The Complexity of Message Delivery in Kleinberg's Small-world Model

C. Martel, V. Nguyen
{cumartel vnknguyen}@ucdavis.edu

Technical Report CSE-2003-23
UC Davis Department of Computer Science

February 22, 2004

Abstract

We prove additional results about Kleinberg's Small-World (KSW) model and its extensions. Kleinberg adds a number of directed long-range random links to an $n \times n$ lattice network (vertices as nodes of a grid, undirected edges between any two adjacent nodes) under a specific non-uniform distribution. He shows that the following phenomenon occurs: between almost any two nodes there exits a short path with length at most $O(log^2 n)$ which can be found using a simple greedy algorithm which has no global knowledge of long-range links. Kleinberg proves that the expected delivery time using his algorithm is $O(log^2 n)$.

We show that indeed Kleinberg's algorithm achieves $\theta(log^2 n)$ delivery time. Moreover, we show that the expected diameter of the graph is $\theta(logn)$, so a $logn$ factor smaller. These results are proved for the general k-dimensional model and our diameter result extends traditional work on the diameter of random graphs which largely focuses on uniformly distributed arcs. Using little additional knowledge of the graph, we show that we can find paths with expected length $O(log^{3/2} n)$ in the basic 2-dimensional model and $O(log^{1+1/k} n)$ in the general k-dimensional model (for $k \geq 1$). Finally, we suggest open problems and initiate the study of further generalized models where grid-related factors (e.g. the use of lattice distance) get weaker roles or are dismissed, and constraints (such as the uniformness of degree distribution) are relaxed.

1 Introduction

Small-world networks (SWN) have been an active and common topic in various disciplines, including both 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), which was discovered by S. Milgram in his pioneering work in the 1960's [14]. A number of network models have been proposed as a framework to study this phenomenon. Recently, J. Kleinberg (1999) has proposed a family of SWNs to study another compelling aspect of Milgram's original findings: a decentralized algorithm operating only on local information can construct short paths.

As Kleinberg has commented, it is striking (yet little studied) that the short paths not only exist but can be found without requiring prior knowledge of the global network. Algorithmic results in this area not only improve our understanding of many practical network structures but also bring in potential applications related to routing problems in

the Internet. Thus, modelling this feature of SWNs with emphasis on algorithmic aspects is well motivated. Kleinberg's model is the first to demonstrate this idea.

Kleinberg's basic model is a structure with a two-dimensional grid as a base and 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 (Kleinberg called this the inverse second-power distribution). In the basic model, from each node, there is a *local* link to each of its four grid neighbors and one *long-range* random link. In this setting Kleinberg shows that a simple greedy algorithm using only local information finds routes between any source and destination using only $O(\log^2 n)$ expected links [12].

There are two major reasons that Kleinberg's model is notable. First, it provides "latent structural cues", what Kleinberg called, which can be used by a routing agent to advance quickly towards the target. Kleinberg's model is unique in its family, and the only one known so far, featuring this capability. If the long-range link instead use the inverse r^{th}-power distribution where $r \neq 2$, it is hard to find short paths even though they still exist. For instance, if we change Kleinberg's model so the long-range links are distributed uniformly ($r = 0$), then the model has diameter $O(logn)$ [4], however there is no decentralized algorithm to find paths with length $o(n^{2/3})$ [12]). Thus, the distribution of long-range random links must correlate with the geometry of the lattice in a precise way to make this special capability emerge.

Second, the lattice is a fundamental network structure, often used in modelling, but its coupling with the addition of random links whose distribution depends on the lattice distance, is new. The main body of random graph theory uses a uniform distribution.

Kleinberg leaves two important issues open in the analysis of routing in his model and we complete this analysis in this paper. First we show that the $O(\log^2 n)$ expected time analysis is tight (thus except for pairs which are quite close, Kleinberg's algorithm uses expected $\Theta(\log^2 n)$ links).

Our second main result shows that the expected diameter of this graph is $\Theta(\log n)$. This extends traditional work on the diameter of random graphs which largely focuses on uniformly distributed arcs [5]. This diameter result shows that an algorithm with global knowledge of the random links can improve on Kleinberg's decentralized algorithm by a log factor.

We now give further details on these results.

Decentralized routing. First, we review the model introduced by Kleinberg [12]. Kleinberg's graph is based on an $n \times n$ two-dimensional grid. The vertices are nodes on the grid, which can be identified as lattices points (i, j) with $i, j \in \{0, 1, \ldots, n-1\}$. The lattice distance between two nodes $u(i, j)$ and $v(k, l)$ is defined as $d(u, v) = |k - i| + |l - j|$. Each node u has local links to all its neighbors within lattice distance $p \geq 1$ and $q \geq 1$ long-range random links with endpoints generated by independent random trials. For any $i \leq q$, the i^{th} directed long-range link from a node u has endpoint v with probability proportional to $d(u, v)^{-2}$ (*inverse second power distribution*).

We study a routing problem in this grid setting that is, for any two nodes s, t with known lattice coordinates to find a reasonably short path from s to t. As mentioned before, we do not know the global topology of the network and at any step a routing algorithm can only use local information, which includes the random long-range links from the current node and perhaps from other nodes in a small neighborhood. Algorithms which operate on this restricted condition are called decentralized algorithms. We call such a neighborhood around a node u (when becomes the current node) *the view at* u.

Kleinberg studied the situation that the view at any node contains only this node and suggested a simple greedy algorithm which achieves *'delivery time'* $O(log^2 n)$ in the basic setting (when $p = 1$ and $q = 1$). That is the resulting $s - t$ paths have $O(log^2 n)$ expected length [12]. The basic idea is, when standing at a node u (i.e. the part from s to u has

been drawn up), the algorithm chooses the next node to be the closest to t (with respect to lattice distance) amongst u's local/long-range contacts. This simple rule is iteratively applied until t is reached and no 'backtracking' is used.

We show that $O(log^2 n)$ is indeed a tight bound in Kleinberg's algorithm. In order to do that we look carefully at the correlation between the geometry of the lattice and the inverse second power distribution of the random links; *the main result is to show that each step will reduce the remaining distance to the destination by an expected factor of* $c/logn$ where c is a constant. We also introduce a new decentralized algorithm. Our algorithm will result in significantly shorter expected $s-t$ paths but uses 'more resources': the view at each node is larger. Moreover, we extend our results so that they apply to the general k-dimensional model.

Bounding the expected diameter. The main part of our work is here. We first show that the expected diameter of one of the 'easiest' models, where $p = 1$ and $q = 4$, is $O(logn)$ then extend our result to more general classes of graphs, especially ones with higher dimensions. We now briefly discuss our main idea.

At a first glance, $O(logn)$ seems a natural bound. Suppose we start at a node s and want to form paths to go to all the other nodes. Applying a 'tree' view, we may name the four long-range contacts of s as s's children. Now, let us define a function χ which takes a node as input and outputs its child nodes. Thus, we can build a tree of s's descendants generated by several steps of applying χ on the nodes of the preceding level (the first level contains s only). Thus, if we (probably falsely) assume the set of children of two separate nodes are almost disjoint, the 'tree' will grow exponentially and very quickly fill up the whole grid in about $log_4 n^2$ χ-steps. Note that the exponential growth does occur if the long-range random contacts are uniformly distributed (see [4]).

However in our setting, the random links generated by this inverse second power distribution tend to jump to nodes not too far from the source nodes; many 'random' descendant nodes can coincide. This possibility may lead to some portion of 'the ground' is very crowded while some other parts, possibly large, are unoccupied. Thus, we must justify thinking of a tree-like, exponentially growing process.

Our basic idea to prove the bound of $O(logn)$ is to look at the early stage of the development of this graph. In this early stage, the tree still has a modest size, and coincidence of descendant nodes is not too common (say, the probability of this is less than a half); thus our tree still grows rapidly. Indeed, at each step of growing this 'tree', we can trim off some portion of it to make it remain a true tree. We show that the growth rate per each χ-step is at least a constant $\gamma > 1$ with overwhelming probability. Later, after the tree size exceeds a critical threshold, the late descendants start to coincide often, and the growing rate starts to drop quickly. Thus it will not be useful to look further in this direction and we stop growing our tree after this turning point.

We now continue by adapting a relatively well-know technique. We also consider another similar tree which roots at t, an arbitrary node different than s. The only difference here is that χ will be replaced by a function, outputting the nodes *from* which there is a random link to the input node. If both trees are large enough (say, $\Omega(nlogn)$ for an $n \times n$ grid) then we can comfortably prove that the two trees will connect with overwhelming probability. Therefore, the key point lies in how to give an $O(logn)$ bound on the heights of the two trees of size $\Omega(nlogn)$.

We also generalize our approach so that we can solve diameter problems in many additional classes of graphs.

The continuum. Kleinberg's delivery time seems to be an important parameter, comparable to the diameter, in a network structure. The two parameters present our knowledge of a network structure in terms of transmission rate and capability of routing algorithms, since they let us know the expected time of delivering a message in two extreme scenarios:

when we know the full network topology and when we know only the links from the current node. *The two parameters present a continuum on the delivery time*: between them many different rates are possible when our knowledge of the network graph and resources availability varies. Thus, we face a trade-off between time and resources.

Kleinberg's model has a log factor continuum while the uniform model with $r = 0$ has a polynomial one. It would be worth thinking of potential scenarios (such as routing in peer-to-peer networks) knowledge about this measure in practical network systems can be of significant benefit.

Related work. There has been considerable work on the small-world phenomenon. See [13] for early surveys and [12] for a more recent account on modelling small-world networks. Before Kleinberg's model, Watts and Strogatz [15] have proposed one of the most refined models, constructed by randomly rewiring the edges of a ring lattice each with a probability parameter p. With p varying, the model's structural properties can be considered through the two key characteristics: path length $L(p)$, measuring "the typical separation between two vertices in the graph", and clustering coefficient $C(p)$, measuring "the cliquishness of a typical neighborhood". Watts and Strogatz observed that for small p the model reflects many practical small-world networks with small typical path length $L(p)$ and non-negligible clustering coefficient $C(p)$.

Applications have been found using Kleinberg's small-world model or the ideas behind it, such as decentralized search protocols in peer-to-peer systems [16], gossip protocols for spreading information in a communication network [9], and secure distributed protocols in cryptographic scenarios [8]. See also [10] for a generalization that encompasses both lattice-based and tree-based ("taxonomic" or "hierarchical") small-world networks.

The diameter of random graphs has been an important issue in random graph theory for quite a while [2, 3, 4, 5, 6] but there has also been a tradition of using uniformly distributed arcs. The work which is closest to our diameter result is Bollobas and Chung's [4], where the graph model is very similar to that in [15] with the nodes of a cycle (or a "ring") randomly matched to form additional long-range links. See also [1] for a practical approach to measure the diameter of the World Wide Web using a simulation model based on the power law distribution [7].

The structure of this paper. Section 2 presents our definitions and supporting facts on Kleinberg's grid setting (we call it Kleinberg's Small-world (KSW) setting sometimes). Section 3 discusses our bound on Kleinberg's delivery time and introduces our alternative algorithm. Sections 4 and 5 study the diameter of graphs based on Kleinberg's models and possible extensions. We finish with our conclusions and future work in the final section.

2 Basic facts on Kleinberg's grid setting

We now present our notation and basic facts on the KSW setting. Let V denote the set of all nodes, the size of which is n^2. We use $\mathcal{K}(n, p, q)$ to refer to the class of all random graphs based on Kleinberg's model (using an $n \times n$ grid, local links to all nodes within distance p, and q long range links/node). We also use $\mathcal{K}^*(n, p, q)$ for a similar class, defined with respect to the lattice distance with wraparound: $d(u, v) = min\{|k - i|, n - |k - i|\} + min\{|l - j|, n - |l - j|\}$. For simplicity we sometimes start with $\mathcal{K}^*(n, p, q)$ to show our main results and then adapt our solution to $\mathcal{K}(n, p, q)$.

Let $B_l(u)$ denote a 'ball' of radius l and center u (actually shaped as a diamond in a two dimensional grid), the set of all the nodes within lattice distance l from u. Define $b_l(u)$ as the number of nodes on the 'surface' of that ball $B_l(u)$, or the number of nodes of distance l from u. Also call a set S a *like-a-ball* with center u if $B_l(u) \subseteq S \subseteq B_{l+1}(u)$ for some node u and integer l.

We define $\chi(u)$ with $u \in V$ as the set of nodes which can be reached from u by a random link. Also let $\chi_i(u)$, for $i = 1, 2, \ldots, q$, be the node incident to each of the q random links from u. Note that these $\chi_i(u)$ may coincide, so $|\chi(u)| \leq q$. We also define $\chi(S) = \bigcup_{u \in S} \chi(u)$, for any $S \subset V$. For any two nodes u and v, if $\chi_i(u) = v$ for some $i = 1..q$ then, we say, the random link (u, v) has label i.

We now consider some basic facts in the special setting where $q = 1$ which can be easily extended to ones with arbitrary $q \geq 1$. For simplicity, let us refer to the two classes above with $q = 1$ as $\mathcal{K}^*$ and $\mathcal{K}$.

For any two nodes u and v, let $p(u, v)$ denote the probability that there is a random link from u to v. Since $p(u, v)$ is proportional to $d^{-2}(u, v)$ we have $p(u, v) = C_u d^{-2}(u, v)$, where C_u is the normalized coefficient, a constant only depending on the position of u. Meanwhile, define the inverse normalized coefficient $c_u = C_u^{-1}$, then clearly, $c_u = \sum_{\forall v \neq u} d^{-2}(u, v) = \sum_{j=1}^{2n-2} b_j(u) j^{-2}$. Note that we always have $p(u, v) = p(v, u)$ if the graph is from $\mathcal{K}^*$, and so c_u is the same for all u; call this value by c^* and its inverse by C^*.

In a graph from $\mathcal{K}(n, p, q)$, note that the nodes at lattice distance j from u actually form a diamond (see figure 1, part a) or a part of a diamond (since some part may lie outside of the grid). If this diamond is totally inside then obviously, $b_j(u) = 4j$. If it is totally outside then $b_j(u) = 0$. Otherwise, as often, a part of this diamond is inside and the remainder outside, and when a complete edge is inside then obviously, $b_j(u) \geq j$. It is easy to observe that in a graph from $\mathcal{K}(n, p, q)$,

$$\begin{aligned} \forall j \geq 1 : b_j(u) \leq 4j \\ \forall j = 1..\lceil n/2 \rceil - 1 : b_j(u) \geq j \end{aligned} \tag{1}$$

Equivalently, in a graph from $\mathcal{K}^*(n, p, q)$,

$$\begin{aligned} \forall j \geq 1 : b_j(u) \leq 4j \\ \forall j \leq \lceil n/2 \rceil : b_j(u) = 4j \end{aligned} \tag{2}$$

as a result we have

$$c^* = \sum_{j=1}^{\lceil n/2 \rceil} (4j)(j^{-2}) = 4 \sum_{j=1}^{\lceil n/2 \rceil} j^{-1} \tag{3}$$

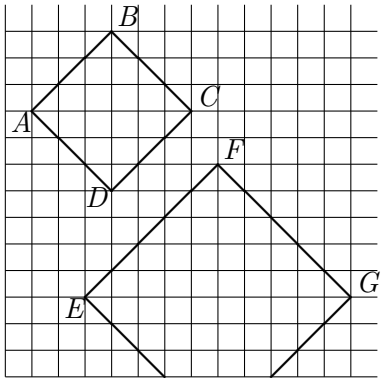
Therefore, $4ln(n/2) < c^* < 4ln(3n/2)$, so $c^* = 4lnn + O(1)$. Meanwhile, if the graph is from $\mathcal{K}$ then from (1), $lnn \leq c_u \leq 4ln(6n)$. Thus for all u, c_u (or c^* for graphs in $\mathcal{K}^*$) is $\theta(logn)$. For any two nodes u and v, obviously $d(u, v) \leq 2n$ therefore $p(u, v) = C_u d^{-2}(u, v) = \Omega((n^2 logn)^{-1})$. We summarizes these by the following fact.

Fact 1. *For graphs from $\mathcal{K}^*$ or $\mathcal{K}$, the inverse normalized coefficient $c_u = \theta(logn)$ and for any two distinct nodes u and v, $p(u, v) = \Omega((n^2 logn)^{-1})$.*

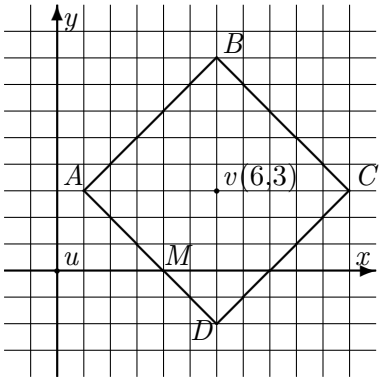
2.1 Links into or out of a ball

Our analysis on the expected diameter (section 4) will be based on this pair of elementary experiments: to see if a random link from the center of a ball is to a node outside the ball and to see if there is a random link from outside the ball to its center.

Fact 2. *On a graph from $\mathcal{K}^*$, given any positive $\theta < 1$, any integer $1 \leq l \leq n^\theta$, for n large enough: i) the probability that a random link from a given node u goes to a node outside of $B_l(u)$ is greater than $1 - \theta - o(1)$; ii) the probability that there is a random link to u from a node outside of $B_l(u)$ is greater than $1 - e^{\theta + o(1) - 1}$ (i.e. almost $1 - e^{\theta - 1}$).*



a)



b)

Figure 1: a) a 'ball' is a full diamond ($ABCD$) while the other is not (KEFGH); b) diamond $ABCD$, determined by $|x-6|+|y-3|=5$, can be decomposed into $AB(x-y=-2)$, $BC(x+y=14)$, $CD(x-y=8)$ and $DA(x+y=4)$. Note that DA has $x\in[1,6], y\in[-2,3]$, and thus it will be split into $AM(x\in[1,4], y\in[0,3])$ and $MD(x\in[4,6], y\in[-2,0])$.

Note that when θ is about 0.5 (the values we will be mostly interested in), these two events may happen with probabilities about 0.5 $(1-\theta)$ or more and about 0.39 $(1-e^{\theta-1})$ or more, respectively.

Proof. For i) let E be the event that u has a random link going to a node $v\notin B_l(u)$. We have $Pr[\overline{E}] = C^*\sum_{v\in B_l(u)} d^{-2}(u,v) \le \sum_{j=1}^{l} b_j(u)(j^{-2})/c^*$

From (2) we have $b_j(u)\le 4j$, then $\sum_{j=1}^{l} b_j(u)(j^{-2}) \le 4\sum_{j=1}^{l} j^{-1} \le 4ln(3l) \le 4ln(3n^\theta)$. From the proof of fact 1 above note that $c^* > 4ln(n/2)$ so, $Pr[\overline{E}] \le \frac{4ln(3n^\theta)}{4ln(n/2)} \le \frac{\theta lnn+ln3}{lnn+ln(1/2)} = \theta + o(1)$. Thus $Pr[E]\ge 1-\theta-o(1)$.

For ii) let F be the event that there is a random link coming to u from a node v outside $B_l(u)$; we have

$$Pr[F] = 1 - Pr[\overline{F}] = 1 - \prod_{v\notin B_l(u)} (1-p(v,u)) \ge 1 - \prod_{v\notin B_l(u)} e^{-p(v,u)} = 1 - e^{-\sum_{v\notin B_l(u)} p(v,u)} \quad (4)$$

Note that here we have used the well known fact $e^x \ge 1+x$ to obtain $e^{-p(u,v)} \ge 1-p(u,v)$. Since $p(u,v) = p(v,u) = d^{-2}(u,v)/c^*$, then $Pr[F] \ge 1 - e^{-\sum_{v\notin B_l(u)} p(u,v)} = 1 - e^{-Pr[E]} \ge 1 - e^{\theta+o(1)-1}$. □

Fact 3. *On a graph from $\mathcal{K}$, given any positive $\theta<1$ and integer $1\le l\le n^\theta$, for n large enough: i) the probability that a random link from a given node u goes to a node outside of $B_l(u)$ is greater than $\frac{1-\theta}{1+3\theta}+o(1)$; ii) the probability that there is a random link to u from a node outside of $B_l(u)$ is greater than $1-e^{-(1-\theta)/4+o(1)}$.*

The proof of this fact is in the appendix.

2.2 Extended models

The above definitions and facts can be easily extended to the more general $\mathcal{K}^*(n,p,q)$ and $\mathcal{K}(n,p,q)$, where $q\ge 1$. The only difference is that, if before we were concerned with the

existence of *a random link* from a node u to a node v then now we are instead concerned with the existence of *a random link with label* i (for any fixed given $i = 1..q$) from u to v.

Moreover, as a natural generalization, we can extend all these concepts and basic facts above to the classes of graphs based on a k-dimensional grid for $k \geq 1$. We redefine the distribution rule of the random links: $p(u, v)$ is now proportional to $d^{-k}(u, v)$. Now $\mathcal{K}^*(k, n, p, q)$ and $\mathcal{K}(k, n, p, q)$ are used to refer to such classes of graphs based on this k-dimensional grid (with size n in each dimension) with respect to the lattice distance and with or without wraparound, respectively.

In $\mathcal{K}(k, n, p, q)$, for a node u close to the edge of the grid, a ball centered at u may not be a 'full' ball but it is *easy to extend all our results with balls to cover such a case*; thus for simplicity we still use term 'ball' in general. We essentially only need to use the following fact to upgrade the facts $1 - 3$ and their proofs.

Fact 4. *For each $k \geq 1$ there exist positive constants c_1 and c_2 such that for a k-dimensional grid of size n, $c_1 \leq b_j(u)/j^{k-1} \leq c_2, \forall u \in V, \forall j = 1..n$.*

Proof. Without loss of generality, we assume that u is the origin of this k-dimensional space. Note that $b_j(u)$ is at most $2^k T$ where T is the number of nodes with non-negative coordinates and at distance j from u. We upper bound T using the following well-known result from elementary combinatorics: the number of ordered sequences of k non-negative integers with their sum equal given positive j is $\binom{j+k-1}{k-1}$. Note that each positive node on the surface of $B_j(u)$ correspond to such a distinct ordered sequence. Therefore $b_j(u) \leq 2^k \binom{j+k-1}{k-1} = \frac{2^k}{(k-1)!} \prod_{i=1}^{k-1}(j+i)$, which clearly is $\theta(j^{k-1})$. □

From this fact, $b_j(u)j^{-k} = \theta(j^{-1})$, which means that the sums of this $b_j(u)j^{-k}$ for some certain range of j can still be bounded by some log-expressions, which essentially keeps all our arguments almost as simple as before. For example, similarly as (3), we still have $c^* = \theta(logn)$. Thus it is easy to generalize fact 1 as follows.

Fact 5. *For graphs from $\mathcal{K}^*(k, n, p, q)$ or $\mathcal{K}(k, n, p, q)$, the inverse normalized coefficient $c_u = \theta(logn)$ and for any two distinct nodes u and v, $p(u, v) = \Omega((n^k logn)^{-1})$.*

Similarly, we can generalize both facts $2 - 3$. Note that for simplicity we do not look for exact bounds but just the existence of them, so knowing $b_j(u)j^{-k} = \theta(j^{-1})$ is enough.

Fact 6. *On a graph from $\mathcal{K}^*(k,n,p,q)$ or $\mathcal{K}(k,n,p,q)$, for any given positive $\theta < 0.6$, there exist positive constants ξ_1 and ξ_2 such that for n large enough: i) the probability that there is a random link (with a given label) from a node u to a node outside of $B_l(u)$, where $l \leq n^\theta$, is greater than ξ_1; ii) the probability that there is a random link (with a given label) to u from a node outside of $B_l(u)$ is greater than ξ_2.*

Note that for any given $\theta < 1$ we can find $\xi_1, \xi_2 > 0$ dependent on θ to have this fact satisfied, however we only need θ about 0.5 in our later analysis. Most of the above results are for supporting our work in sections 4 and 5.

2.3 Links to a spherical surface

We now study another probabilistic experiment, namely, if a random link from a given node u goes onto the surface of a given ball B; we only consider the case where u is outside of B. Our work in section 3 will be based on this experiment. Here we use the general models $\mathcal{K}(k, n, p, q)$ and $\mathcal{K}^*(k, n, p, q)$.

Define distance $d(u, B)$ between node u and ball B as the minimum lattice distance between u and the nodes on B's surface. Also call S_B the set of nodes 'on the surface' of ball B. Let $P(u, \mathcal{O})$ denote the probability that the random link from u goes to a node in a set of points (an object) $\mathcal{O}$.

Fact 7. *For any $k \geq 1$ there exists a constant $\hat{c}$ such that $P(u, S_B) \leq \frac{\hat{c}l}{m^2 logn}$, where $B = B_l(v)$ is a ball with center v and radius l on a graph from $\mathcal{K}(k, n, p, q)$ or $\mathcal{K}^*(k, n, p, q)$, u is a node outside of B, $d = d(u, v)$, $m = d(u, B)$ and $d \geq 2l$.*

Proof. From fact 4, there exists constant c such that $|S_B| = b_l(v) \leq cl^{k-1}$. The distance from u to any node in B is at least m. Therefore combining with fact 5, $P(u, S_B)$ is at most $cl^{k-1} \times \frac{c_1}{m^k logn} \leq \frac{\hat{c}l^{k-1}}{m^k logn}$ for some constants c_1 and $\hat{c} = cc_1$. For $k > 1$ and $d \geq 2l$, $l^{k-2} \leq m^{k-2}$ (note that $m = d - l > l$), so $P(u, S_B) \leq \frac{\hat{c}l}{m^2 logn}$. Meanwhile the case of $k = 1$ is trivial. □

Fact 8. *For any $k \geq 1$ there exists a constant $\hat{c} > 0$ such that $P(u, S_B) \leq \frac{\hat{c}}{mlogn}$ where B is a ball on a graph from $\mathcal{K}(k, n, p, q)$ or $\mathcal{K}^*(k, n, p, q)$, node u is outside B and $m = d(u, B)$.*

We prove the fact for $\mathcal{K}(k, n, p, q)$ but it is easy to extend our proof for $\mathcal{K}^*(k, n, p, q)$. Our proof is based on claims and 9 and 10 below. Once these claims are verified, fact 8 immediately follows. We need a few comments before showing the claims.

The bound we give on $P(u, S_B)$ depends on the distance from u to the ball only, and thus, is independent of the size of the ball. If we think of this probability as a measure of 'attractive force' (to u), the force generated by S_B is not stronger than the joint force generated by m nodes at about distance m from u. Thus a small fraction of nodes in S_B, which are at the 'pole' closest to u, generates a dominant term for $P(u, S_B)$.

Let node v and integer $l > 0$ be the center and radius of ball B. Let $m = d(u, B)$ and $d = d(u, v)$; clearly $d = m + l$. Let $\mathcal{Z}^+ = \{i \in \mathcal{Z} : i \geq 0\}$. Assume that the nodes on the grid are the integer points ('wrapped' by a hypercube with size n) in the space $\mathcal{R}^k$. For simplicity, we also assume the origin of $\mathcal{R}^k$ is at u (i.e. u has all coordinates zero) and $v(d_1, d_2, \ldots, d_k)$ is in the 'positive angle' relative to u, i.e. $v \in \mathcal{Z}^{+k}$. Also call $N(\mathcal{L})$ the number of integer nodes within an object $\mathcal{L}$.

Claim 9. *Consider a linear hyper-multilateral (LHM) object $\mathcal{L}$ which is defined by a linear system of the following form*

$$\begin{gathered} s_1x_1 + s_2x_2 + \ldots + s_kx_k = a \\ s_i \in \{-1, 1\}, x_i \in [a_i, b_i], i = 1..k \\ a > 0, a_i \times b_i \geq 0 \end{gathered} \tag{5}$$

where a, a_i, b_i, s_i are constant integers. For $u = (0, 0, \ldots, 0)$, if $|x_1| + |x_2| + \ldots + |x_k| \geq m$ then $P(u, \mathcal{L}) \leq \frac{c}{mlogn}$ for some constant c depending on k only.

Claim 10. *S_B can be decomposed into at most C LHM objects (determined by an instance of (5) above), where C is a constant dependent on k only.*

For $k = 1$, a ball is an interval, so a LHM is simply an endpoint of an interval. Similarly, a LHM is an edge of a diamond when $k = 2$, a face of a cube when $k = 3$, and so on. With claims 9 and 10, fact 8 follows immediately; thus we are just left to show them.

Proof of claim 9. From equation (5), we have either $0 \leq a_i < b_i$ or $0 \geq a_i > b_i$. We now assume $0 \leq a_i < b_i$ for the rest of the proof; it is easy to extend our proof for the other case. By a proper permutation of indexes we can make the characteristic equation become $x_1 + \ldots + x_j - x_{j+1} - \ldots - x_k = a$ for some integer j in $1..k$. Obviously this preserves $P(u, \mathcal{L})$. Also note that $x_1 + x_2 + \cdots + x_k \geq m$ since $x_i \geq a_i \geq 0, \forall i$.

We now consider the special case of $j = k$, i.e. $x_1 + \ldots + x_k = a$. Thus, all points in $\mathcal{L}$ are at distance a from u. From fact 4, $N(\mathcal{L}) \leq b_a(u) \leq c_1a^{k-1}$ for some constant

c_1. Therefore, from fact 5, $P(u,\mathcal{L}) \leq c_1 a^{k-1} \times \frac{c_2}{a^k logn} \leq \frac{c}{mlogn}$ for some constants c_2 and $c = c_1c_2$ (note that $a > m$).

Now consider the case of $j < k$. Set $x = x_1 + \ldots + x_j$ and $y = x_{j+1} + \ldots + x_k$ we then have $x - y = a$. Clearly, for each value $\xi = x + y$, there is an unique pair of x, y. Let $\mathcal{L}_\xi = \mathcal{L} \cap B_\xi(u)$, i.e. the subset of nodes in $\mathcal{L}$ with $x + y = \xi$. Let Z_1 be the $\mathcal{Z}^j$ space of $x_1, \ldots, x_j$ and Z_2 be the $\mathcal{Z}^{k-j}$ space of $x_{j+1}, \ldots, x_k$. We now project $\mathcal{L}_\xi$ onto Z_1 to obtain an image $\mathcal{I}_1$, which has characteristic equation $x_1 + \ldots x_j = x$, and onto Z_2 to obtain an image $\mathcal{I}_2$, which has characteristic equation $x_{j+1} + \ldots x_k = y$. Again by fact 4, we have $N(\mathcal{I}_1) = O(x^{j-1})$ and $N(\mathcal{I}_2) = O(y^{k-j-1})$. Thus, clearly $N(\mathcal{L}_\xi) = N(\mathcal{I}_1) \times N(\mathcal{I}_2) \leq c_3 x^{j-1} y^{k-j-1} \leq c_3 \xi^{k-2}$ for some constant c_3.

Therefore, from fact 5, $P(u,\mathcal{L}_\xi) \leq c_3 \xi^{k-2} \times \frac{c_4}{\xi^k logn} \leq \frac{c_3 c_4}{\xi^2 logn}$ for some constant c_4. Now, summing $\mathrm{P}(u,\mathcal{L}_\xi)$ over all possible values of $\xi \geq m$, we have $P(u,\mathcal{L}) < \sum_{\xi=m}^{\infty} \frac{c_3 c_4}{\xi^2 logn}$. Note that $\sum_{i=m}^{\infty} 1/m^2 \leq 1/m^2 + \sum_{i=m}^{\infty} \frac{1}{i(i+1)} \leq 1/m^2 + \sum_{i=m}^{\infty} (\frac{1}{i} - \frac{1}{i+1}) \leq 1/m^2 + 1/m$; thus $P(u,\mathcal{L}) \leq \frac{c}{mlogn}$ for $c = 2c_3c_4$. □

Proof of claim 10. We need to show that S_B can be decomposed into at most C LHM objects (determined by an instance of (5) above), where C is a constant dependent on k only. Clearly, a node $W(x_1, x_2, \ldots, x_k) \in \mathcal{Z}^k$ belongs to S_B, the surface of $B_l(u)$, if and only if $|x_1 - d_1| + |x_2 - d_2| + \ldots + |x_k - d_k| = l$. Note that $|x_1| + \ldots + |x_k|$ is at least $m = d(u, B)$.

Thus, S_B is composed of 2^k faces, each of which is on a hyper-plane determined by an equation of the form $\pm x_1 \pm x_2 \pm \ldots \pm x_k = a$, where integer constant a may vary for different hyperplanes. (There are 2^k combinations of these $+/-$ so we have 2^k such hyperplanes.) More specifically, such a face can be determined by a linear system similar to (5) but without $a_i b_i \geq 0$, where a, a_i, b_i are integer constants completely determined by $d_1, \ldots, d_k, n$ and l (for instance is it not hard to see that $a = l + s_1 d_1 + s_2 d_2 + \ldots + s_k d_k$). See figure 1 (part b) for an illustration. Call such a linear system a LES.

We can think of S_B as being decomposed into 'pieces'. Let us further decompose this as follows. For each axis-hyperplane $x_i = 0$, we divide the pieces, being 'cut through' by this hyperplane, each into two 'smaller pieces': one 'above' and one 'below' this hyperplane. For example, if a piece has LES with $a_i < 0 < b_i$ for some $i = 1..k$, we then split this object into two, each of which is described by the same LES with additional $x_i \in [0, b_i]$ for the 'the above' and $x_i \in [a_i, 0]$ for the 'the below'. We repeat this process until all pieces are inside a 'right angle': none of them are 'cut through' by any axis-hyperplane. Now each obtained object is described by an instance of (5) and the number of them is less than $2^k \times 2^k = 2^{2k}$. □

3 Decentralized routing

Kleinberg's algorithm uses a simple greedy strategy. Given a source s and a destination t, the algorithm always chooses the next message holder to be the node which is the endpoint of a link (local or long-range random) from the current holder and which is closest to t. Kleinberg proved that the expected number of links used (*Kleinberg's delivery time*) by the algorithm is $O(log^2 n)$ for the two-dimensional case [12], adding that it was straightforward to generalize his proof for the general k-dimensional one. His proof left open the question of whether this expectation is also $\Omega(log^2 n)$. We will prove that indeed Kleinberg's algorithm takes $\Omega(log^2 n)$ and introduce other algorithms with better delivery time.

3.1 Kleinberg's bound is tight

Theorem 11. *For any constant $c_1 > 0$, there exists a constant $c_2 > 0$ such that, for any two nodes s and t on a graph from $\mathcal{K}^*(k,n,1,1)$ or $\mathcal{K}(k,n,1,1)$ with $d(s,t) > c_1 n$,* Kleinberg's delivery time *with respect to endpoints s and t is greater than $c_2 log^2 n$ with probability at least 0.5, when n is large enough.*

This theorem shows that for the majority of $s-t$ pairs, *Kleinberg's delivery time* can not be $o(log^2 n)$, therefore $O(log^2 n)$ is indeed a tight bound. Note that we can actually make the probability above arbitrarily high, but we prove 0.5 for simplicity.

Let $\mathcal{A}$ denote Kleinberg's algorithm. For a node v we call $\kappa_t(v)$ the next node found by $\mathcal{A}$ when we are at v (initially, $v = s$) and seeking a path to the destination node t. Let $\delta_t(v) = d(v,t) - d(\kappa_t(v), t)$. We use $\kappa(v)$ and $\delta(v)$ when the destination t can be determined unambiguously; also let d_v denote $d(v,t)$. The ratio $\delta(v)/d_v$ characterizes the speed that $\mathcal{A}$ converges to t. The following lemma, which follows directly from facts 7 and 8, will help to estimate this speed.

Lemma 12. *For each $k \geq 1$ (the number of dimensions), there exists a constant $\hat{c}$ such that, for any two distinct nodes v and t, for any integer $1 < m < d_v$, $Pr[\delta(v) = m] \leq \frac{\hat{c}}{mlogn} \times min\{1, (d_v - m)/m\}$.*

As seen in [12], when we arrive at a new node v, we have never known about its random link before, hence we may assume that this random link is generated at this moment. Thus we can think of $\kappa(v)$ as an experiment in that we generate v's random link to w and set $\kappa(v) = w$ only if $d_w \leq d_v - 2$, i.e. w is closer to t than u by at least 2, otherwise we choose $\kappa(v)$ as v's grid neighbor closest to t. This is why the above lemma does not apply for $m = 1$. The following lemma is crucial to prove theorem 11.

Lemma 13. *For each $k \geq 1$, there exists a constant c such that $E[Z_v] \leq c/lnn$, where v and t are two distinct nodes at least distance lnn far apart and $Z_v = ln(\frac{d_v}{d(\kappa_t(v),t)})$.*

Proof. Let $d = d_v$. Note that $Z_v = ln(\frac{d}{d-\delta(v)})$. We define the function $f(i) = ln(\frac{d}{d-i})$. We need to show that $\sum_{i=1}^{d} f(i)Pr[\delta(v) = i] \leq c/lnn$ for some constant c. For simplicity, assume that d is an even integer. First, using the common fact that $ln(1+x) < x, \forall x > -1$, we have

$$\sum_{i=1}^{d/2} f(i)Pr[\delta(v) = i] \leq \sum_{i=2}^{d/2} ln(1 + \frac{i}{d-i})Pr[\delta(v) = i]$$

$$\leq \frac{1}{d-1} + \sum_{i=2}^{d/2} \frac{i}{(d-i)} \frac{c'}{ilnn} = \frac{1}{d-1} + \frac{c'}{lnn} \sum_{i=d/2}^{d-1} 1/i \leq \frac{c'+1}{lnn}$$

for some constant c'. Note that $Pr[\delta(v) = i] \leq \frac{c'}{ilnn}$ is due to lemma 12. Second, using another simple fact that $lnx/x < 1, \forall x > 1$ (using $x = d/(d-i)$),

$$\sum_{i=d/2}^{d} f(i)Pr[\delta(v) = i] \leq \sum_{i=d/2}^{d} ln(\frac{d}{d-i}) \frac{d-i}{d} \frac{c''}{ilnn} \leq \frac{c''}{lnn} \sum_{i=d/2}^{d} 1/i \leq \frac{c''}{lnn}$$

for some constant c''. Note that $Pr[\delta(v) = i] \leq \frac{d-i}{d} \frac{c''}{ilnn}$ is from lemma 12 and $d/2 \leq i$.

Choose $c = c' + c'' + 1$ then the lemma is shown. □

Proof of theorem 11. We first present our main idea to prove this theorem. We think of doing a series of random trials, in each of them, given a node v, we do an experiment to

find $\kappa(v)$ (to be assigned to v for the next trial) and measure $X_v = \frac{d_v}{d(\kappa_t(v),t)}$, which reflects the ratio between the distance to t before and after this trial. Let $\{v_1 = s, v_2, \ldots, v_k\}$ be a chain of 'stations' made by a run of $\mathcal{A}$ after k such steps (thus $v_{i+1} = \kappa(v_i)$). $\prod_{i=1}^{k-1} X_{v_i}$ will reflect the ratio between the distance at the initial and at the current state.

However, in order to make this ratio always meaningful we need to consider a few exceptions. In the case $\kappa(v) = t$ we define $X_v = d_v$ and if v itself happens to be just t before the trial then we define $X_v = 1$. The ratio is, therefore, always well defined and it is at most equal to the initial distance $d_s = d(s,t)$.

It is clear that the product $\prod X_{v_i}$ is always between 1 and d_s. Observe also that $\prod X_{v_i} = d_s$ if and only if v_k is t or one of its adjacent neighbors. Then we are about to reach t when $\prod X_{v_i} = d_s$, or $\sum Z_{v_i} = ln(d_s)$ since $Z_v = lnX_v$. Thus, the main idea is to work on bounding the sum $\sum Z_{v_i}$ under some desired circumstance, where lemma 13 is clearly useful.

Consider $Z^k = Z_{v_1} + Z_{v_2} + \ldots + Z_{v_k}$. If we reach t with less than k steps, or say, $v_j = t$ for some $j < k$, then as mentioned before, we have $Z_{v_j} = Z_{v_{j+1}} = \ldots = Z_{v_{k-1}} = 0$. Basically, we need to prove that there exists a constant c_2 such that for $k = c_2 log^2 n$, $Pr[Z^k < ln(d_s)] \geq 0.5$.

From lemma 13, there exists a constant c such that $E[Z_{v_i}] < c/lnn$ if $d_{v_i} > lnn$. To get rid of this condition, $d_{v_i} > lnn$, let modify $\mathcal{A}$ a bit, that is whenever we reach to within a distance lnn from t (i.e. $d_v \leq lnn$), set $\kappa(v) = v$ and $Z_v = 0$ (this will not weaken our proof since when we get there we can simply 'walk' to t by at most $\lceil lnn \rceil$ local links). Thus we always have $E[Z_{v_i}] \leq c/lnn, \forall i = 1..k$. We now show that

$$Pr[Z_{v_k} < M] \geq 0.5 \tag{6}$$

where $k = \lfloor \frac{ln^2 n}{4c} \rfloor$ and $M = ln(\frac{d_s}{lnn}) \geq lnn + lnc_1 - ln(lnn)$.

However, from $E[Z_{v_i}] < c/lnn$, we have $E[Z^k] \leq kc/lnn \leq lnn/4 < M/2$ for n large enough. Thus, $Pr[Z^k \geq M] < 0.5$ since otherwise $E[Z^k] \geq 0.5M = M/2$. Hence, $Pr[Z^k < M] \geq 0.5$. From (6), choose $c_2 = \lfloor \frac{1}{4c} \rfloor$ then by at least half of the time, $\mathcal{A}$ can not finish with less than $c_2 log^2 n$ steps; the theorem follows. □

3.2 Algorithms for improved delivery time

We now consider variants of Kleinberg's greedy algorithm which use additional knowledge of the graph to improve expected path length. Our basic algorithm operates on graphs from class $\mathcal{K}^*(n,1,1)$ but it can be easily extended to more general classes. So, for a node u define $\chi(u) = v$ if the only long-range random link from u goes to v. We assume that each node u knows the long range links of the $\lceil logn \rceil$ neighbor nodes closest to u in the grid. Call W_u, the set of these neighbors and their long range contacts, the view at u. During a basic step of routing, if u is the current message router, just as in Kleinberg's greedy algorithm, we go next to v, the node in W_u closest to t (we may need to follow several local links and, possibly one final random link to v and will start with v for the next step).

Note that in Kleinberg's algorithm W_u includes only the nodes incident to links from u (at most 5 in a two-dimensional grid) and a routing decision is made each node on the path. In our algorithm, once we make a decision for a step (at each current message router) we just follow the sub-path to the next router node. When we are at an intermediate node (e.g. a neighbor of a router node) we follow local links (up, down, left or right) except for the final link which may be a random long-range link). We can describe this sub-path using only $O(loglogn)$ extra bits to specify the last node on the sub-path reached by local links only. Alternately, we can use $2log^{1/2} n$ bits to describe the sequence of local moves (two bits per local move).

We now claim that *the expected number of links used in this algorithm is* $O(log^{3/2}n)$. We show this using the phases introduced by Kleinberg in [12]. Suppose that we need to find a path from s to $t \neq s$. Let $\mathcal{B}_i$ denote the ball $B_{2^i}(t)$. Our algorithm, called $\mathcal{A}$, contains several phases. Similarly as in [12], we say $\mathcal{A}$ is in phase i if the current node u is in $\mathcal{B}_{i+1}$ but $u \notin \mathcal{B}_i$. If k is a number such that $s \in \mathcal{B}_{k+1}$ but $s \notin \mathcal{B}_k$ then $\mathcal{A}$ will start in phase k and end up in phase 0.

Kleinberg has shown that [12] random variable X, the number of nodes whose random links are explored during a phase in his algorithm (approximate the number of links used in a phase), has expectation in $O(logn)$. The intuition for that is to think of an comparative experiment, where given two arbitrary distinct nodes u and v, generate a fresh random link from u and check if it goes to within halfway from v. It is not hard to see that the expected number of times of repeating this experiment until having a success (compare it with X) is $\theta(logn)$ (partly similar to work in section 2). Using a similar approach, we can easily see that in our algorithm $\mathcal{A}$, X is also $O(logn)$, although here X must be a multiple of $\lceil logn \rceil$.

On other hand, it is not hard to see that Y, the number of links used during a phase in $\mathcal{A}$, is upper bounded by $cX/\sqrt{logn}$ (X of the same phase) for some constant c. Note that during a basic step of routing (find v for u), the (fresh) nodes with their random links explored occupy at least a constant fraction of a ball centered at u with radius $\approx \sqrt{logn}$; call the number of these nodes by x. Meanwhile the number y of links we follow from u to v is at most this radius plus 1; thus, the ratio x/y is at least $\hat{c}\sqrt{logn}$ for some constant $\hat{c}$.

Therefore, $E[Y] = O(\sqrt{logn})$, and so, with the number of phases is at most $logn$, our claim is justified.

Note that, the expected number of links used in $\mathcal{A}$ is probably improved a bit (but still asymptotically the same) if we look for v such that its predecessor (just before it in the path from u) is closest to u yet v is in the next inner ball (if some $v \in W_u$ is in the next inner ball).

More careful work (using tail inequalities) gives us a slightly stronger result: the longest path found by $\mathcal{A}$ (over all possible $s - t$) has length in $O(log^{3/2}n)$ with overwhelming probability. The underlying intuition is to think of doing a chain of such experiments mentioned above until we have about $logn$ successes.

As mentioned earlier, our algorithm can be easily extended to the classes $\mathcal{K}^*/\mathcal{K}(n,p,q)$ where $p, q \geq 1$. Moreover, it is relatively easy to show a slightly improved bound $O(\frac{log^{3/2}n}{p\sqrt{q}})$. So, the bound is just $O(logn)$ if for instance, $p = \Omega(log^{1/2}n)$ and q is a constant, or p is a constant and $q = \Omega(logn)$. It is also easy to extend our algorithm to $\mathcal{K}^*/\mathcal{K}(k,n,p,q)$. *For* $p = 1, q = 1$*, we obtain the upper bound* $O(log^{1+1/k}n)$, which is close to the ideal $O(logn)$ when k is large.

4 Tight bound for the diameter of $\mathcal{K}^*(n,1,4)$

In this section we will show that the expected diameter of a graph from $\mathcal{K}^*(n,1,4)$ is $O(logn)$ and we will deal with more general classes of graphs in the next section. We now consider a source s and a destination $t \neq s$, both chosen arbitrarily from V, and want to show that with high probability there is an $O(logn)$ length path from s to t. As mentioned in the introduction section, our basic idea is to 'grow two trees' with roots from s and t, and try to show that they will intersect with overwhelming probability after $O(logn)$ χ-steps of growing.

Now, to simulate an $s-$tree with a height μ, we construct a chain of disjoint subsets $\{S_k\}_{k=0}^{\mu}$ with $S_0 = B_r(s)$ where $r = r_0\sqrt{logn}$ for some constant r_0. Let $C_k = \bigcup_{i=0}^{k} S_i$ and define $S_{k+1} = \chi(S_k) - C_k$. Thus, S_{k+1} is built by iteratively applying χ on elements of S_k and taking only 'fresh' nodes, which have not been in any preceding subsets. The subset

chain has the following properties:

1. By definition, all the nodes in S_0 can be reached from s by $O(logn)$ local links.

2. We show that for a given constant α, there exists μ such that μ is in $O(logn)$ yet $|S_\mu| \geq = \alpha nlogn$ with overwhelming probability.

Later we will also construct the similar subset chain $\{T_k\}_{k=0}^{\nu}$ to simulate $t-$tree as well. The rest of this section will focus on proving the second property. The second property will help us to later establish that the last subsets (or last levels), S_μ and the counterpart T_ν, very likely intersect or will intersect if the $s-$tree is extended further by one level. Here, 'almost surely' or 'very likely' will be refined later and can be made as high as needed by choosing the constant r_0 sufficiently large. All these will allow us to show the $O(logn)$ bound on the expected diameter of a graph in $\mathcal{K}^*(n, 1, 4)$.

In order to obtain property 2, we will show that the subset chain will almost surely grow exponentially in cardinality, i.e. there is some constant $\gamma > 1$ such that $Pr[|S_{k+1}|/|S_k| > \gamma]$ is almost one for all $k \leq \mu$. We make the matter easier by the following crucial observation. The main concern is on how many fresh nodes we get from $\chi(S_k)$ to include in S_{k+1}. In fact, we can take any order in scanning the elements of S_k and for each $u \in S_k$, consider a node $v \in \chi(u)$ fresh if it has not been included in a subset G, which keeps track of the union of the current S_{k+1} and all the preceding (already complete) subsets S_i, $i \leq k$.

To analyze this we consider an experiment where we make a trial under the inverse second power distribution and get a node v matched with u and consider if v has not been included in G; let $X(u, G)$ denote the indicator random variable of this happening and $\mathcal{E}(u, G)$ denote this experiment. We do this experiment 4 times for each $u \in S_k$ and $|S_{k+1}|$ will be the sum of these $4|S_k|$ random variables $X(u, G)$; note that these variables are not identical since G keeps growing larger. Let $\mathcal{G}_k$ denote the whole process. *If we could move the elements in G around, how would we minimize $Pr[X(u, G) = 1]$?* The nature of the inverse second power distribution makes it clear: this can be done by moving elements of G closer to u; in fact, $Pr[X(u, G) = 1]$ is minimized (for a fixed size for G) *when G is a like-a-ball with u as its center.*

This observation leads to a way to lower bound $|S_{k+1}|/|S_k|$ by using a worse scenario which is easier to analyze. Let H be a ball with center s and define the random variable $f(m, H)$ as the number of times we get $X(s, H) = 1$ when doing the $\mathcal{E}(s, H)$ experiment $4m$ times independently. Call this experiment $\mathcal{F}(m, H)$. The above observation shows that if H is chosen such that $|H| \geq |C_{k+1}|$ (i.e. always $\geq |G|$) then $\mathcal{F}(|S_k|, H)$ will generally under-perform $\mathcal{G}_k$ in term of output quantity (more precisely, for any $a > 0$, $Pr[f(|S_k|, H) \geq a] \leq Pr[|S_{k+1}| \geq a]$). Thus, any probabilistic lower bound on $f(|S_k|, H)/|S_k|$ can also hold on $|S_{k+1}|/|S_k|$.

We show a lower bound with the ball experiment in the following fact and then as a result of this, we establish a growth rate, which is greater than one, in each step of growing S_{k+1} from S_k. Then by applying this result iteratively, we can show that the subset chain grows exponentially in size before bypassing the threshold $\alpha nlogn$.

Fact 14. *Let H be a ball with $|H| = O(nlogn)$. If $m < |H|$ and $m = \Omega(logn)$ then there exists a constant $\gamma > 1$ such that for n large enough almost surely $f(m, H)/m > \gamma$; more precisely:*

$$\exists \gamma > 1, \exists\, \eta > 0 : Pr[f(m, H)/m > \gamma] = 1 - O(n^{-\eta}) \tag{7}$$

Proof. Since $|H| = O(nlogn)$ then H's radius is smaller than $n^{0.5+\xi}$ for any constant $\xi > 0$, when n is large enough. Thus, by fact 2, $E[X(s, H)] > \beta$ for any fixed $\beta < 0.5$ when n large enough. Thus $f(m, H)$ is the sum of $4m$ independent Bernoulli random variables each with expectation greater than β. Thus we have $E[f(m, H)] \geq 4m\beta$ and applying

Chernoff's inequality, we have $Pr[f(m,H) \leq (4m\beta)(1-\delta)] \leq e^{-2m\beta\delta^2}$ where $0 < \delta < 1$. Since $m = \Omega(logn)$ then $e^{-2m\beta\delta^2} = O(n^{-\eta})$ for some $\eta = \eta(\delta) > 0$. This is due to a simple fact that $e^{xlogn} = n^x$ for any x. Thus, $Pr[f(m,H) \leq 4\beta(1-\delta)m] = O(n^{-\eta})$ or $Pr[f(m,H)/m > 4\beta(1-\delta)] = 1 - O(n^{-\eta})$. This assures that when n is large enough, $f(m,H)/m$ is almost surely greater than $\gamma = 4\beta(1-\delta)$. □

Indeed we can always choose $\delta > 0$ small enough so that $\gamma = 4\beta(1-\delta)$ can be arbitrarily close to 4β, and therefore, arbitrarily close to 2 (since $\beta = 0.5 - o(1)$). When we have chosen δ, *given any* $\eta > 0$, we can choose $\hat{c}$ such that when $m \geq \hat{c}logn$ we have $Pr[f(m,H)/m > \gamma] = 1 - O(n^{-\eta})$. The following fact will show how we can find such $\hat{c}$ when given an arbitrary η. See the appendix for its proof.

Fact 15. *Define function* $\tau(\beta,\eta,\gamma) = 8\eta\beta/(4\beta-\gamma)^2$ *and let* $\beta_s = 0.5$. *Let* $\eta > 0$, $\gamma \in (1,2)$ *and* $\hat{c} > 0$ *be constants. Let* H *be a ball with* $|H| = O(nlogn)$ *and* m *be an integer such that* $\hat{c}logn \leq m < |H|$. *If* $\hat{c} > \tau(\beta_s,\eta,\gamma)$ *then* $Pr[f(m,H)/m > \gamma] = 1 - O(n^{-\eta})$ *for* n *large enough.*

As noted earlier, a similar result can be obtained with $|S_{k+1}|/|S_k|$. Intuitively, for any value of $\gamma \in (1,2)$, we know what size for S_k is big enough in order to make the probability of $|S_{k+1}|/|S_k| > \gamma$ arbitrarily close to one.

Lemma 16. *Let* $\eta > 0$, $\gamma \in (1,2)$, $\hat{c} > \tau(\beta_s,\eta,\gamma)$ *be constants and integer* $k \geq 0$. *Suppose that* $|S_k| \geq \hat{c}logn$ *and* $|C_{k+1}| = O(nlogn)$ *then* $Pr[|S_{k+1}|/|S_k| > \gamma] = 1 - O(n^{-\eta})$ *for* n *large enough.*

Given fixed constants η, γ and $\hat{c}$ as above, and given $|S_0| \geq \hat{c}logn$, we can expect the cardinality series $\{|S_i|\}$ to grow exponentially before reaching or exceeding the threshold $\alpha nlogn$ during the first $k+1$ terms, where $k = \lceil log_\gamma(\alpha n/\hat{c})\rceil$, which is $\Omega(logn)$. For simplicity, we define $S_{i+1} = S_i$ when $|S_i| \geq \alpha nlogn$, then the next lemma claims that almost surely $|S_k| \geq \alpha nlogn$. To show this lemma, the main idea is to iterate lemma 16 for each step of growing S_{i+1} from S_i (when $|S_i| < \alpha nlogn$). Thus the major task is to assure that we have the necessary condition to use it in each step.

Lemma 17. *Let* $\theta > 0$, $\gamma \in (1,2)$, $\hat{c} > \tau(\beta_s,\theta,\gamma)$ *be constants. Let* $k = \lceil log_\gamma(\alpha n/\hat{c})\rceil$. *If* $|S_0| \geq \hat{c}logn$ *then* $|S_k| > \alpha nlogn$ *with probability* $1 - O(n^{-\theta})$.

Proof. First, note that $|S_k| > \alpha nlogn$ if $|S_k|/|S_0| > \gamma^k$, which is true if we have $|S_{i+1}|/|S_i| > \gamma$ for all $i = 0..k-1$. Define event A_i as "$|S_{i+1}|/|S_i| > \gamma$ or $|S_i| \geq \alpha nlogn$" for integer $i \geq 0$, and define $B_j = A_0\, and\, A_1\, and \ldots and\, A_j$ for integer $j \geq 0$. Clearly, $|S_k| > \alpha nlogn$ if B_{k-1} occurs, the probability of which equals $Pr[A_0] \times Pr[A1|B_0] \times Pr[A_2|B_1] \times \ldots \times Pr[A_{k-1}|B_{k-2}]$.

We now consider $A_j|B_{j-1}$, i.e. the event A_j when B_{j-1} occurs. Note that if $|S_{j-1}| \geq \alpha nlogn$ then immediately $|S_j| = |S_{j-1}| \geq \alpha nlogn$, so A_j occurs. Otherwise, $B_{j-1}\, and\, |S_{j-1}| < \alpha nlogn$ means $\{|S_0|, |S_1| \ldots, |S_j|\}$ growing exponentially and clearly, $|S_j| \geq |S_0| \geq \hat{c}logn$ and $C_{j+1} = O(nlogn)$. For the later, note that when a series is growing exponentially, the current sum value will always be within a constant multiplier of the last term. Hence, from lemma 16, $Pr[A_j|B_{j-1}] \geq Pr[A_j|B_{j-1}\, and\, |S_{j-1}| < \alpha nlogn] = 1 - O(n^{-\eta})$, where η is chosen slightly greater than θ so that $\hat{c} > \tau(\beta_s,\eta,\gamma)$. This can be made because of τ's continuity.

Combining the probabilities of all these events $A_j|B_{j-1}$, we then have $Pr[|S_k| \geq \alpha nlogn] \geq \left(1 - c_1 n^{-\eta}\right)^k \geq 1 - kc_1 n^{-\eta}$ for some constant $c_1 > 0$. Note that we have used a basic calculus fact that $(1+x)^n \geq 1 + nx$ for any $x > -1$ and $n \geq 1$. Since $\theta < \eta$ and clearly, $k = O(logn)$, it is easy to see that $kc_1 n^{-\eta} = O(n^{-\theta})$ and hence, $Pr[|S_k|/|S_0| > \gamma^k] \geq 1 - O(n^{-\theta})$. □

Therefore, if we choose $\mu \leq k = \Omega(logn)$ such that S_μ is the first term to exceed $\alpha nlogn$, we then have property 2 satisfied. This can be summed up in the following fact.

Fact 18. *For any given node $s \in V$, any $\theta > 0$ and any $\alpha > 0$, by constructing s's subset chain as above, we will obtain subset S_μ where any node in S_μ can be reached from s by an $O(logn)$ path and $Pr[|S_\mu| \geq \alpha nlogn] = 1 - O(n^{-\theta})$.*

Note that S_μ contains only 'fresh' nodes such that we have not yet considered their random long-range contacts.

We now discuss a similar result with respect to a subset of nodes with $O(logn)$ length paths to a given destination t. Consider a tree of nodes with paths to t. Instead of χ we now use a function $\hat{\chi}$ which, given an input node u, outputs the nodes with a random link *to* u. Similarly as before, we construct a subset chain $\{T_k\}_0^\nu$ by having $T_{k+1} = \hat{\chi}(T_k) - \hat{C}_k$ where $\hat{C}_k = \bigcup_{i=0}^{k} T_i$. Thus, we include into T_{k+1} all the 'fresh' nodes which have a random link to any node in T_k. We wish to show properties similar to the two above with s's subset chain.

It turns out that we can use much the same approach as before with some modifications and additional details. We still use a state-variable G to denote the set of all nodes in the tree we have reached so far. Note that we have $G = \hat{C}_k$ if we have just finished the first k subsets, otherwise G is the sum of $\hat{C}_k$ and the developing T_{k+1}. Let $\overline{G} = V - G$, the set of nodes not in the tree yet.

We now look closer at process $\hat{\mathcal{G}}_k$, the growing step of T_{k+1} from T_k, which is more complicated than the counterpart in s's subset chain. A long-range random link from a node u to a node v is considered having label i if $\chi_i(u) = v$. Let $\hat{\mathcal{E}}(u, i, G)$ denote an experiment which has each node $w \in \overline{G}$ look at its random link labelled i, and if this link hits u we add w to G (for simplicity we may also use $\hat{\mathcal{E}}_u$ instead). Process $\hat{\mathcal{G}}_k$ simply repeats the $\hat{\mathcal{E}}(u, i, G)$ experiment for each node $u \in T_k$ and for $i = 1, 2, 3, 4$. Note that, G is changing as we discover new fresh nodes. Also the order we scan T_k is not important and we can initially fix a unique order.

As with the case of the $s-$tree, we now consider an artificial experiment which should be 'outperformed' by the real $\hat{\mathcal{G}}_k$ but is easier to analyze. Let H be a ball centered at t with size at least $|\hat{C}_{k+1}|$. The basic operation is to have all nodes not in H each generate a random link (under the inverse square power distribution) and consider if t is hit. Let $\hat{f}(m, H)$ be a random variable recording the number of times that t is hit when we repeat the above operation for $4m$ times.

We now compare the real and artificial processes above. Looking closely at each real experiment $\hat{\mathcal{E}}_u$, observe that it is slightly preconditioned by the results of the previous ones. A random link from a node $w \in \overline{G}$ must not go to a node in $\hat{C}_k$, because if w 'hits' a node $v \in \hat{C}_k$, then w must have been included in G earlier during $\hat{\mathcal{E}}_v$. With $\hat{C}_k \subset G$, in fact, this precondition makes this random link more likely to hit u. Thus in the artificial process, each link is less likely to hit a target than in the real process, noting also that, as before, a ball is the worst shape for G, and we start with $|H|$ at least the maximum size for G. Thus, any probabilistic lower bound on $\hat{f}(|T_k|, H)/|T_k|$ can also hold for $|T_{k+1}|/|T_k|$.

It is easy to see that $\hat{f}(m, H)$ is lower bounded by the sum of $4m$ independent Bernoulli random variables, and the expectation of each can be lower bounded by a constant $\beta > 0$, using the second proposition of fact 2. Thus as in fact 2, using Chernoff's inequality, we have the following lemma for the new ball experiment.

Fact 19. *Let H be a ball with $|H| = O(nlogn)$. If $m < |H|$ and $m = \Omega(logn)$ then there exists a constant $\hat{\gamma} > 1$ such that for n large enough, almost surely $\hat{f}(m, H)/m > \hat{\gamma}$:*

$$\exists \hat{\gamma} > 1, \exists\, \eta > 0 : Pr[\hat{f}(m, H)/m > \hat{\gamma}] = 1 - O(n^{-\eta}) \tag{8}$$

Hence, we can obtain analogues of the fact 15 and lemma 16 for t's subset chain. There are a few differences. Since we use the second proposition in fact 2 instead, the range of β is now $(0, \beta_t)$, where $\beta_t = 1 - e^{-0.5} \gtrapprox 0.39$; then, the range of growing character $\hat{\gamma}$ is $(1, 4\beta_t)$ instead. Thus, we need $\hat{c} > \tau(\hat{\gamma}, \beta_t, \eta)$ in the new versions of results 15 and 16 instead.

Finally, we can obtain an analogue of lemma 17 for t's subset chain but we need to condition that $\hat{c} > \tau(\hat{\gamma}, \beta_t, \theta)$ instead. We can prove this result the same way as in the proof of lemma 17 by iteratively applying the new version of fact 14 to claim a growth rate at least a constant $\hat{\gamma} > 1$ in each step of growing T_{k+1} from T_k, then combining all these growing steps. Note that this combination does not require the independence between these steps. Thus, similarly as before we have the following fact.

Fact 20. *For any node $t \in V$, any $\theta > 0$ and any $\alpha > 0$, by constructing $\{T_i\}_{i=0}^{\nu}$, we will obtain a subset T_ν such that t can be reached from each node in T_ν by a path of length $O(logn)$ and $Pr[|T_\nu| > \alpha nlogn] = 1 - O(n^{-\theta})$.*

Now we are going to show that for any two distinct nodes s and t, there is a path of length $O(logn)$ from s to t with overwhelming probability. Using facts 18 and 20, we will show that there exists with overwhelming probability two such subsets S_μ and T_ν with mentioned properties and either they intersect or S_μ is one-link separate from T_ν, i.e. there exists $u \in S_\mu$ such that $\chi(u) \cap T_\nu \neq \emptyset$. This means, with overwhelming probability there exists an $O(logn)$ path from s to t. Moreover, by choosing appropriate constants θ, γ and α, we can make this probability arbitrarily close to one and as a result of that, we show that the expected diameter is $\Omega(logn)$.

Choose $\theta = 6$ in fact 18 then, given any constant $\alpha > 0$, we can find a subset S_μ with at least $\alpha nlogn$ nodes, *all of which can be reached from s by a path of length $O(logn)$*, with probability $1 - O(n^{-6})$. Also, using fact 20, we can find a subset T_ν with at least $\alpha nlogn$ nodes, *all of which can reach to t by a path of length $O(logn)$*, with probability $1 - O(n^{-6})$ again. Let E_1 denote the event of having both such subsets S_μ and T_ν. We can think of constructing subset S_μ first and then T_ν. Note that it is not straightforward to say that the $Pr[E_1] = (1 - O(n^{-6}))^2$, since when we are constructing T_ν, we have already been conditioned by the existence of s's subset chain. In fact, the state-variable subset G, which includes all the 'non-fresh' nodes, will now be larger, containing all the nodes in the completed $s-$tree and the developing $t-$tree. However, since the number of nodes in $s-$tree $|C_\mu| = O(nlogn)$, we still have $G = O(nlogn)$, which means all the results for t's subset chain still apply. Thus $Pr[E_1] = (1 - O(n^{-6}))^2$, which is also $1 - O(n^{-6})$ due to the common fact that $(1 + x)^a \geq 1 + ax$ when $x > -1, a \geq 1$.

Assuming the occurrence of E_1, consider the event E_2 that either S_μ and T_ν intersect, called event E_{2a}, or S_μ is one-link separate from T_ν, called event E_{2b}. Note that for S_μ we have not yet looked at its out-going random links and for T_ν we have not yet considered its incoming arcs. Observe that the probability of event E_2, the sum of E_{2a} and E_{2b}, is at least the probability of event E_{2b} given E_{2a} does not occur; so, we lower bound the probability of E_2 given E_1 using the probability of E_{2b} given E_1 and not E_{2a}. Given an arbitrary node $u \in S_\mu$, let ξ denote the probability that u 'misses' T_ν, i.e. none of the 4 random links from u goes to any node in T_ν. From fact 1, the probability of a random link from u to any other node is at least $\epsilon = C_u(2n)^{-2} = \Omega(n^{-2}log^{-1}n)$, therefore $\xi \leq (1-\epsilon|T_\nu|)^4$, but $|T_\nu| \geq \alpha nlogn$, so there exists a constant c_1 such that $\xi \leq (1 - \alpha c_1 n^{-1})^4 \leq e^{-4\alpha c_1 n^{-1}}$. Now the probability that all the nodes in S_μ 'miss' T_ν is at most $(e^{-4\alpha c_1 n^{-1}})^{\alpha nlogn} \leq e^{-4\alpha^2 c_1 logn} = n^{-4\alpha^2 c}$ where $c = c_1 loge$. Thus, $Pr[E_2|E_1] \geq 1 - n^{-4\alpha^2 c}$ and when we choose α large enough:

$$Pr[E_2|E_1] \geq 1 - O(n^{-6}) \tag{9}$$

Thus (by choosing α large enough), the probability of having the shortest path from s to t with length $O(logn)$ is at least $Pr[E_2|E_1]Pr[E_1] \geq (1 - O(n^{-6}))^2 = 1 - O(n^{-6})$. In other

words, there exist constants c and $\hat{c}$ such that for any two nodes s and t the probability, that the shortest path from s to t is greater than $\hat{c}logn$, is at most cn^{-6} when n is large enough. Note that from the way we have formed the subset chains and hence, a short path from s to t, it is easy to see that with overwhelming probability this path has length at most $\phi(n) = \lceil log_\gamma(\alpha n/c_1)\rceil + \lceil\sqrt{c_1 logn/2+1}\rceil + \lceil log_{\hat{\gamma}}(\alpha n/c_2)\rceil + \lceil\sqrt{c_2 logn/2+1}\rceil + 1$, where $\gamma = 2 - o(1)$,$\hat{\gamma} = 4\beta_t - o(1)$,$c_1 = \tau(\gamma, \beta_s, 6)$ and $c_2 = \tau(\hat{\gamma}, \beta_t, 6)$. When n is large enough, clearly $\phi(n) = log_\gamma n + log_{\hat{\gamma}} n + o(1)logn$. We can choose γ arbitrarily close to 2 $(= 4\beta_s)$ and $\hat{\gamma}$ arbitrarily close to $4\beta_t$ then $\phi(n)/logn = log_\gamma 2 + log_{\hat{\gamma}} 2 + o(1)$. This means we can choose $\hat{c} > 1 + log_\lambda 2$ but arbitrarily close to $1 + log_\lambda 2$ ($\lessapprox 2.53$), where $\lambda = 4\beta_t = 4(1-e^{-0.5}) \approx 1.574$. Thus we can choose $\hat{c} = 3$.

The diameter is greater than $\hat{c}logn$ if there is any pair (s,t) such that the $s-t$ shortest path is greater than $\hat{c}logn$. The number of different pairs (s,t) on the KSM grid is less than $(n^2)^2 = n^4$, therefore $Pr[diameter \geq \hat{c}logn]$ is at most $n^4 \times (cn^{-6}) = cn^{-2}$. That is almost all graphs in $\mathcal{K}^*(n,1,4)$ have diameter less than $\hat{c}logn$. Since the diameter of a graph in $\mathcal{K}^*(n,1,4)$ can not exceed $2n$, its expectation is at most $\hat{c}logn \times (1 - cn^{-2}) + 2n \times (cn^{-2}) = \hat{c}logn - \hat{c}c\frac{logn}{n^2} + \frac{2c}{n} = O(logn)$. *Thus, the expected diameter of the graphs in* $\mathcal{K}^*(n,1,4)$ *is* $\theta(logn)$. (As mentioned earlier, the bound $\Omega(logn)$ is obvious)

Because of computation simplicity, we have chosen to start with this class with $q = 4$, however, it is easy to see that the same result still holds when q takes any value not smaller than 3. When we grow S_{k+1} from S_k, the probability that a certain random link from a given node $u \in S_k$ contributes a new 'fresh' node is at least $\beta_s - o(1)$ (arbitrarily close to β_s when n is large enough); therefore the q random links from u contribute expected at least $q\beta_s - o(1)$ fresh nodes. Clearly, $q\beta_s > 1$ is the sufficient condition to have s's subset chain to grow exponentially in size (before passing the turning threshold). Similarly, $q\beta_t > 1$ is sufficient to have t's subset chain to grow exponentially in size, and both of these happen if $q > min\{1/\beta_s, 1/\beta_t\}$, or $q \geq 3$. Thus, $q \geq 3$ is enough to claim that the expected diameter of graphs in $\mathcal{K}^*(n,1,q)$ is $\theta(logn)$. However, as we will show in the next session, the same result still holds for any $q \geq 1$ indeed.

5 Further extension

Based on the approach we have used above to analyze the expected diameter of graphs in $\mathcal{K}^*(n,1,4)$, we now study other and more general classes which also have expected diameter $\theta(logn)$. We focus only on the main ideas and often skip unimportant details.

The diameter of $\mathcal{K}^*(n,1,1)$. We now show that indeed we just need each node to have at least one out-going random link, i.e. we show that the diameter of graphs in $\mathcal{K}^*(n,1,1)$ is also $O(logn)$. It seems that we can transform this problem to $\mathcal{K}^*(n,1,4)$ above by using a simple trick. That is to 'shrink' the $n \times n$ grid by a factor of two in each dimension in such a way that groups of four adjacent nodes, vertices of a 2×2 square cell, collapse into single 'big nodes'. It is easy to see that we can obtain a new graph (based on a grid with size roughly $n/2$ in each dimension) with each node having four out-going random links. However, it is easy to see that the distribution of the random links after this transformation is not exactly the inverse second power distribution, but only an approximation. Thus, this idea is a simple way to see that the diameter of $\mathcal{K}^*(n,1,1)$ is also $O(logn)$, yet not make a complete proof.

The key idea of doing that is to modify the subset chain construction so that $\chi(u)$ still has multiple outputs (at least 3 as shown above), despite the fact that there is only one random link going out from each node.

We now redefine function χ such that it still has four outputs, that is for each $u \in V$, $\chi(u)$ consists of *nodes which can be reached from a neighbor of* u *by a random link.* We also

define $\chi_i(u)$, for $i = 1, 2, 3$ or 4, as the node that the random link of the left, right, above or below neighbor of u goes to, accordingly. Thus $\chi(u)$ still has the potential to contribute more than one expected fresh node during a χ-step. As can be seen easily, we 'trim off' a lot of s's 'descendants' to give the $s-$tree a shape which is easier to estimate. In a sense, one level in this new tree is formed by merging and trimming two consecutive levels in the version with full descendants.

However, this redefinition also requires a bit of extra care for us. That is, when any two nodes u and v are at distance 2 apart, they have direct neighbors in common and hence, $\chi(u)$ and $\chi(v)$ are not independent. However, it is easy to see that $\chi(u)$ and $\chi(v)$ are independent (the formation of one does absolutely not affect that of the other) if $d(u, v) \geq 3$. Thus we define a node $w = \chi_i(u)$ as a fresh node if and only if w is at least distance 3 from any node in G (which keeps track of all nodes already in the tree). Thus, in order to see if v can be added to the tree, we need to check if $v \in B_2(G)$, where $B_2(S) = \cup_{u \in S} B_2(u)$ for any subset S. It is obvious to note that G now consists of nodes which are at least distance 3 from each other.

Note that we now have more than one way to grow S_{k+1} from S_k, since we can scan the nodes in S_k in different orders, which will affect the list of fresh nodes obtained. But we can initially sequence all the nodes in V (say, by using their coordinates) to shape the growing process uniquely.

However, this modification of s's subset chain construction will not affect the fact that when the subset chain has not grown beyond the size $\theta(nlogn)$, the subset $B_2(G)$ (which a node 'needs to avoid to hit' to be seen fresh) still has size $O(nlogn)$. It is clear because $|B_2(G)|$ and $|G|$ are not different by more than a constant factor. Therefore, results 14-17 still apply here.

We can also argue similarly with t's subset chain (with $\hat{\chi}(u)$ is defined as the nodes having a random link to one of u's four direct neighbors); thus, we can continue to analyze similarly as in section 4 to show that the expected diameter of graphs in $\mathcal{K}^*(n, 1, 1)$ is also $\theta(logn)$.

Grouping adjacent nodes to form 'super nodes' (with enough out-going random links) is the common idea in both the 'shrinking approach' before and this approach but is deployed more dynamically, hence more successfully, in the latter. More importantly, we can generalize this idea to find more general classes, even wider than $\mathcal{K}^*/\mathcal{K}(n, p, q)$, which also have expected diameter $\theta(logn)$. For constructing the subset chains, when we need to make a node u a super node, we can 'collect' the random links of many of u's neighbors (the number of which must be lower bounded by a constant C which we choose arbitrarily), and make a virtual re-assignment of these links to u, i.e. $\chi(u)$ contains the nodes incident to these links. For instance with $\mathcal{K}^*(n, 1, 1)$, as suggested at the end of section 4, we need to collect at least 3 such random links to form a super node, so that it allows the subset chains to grow exponentially in size.

On a general approach. As suggested above, the work presented in section 4 can be used as a framework to approach the diameter problem on many other similar lattice-based settings. We only need to maintain the principle of always having enough out-going random links from any small neighborhood with a determined size (but arbitrarily chosen) in order to form a super node. Thus, in fact, we can relax many conditions in Kleinberg's original model yet we can still analyze the new graphs using our approach. Thus, we can include more practical graphs: distribution of links are less uniform, some nodes may have many long-range contacts some may have none, and local links may be 'broken' or missing.

In such a new setting, the above principle of forming super nodes will help to create the $s-$ and $t-$ trees with exponential growth before their intersecting, which results in $\theta(logn)$ length paths from s to t. Note that fact 1 (clearly, a direct result of using the inverse second

power distribution) will again result in equation (9), i.e. the two trees will almost surely intersect.

The issue of missing local links may affect the view of the graph as a grid with added random links and then affect the definition of distance $d(u, v)$. However we still assume that the locations of all the vertices form a geographic grid though each vertex may or may not have local links to its neighbors; thus we still be able to use lattice distance. Wider generalizations are possible but are not discussed here.

We leave a thorough study of a more general approach to our future work but we suggest here important conditions of a lattice-based setting, wherein our current approach could be applicable. To adapt this approach to such a new setting, our main idea is to suitably formulate underlying properties, analogues of facts 2 and others, with respect to the new scenario. Thus we will be able to simulate the work in section 4 in this new scenario.

1. The setting is rich enough in local links so that it is easy to construct the starting subset S_0 (with size $\Omega(logn)$) using short paths. An alternative is the graph's being strongly connected (for any two distinct nodes u and v there is a path from u to v and vice versa).

2. There exists a mechanism of producing super nodes (wherever we need) so that we can get enough long-range random links per super node. We suggest the following proposition but other similar ones can be used instead:

   Sufficiency of random links everywhere: for any constant $C > 1$, there exists a constant L such that, for n large enough, for any node u, there exist, with overwhelming probability, a set of nodes with at least C out-going random links and reachable from u by no more than L local links.

   In fact we just look at nodes in ball $B_L(u)$ to collect that many random links. Let $N_L(u)$ denote the number of random links going out from nodes of ball $B_L(u)$; basically, for a given $C > 1$, we need to find L such that we almost always have $N_L(u) \geq C$ for any node u. Alternately, we have $N_L(u) \geq C$ almost everywhere; thus we call this "sufficiency of random links everywhere".

   The proposition above is only for constructing the tree from the source node; a similar condition is also needed with respect to the destination node, and for simplicity, constants C and L can be chosen to be the same in both cases. These conditions also reflect a sufficient ratio of random links per node.

3. An analogue to fact 2: there exist positive constants ξ_1 and ξ_2 such that for any positive $\theta < 0.6$, for n large enough, for any node u with an out-going random link (if more then pick one arbitrarily), the probability that this random link goes to a node outside of $B_l(u)$, where $l = n^\theta$, is greater than ξ_1. Also, the probability that there is a random link to u from a node outside of $B_l(u)$ is greater than ξ_2.

These conditions can be stated either as facts to be proved in a new setting or as assumptions. In order to have the two trees grow exponentially before they connect (after $O(logn)$ steps of growing), we need to find proper C (and hence, L) such that $C\xi_1 > 1$ and $C\xi_2 > 1$ when $\theta > 0.5$. We will illustrate these with some notable classes of graphs below.

The diameter of $\mathcal{K}(n, p, q)$ where $p, q \geq 1$. It is easy to see that condition 1 and 2 are automatically met since the local links are 'full' and condition 3 is met because of fact 3. Similarly as in section 4 we use fact 3 with $\theta = 0.5 + o(1)$ and hence obtain $\xi_1 = 0.2 + o(1)$ and $\xi_2 = 1 - e^{0.125+0(1)}$ (roughly, $0.117 + o(1)$). Thus, we need to specify constant C (as in the above proposition) so that $C\xi_1 > 1$ and $C\xi_2 > 1$, which help to establish the exponential

growth of the two trees. Therefore any $C \geq 9$ will work and if we choose $L = 2$ we can take $C = 9..13$ (since $|B_2(.)| = 1 + 4 + 8 = 13$). Thus, when we want to make a node u a super node we need to 'collect' the random long-range contacts of u's neighbors within distance 2 and assign them to $\chi(u)$ (similarly for $\hat{\chi}(u)$). If u is too close to the edges of the grid, we can simply drop it as the fraction of such u is negligible (note that we do not need to care about this case with $u = s$ or t, since this is only relevant to that we need a big enough initial S_0 or T_0, which is always trivial in this setting). Then we can continue similarly as before to show that the diameter of $\mathcal{K}(n, p, q)$, for $p, q \geq 1$, is $\theta(logn)$.

Higher dimensions: the diameter of $\mathcal{K}(k, n, p, q)$ where $k, p, q \geq 1$. All these above results leave us with a simple task. It is again clear that conditions 1 and 2 are met. Fact 6 meets condition 3, albeit we can not give exact values for the constant parameters as above. The model's connectedness and sufficiency of random links (everywhere) makes it as easy as above to find C and L such that $C\xi_1 > 1$ and $C\xi_2 > 1$ for any given positive constants ξ_1 and ξ_2. Our approach is then applicable as before which results in the following theorem, solving the general diameter problem in Kleinberg's model.

Theorem 21. *For any $k, p, q \geq 1$, the expected diameter of the graphs from $\mathcal{K}(k, n, p, q)$ is $\theta(logn)$.*

The case of lacking local links. We now study the classes $\mathcal{K}(n, 0, q)$ and $\mathcal{K}^*(n, 0, q)$ with $q > 0$, a special case when we do not assume the existence of local links. Note that this does not rule out the possibility of two adjacent nodes being connected by a random link. We now show that, if q is a constant then no matter how large q is, a graph in $\mathcal{K}^*(n, 0, q)$ is not strongly connected with probability tending to one when n goes to infinity.

Let l be a constant integer such that $l^2 > q+1$. Consider a group of l^2 nodes forming an $l \times l$ square in the grid. We show that the probability that the group forms a 'semi-island', i.e. there is no link from any element to the rest of V, is $\Omega(log^{-ql^2} n)$. Clearly, in this case the graph is not connected. First, the probability that a particular random link from a group member u goes to another group member v is $\frac{1}{c_u d^2(u,v)} \geq \frac{1}{c_u (2l)^2}$, which is $\theta(log^{-1} n)$ since $c_u = \theta(logn)$ from fact 1. Therefore the probability that the q random links of u go to q specified nodes in the group is $p_1 \geq (\frac{1}{4c_u l^2})^q$, which is obviously $\Omega(log^{-q} n)$. Thus, the probability that this group of l^2 nodes is a 'semi-island' is at least $p_2 = p_1^{l^2}$, which is $\Omega(log^{-ql^2} n)$, and not a 'semi-island' is at most $1 - p_2 = 1 - \Omega(log^{-ql^2} n)$.

If we divide the $n \times n$ grid into separate such groups we then have at least $a = \lfloor n/l \rfloor^2$ groups. Also, the events that each such group is a 'semi-island' are independent, therefore the probability that there exists at least one such 'semi-island' is at least $1 - (1 - p_2)^a \geq 1 - e^{-ap_2}$, which tends to 1 when n goes to infinity since clearly $ap_2 \geq \frac{c_1 n^2}{log^{-ql^2} n}$ for some constant $c_1 > 0$, which tends to infinity when n goes to infinity (note that l is a constant). Thus the graph tends to be not connected when n goes to infinity. Note that it is easy to extend the above argument to show the same fact but assuming that the q random links from a node are chosen randomly but without replacement (i.e. no two links to the same node).

Moreover, we can extend the above argument to show the existence of an 'island', i.e. no link in any direction between this group of nodes and the outside, with probability tending to 1 when n goes to infinity. Thus the graph is even not semi-connected if we consider all the directed links as indirect links. We now discuss this briefly. We observe that for a group of l^2 nodes mentioned above, the probability that the group has no in-coming link (from a node outside) is lower bounded by a constant greater than 0. We omit the proof of this fact but comment that the idea underlying is similar to that in the proof of the second proposition in fact 2 but we use the common fact $1 - x \geq e^{-2x}$ (where $0 < x < 0.5$) instead of $1 - x < e^{-x}$ (to give a lower bound instead of an upper bound). Therefore, the

probability that the group is an 'island' is at least $c_2 p_2$ (for some constant $c_2 > 0$) which is still $\Omega(log^{-ql^2} n)$. We then continue the same way as before to show that there exists an island with probability tending to 1 when n goes to infinity.

Similarly, the same results can be obtained with the classes $\mathcal{K}(k, n, 0, q)$ and $\mathcal{K}^*(k, n, 0, q)$ where $k, q \geq 1$.

6 Open problems

There are other possibilities to extend our diameter result. Condition *sufficiency of random links everywhere* requires us to have $N_L(u) \geq C$ almost everywhere. However it is also sensible to consider settings where a non-negligible fraction of nodes have low N_L ($< C$), yet we also have a significant volume of nodes with expected N_L high enough to compensate for this. Thus, we again suggest to further reduce the uniformness of our existing models, such that both 'rural places' and 'crowded cites' can exist at the same time as in practice.

So far the model is still heavily based on a grid structure while in practice (for example, over the Internet) nodes would never locate in such an ideal way (but often with unequal distances between neighbors and different number of neighbors per node). This and the issue of 'missing local links' make it hard to keep using the concept of lattice distance. This suggests considering other kinds of distance measures (but still be in accordance with the distribution rule for random links). It would be interesting to try group-induced models like the ones in [10]. Note that with such a new model not based on a grid structure (and possibly based on a different kind of distance measure), we also need a successful conversion of condition 3 above as well.

Such further generalizations may also need to cope with the issue of 'lacking local links' (in order to construct a sufficiently large initial subset S_0). Kleinberg's hierarchy models [10] do not have local links, but only random links. To compensate for no local links, it would be worth considering models with another kind of random links, more frequently appearing but rather 'short'. For example, we can introduce at least one more out-going link per node using an r-th inverse power distribution, where $r > 2$. This would also reflect social relations: business people may have world-wide, national and regional contacts. It is also worth noting that we still do not know any good upper bound for the expected diameter of a customized Kleinberg's grid model with using an r-th inverse power distribution of random links where $r > 2$ (our diameter result is for the case of $r = 2$ but can be easily extended for $0 \leq r \leq 2$ also).

Modelling SWNs with both 'social' super nodes ('crowded towns') and 'unsocial' super nodes ('rural places') would also lead to generalizations of Kleinberg's algorithm for decentralized routing mentioned in section 3. If the current random link leads us to an unsocial super node we may consider to use backtracking to find another better branch which leads to a social super node instead. Thus, a routing decision would be a function of two variables: the distance from the considered node to the destination and the prospect of reaching from this node's neighborhood to the world outside.

References

[1] R. Albert, H. Jeong, A.-L. Barabasi, "The diameter of the World Wide Web" Nature 401, 130 (1999).

[2] B. Bollobas, "Random Graphs", Proc. of Symposia in Applied Math. 44,1 (1991).

[3] B. Bollobas, "The Diameter of Random Graphs", IEEE Trans. Inform. Theory 36 (1990), no. 2, 285-288.

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

[5] B. Bollobas, *Random Graphs* (Academic Press, London, 1985)

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

[7] C. Faloutsos, P. Faloutsos, and M. Faloutsos, "On Power-Law Relationships of the Internet Topology", in Proc. ACM SIGCOMM, Sept. 1999.

[8] M. Franklin and C. Martel, "Small-World Networks and Secure Distributed Protocols", Tech. Report UC-Davis CS, 2002.

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

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

[11] J. Kleinberg, "Navigation in a Small World", Nature 406(2000).

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

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

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

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

[16] H. Zhang, A. Goel and R. Gividan, "Using the Small-World Model to Improve Freenet Performance" Proc. IEEE Infocom, 2002.

A Proof of supporting facts on KSW grid setting

Proof of fact 3. We restate the fact:

On a graph from $\mathcal{K}$, given any positive $\theta < 1$ and integer $1 \leq l \leq n^\theta$, for n large enough: i) the probability that a random link from a given node u goes to a node outside of $B_l(u)$ is greater than $\frac{1-\theta}{1+3\theta} + o(1)$; ii) the probability that there is a random link to u from a node outside of $B_l(u)$ is greater than $1 - e^{-(1-\theta)/4+o(1)}$.

We still keep the denotation of E and F as in fact 2's proof. Let $a = \sum_{v \notin B_l(u)} d^{-2}(u,v)$ and $b = \sum_{v \in B_l(u)} d^{-2}(u,v)$. Obviously, $Pr[E] = \frac{a}{a+b}$.

From (1) we have

$$a = \sum_{j=l+1}^{2n-2} b_j(u)(j^{-2}) \geq \sum_{j=l+1}^{\lceil n/2 \rceil - 1} j(j^{-2}) \geq \sum_{j=n^\theta+1}^{\lceil n/2 \rceil - 1} j^{-1} = (1-\theta)lnn + c_1$$

for some $c_1 = O(1)$. On the other hand, from (1), $b = \sum_{j=1}^{l} b_j(u) j^{-2} \leq \sum_{j=1}^{l} 4j^{-1} \leq 4\theta lnn + c_2$ for some $c_2 = O(1)$. Thus $Pr[E] = \frac{a}{a+b} \geq \frac{(1-\theta)lnn+c_1}{(1-\theta)lnn+c_1+4\theta lnn+c_2}$, which converses to $\frac{1-\theta}{1+3\theta}$ when n goes to the infinity. So, $Pr[E] \geq \frac{1-\theta}{1+3\theta} + o(1)$.

On the other hand, from (4) we have $Pr[F] \geq 1 - e^{-\sum_{v \notin B_l(u)} p(v,u)}$; thus

$$Pr[F] \geq 1 - e^{-\sum_{v \notin B_l(u)} d^{-2}(u,v)/c_u} \geq 1 - e^{-a/(4ln6n)} = 1 - e^{-(1-\theta)/4+o(1)}.$$

Note that $c_v \leq 4ln(6n)$ as from fact 1's proof. The fact has been shown. □

Proof of fact 15. We restate the fact: *Define function* $\tau(\beta, \eta, \gamma) = 8\eta\beta/(4\beta - \gamma)^2$ *and let* $\beta_s = 0.5$. *Let* $\eta > 0$, $\gamma \in (1, 2)$ *and* $\hat{c} > 0$ *be constants. Let* H *be a ball with* $|H| = O(nlogn)$ *and* m *be an integer such that* $\hat{c}logn \leq m < |H|$. *If* $\hat{c} > \tau(\beta_s, \eta, \gamma)$ *then* $Pr[f(m, H)/m > \gamma] = 1 - O(n^{-\eta})$ *for* n *large enough.*

From the proof of fact 14, this can be made ($Pr[f(m, H)/m > \gamma] = 1 - O(n^{-\eta})$) by choosing $\hat{c}$ such that $e^{-2\hat{c}logn\beta\delta^2} = O(n^{-\eta})$ or $2\hat{c}\beta\delta^2 \geq \eta$ or $\hat{c} \geq \frac{\eta}{2\delta^2\beta}$. From $\gamma = 4\beta(1 - \delta)$ we have $\delta = 1 - \gamma/4\beta$, then we need $\hat{c} \geq \tau(\beta, \eta, \gamma) = 8\eta\beta/(4\beta - \gamma)^2$.

On the other hand, from any given $1 < \gamma < 2$, we can choose any β ($= \frac{\gamma}{4(1-\delta)}$) from the interval $(\gamma/4, \beta_s = 0.5)$ then choose $\hat{c} > \tau(\beta, \eta, \gamma)$ with η also given.

Moreover, if we choose $\hat{c} > \tau(\beta_s, \eta, \gamma) = \frac{4\eta}{(2-\gamma)^2}$ we then can always find β close enough to β_s such that $\hat{c} > \tau(\beta_s, \eta, \gamma)$ also holds because of τ's continuity. Thus, for given such γ and η, by choosing $\hat{c} \geq \tau(\beta, \eta, \gamma)$ we have $Pr[f(m, H)/m > \gamma] = 1 - O(n^{-\eta})$. □