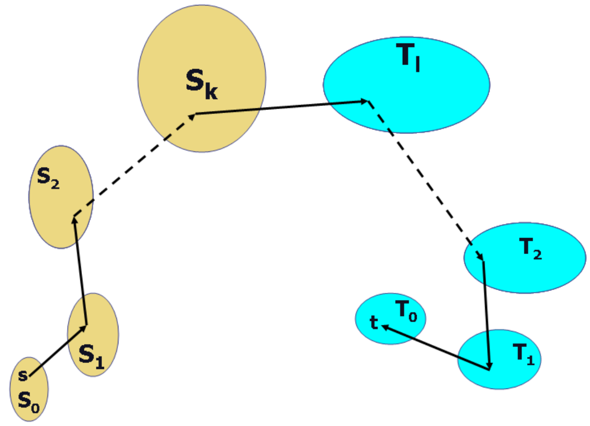

Figure 2: A path from s to t

3.3 Extended settings

We can extend our results to the setting without wraparound and to the general k-D setting for $k = 1, 2, 3 \ldots$ *The general k-D setting is still a small-world when $r < 2k$ but a 'large-world' when $r > 2k$.* We only need to adapt our basic proof above so that equation (3.1) is still maintained, and hence the rest of our arguments still apply. For the general k-D setting, we use k-D hypercubes instead of segments as in 1-D setting. Similarly, we also introduce a decomposition of a hypercube of size x(in each dimension) into smaller sub-cubes (of size x^ξ) and thus call the hypercube ξ-complete if there is a random edge from any sub-cube to any other. Thus, we can reuse most of the proof above except some extra calculation (say, for the sizes of some k-D cubes). See [23] for full proofs.

Note that the case $r = 2k$ is open, however initial experiments (for the 1-D setting only) suggest that the setting has polynomial expected diameter.

4 Constructing $O(\log n)$ diameter graphs with non-uniform random links

To analyze the shortest path between a source node s and a destination node t, we construct two subset chains, which can be viewed as two trees rooted at s and t, and then show they intersect. Each subset in s's subset chain contains nodes which can be reached directly from the preceding subset, and hence, can be reached from s. The subset chain from t is similar, but contains nodes with links towards t. To show that the shortest $s - t$ path has length $O(\log n)$, the main idea is to show that each subset chain grows exponentially in size before they intersect[7] (see figure 2).

[7]Alternatively, each subset chain grows exponentially to a threshold, so they intersect with high probability.

Exponential growth will be likely if each time we grow a new subset, with high probability more than one link from each node leaves the current subset. This was true in Kleinberg's grid setting [20] (we called this: "link into or out of a ball" property). We now include this feature to refine our basic class $\mathcal{FRG}(\mathcal{H}, \tau, q)$. Recall that, a family of random graphs $\mathcal{FRG}(\mathcal{H}, \tau, q)$ consists of graphs, each of which is a base graph $H \in \mathcal{H}$ plus at least q out-going random links generated under distribution τ for each node.

DEFINITION 3. *For constants $\mu > 0$ and $\xi > 0$, family $\mathcal{F} = \mathcal{FRG}(\mathcal{H}, \tau, q)$ meets 'the (μ,ξ) expansion criterion', or $\mathcal{F}$ is (μ,ξ)-EXP, if $\forall H = (V, E) \in \mathcal{H}$, with $n = |V|$:*

$$(4.3)\quad \forall u \in V, \forall \mathcal{C} \subset V, |\mathcal{C}| < n^\mu : Pr[v \overset{R_\tau}{\leftarrow} u : v \notin \mathcal{C}] \geq \xi$$

For example, from [19], it is easy to verify that Kleinberg's grid setting with wrap-around distance is (μ,$1 - \mu - o(1)$)-*EXP* for any fixed positive constant $\mu < 1$. *This criterion supports diversity and fairness in the distribution of random links:* For a random link from any node, no small set of vertices (size $\leq n^\mu$) can take most of the chance to have this link come into it.

DEFINITION 4. (TYPE *μ-Expansion*) *For a constant $\mu > 0$, type μ-Expansion contains all the families $\mathcal{FRG}(\mathcal{H}, \tau, q)$ which meet (μ,ξ)-EXP for some $\xi > 1/q$.*

We define χ, called an *'expansion function'*, as follows. Given any $u \in V$, this operation will call operation R_τ q times. Also, let $\chi(u)$ denote the set of vertices from these q R_τ calls. Thus the random links for graph G are formed by performing operation χ on each node. For any set S: $\chi(S) = \bigcup_{u \in S} \chi(u)$.

Consider a family $\mathcal{F}$ of type *μ-Expansion*. Let $\beta = q\xi$ (so $\beta > 1$). For any node u and set $\mathcal{C}$ of size less than $n^\mu - q$, which is determined before $\chi(u)$ is known, the expected number of fresh elements generated by $\chi(u)$ that do not belong to $\mathcal{C}$ is greater than β: $E[\,|\,\chi(u) - \mathcal{C}\,|\,] > \beta > 1$. Since $\chi(u)$ 'contributes' more than one expected fresh element outside of $\mathcal{C}$, χ can be used to generate a chain of subsets from a small initial subset such that with high probability, the subsets will quickly grow to size $\Theta(n^\mu)$.

4.1 The out-going subset chain Let $\mathcal{F}$ be a *μ-Expansion* family, and $G = (V, E)$ be an arbitrary graph from $\mathcal{F}$. Now, from an arbitrary initial set $S_0 \subset V$, we construct a chain of subsets $\{S_k\}$, namely the *out-going subset chain* with respect to the initial set S_0, s.t. $S_{k+1} = \chi(S_k) - \cup_{i=0}^k S_i; k = 1, 2, 3, \ldots$ Thus, S_i is the nodes at distance i from S_0 using random links. The following results for *μ-Expansion* families show the subset chain grows rapidly if S_0 is large enough.

LEMMA 4.1. *$\forall \mathcal{C}, S \subset V$ s.t. $S \subset \mathcal{C}$, $|\mathcal{C}| \leq \alpha = \theta(n^\mu)$: if $|S| = \Omega(\log n)$, almost surely $|\chi(S) - \mathcal{C}|/|S| > \gamma$ for a constant $\gamma > 1$. Also, $\exists \gamma > 1, \forall \theta > 0, \exists c > 0$:*

$$|S| > c \log n \Rightarrow Pr[\tfrac{|\chi(S) - \mathcal{C}|}{|S|} > \gamma] = 1 - O(n^{-\theta})$$

The above lemma (see [23] for proof) provides a probabilistic lower bound γ on the growth rate of the subset chain in each early step (by choosing $\mathcal{C} = \cup_{i=0}^k S_i$ to apply the lemma in each step). This growth rate can be maintained as long as the subset sizes are still under a threshold. *For any $S_0 \in V$ with size $\Omega(\log n)$, the subset chain originating from S_0 will almost surely grow exponentially in size until it reaches size $\alpha = \theta(n^\mu)$.* Also, for any $\theta > 0$, by choosing a sufficiently large constant c s.t. $|S_0| > c \log n$, $Pr[|S_k| \geq \alpha] = 1 - O(n^{-\theta})$ for some $k = O(\log n)$. Moreover, this can be true for any given $\theta > 0$ by choosing c large enough.

4.2 The in-coming subset chain We now construct a subset chain, based on the random links coming to the sets of the chain. We use an 'expansion function' ψ, which is a counterpart of χ, so we can reuse the formalism used in §4.1 on the out-going subset chain and obtain similar results. Function ψ is not state-less as χ was. For any subset of vertices $\mathcal{D}$ and a node $u \in V$ we define $\psi(u, \mathcal{D})$ to return the set of all nodes $v \notin \mathcal{D}$ s.t. v has a random link to u. As before, $\psi(T, \mathcal{D}) = \bigcup_{u \in T} \psi(u, \mathcal{D})$ for any subset T. Now, from an arbitrary subset $T_0 \subset V$, we can construct a chain of subsets $\{T_k\}$, namely the *in-coming subset chain* with respect to the initial set T_0, s.t. $T_{k+1} = \psi(T_k, \mathcal{D})$ for $k = 1, 2, 3, \ldots$, where $\mathcal{D} = \cup_{i=0}^k T_k$. Similar to definition 3, we have:

DEFINITION 5. *For constants $\mu > 0$ and $\xi > 0$, family $\mathcal{F}$ meets 'the (μ,ξ) incoming expansion criterion', or $\mathcal{F}$ is (μ,ξ)-IE, if the following is satisfied.*

$$(4.4)\quad \forall \mathcal{D} : |\mathcal{D}| < n^\mu, \forall u \in \mathcal{D} : Pr[\exists v \notin \mathcal{D} : R_\tau(v) = u] > \xi$$

Similarly as with *μ-Expansion*, for a fixed $\mu > 0$, we define type *μ-IncExpansion*, which includes all the $\mathcal{FRG}(\mathcal{H}, \tau, q)$ families which meet (μ,ξ)-*IE* where $\xi > 1/q$. For a *μ-IncExpansion* family, lemma 4.1 holds if we replace the use of function χ by that of function ψ and subset $\mathcal{C}$ by subset $\mathcal{D}$ ([8]). There is an interesting implication between these two expansion criteria for a large class of families. We call a family of random graphs, using a distribution τ, δ-symmetric (or just symmetric if $\delta = 1$) for some constant $\delta \geq 1$, if $\frac{Pr[R_\tau(v)=u]}{Pr[R_\tau(u)=v]} \leq \delta$ for all pairs of nodes (u, v). It is easy

[8]The constructions of both subset chains share the same formalism

to see that Kleinberg's grid settings (using the inverse power distributions) have this property, and they are symmetric if wrap-around distance is used.

LEMMA 4.2. *If family $\mathcal{F}$ is (μ,ξ)-EXP, for $0 < \mu, \xi < 1$, and is δ-symmetric for some $\delta \geq 1$ then $\mathcal{F}$ is (μ,$1 - e^{-\xi/\delta}$)-IE.*

Proof. [Proof(sketch)] We need to prove (4.4) holds. Let $p(u, v) = Pr[R_\tau(u) = v]$ and F be the event that $\exists v \notin \mathcal{D} : R_\tau(v) = u$. The lemma is shown as $Pr[\overline{F}] = \prod_{v \notin \mathcal{D}} (1 - p(v, u)) \leq \prod_{v \notin \mathcal{D}} e^{-p(v,u)} = \exp\{-\sum_{v \notin \mathcal{D}} p(v, u)\} \leq \exp\{-\frac{1}{\delta} \sum_{v \notin \mathcal{D}} p(u, v)\} \leq e^{-\frac{\xi}{\delta}}$. Note that $\sum_{v \notin \mathcal{D}} p(u, v) = Pr[\exists v \notin \mathcal{D} : R_\tau(u) = v] \geq \xi$.

4.3 Abstract classes of small-world graphs We refine the above families by adding conditions to obtain small-world graphs. If our graph is from a family of type μ_1-*Expansion* and μ_2-*IncExpansion* for some $0 < \mu_1, \mu_2 < 1$ then, given any source s and destination t, we can use the following strategy to construct a $\log n$-length path from s to t (see figure 2). First, we want a connected subset S_0 containing s and T_0 containing t of $\Omega(\log n)$ size in the base graph H. We then construct the out-going subset chain from S_0 and the in-coming subset chain from T_0. Our above results show that, with overwhelming probability, there exist subsets S_k with size $\theta(n^{\mu_1})$ and T_l with size $\theta(n^{\mu_2})$ s.t. any node in S_k can be reached from S_0 by $O(\log n)$ links, and T_0 from T_l by $O(\log n)$ links. We now consider proper conditions so we can easily reach T_l from S_k.

If $\epsilon = \epsilon(\tau)$, the minimum value of $Pr[R_\tau(u) = v]$ for all $u \neq v$, is large enough, then almost surely there is an arc from S_k to T_l (or they intersect).

DEFINITION 6. (EXPANSION FAMILY) *A $\mathcal{FRG}(\mathcal{H}, \tau, q)$ is an Expansion family if it is (μ_1,ξ)-EXP and (μ_2,ξ)-IE for some constants $\xi > 1/q, \mu_1, \mu_2 > 0$, and $\epsilon(\tau) = \Omega(n^{-\mu_3})$ for a constant $\mu_3 < \mu_1 + \mu_2$.*

We now show that a graph from an *Expansion family* almost always has an arc from S_k to T_l (or they already intersected). We can assume all the nodes in S_k are fresh (we do not know their random links yet[9]) and hence, using lemma 2.1, $Pr[S_k \to T_l] \geq 1 - e^{-q\epsilon|T_l||S_k|} \geq 1 - e^{-\Omega(n^{\mu_1+\mu_2-\mu_3})} \geq 1 - O(n^{-1})$, which tends to 1 when n goes to the infinity.

[9] We omit a conditioning issue: if we construct the s subset chain (s-SSC) first then the growth of the t subset chain (t-SSC) is conditioned on the existence of s-SSC and vice versa. Thus, we need to add $\cup_{i=0}^{k-1} S_i$ to $\mathcal{D}$ (§4.2) or $\cup_{i=0}^{l-1} T_i$ to $\mathcal{C}$ (4.1). Therefore, if $\mu_1 > \mu_2$ then we construct t-SSC first, otherwise s-SSC first.

The graphs from an *Expansion family*[10] are small-worlds, i.e. their expected diameter is poly-log in n, as long as each node is rich enough in neighbors in the base graph to form large enough initial subsets (i.e. S_0, T_0). Without this final condition, however, often these graphs are not connected. If there are no edges in the base graph ($E = \emptyset$) then even with the added random edges, the graphs can be unconnected; an example will be presented in the next subsection.

We now add the notion of neighboring in the base graphs. A node u is called *k-neighbored* for some $k \in \mathbf{N}$ if u belongs to a connected component of size k in the base graph. A base graph $H = (V, E)$ is called k-neighbored if all the nodes are k-neighbored. A connected graph is k-neighbored for all $k \leq |V| - 1$. For k large enough, k-neighbored graphs allow us to construct large enough initial subsets. The next theorem now follows fairly directly[11].

THEOREM 4.1. *For any two nodes s, t in a graph of an Expansion family, if s and t are $c \log n$-neighbored for any constant $c > 0$ then there almost surely exist $O(\log n)$-length paths between s and t. An Expansion family, using ($c \log n$)-neighbored base graphs where $c > \frac{6q\xi}{(q\xi-1)^2}$, has expected diameter $O(\log n)$.*

Thus, a graph from an *Expansion family* almost always consists of a giant component with diameter $O(\log n)$ and perhaps some small components of size $O(logn)$. There are perhaps random (directed) links between the components (but only in one direction between a given pair).

Using super-nodes. We now consider random graphs which use $\log n$-neighbored base graphs.

THEOREM 4.2. *Consider a family $\mathcal{FRG}(\mathcal{H}, \tau, q)$, which is ($\mu_1$,$\xi$)-EXP and ($\mu_2$,$\xi$)-IE for some constants $\xi, \mu_1, \mu_2 > 0$, where $\epsilon(\tau) = \Omega(n^{-\mu_3})$ for some constant $\mu_3 < \mu_1 + \mu_2$, and all base graphs in $\mathcal{H}$ are $\log n$-neighbored. There almost surely exists a path of length $O(\log n)$ between any two nodes (for n large enough).*

Proof. This theorem is a simple corollary of the previous theorem if q is s.t. $\xi > 1/q$. However, for $q < Q = \lceil 1/\xi \rceil$ the theorem still holds. The main idea is to form super-nodes with Q random links. The $\log n$-neighbored property assures that we can always partition the graph into super-nodes each of which is a subgraph of constant diameter and has at least Q random links. The length of a path constructed here differs by only a constant from before (when we have $q \geq Q$).

[10] Note that we can construct similar classes by using μ-*Expansion* and δ-symmetric property instead.

[11] In fact, a full proof of it is very similar to that of theorem 14 in our previous work [20].

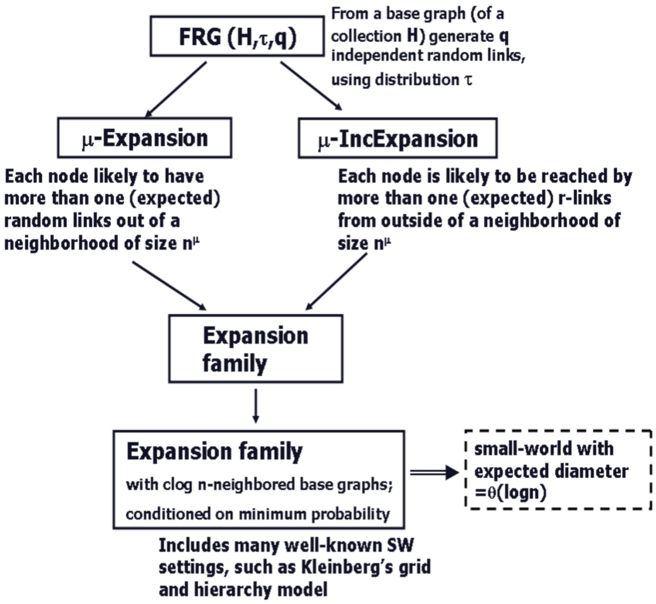

Figure 3: The hierarchy of classes

These abstract classes for (almost) small-world graphs are broad enough to accommodate many different well-known small-world models: Bollobas and Chung's [6], Watts and Strogatz's [24], Kleinberg's grid-based [16], tree-based, and group-induced models [15]. Kleinberg describes his group-induced model with two abstract properties, and it is not hard to see that the second property implies our (μ,ξ)-IE for some $0 < \mu, \xi < 1$. We show that our results apply to Kleinberg's tree-based model in the following section. It is relatively straightforward to extend this case for similar results in the group-induced model. Figure 3 shows how the classes relate.

4.4 The diameter of a tree-based random graph

We now use our framework to analyze the diameter of Kleinberg's tree-based model [15] and its variants. Kleinberg shows that decentralized routing can be applied in more settings (not only the grid-based [16]), but even when no lattice structure appears at all (say, the network of the Web's hyper-links). Kleinberg also introduces a group-induced model, a generalization of both grid-based and tree-based models [15]. He shows that using these models, greedy routing takes expected time $O(\log n)$ if nodes have out-degree $\theta(\log^2 n)$, and $O(\log^4 n)$ if the degrees are bounded by a constant.

In Kleinberg's tree-based model nodes are the leaves of a complete (for simplicity) b-ary tree T, where b is a constant. Let $h(u, v)$ denote the height of the least common ancestor of u and v in T. There are no local links in this setting but there are a number of directed random links leaving each node u, under a distribution τ, where a link is to v with probability proportional to $b^{-h(u,v)}$.

If there are exactly q directed random links leaving each node, the graphs in this tree-based setting are very likely unconnected (similar to the case of lacking local links in the grid-based setting [19]), however, the setting can still be an *Expansion family* by adding proper conditions. From [15], the normalizing coefficient of this link distribution is $\theta(\log^{-1} n)$. So, $\epsilon(\tau) = \theta(n^{-1} \log^{-1} n)$; thus, to have an *Expansion family* we need this setting to meet (μ_1,ξ)-EXP and (μ_2,ξ)-IE for some $\xi > 1/q$ and $\mu_1 + \mu_2 > 1$. Consider the following fact which holds even if $q = 1$.

FACT 4.1. *For Kleinberg's tree graphs with any $q \geq 1$, given a positive $\theta < 1$, a node u and $\mathcal{C} \subset V$ with size at most n^θ, the probability that a random link from u hits a node outside of $\mathcal{C}$ is more than $1-\theta-o(1)$ when n is large enough. Also, the probability that there is a random link to u from outside of $\mathcal{C}$ is more than $1 - e^{\theta+o(1)-1}$ (i.e. almost $1 - e^{\theta-1}$) when n is large enough.*

See [20] for a proof of a similar fact. It is easy to see that the setting meets ($x - o(1)$,$1 - x$)-EXP and ($y - o(1)$,$1 - e^{y-1}$)-IE for any $0 < x, y < 1$. Therefore, given q, we need to find x, y s.t.

$$x + y > 1;\ q(1 - x) > 1;\ q(1 - e^{y-1}) > 1$$

Solving this system of equations, we find $q \geq 3$.

THEOREM 4.3. *For $q \geq 3$, Kleinberg's tree-based setting is an Expansion family.*

We can add in local links to make the base graph connected or make the base graph $c \log n$-neighbored: ring all the nodes in the base graph H or alternately, ring all the subtrees of height at most $\log_b(c \log n)$. With c determined as in theorem 4.1, this setting will have expected diameter $O(\log n)$.

5 Random graphs induced by semi-metric or metric functions

We have abstracted away topological features of Kleinberg's grid setting with our expansion criteria to create classes where the strongest has $O(\log n)$ expected diameter. We now generalize the use of a distance measure in the distribution of random links, and this makes greedy-like routing (defined later) work. We design classes of random graphs using distributions based on semi-metric functions: we define a semi-metric function $d(u, v)$ and generate random links between any two nodes u and v with probability $\propto d^{-r}(u, v)$. We omit proofs in this section, which can be found in [23].

Consider a pair (G, d): a graph $G = (V, E)$ and a function $d = d_G : V^2 \to \mathbf{R}^+$ associated with G. We define d to be a semi-metric function if for any $u, v \in V$, $d(u, v) = 0 \Leftrightarrow u = v$; and $d(u, v) = d(v, u)$. We define $N_k(u) = \{v \in V | d(u, v) \leq k\}$, the nodes within 'distance' k of u. For $c_1, c_2 > 0$, graph G is called (c_1, c_2) *linear-expanded* with respect to d if $\forall u \in V, k = 1, 2 \ldots : c_1 \leq \frac{|N_k(u)|}{k} \leq c_2$ if $N_{k-1}(u) \neq V$, i.e. $|N_k(u)|$ grows nearly-proportionally to k before $N_k(u)$ becomes V.

DEFINITION 7. ($\mathcal{I}nv\mathcal{D}ist$ FAMILY) *An $\mathcal{I}nv\mathcal{D}ist(r)$ is a $\mathcal{FRG}(\mathcal{H}, \tau, q)$ family where each base graph $H \in \mathcal{H}$ has an associated metric-function d and there exists constants $c_1, c_2 > 0$ s.t. H is (c_1, c_2) linear-expanded* w.r.t *d, and where* [12]*: $Pr[R_\tau(u) = v] \propto d^{-r}(u, v)$.*

All Kleinberg's small-world models (grid-based, tree-based and group-induced) fall into $\mathcal{I}nv\mathcal{D}ist(1)$ for an appropriate d. For example, for Kleinberg's 2-D grid model [16], we define $d(u, v)$ as the square of the lattice distance between u and v; for Kleinberg's group-induced model [15], we define $d(u, v)$ as the size of the minimum set containing both u and v [13].

THEOREM 5.1. *$\forall r : 0 < r < 1, \delta > 0, c_2 > c_1 > 0$, $\exists q \geq 1$ s.t. any δ-symmetric $\mathcal{I}nv\mathcal{D}ist(r)$ family specified by c_1, c_2 and q (as in definition 7) is an Expansion family .*

For any graph from a δ-symmetric $\mathcal{I}nv\mathcal{D}ist(r)$ family using $\log n$*-neighbored base graphs, there almost surely exists an $O(\log n)$ length path between any two nodes*[14].

5.1 Greedy-like routing.

This section constructs a new class of graphs where most pairs of nodes have shortest paths of length $O(\log n)$ and *greedy-like paths* (defined below) with expected length $O(\log^2 n)$. Inspired by Kleinberg's idea of greedy routing using only local information [16], we assume that each node u knows the random links which leave nodes in a small *neighborhood* near u (e.g. the $\log n$ nodes closest to u in the base graph). *Greedy-like paths* are paths found by a *greedy-like algorithm* which is defined as follows: if the current node is u, choose the random link (w, v) where w is in $u's$ neighborhood and v is the closest such node to the destination. Route to w using local links and then take link (w, v). Update v to be the current node. We now present new definitions and then our theorems for this routing strategy.

We restrict $d(u, v)$ to be a *'light'* metric by adding the condition that $d(u, v) \leq \alpha(d(u, w) + d(w, v))$ for any nodes u, v, w and for a constant α (so less strict than the triangle inequality). We *define class $\mathcal{METR}(r)$ as class $\mathcal{I}nv\mathcal{D}ist(r)$ but each function d is a light metric function instead.* All Kleinberg's small-world models (grid-based, tree-based and group-induced) are $\mathcal{METR}(1)$ families with the function $d(u, v)$ derived naturally from each model's context. Except for the 1-D and the tree-based setting, this function is not a metric. For the tree-based setting, let $d(u, v)$ be the number of leaves in the smallest subtree containing u and v (this satisfies the triangle inequality). For the group-induced model, we let $d(u, v)$ be the size of the smallest group containing nodes u and v. This generally doesn't satisfy the triangle inequality, but satisfies ours for a proper α.

We now add neighboring conditions so our greedy-like routing strategy can be used. An undirected base graph $H(V, E)$ is called *k-strongly neighbored* if for each $u \in V$, the sub-graph induced by the set of nodes v such that $d(u, v) \leq k$ is connected.

THEOREM 5.2. *For any graph from a $\mathcal{METR}(1)$ family using* $\log n$*-strongly neighbored base graphs, a greedy-like algorithm will find paths of expected length $O(\log^2 n)$ between any two nodes.*

Combining theorems 5.1 and 5.2, we have:

THEOREM 5.3. *For any graph from a $\mathcal{METR}(1)$ δ-symmetric family using* $\log n$*-strongly neighbored base graphs, there almost surely exists a path of length $O(\log n)$ and a greedy-like path of expected $O(\log^2 n)$ between any two nodes.*

This theorem easily applies to Kleinberg's grid, tree-based and group-induced models with proper local links (to make the base graphs $\log n$-strongly neighbored).

5.2 Diameter of $\mathcal{METR}(r)$ for $1 < r < 2$.

We now present a natural generalization of our results in §3. We consider the diameter of $\mathcal{METR}(r)$, where $1 < r < 2$.

THEOREM 5.4. *For $1 < r < 2$, for a $\mathcal{METR}(r)$ family, there exists a constant $\hat{c}$ s.t. if the base graphs are x_0-strongly neighbored, where $x_0 = \hat{c} \log^{\frac{2}{2-r}} n$ (n: number of vertices), then almost surely the expected diameter of this family is upper bounded by a poly-log function.*

---

[12] Note that any family satisfying all these criteria except having $c_1 \leq \frac{|N_k(u)|}{k^\beta} \leq c_2$ instead (for some given constant $\beta > 0$) can be normalized by using function $d'(u, v) = d^\beta(u, v)$ instead, and hence becomes an $\mathcal{I}nv\mathcal{D}ist(r')$ family where $r' = r/\beta$.

[13] It is not hard to see that the second property (of the two abstract properties Kleinberg uses to describe his group-induced model) implies that $|N_k(u)|$ grows nearly-proportionally to k.

[14] Note that if we only use undirected random links then the condition of δ-symmetry is not necessary.

For example, we can modify Kleinberg's tree-based setting to become a small-world graph with poly-log expected diameter as follows. We connect all the nodes together (say, order the nodes from left to right and connect them with undirected edges), and for any two nodes u and v, we add an arc from u to v with probability proportional to $b^{-rh(u,v)}$ instead of $b^{-h(u,v)}$ with $1 < r < 2$. Note that, under the context of this section we define $d(u, v)$ as the number of leaves in the smallest subtree [15] which contains both u and v: $b^{h(u,v)}$.

6 Concluding remarks

We consider a general construction of random graphs: a base graph plus random links added to each node. By gradually adding properties to the base graphs and/or the distribution of random links, we build a hierarchy of classes of random graphs with the finest ones featuring small-world properties (small diameter and greedy-like routing using local information only). Thus, we propose a framework for analyzing and characterizing small-world graphs.

There are still some open questions in our study of 'adding links with probability $\propto$ the inverse distance'. As noted before, the case $r = 2k$ in the k-D grid setting is still open. We also expect to extend our results in §5 for base graphs with restricted growth rate[16], a general class of graphs which can be used to model many real networks [12]. Thus our work can be useful for a practical design problem, where we want to add in "long links" to a given network to shrink its diameter.

References

[1] J. Aspnes, Z. Diamadi, G. Shah, " Fault-tolerant Routing in Peer-to-peer Systems", Proc. 22nd ACM Symp. on Princ. of Dist. Comp., PODC'02, p. 223-232

[2] L. Barriere, P. Fraigniaud, E. Kranakis, D. Krizanc, "Efficient Routing in Networks with Long Range Contacts", Proc. 15th Intr. Symp. on Dist. Comp., DISC'01, p. 270-284

[3] I. Benjamini and N. Berger, "The Diameter of a Long-Range Percolation Clusters on Finite Cycles" Random Structures and Algorithms, 19,2:102-111 (2001).

[4] M. Biskup, "Graph Diameter in Long-range percolation", submitted to Electron. Comm. Probab.

[5] B. Bollobas, "The Diameter of Random Graphs", IEEE Trans. Inform. Theory 36 (1990), no. 2, 285-8.

[6] B. Bollobas, F.R.K. Chung, "The diameter of a cycle plus a random matching", SIAM J. Discrete Math. 1,328-333 (1988)

[15] Or we can use the size of the subtree: $\approx \frac{b}{b-1} b^{h(u,v)}$

[16] Where, informally, any 'ball' with radius r around a node u has at most $O(r^{\beta})$ nodes for some fixed constant $\beta > 0$.

[7] F. Chung and L. Lu, "The Diameter of Random Sparse Graphs", Adva. in App. Math. 26 (2001), 257-279.

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

[9] P. Dodds, R. Muhamad, and D. Watts, "An experimental study of search in global social networks", Science, 301:827-9, 2003.

[10] P. Erdos and A. Renyi, "On random graphs", Publicationes Mathematicas 6, 290-7 (1959).

[11] P. Fraigniaud, C. Gavoille and Christophe Paul, "Eclecticism Shrinks the World", Proc. 23rd ACM Symp. on Princ. of Dist. Comp. PODC'04, p. 169-178.

[12] J. Gao and L. Zhang, "Tradeoffs between Stretch Factor and Load Balancing Ratio in Routing on Growth Restrict Graphs", Proc. 23rd ACM Symp. on Princ. of Dist. Comp. PODC'04, p. 189-196.

[13] R. Karp, "Probabilistic Recurrence Relations", in *Journal of the ACM* 41(6):1136-1150, Nov. 1994.

[14] D. Kemper, J. Kleinberg, and A. Demers, "Spatial gossip and resource location protocols", in Proc. ACM Symp. Theory of Computing, 163-172, 2001.

[15] J. Kleinberg, "Small-World Phenomena and the Dynamics of Information", in *Advances in Neural Information Processing Systems (NIPS)* 14, 2001.

[16] J. Kleinberg, "The Small-World Phenomenon: An Algorithmic Perspective", in Proc. 32nd ACM Symp. Theory of Comp., 163-170, 2000.

[17] M. Kochen, Ed., *The Small-World* (Ablex, Norwood, 1989)

[18] G. S. Manku and M. Naor and U. Wieder, "Know Thy Neighbour's Neighbour: The Role of Lookahead in Randomized P2P Networks", Proc. 36th ACM Symp. on Theory of Comp. STOC'04, p. 54-63.

[19] C. Martel and V. Nguyen, "The complexity of message delivery in Kleinberg's Small-world Model", Tech-Report CSE-2003-23, CS-UCDavis, 2003.

[20] C. Martel and V. Nguyen, "Analyzing Kleinberg's (and other) Smallworld Models", Proc. 23rd ACM Symp. on Princ. of Dist. Comp. PODC'04, pp. 179-188.

[21] D. Malkhi, M. Naor, D. Ratajczak, "Viceroy: A Scalable and Dynamic Emulation of the Butterfly", Proc. 21st ACM Symp. on Princ. of Dist. Comp., p. 183-192.

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

[23] V. Nguyen and C. Martel, "Analysis and Models for Small-World Graphs", Tech-Report CS-UCDavis, 2004. www.cs.ucdavis.edu/~martel/main/smallworld.tech.pdf

[24] D. Watts and S. Strogatz, "Collective Dynamics of small-world networks" Nature 393,p. 440-2 (1998).

[25] H. Zhang, A. Goel and R. Gividan, "Using the Small-World Model to Improve Freenet Performance", IEEE INFOCOM'02, vol. 21(1), pp. 1228-37.