Laboratoire Bordelais de Recherche en Informatique
Université Bordeaux I, 351, cours de la Libération,
33405 Talence CEDEX, France

Research Report RR-1195-98

Compact Routing Schemes With Low Stretch Factor

by Tamar Eilam, Cyril Gavoille and David Peleg

January, 1998

Compact Routing Schemes With Low Stretch Factor

Tamar Eilam* Cyril Gavoille† David Peleg‡

Abstract

This paper presents a routing strategy called *Pivot Interval Routing (PIR)*, which allows message routing on every weighted n-node network along paths whose *stretch* (namely, the ratio between their length and the distance between their endpoints) is at most five, and whose average stretch is at most three, with routing tables of size $O(n^{3/2}\log^{3/2} n)$ bits in total. A similar routing strategy for unweighted networks which guarantees the same bounds on the stretch factor and in addition a bound of $\lceil 1.5D \rceil$ on the route lengths, where D is the diameter of the network, is also presented. Moreover, it is shown that the PIR strategy can be implemented so that the generated scheme is in the form of an *interval routing scheme (IRS)*, using at most $2\sqrt{n(1+\ln n)}$ intervals per link in the first case and $3\sqrt{n(1+\ln n)}$ in the second case. As a result, the schemes are simpler than previous ones and they imply that paths of messages are loop-free. Finally, it is shown that there is no *loop-free* routing strategy guaranteeing a memory bound of $\sqrt{n}$ bits per router for all networks, regardless of the route lengths.

*Department of Computer Science, Technion, Haifa 32000, Israel. *eilam@cs.technion.ac.il*

†LaBRI, Université Bordeaux I, 351, cours de la Libération, 33405 Talence Cedex, France. *gavoille@labri.u-bordeaux.fr*

‡Department of Applied Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot, 76100 Israel. *peleg@wisdom.weizmann.ac.il*. Supported in part by a grant from the Israel Science Foundation and by a grant from the Israel Ministry of Science and Art.

1 Introduction

1.1 Background

In point-to-point communication networks, a routing scheme is employed in order to deliver messages between processors. As networks grow in size, it becomes important to reduce the amount of memory kept in each node for routing purposes. At the same time, it is essential to route messages along paths that are as short as possible. The efficiency of a routing scheme is often measured in terms of its *stretch factor*, namely, the maximum ratio between the length of the path traversed by a message and that of the shortest path between its source and destination.

A *universal routing strategy* is an algorithm which generates a routing scheme for every given network. One type of trivial universal routing strategy is based on schemes that keep in each node a full routing table which specifies an output port for every destination. Though this strategy can guarantee routing along shortest paths, each router has to store locally $\Theta(n \log d)$ bits of memory, where d is the degree of the router (i.e., the number of output ports) and n is the number of nodes in the network. Therefore, this scheme is impractical when dealing with large networks.

It was shown in a series of papers (see, e.g., [2, 3, 4, 26]) that there is a tradeoff between the memory requirements of a routing scheme and the worst case stretch factor it guarantees. In [26] it was shown that any universal routing strategy achieving stretch factor $s \geq 1$ must use a total of $\Omega(n^{1+\frac{1}{2s+4}})$ bits of routing information in the network. Stronger lower bounds hold for small stretch factors: any routing scheme with stretch factor $s < 3$ must use a total of $\Omega(n^2)$ bits [9, 17], and an optimal bound of $\Theta(n^2 \log n)$ bits holds for stretch $s = 1$ [20].

On the positive side, a number of routing strategies have been proposed that achieve almost optimal efficiency-space tradeoff. Specifically, in [26] it was shown that for every graph and every integer $k \geq 1$ it is possible to construct a hierarchical routing scheme with stretch factor $O(k)$ which uses a total of $\Omega(k^3 n^{1+1/k} \log n)$ bits and names each node with $O(\log^2 n)$ bits. Though this routing scheme is almost optimal in terms of its efficiency-space tradeoff, it has a few drawbacks. First, it is not *name-independent*, that is, it relabels the nodes with new names. Secondly, this scheme does not bound the local memory requirement of a node, and finally, it assumes a unit cost on the links of the network. Other methods that overcome these problems achieve an inferior efficiency-space tradeoff; the hierarchical routing scheme presented in [2] uses $O(k\, n^{1/k} \log n)$ bits of memory per node and guarantees a stretch factor $O(k^2\, 9^k)$. The scheme presented in [4] keeps the advantages of the former one, while regaining the polynomial tradeoff. In particular it guarantees, for every $k \geq 1$, stretch factor $O(k^2)$, while using $O(k\, n^{1/k} \log n \log D)$ bits per node, where D is the weighted diameter of the network.

The major disadvantage of all the proposed hierarchical routing strategies is that they are rather complex, and thus are impractical especially for high-speed networks for which the latency in each node must be very short. Broadly speaking, the hierarchical routing strategies are based on techniques for generating a sparse cover of clusters for the underlying graph. In order to implement the routing scheme, there are $O(\log n)$ levels of covers with increasing radii. Intuitively, higher levels are responsible for routing to farther destinations. A message is first sent on the lowest level, on the hope that the destination is nearby. If this is not the case, then the transmission might fail, in which case the message will bounce back to the originator. This step may be viewed as one "phase" of the routing process. Once receiving the message back (with a notification of failure), the originator tries to send it on a higher level, and so on until the routing process succeeds in

delivering the message. Clearly, the implementation of this procedure in every router results in a complex decision function; the number of the present phase has to be coded into the header of the message, thus the message header must be rewritten upon each retransmission. Moreover, intermediate routers must also change the message header, for example, in order to notify of a failure. (Another possibility is to send only a failure notification, but in this case the message originator must keep a copy, and an intermediate router has to generate additional messages). In any case, the originator has to compute a new header for the message in every phase of the algorithm. In addition, those strategies treat the nodes of the network non-uniformly, in the sense that different nodes play different roles, thus the decision function could be substantially different at different nodes. As high-speed networks gain popularity and increase in size, these drawbacks become crucial, since the main routing bottleneck in these networks is often the decision function in the nodes and not the propagation delay. Therefore, simple routing schemes which could be implemented in hardware may be preferable in practice.

Subsequently, considerable attention is given recently to an opposing design philosophy, focusing on *simple* and *uniform* compact routing strategies. Many compact routing strategies were proposed in the last decade (see, e.g., [7, 10, 11, 14, 27, 29]). An important property common to all of them is that they are *direct*. Namely, they employ a simple "shoot and forget" type decision function in the nodes, depending only on the destination of the message, and the destination is the only information coded in the message header (which is determined once and for all by the originating router, and is never changed afterwards.) As a consequence, it is easy to verify that every message traverses a loop-free path from its source to its destination. (Schemes enjoying this property are referred to in what follows as *loop-free* routing schemes.) All of these strategies involve relabeling of the nodes in order to reduce the amount of space needed for the routing, thus they are not name-independent. The most popular such scheme is the *interval routing scheme (IRS)*, presented in [27, 29] and implemented in the T9000 Transputer router chip of INMOS [22]. The idea of this scheme is to label nodes with unique integers from a contiguous range, and to label the outgoing arcs in every node with a set of intervals forming a partition of the name range. When invoking the delivery protocol, a message is sent on the unique outgoing arc labeled by an interval that contains the destination label. While the preprocessing stage of such a routing scheme (which is performed once in the initialization of the network) might be complex, the delivery protocol consists of simple "shoot and forget" decision functions in every node that depend only on the destination and thus imply loop-free routing paths.

One of the desirable goals in interval routing is to minimize the maximum number of intervals that label an arc. Unfortunately, while many lower bounds are known for this problem (see, e.g., [6, 8, 15, 18, 23, 28]), few trade-off results between efficiency and space are known for any of these strategies for general graphs. For interval routing, a universal strategy which is based on routing on a BFS-tree is presented in [27, 29]. This scheme uses only one interval per edge, thus the memory in a node with degree d is $O(d \log n)$, but it implies no non-trivial upper bound on the stretch factor. On the other extreme it was shown in [19] that it is possible to generate for every network an interval routing scheme which uses at most $n/4 + o(n)$ intervals per edge and guarantees stretch factor $s = 1$. Lately, [23] showed that for every graph there exists an interval routing scheme under which every message traverses a path of length at most $\lceil 1.5D \rceil$, where D is the diameter of the network, and which labels every arc with at most $\left\lceil \sqrt{n \ln n} \right\rceil + 1$ intervals. While this result implies an upper bound on the *dilation*, i.e., on the length of paths traversed by messages, it does not imply any non-trivial upper bound on the stretch factor. Moreover, the paper does not present any efficient (say, polynomial time) preprocessing algorithm for generating such an interval routing

scheme for a given graph.

This unfortunate situation holds not only for interval routing but for all loop-free, compact routing strategies proposed in the literature; while there are many compact routing strategies for specific topologies (see, e.g., [12, 13]), to the best of our knowledge there is no universal strategy which generates for every graph a loop-free routing scheme with a constant stretch factor that uses significantly fewer than $O(n^2)$ memory bits.

1.2 Our Results

A basic question which arises from the above discussion is whether it is possible to combine the benefits of both worlds, and design "simple" near-optimal routing strategies, which are still compact. In particular, can "simple" routing schemes achieve a similar tradeoff between stretch factor and memory as the hierarchical routing schemes? In this paper we analyze the power of "simple" routing strategies and give both positive and negative results for this question. To do that, we must first give a precise definition that captures our intuition about "simple" routing schemes. Define 1-*phase routing schemes* as schemes in which the header is not allowed to be re-written. *Oblivious routing schemes* pose further restrictions which make the decision function in the nodes even simpler. Finally, a *direct* routing scheme is defined as one in which the decision functions depend only on the destination, which is the only information that is kept (and never changed) in the header. Note that all the proposed compact routing strategies (e.g., interval routing) are direct. Our main results are two direct universal routing strategies, termed *pivot interval routing* (*PIR*, in short). The first one, PIR1, generates for every weighted graph with arbitrary link costs, a routing scheme with stretch factor $s \leq 5$ and average stretch factor $\overline{s} \leq 3$ that uses $O(n^{3/2} \log^{3/2} n)$ bits in total. Moreover, the PIR1 preprocessing algorithm actually generates for every graph an interval routing scheme which labels every arc with at most $2\sqrt{n(1+\ln n)}$ intervals. The dilation guaranteed by the PIR1 strategy is $2D$, where D is the weighted diameter of the network. For the unweighted case, we present a slightly different universal routing strategy named PIR2 (which also generates for every graph an interval routing scheme, but with $3\sqrt{n(1+\ln n)}$ intervals per arc), achieving the same memory requirements, stretch factor and average stretch factor as the PIR1 algorithm while guaranteeing a better dilation bound of $\lceil 1.5D \rceil$. We note that our routing strategies provide the first constructive method for designing for every graph an interval routing scheme with small stretch factor and $o(n)$ intervals per arc. We also prove that the bound of $\overline{s} \leq 3$ on the average stretch factor of the proposed routing strategies is tight, and yet that for almost all graphs (specifically, all but a $1/n^3$ fraction of all the n-node graphs), the average stretch factor guaranteed by the PIR1 and PIR2 routing strategies is asymptotically equal to 1.875.

When comparing the proposed strategies with the hierarchical routing strategies (which are not loop-free, direct, oblivious or 1-phase), a natural question is whether it is possible to design a family of loop-free routing strategies spanning the entire range of possible tradeoffs between different values of stretch factors and memory requirements. In other words, is it possible using loop-free routing strategies to decrease the memory requirements at the expense of a larger stretch factor? We answer this question negatively by showing that no matter which stretch factor one aims at guaranteeing, if the routing scheme is loop-free then there must be some graphs for which it requires $\sqrt{n}$ memory bits in at least one of the nodes. Hence there is no universal loop-free routing strategy that can guarantee $o(\sqrt{n})$ bits in each node, whatever the stretch factor is.

Finally, the PIR1 and PIR2 routing strategies are also compared with one of the hierarchical

routing strategies proposed in [2]. It is shown that while the difference between the average stretch factor achieved by both strategies on every graph does not exceed $o(1)$, there are concrete examples in which our strategy is superior in terms of the average stretch factor.

Following is a list of the contributions in this paper.

1. A universal routing strategy PIR1 that generates for every weighted network, a direct routing scheme which guarantees stretch factor $s \leq 5$ and average stretch factor $\overline{s} \leq 3$ and uses $O(\sqrt{n}\log^{3/2} n + d\log n)$ bits of memory in a node of degree d and $O(n^{3/2}\log^{3/2} n)$ bits in total. The dilation guaranteed by PIR1 is $2D$, where D is the weighted diameter of the network.

2. A similar universal routing strategy PIR2 for the unweighted case, that generates a direct routing scheme with the same memory requirements, maximum and average stretch factors as the PIR1 strategy, guarantees a better dilation of $\lceil 1.5D \rceil$.

3. The PIR1 and PIR2 universal strategies actually generate in polynomial time interval routing schemes with the same efficiency measures (stretch, average stretch, and dilation) as stated above and use up to $2\sqrt{n(1+\ln n)}$ intervals per arc in the first case and up to $3\sqrt{n(1+\ln n)}$ intervals per arc in the second.

4. A proof that the bound of $\overline{s} \leq 3$ on the average stretch factor of the proposed routing strategies is tight.

5. A proof, using the family of random graphs $\mathcal{G}_{n,p}$ (with $p = \frac{1}{2}$), that for almost all graphs the average stretch factor guaranteed by the PIR1 and PIR2 routing strategies is asymptotically equal to 1.875.

6. A lower bound on the power of loop-free routing schemes; specifically, it is shown that there exists an n-node graph G such that for any loop-free routing scheme on G there is a node that uses $c\sqrt{n}$ bits, $c \approx 3.71$, for any stretch factor guaranteed by the scheme. This result implies that, unlike the situation with hierarchical routing strategies (that are not 1-phase and that guarantee $O(k\, n^{1/k}\log n)$ bits per node [2], for every $k \geq 1$) the PIR1 and PIR2 routing strategies cannot be generalized to construct a family of routing schemes with different bounds on the stretch factor and memory.

7. The PIR1 and PIR2 routing strategies are compared with one of the hierarchical routing strategies proposed in [2]. It is shown that while the difference between the average stretch factor achieved by both strategies on every graph does not exceed $o(1)$, there are concrete examples in which our strategy is superior in terms of the average stretch factor.

The rest of the paper is organized as follows. In Section 2 we give a precise definition of routing schemes and efficiency measures. We formally define interval routing and in addition we overview some covering techniques, adopted from [2]. We make use of these techniques for the proposed routing strategy. In Section 3 we present our PIR routing strategies, prove their correctness and analyze their complexity measures. The lower bound on the power of loop-free routing schemes is proved in Section 4. A comparison between PIR and HCP_k, the hierarchical routing strategy presented in [2], is given in Section 5. Conclusions and discussion are offered in Section 6. Proofs are all deferred to the appendix.

2 Model and Definitions

2.1 Routing Schemes and Complexity Measures

A point-to-point communication network is modeled as a symmetric, weighted, finite digraph $G = (V, E, W)$, $|V| = n$, where the set of nodes represent the processors of the network and every pair of two opposite arcs represents a bidirectional communication link. Every arc of the network $e \in E$ is associated with a non-negative weight $w(e)$ (i.e., its cost). We assume that for every two opposite arcs e_1 and e_2, $w(e_1) = w(e_2)$. In the special case of uniform unit weight links, we say that the graph is *unweighted*, and denote it simply by $G = (V, E)$. Graphs are connected and do not contain self-loops or multi-arcs. We assume that every node v is named with a unique identification integer $\mathrm{ID}(v)$. In what follows, we informally confuse between the node v and its name $\mathrm{ID}(v)$. Note that the identifications induce a total order on the nodes, thus for every two nodes $u, v \in V$ either $u < v$ or $v < u$.

The length $|P|$ of a directed path P in the graph is the sum of weights of the arcs in P. The distance $\mathrm{dist}_G(u, v)$ between two nodes $u, v \in V$ is the length of a shortest path connecting them. The *diameter* of the graph G is defined as $D(G) = \max\{\mathrm{dist}_G(u, v) \mid u, v \in V\}$. For a node $v \in V$, let E_v denote its set of outgoing arcs, and denote its *degree* by $\deg(v) = |E_v|$.

A *routing scheme* R is a distributed algorithm whose role is to deliver messages between nodes of the network. The routing scheme consists of certain *distributed data structures* in the network, and a *delivery protocol*, which can be invoked in any node u with two parameters: a *routing label* $\mathcal{L}(v)$ of the destination node v, and the message's information field. A routing scheme is *name-independent* if node "names" do not change in the preprocessing, that is, the destination's routing label $\mathcal{L}(v)$ is simply its original name, $\mathrm{ID}(v)$. Clearly, this property is desirable in terms of transparency of the routing process. On the other hand, an appropriate labeling of the nodes may considerably reduce the size of the routing tables. In a name-dependent routing scheme, nodes are relabeled by unique integers in the preprocessing stage, and the new labels are then used when invoking the delivery protocol. The message is delivered to v via a sequence of transmissions determined uniquely by the distributed data structure. The length of the route traversed by a message from u to v in the graph G according to the routing scheme R is denoted by $\mathrm{dist}_G(R, u, v)$.

Specifically, the delivery protocol of the routing scheme R uses a pair of integer functions (P, H) where P is the *port function* and H is the *header function*. For any two distinct nodes u and v, R produces a path of nodes $(u = u_0, u_1, \ldots, u_k = v)$, a sequence $h_0, h_1, \ldots, h_k$ of headers, and a sequence $p_0, p_1, \ldots, p_k$ of output port numbers. A message with header h_i arriving at node u_i through input port p_i is given a new header $H(u_i, p_i, h_i) = h_{i+1}$, and is forwarded to the output port $P(u_i, p_i, h_i) = p_{i+1}$. Thus, we require that for every $i \in \{1, \ldots, k-1\}$, $H(u_i, p_i, h_i) = h_{i+1}$, $P(u_i, p_i, h_i) = p_{i+1}$ and that p_{i+1} corresponds to the output port number of the link (u_i, u_{i+1}). On each router, the input and output ports numbered 0 are associated with the special link between the router and its host. This allows us to complete the description by imposing the constraints that $p_0 = 0$, $P(u_k, p_k, h_k) = 0$, as well as $h_0 = \mathcal{L}(v)$, thus fixing the initial header, which is provided to the router from a higher layer. (In a name-independent routing scheme, h_0 is set to $\mathrm{ID}(v)$, the original name of the destination node v.) Since we are interested in distributed routing strategies, we will express the delivery protocol as a collection of local routing functions $R_x = (P_x, H_x)$, one for each node x.

A 1-*phase routing scheme* is a routing scheme in which the header of the message is determined

by the originator and is never modified by any intermediate node. Formally, the first header in the sequence, h_0, is equal to the destination. This is simply the mathematical formulation of the command to invoke the delivery protocol which is given from a higher layer. When the delivery protocol is invoked, the first router computes the header h_1 of the message (which is not necessarily equal to the destination). Once the header is fixed by the first router, it is never changed, i.e., 1-phase routing schemes satisfy $h_0 = \mathcal{L}(v)$ and $h_1 = h_2 = ... = h_k$. The fact that there is no need to recompute and rewrite the header once it is computed by the first router when invoking the delivery protocol, implies a simple and fast decision function in the nodes. This is important especially when dealing with high-speed networks.

An *oblivious routing scheme* is a 1-phase routing scheme in which for every node x, two input ports p and p' and header h, $P(x, p, h) = P(x, p', h)$, namely, the decision function in a node x depends only on the header of the message (and not on its entry port). Note that 1-phase routing schemes may route messages along paths with loops: a message might arrive at the same node of the network twice but from different input ports. In the case of oblivious routing schemes the routing path may cycle once on the originator node only, since the routing decision depends only on the header, which is changed at most once (at the first router).

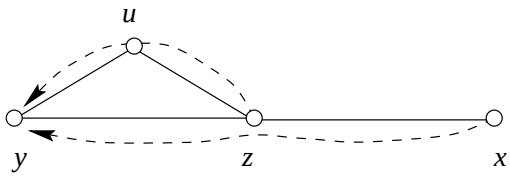

Figure 1: An oblivious routing scheme which is not direct.

We advance one step further and define a *direct routing scheme* as an oblivious routing scheme in which the decision functions in the nodes depend only on the destination of the message, namely, $h_0 = h_1 = ... = h_k = \mathcal{L}(v)$ (where $\mathcal{L}(v)$ is the destination's label). Note that every direct routing scheme is also a *loop-free* routing scheme, i.e., it guarantees that a message always traverse a loop-free path, since otherwise the message will cycle forever without reaching its destination, contradicting the definition of a routing scheme. The classical full routing tables are direct routing schemes. Another example of a direct routing scheme is interval routing, a well-known routing method presented in [27, 29]. Figure 1 is an example of an oblivious routing scheme which is not direct: a message destined to a node y, entering z on the arc from x is forwarded on one output port, while a message entering z on input port 0 (i.e., the originator is z) will be forwarded on different output port. This is possible with oblivious routing schemes since the header used for a message from z to y might be different from the header used for a message from x to y.

A *universal routing strategy* is a function that returns, for every n-node graph G, a routing scheme on G. It is implemented by a *preprocessing algorithm*, performed during set-up time in order to construct the distributed data structures required for the routing scheme. A routing strategy is *name-independent* if it produces for every graph a name-independent routing scheme.

Given a graph G, a routing scheme R on G, and a node x, the *memory requirement* of x, denoted by $\text{Memory}_G(R, x)$, is the smallest number of bits that are required in order to code R_x. The *total memory requirements* of R, are

$$\text{Memory}_G(R) \;=\; \sum_x \{\text{Memory}_G(R, x)\} \;.$$

Let R be a routing scheme on an n-node graph G. The maximum and average stretch factors of R are respectively defined as follows.

$$\mathrm{Stretch}_G(R) = \max_{x \neq y} \frac{\mathrm{dist}_G(R,x,y)}{\mathrm{dist}_G(x,y)} \quad \text{and} \quad \mathrm{AvStr}_G(R) = \frac{1}{n(n-1)} \sum_{x \neq y} \frac{\mathrm{dist}_G(R,x,y)}{\mathrm{dist}_G(x,y)}.$$

A routing scheme of stretch factor 1 is termed a *shortest path* routing scheme. The *dilation* of a routing scheme R is the maximal length of a path traversed by a message. Formally,

$$\mathrm{Dilation}_G(R) = \max_{x \neq y} \{\mathrm{dist}_G(R,x,y)\}.$$

Given a graph G and a set of routing schemes $\mathcal{RS}$ on G, denote

- $\mathrm{Memory}_G(\mathcal{RS}) = \min_{R \in \mathcal{RS}} \{\mathrm{Memory}_G(R)\}$
- $\mathrm{Stretch}_G(\mathcal{RS}) = \min_{R \in \mathcal{RS}} \{\mathrm{Stretch}_G(R)\}$
- $\mathrm{AvStr}_G(\mathcal{RS}) = \min_{R \in \mathcal{RS}} \{\mathrm{AvStr}_G(R)\}$
- $\mathrm{Dilation}_G(\mathcal{RS}) = \min_{R \in \mathcal{RS}} \{\mathrm{Dilation}_G(R)\}$

Hereafter, we sometimes omit the subscript G, when the graph is clear from the context.

2.2 Direct Routing Schemes and Interval Routing

In this section we give a definition of interval routing, and show its equivalence to direct routing.

An *interval routing scheme* R on G is a routing scheme consisting of a pair $(\mathcal{L}, \mathcal{I})$, generated in the preprocessing step, where $\mathcal{L}$ is a *node-labeling*, $\mathcal{L} : V \to \{1, \ldots, n\}$, and $\mathcal{I}$ is an *arc-labeling*, $\mathcal{I} : E \to 2^{\mathcal{L}(V)}$, that satisfy the following condition. For any node x, the collection of sets that label all the outgoing arcs of x forms a partition of the name range (possibly excluding x itself). Formally, for every $x \in V$:

1) $\bigcup_{e \in E_x} \mathcal{I}(e) \cup \mathcal{L}(x) = \{1, \ldots, n\}$,
2) $\mathcal{I}(e) \cap \mathcal{I}(e') = \emptyset$, for every two distinct arcs $e, e' \in E_x$.

The delivery protocol is defined as follows. Given a destination node v, set the first header to be $h_1 = \mathcal{L}(v)$. Also, for every node y and port p, $H(y,p,h) = h$ and $P(y,p,h) = p'$ if and only if $h \in \mathcal{I}(x,y)$ and p' is the output port number of the arc (x,y). Namely, the message is sent on the arc which is labeled by a set that contains the destination. For any set $V' \subseteq V$, denote $\mathcal{L}(V') = \{\mathcal{L}(v) \mid v \in V'\}$.

We denote by IRS the class of all the interval routing schemes on arbitrary graphs. Clearly, every interval routing scheme is a direct routing scheme. Conversely, we argue that every direct routing scheme could be implemented using interval routing. This is done by labeling every outgoing arc in a node with the set of destinations for which a message will be sent on that arc. This is always possible since with direct routing schemes, in every node, two messages with the same destination are forwarded on the same arc (even if their originators are different). Note that this is not the case with oblivious routing, as implied by Figure 1.

Given an integer n and a subset $I \subseteq \{1, \ldots, n\}$, define the *compactness of I w.r.t. n*, denoted $c(I)$, as the smallest integer k such that I can be represented by the union of k intervals $[a,b]$ of

consecutive integers from $\{1,\ldots,n\}$, n and 1 being considered as consecutive (cyclically). The compactness of an interval routing scheme $R=(\mathcal{L},\mathcal{I})$ on G, denoted by $\mathrm{Comp}_G(R)$, is the maximum, overall arcs $e\in E$, of the compactness $c(\mathcal{I}(e))$ of the set $\mathcal{I}(e)$ labeling e. Intuitively, smaller compactness and degrees imply smaller routing tables. The interval routing strategy presented in [27], based on routing on a minimum spanning tree, has compactness 1 (for every graph) but unbounded stretch factor. Another example is the ring C_n which admits an interval routing scheme R of compactness 1 with stretch factor 1 ([27]). Therefore, $\mathrm{Memory}_{C_n}(R,x)=O(\log n)$ for every node x, as it suffices for x to store its label and the intervals assigned to the two arcs. Actually, every graph G supporting an IRS R of compactness k satisfies $\mathrm{Memory}_G(R,x)=O(\deg(x)k\log(n/k))$ [16].

2.3 Balls, Neighborhoods and Covers

Our routing scheme constructions are based on the notions of neighborhoods, balls and covers. For every node v we can order all the nodes of the graph w.r.t. v by increasing distance from v, breaking ties by increasing node identifications. Formally, $x\prec_v y$ if and only if either $\mathrm{dist}_G(x,v)<\mathrm{dist}_G(y,v)$ or $\mathrm{dist}_G(x,v)=\mathrm{dist}_G(y,v)$ and $x<y$. The *t-ball* $\mathcal{B}_v(t)$ of v, is the set of the first t nodes according to the node ordering w.r.t. v. The *r-neighborhood* of a node $v\in V$ is defined as $\Gamma_v(r)=\{u\in V\mid \mathrm{dist}_G(v,u)\leq r\}$. Hence intuitively, a ball is a neighborhood defined by volume rather than by radius.

Following is a simple fact which holds for both neighborhoods and balls. The proof can be found in [25], Page 53.

Fact 2.1 (monotonicity) *If $u\in\mathcal{B}_v(t)$ (resp., $u\in\Gamma_v(r)$) then for every node x on a shortest path from v to u, $u\in\mathcal{B}_x(t)$ (resp., $u\in\Gamma_x(r)$).*

Consider a collection $\mathcal{H}$ of subsets of size t of elements from a set B. A set $P\subseteq B$ is said to *cover* the collection $\mathcal{H}$ if for every $A\in\mathcal{H}$, $A\cap P\neq\emptyset$. We review two techniques presented in [2] for generating relatively small covers for a given collection of sets of equal size.

The first technique is by using a greedy algorithm that starts with $P=\emptyset$ and iteratively adds to the set P an element in B occurring at the most uncovered sets. The algorithm stops when P becomes a cover. The set P is termed a *greedy cover* for $\mathcal{H}$.

Lemma 2.2 (Lovász [24]) *Let P be a greedy cover for $\mathcal{H}$. Then $|P|<|B|(\ln|\mathcal{H}|+1)/t$.*

The second method is randomized, and takes each element of B to the set P with probability $(c\ln|\mathcal{H}|)/t$, for some constant $c>1$.

Lemma 2.3 (Awerbuch et al. [2]) *Let P be the set constructed by the randomized cover algorithm under the assumptions that $|B|\geq 2t$ and $\ln|\mathcal{H}|=o(t)$. Then with probability at least $1-1/|\mathcal{H}|^{c-1}$, P is a cover for $\mathcal{H}$ and $|P|\leq(2c|B|\ln|\mathcal{H}|)/t$.*

For our needs, $|\mathcal{H}|=|B|=n$. In this paper, we are interested in the r-neighborhoods of nodes only for the case $r=\lceil D/2\rceil$, where D is the diameter of the graph, and only for unweighted graphs. In particular, we would like to use such neighborhoods in order to construct a small $\lceil D/2\rceil$-dominating set for our graph.

For every node v_i, $1 \leq i \leq n$, let $W_i = \Gamma_{v_i}(\lceil D/2 \rceil)$. Note that a cover for the set family $\mathcal{W} = \{W_1, \ldots, W_n\}$ (i.e., a set $X \subseteq V$ whose intersection with each W_i is nonempty) is a $\lceil D/2 \rceil$-dominating set for the graph. Also, note that (since the weights on the graph arcs are uniform) $\mathcal{W}$ is an intersecting hypergraph, namely, the intersection $W_i \cap W_j$ is nonempty for every $i \neq j$, since otherwise $\mathrm{dist}(v_i, v_j) > D$. It is therefore easy to verify that $\mathcal{W}$ has a cover of cardinality $O(\sqrt{n \ln n})$. Note that this cannot be deduced directly from Lemma 2.2, since $|W_i|$ is not necessary exactly $\sqrt{n(1 + \ln n)}$. Nevertheless, such a cover X can be constructed as follows.

First, note that in an intersecting hypergraph, each hyperedge is in itself a cover for the hypergraph. Hence if there exists some $1 \leq i \leq n$ such that the set W_i of cardinality $|W_i| \leq \sqrt{n(1 + \ln n)}$ then we are done. So now suppose that $|W_i| > \sqrt{n(1 + \ln n)}$ for every $1 \leq i \leq n$. In this case, the result follows from observing that Lemma 2.2 holds also for a collection of sets of non-equal sizes, provided t is a *lower bound* on the size of the sets in the collection. Hence setting $t = \sqrt{n(1 + \ln n)}$ we have the following.

Lemma 2.4 *For every n-node unweighted graph G there exists a $\lceil D/2 \rceil$-dominating set X of cardinality $|X| \leq \sqrt{n(1 + \ln n)}$. Moreover, this set can be selected by a straightforward greedy algorithm.*

3 The Pivot Interval Routing Strategy

We present two routing strategies, termed *Pivot Interval Routing*. The first one, PIR1, generates a direct routing scheme for every weighted graph (Theorem 3.1) and the second one, PIR2, generates a direct routing scheme for unweighted graphs with improved dilation (Theorem 3.11).

Theorem 3.1 *For every n-node weighted graph $G = (V, E, W)$ with weighted diameter D there exists a (direct) interval routing scheme $R = (\mathcal{L}, \mathcal{I})$ on G such that:*

- $\mathrm{Memory}_G(R) = O(n^{3/2} \log^{3/2} n)$,
- $\mathrm{Memory}_G(R, v) = O(n^{1/2} \log^{3/2} n + \deg(v) \log n)$, *for every $v \in V$,*
- $\mathrm{Stretch}_G(R) \leq 5$,
- $\mathrm{AvStr}_G(R) \leq 3$,
- $\mathrm{Dilation}_G(R) \leq 2D$,
- $\mathrm{Comp}_G(R) \leq 2\sqrt{n(1 + \ln n)}$.

Moreover, R can be constructed in time polynomial in n.

Intuitively, the idea of the PIR1 strategy is first to find the collection of t-balls of all the nodes of the graph, then to cover this collection by a (comparatively small) set of nodes termed *pivots*, and finally to label the nodes and arcs of the graph in such a way that a message with source u and destination v will be routed as follows. If the destination v is in u's t-ball, then the message will traverse a shortest path from u to v. Otherwise, it will traverse (in the worst case) a shortest path to the pivot nearest to v, and then a shortest path from that pivot to v itself. We present the preprocessing algorithm of the PIR1 strategy for any given ball size t and for any given cover P of the collection of t-balls of nodes in the graph. The value of t is determined later, in the proof of Theorem 3.1, where it is also shown how to construct a cover P such that PIR1 will satisfy the properties in Theorem 3.1. We now give a formal description of the PIR1 strategy.

The preprocessing algorithm consists of three parts. In the first part, some preliminary structures are constructed which are used later to define the node and arc labeling functions $\mathcal{L}$ and $\mathcal{I}$. The delivery protocol of PIR1, is the usual delivery protocol for interval routing schemes.

Strategy PIR1

Input: A weighted graph $G = (V, E, W)$.

The preprocessing algorithm

a. Preliminary constructions

1. Let P be a cover for the collection of t-balls of the nodes, $\{\mathcal{B}_v(t) \mid v \in V\}$.
2. Assign to every node v its pivot $p(v) \in P$, where $p(v)$ is the nearest node to v in P (breaking ties by increasing node identifications).
3. For every pivot $p \in P$, let $S_p = \{v \mid p(v) = p\}$ be its set of "clients".
4. For every pivot $p \in P$, construct a minimum BFS spanning tree T_p rooted at p and spanning the entire graph G, and let T'_p be the subtree of T_p induced by p's set of clients S_p.

b. Labeling the nodes
Assume that $P = \{p_1, p_2, \ldots, p_l\}$. We start by labeling the nodes in S_{p_1}. The labeling is performed by traversing the tree T'_{p_1} (i.e., the subtree spanning S_{p_1}), and assigning the nodes of T'_{p_1} a DFS (pre-order) numbering in sequential ascending order, starting from 1. Assuming that n_1 is the largest label assigned to a node in S_{p_1}, we continue by labeling the nodes in S_{p_2} in the same manner (traversing the tree T'_{p_2}), starting from the integer $n_1 + 1$. Then we label the nodes in $S'_{P_3}, \ldots S'_{p_l}$ in the same way.

c. Labeling the arcs
For every node $u \in V$, we label every arc $e \in E_u$ by a set of destinations $\mathcal{I}(e) \subseteq \{1, \ldots, n\}$ in three steps. We start by fixing $\mathcal{I}(e) = \emptyset$ for every $e \in E_u$, and then at each step we add some node labels to the sets $\{\mathcal{I}(e)\}_{e \in E_u}$, such that labels are never deleted from the sets $\mathcal{I}(e)$, and the sets are mutually disjoint throughout the process. Formally, we label the arcs E_u of every node u as follows.

0. fix $I(e) = \emptyset$, for every $e \in E_u$.
1. If u is not a leaf in the tree $T'_{p(u)}$, let $s_1, ..., s_j$ be its successors in $T'_{p(u)}$ and let T'_{s_i} be the subtree of $T'_{p(u)}$ rooted at s_i, for $1 \leq i \leq j$. Assign $\mathcal{I}(u, s_i) = \{\mathcal{L}(v) \mid v \in T'_{s_i}\}$, for every $1 \leq i \leq j$. Define $M_1 = \bigcup_{1 \leq i \leq j} \mathcal{I}(u, s_i)$.
2. Define $M_2 = \mathcal{L}(\mathcal{B}_u(t)) \setminus M_1$. For every node $v \neq u$ such that $\mathcal{L}(v) \in M_2$, let $e \in E_u$ be an arc on a shortest path from u to v (if there is more than one such arc, choose one arbitrarily). Assign $\mathcal{I}(e) = \mathcal{I}(e) \cup \{\mathcal{L}(v)\}$.
3. Let $M_3 = \mathcal{L}(V) \setminus (M_1 \cup M_2)$. For every $p \in P$, let $e_p \in E_u$ be the arc from u to its predecessor on the tree T_p. Assign $\mathcal{I}(e_p) = \mathcal{I}(e_p) \cup (\mathcal{L}(S_p) \cap M_3)$.

The delivery Protocol
In a node u, a message M with destination $v \neq u$ is sent on the unique arc $e \in E_u$, such that

$\mathcal{L}(v) \in \mathcal{I}(e)$.

Figure 2 depicts the path of a message on the spanning tree $T_{p(v)}$ from a source u to a destination v which is not in u's ball.

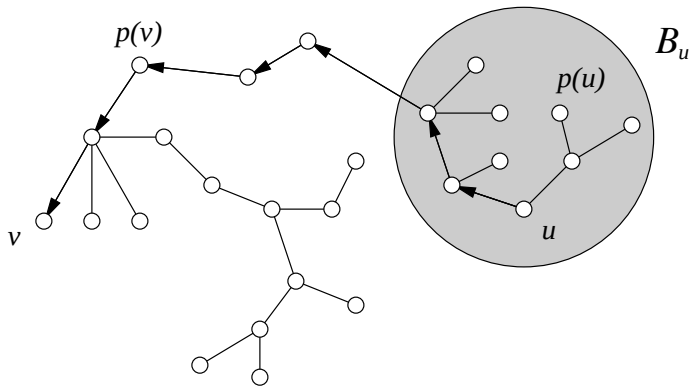

Figure 2: The route from u to a destination v which is not in u's ball.

We now prove that the PIR1 preprocessing algorithm generates a (direct) interval routing scheme that satisfies (for specific possible values of t and $|P|$) the properties stated in Theorem 3.1.

We start with some simple observations.

Observation 3.2

1. *For every node $v \in V$, $p(v) \in \mathcal{B}_v(t)$.*
2. *For every node $v \in V$, and for every node $u \in V$ on a shortest path from v to its pivot $p(v)$, $p(u) = p(v)$.*
3. *$\{S_p\}_{p \in P}$ is a partition of the set of nodes V.*
4. *For every pivot $p \in P$, T'_p is a connected tree.*

Proof. The first observation follows from the definition of P as a cover of the collection of t-balls and the choice of the pivot $p(v)$ as the nearest pivot for every node v. The second one is easily verified by the monotonicity property (Fact 2.1), and the choice of $p(v)$ for every v. The third observation follows from the definition of the sets S_p, and the last observation is easily verified using the definition of the tree T'_p, the sets S_p, and Observation 2. □

Define a priority function on the labels of nodes in every node u. A label of a node $\mathcal{L}(x)$ has priority i in u, for $i \in \{1, 2, 3\}$, if and only if it was inserted to a set $\mathcal{I}(e)$, for some $e \in E_u$, in Step i of the arc labeling algorithm. The *priority of a message M in a node u* is the priority of the label of its destination in u. We first prove that the output of the PIR1 preprocessing algorithm for every weighted graph G is an interval routing scheme on G, and then analyze the routes taken by the messages.

Proposition 3.3 *R is an interval routing scheme on G. Specifically,*

a. For every $x \in V$, $\{\mathcal{I}(e) \mid e \in E_x\} \cup \mathcal{L}(x) = \{1, \ldots, n\}$,

b. *For every two distinct arcs* (x,y) *and* (x,z), $\mathcal{I}(x,y) \cap \mathcal{I}(x,z) = \emptyset$.

c. *For every two nodes* $u, v \in V$, *there exists a sequence of nodes* $u = x_1, x_2, ..., x_l = v$ *such that* $\mathcal{L}(v) \in \mathcal{I}(x_{i-1}, x_i)$, *for every* $2 \leq i \leq l$ (*namely, every message eventually arrives at its destination*).

Moreover, for every two nodes $u, v \in V$, *a message* M *with source* u *and destination* v, *will traverse a shortest path from* u *to* v *if* $v \in \mathcal{B}_u(t)$, *otherwise it will traverse a shortest path to* $p(v)$ *(or a prefix of that path) and then a shortest path to* v.

Proof. In every step of the arc labeling algorithm, every label which is added to one of the sets $\mathcal{I}(e)$ did not belong to any of the sets in any former step. In addition, it is clear that in every step a label is inserted only to one of the sets $\mathcal{I}(e)$. Note that in Step 3 we rely on the fact that $\{S_p\}_P$ is a partition of V. Proposition 3.3b follows. Since every label of a node belongs to a (unique) set S_p for some $p \in P$, Step 3 of the arc labeling algorithm guarantees that every node label will be contained in one of the sets $\mathcal{I}(e)$. Proposition 3.3a follows.

In order to prove Proposition 3.3c, we show that the sequence of priorities of a message on the path traversed by it is non-increasing. In addition, if the priority of a message in a node is 1 then eventually the destination is reached and if it is 2 or 3 then eventually the destination is reached or the priority of the message decreases.

Consider a message M with source u and destination v. Assume that its priority in a node x is 1. Let s_i be the successor of x in the tree $T'_{p(x)}$ such that v belongs to the subtree T'_{s_i} of the tree $T'_{p(x)}$. The message is delivered to s_i which is strictly closer to v than x and the same condition holds for s_i. Thus, M has priority 1 also in s_i. It can be easily proved by induction (using Observation 3.2) that the priority of M will remain 1 and that eventually v is reached.

Now assume that M has priority 2 in a node x. It follows that $v \in \mathcal{B}_x(t)$. Assume that $\mathcal{L}(v) \in (x,y)$. Then since (x,y) is on a shortest path from x to v, by Observation 3.2, $v \in \mathcal{B}_y(t)$ and y is strictly closer to v than x. It follows that either the priority of M in y is 1, and the claim follows by reduction to the former case, or the priority remains 2. In any case, since we advance towards v, it is easy to verify by induction that eventually v is reached in this case as well.

Finally, assume that M has priority 3 in a node x. Let $e = (x,y)$ be the arc such that $\mathcal{L}(v) \in \mathcal{I}(e)$. Then it is clear by the PIR1 preprocessing algorithm, that e is the arc in the tree $T_{p(v)}$ from x to its predecessor. It follows that y it is strictly closer to $p(v)$ and since the priority of M in $p(v)$ is 1, eventually the priority of M decreases (to either 2 or 1), and thus, by reduction to the former cases, it will eventually arrive at its destination.

Note that if the priority of a message in a node x is 2 or 1 then it is sent along an arc which is on a shortest path to the destination. Thus in case the destination belongs to the source's t-ball, the message will traverse a shortest path. If the priority of a message in a node is 3 then the message is sent on an arc which is on a shortest path to the destination's pivot, thus in this case a shortest path to the destination's pivot is traversed until the priority decreases. Since when $p(v)$ is reached the priority of M decreases to 1, M will traverse in this case a shortest path to $p(v)$ or some prefix of that path, and then it will traverse a shortest path to v. Proposition 3.3 follows. □

We now analyze the properties of the PIR1 routing algorithm. We denote by $PIRS1(t,l)$ the set of all routing schemes R generated by algorithm PIR1, where t is the size of the balls used and l is the size $|P|$ of the cover P for the collection of t-balls.

Proposition 3.4 $\text{Dilation}_G(R) \leq 2D$, *for every weighted graph* G *and* $R \in PIRS1(t,l)$.

Proof. By Proposition 3.3, $\text{dist}(R,u,v) \leq \text{dist}(u,p(v)) + \text{dist}(p(v),v) \leq 2D$. □

Proposition 3.5 $\text{Stretch}_G(R) \leq 5$, *for every weighted graph* G *and* $R \in PIRS1(t,l)$.

Proof. Consider a message M with source u and destination v. By Proposition 3.3, in the case where $v \in \mathcal{B}_u(t)$ the stretch factor is 1, thus we only have to bound the stretch factor in case $v \notin \mathcal{B}_u(t)$. The length of the path traversed by the message M in this case is bounded by the distance between u and $p(v)$ plus the distance between $p(v)$ and v. Formally,

$$\text{dist}(R,u,v) \leq \text{dist}(u,p(v)) + \text{dist}(p(v),v).$$

Since $p(u) \in \mathcal{B}_u(t)$, $\text{dist}(u,p(u)) \leq \text{dist}(u,v)$. Thus

$$\text{dist}(v,p(u)) \leq \text{dist}(v,u) + \text{dist}(u,p(u)) \leq 2\,\text{dist}(u,v).$$

Since $p(v)$ is the pivot minimizing the distance to v among all pivots in P,

$$\text{dist}(v,p(v)) \leq \text{dist}(v,p(u)) \leq 2\,\text{dist}(u,v).$$

Finally,

$$\text{dist}(u,p(v)) \leq \text{dist}(u,v) + \text{dist}(v,p(v)) \leq 3\,\text{dist}(u,v).$$

It follows that

$$\text{dist}(R,u,v) \leq \text{dist}(u,p(v)) + \text{dist}(p(v),v) \leq 5\,\text{dist}(u,v).$$

□

Proposition 3.6 $\text{AvStr}_G(R) \leq 3$, *for every weighted graph* G *and* $R \in PIRS1(t,l)$.

Proof. The claim is established by showing that for every two nodes $u, v \in V$,

$$\frac{\text{dist}(R,u,v)}{\text{dist}(u,v)} + \frac{\text{dist}(R,v,u)}{\text{dist}(v,u)} \leq 6.$$

We consider three cases.
Case 1: $u \in \mathcal{B}_v(t)$ and $v \in \mathcal{B}_u(t)$. Then a message from v to u and a message from u to v will both traverse a shortest path. Thus, the sum of the stretch factors of both paths is 2.
Case 2: $u \in \mathcal{B}_v(t)$ but $v \notin \mathcal{B}_u(t)$ (or vice-versa). Then a message from v to u will traverse a shortest path, and the message from u to v will traverse, by Proposition 3.5, a path of length at most $5\,\text{dist}(u,v)$. Thus, the sum of the stretch factors is at most 6.
Case 3: $u \notin \mathcal{B}_v(t)$ and $v \notin \mathcal{B}_u(t)$. We bound the stretch factor of the path of a message from, say, u to v, and the same bound holds symmetrically for the path of a message from v to u. Since $u \notin \mathcal{B}_v(t)$, and recalling that $p(v) \in \mathcal{B}_v(t)$, we have that $\text{dist}(p(v),v) \leq \text{dist}(u,v)$. It follows that $\text{dist}(u,p(v)) \leq \text{dist}(u,v) + \text{dist}(v,p(v)) \leq 2\,\text{dist}(u,v)$. Thus, $\text{dist}(R,u,v) \leq \text{dist}(u,p(v)) + \text{dist}(v,p(v)) \leq 3\,\text{dist}(u,v)$. Symmetrically, $\text{dist}(R,v,u) \leq 3\,\text{dist}(u,v)$. Thus, in this case as well,

the sum of the stretch factors of the path of a message from u to v and from v to u is at most 6. It follows that

$$\begin{aligned}
\mathrm{AvStr}(R) &= \frac{1}{n(n-1)} \sum_{u \neq v} \frac{\mathrm{dist}(R,u,v)}{\mathrm{dist}(u,v)} \\
&= \frac{1}{n(n-1)} \sum_{u<v} \left(\frac{\mathrm{dist}(R,u,v)}{\mathrm{dist}(u,v)} + \frac{\mathrm{dist}(R,v,u)}{\mathrm{dist}(v,u)} \right) \\
&\leq \frac{1}{n(n-1)} \sum_{u<v} 6 \;\leq\; 3 \;.
\end{aligned}$$

□

Proposition 3.7 $\mathrm{Comp}_G(R) \leq t + l$, *for every weighted graph* G *and* $R \in PIRS1(t,l)$.

Proof. We use the following facts, which are easily verified, in order to bound the compactness of a set $\mathcal{I}(e)$ that labels an arc e.

Fact 3.8 *For every two sets* U_1 *and* U_2,

a. $c(U_1) \leq |U_1|$.

b. $c(U_1 \cup U_2) \leq c(U_1) + c(U_2)$.

c. $c(U_1 \setminus U_2) \leq c(U_1) + c(U_2)$.

Consider a node $u \in V$. Let V_i, $i \in \{1,2,3\}$, be the set of labels with priority i in the node u. Define $V_i(e) = \mathcal{I}(e) \cap V_i$ for every outgoing arc $e \in E_u$, $i \in \{1,2,3\}$. Clearly, since every tree T_p is labeled by a DFS order, $c(V_1(e)) = 1$ for every $e \in E_v$ and also $c(V_1) \leq 2$. It is also clear, by Fact 3.8[a], that $c(V_2(e)) \leq t-1$, for every $e \in E_u$, and moreover, $c(V_2) \leq t-1$. Assume that e is the arc from u to its predecessors in the trees $T_{p_{i_1}}, T_{p_{i_2}}, \ldots, T_{p_{i_j}}$, $p_{i_x} \in P$. Note that $j \leq l$ and let $S_e = \cup_{1 \leq x \leq j} S_{p_{i_x}}$. Clearly, $c(S_p) = 1$ for every $p \in P$, thus, by Fact 3.8[b], $c(S_e) \leq l$, and by Fact 3.8[c], $c(V_3(e)) = c(S_e \setminus (V_1 \cup V_2 \cup \{\mathcal{L}(u)\})) \leq l + t$.

Now, naively, by Fact 3.8[b],

$$c(\mathcal{I}(e)) = c(V_1(e) \cup V_2(e) \cup V_3(e)) \leq 1 + (t-1) + (l+t) = 2t + l + 1.$$

We can give however a better bound as follows. Assume that $|V_2(e)| = x$. Then clearly, $c(V_2(e)) \leq x$ and $c(V_2 \setminus V_2(e)) \leq t - 1 - x$. We get,

$$\begin{aligned}
c(\mathcal{I}(e)) &= c(V_1(e) \cup V_2(e) \cup V_3(e)) \;=\; c(V_1(e) \cup V_2(e) \cup (S_e \setminus (V_1 \setminus V_1(e)) \setminus (V_2 \setminus V_2(e)))) \\
&\leq 2 + x + l + 2 + (t - 1 - x) \;=\; l + t + 3 \;.
\end{aligned}$$

In order to further improve the bound, we note that e cannot be in the same time both an arc from u to its successor in the tree $T_{p(u)}$ and the arc from u to its predecessor in the same tree. In the first case we get $S_e \setminus (V_1 \setminus V_1(e)) = S_e$, and $|S_e| \leq l - 1$. Thus,

$$c(\mathcal{I}(e)) \;\leq\; 2 + x + (l-1) + (t - 1 - x) \;=\; l + t \;.$$

In the second case, $c(V_1(e)) = 0$ and $c(V_1 \setminus V_1(e)) = 1$. We get in a similar way a bound of $l + t$ on the compactness of $c(\mathcal{I}(e))$. □

Proposition 3.9 *Let $G = (V, E, W)$ be a weighted graph and $R \in PIRS1(t, l)$. Then for every node $u \in V$,* $\text{Memory}_G(R, u) = O((t + l + \deg(u)) \log n)$.

Proof. Clearly, knowing n, $2 \log n$ bits are sufficient to code any interval. The number of intervals in u due to labels of priority 1 is at most $\deg(u)$. The number of intervals in u due to labels of priority 2 is at most $t - 1$ and the number of intervals due to labels of priority 3 is at most $l + t$ (see proof of Proposition 3.6). It follows that $2 \log n \cdot (\deg(u) + 2t + l - 1)$ bits are sufficient to code all the intervals on the outgoing arcs of u. Additional $O(\log n)$ bits are required to code u's label. □

Proposition 3.10 $\text{Memory}_G(R) = O((t+l)n \log n)$, *for every weighted graph G and $R \in PIRS1(t, l)$.*

Proof. A trivial upper bound of $O(n^2 \log n)$ on the total memory can be derived from Proposition 3.9, bounding $\deg(u)$, t and l by n. We note however that the additive term $\deg(u)$ in the bound of Proposition 3.9 is due to the intervals that label the downwards arcs of the tree $T'_{p(u)}$. Since exactly one such interval is required in every downwards arc of a tree T'_p, and since the trees $\{T'_p\}_{p \in P}$ are node-disjoint, the total number of such intervals in the graph is bounded by n. We get, $\text{Memory}_G(R) \leq 2n \log n + 2n \log n(2t + l - 1) + n \log n = O((t + l)n \log n)$. □

Proof of Theorem 3.1.
It remains to choose the size t of the balls so as to minimize the memory requirements and the compactness of the routing scheme R, and to determine the size of the cover P. Fix $t = \sqrt{n(1 + \ln n)}$, and construct the cover P by the greedy algorithm described in Section 2.3. By Lemma 2.2, where $|B| = |\mathcal{H}| = n$, $l = |P| \leq \sqrt{n(1 + \ln n)}$. We note that the randomized algorithm described in Section 2.3 could also be used in order to construct a cover P of (asymptotically) the same size. Let R be a routing scheme constructed by the PIR1 preprocessing algorithm, where the size of the balls $t = \sqrt{n(1 + \ln n)}$ and P is a greedy cover for the collection of t-balls (thus $|P| \leq \sqrt{n(1 + \ln n)}$). Then by Propositions 3.7, the compactness of R is $\text{Comp}_G(R) \leq 2\sqrt{n(1 + \ln n)}$. By Proposition 3.9 the number of bits required for R is $\text{Memory}_G(R, v) = O(n^{1/2} \log^{3/2} n + \deg(v) \log n)$, for every $v \in V$, and the total number of bits is $\text{Memory}_G(R) = O(n^{3/2} \log^{3/2} n)$ by Proposition 3.10. Clearly, the time complexity of the PIR1 preprocessing algorithm, where the cover P is constructed by the greedy algorithm, is polynomial in n. □

Theorem 3.11 *For every n-node unweighted graph $G = (V, E)$ with diameter D there exists a (direct) interval routing scheme $R = (\mathcal{L}, \mathcal{I})$ on G such that:*

- $\text{Memory}_G(R) = O(n^{3/2} \log^{3/2} n)$.
- $\text{Memory}_G(R, v) = O(n^{1/2} \log^{3/2} n + \deg(v) \log n)$, *for every $v \in V$.*
- $\text{Stretch}_G(R) \leq 5$.
- $\text{AvStr}_G(R) \leq 3$.
- $\text{Dilation}_G(R) \leq \lceil 1.5D \rceil$.
- $\text{Comp}_G(R) \leq 3\sqrt{n(1 + \ln n)}$.

Moreover, R can be constructed in time polynomial in n.

Strategy PIR2

Strategy PIR2 is very similar to PIR1, except that the preprocessing algorithm constructs a pivot collection covering simultaneously the collection of t-balls and the collection of $\lceil D/2 \rceil$-neighborhoods around the nodes of G. Formally, the only necessary change is to replace Step 1 of the preprocessing algorithm of PIR1 by the following step.

1. Let P_1 be a cover for the collection of t-balls of the nodes, $\{\mathcal{B}_v(t) \mid v \in V\}$.
   Let P_2 be a cover for the collection of the $\lceil D/2 \rceil$-neighborhoods of the nodes, $\{\Gamma_v(\lceil D/2 \rceil) \mid v \in V\}$.
   Set $P = P_1 \cup P_2$.

This change is responsible for the improvement in the dilation bound, as proved next. Clearly, since P is a cover for the collection of t-balls of the nodes, the PIR2 strategy is a special case of the PIR1 strategy, thus Observation 3.2 and all the propositions in the proof of Theorem 3.1 hold also when the PIR2 strategy is considered.

We now prove a better upper bound on the dilation of PIR2. We denote by $PIRS2(t,l)$ the set of routing schemes for every unweighted graph, generated by the PIR2 strategy, where t is the size of the balls used and l is the size $|P|$ of the cover P (for the t-balls and the $\lceil D/2 \rceil$-neighborhoods of the nodes).

Proposition 3.12 $\mathrm{Dilation}_G(R) \leq \lceil 1.5D \rceil$, *for every graph* $G = (V,E)$ *and* $R \in PIRS2(t,l)$.

Proof. Clearly, the pivot $p(v)$ of every node $v \in V$, satisfies both $v \in \mathcal{B}_v(t)$ and $v \in \Gamma_v(\lceil D/2 \rceil)$. It follows that the distance between every node v and its pivot $p(v)$ is at most $\lceil D/2 \rceil$. By Proposition 3.3, for every two nodes $u, v \in V$,

$$\mathrm{dist}(R,u,v) \;\leq\; \mathrm{dist}(u,p(v)) + \mathrm{dist}(p(v),v) \;\leq\; \lceil 1.5D \rceil \ .$$

□

Proof of Theorem 3.11.
It remains to determine the size of the balls t and to show how to construct "good" covers P_1 and P_2. First we determine t and construct the cover P_1 as in the PIR1 preprocessing algorithm. That is, $t = \sqrt{n(1+\ln n)}$, and P_1 is a greedy cover of size at most $\sqrt{n(1+\ln n)}$ (by Lemma 2.2). As in the PIR1 preprocessing algorithm, here too we can alternatively use the randomized algorithm in order to construct the cover P_1. Next, we use the same greedy algorithm in order to find the cover P_2 for the collection of $\lceil D/2 \rceil$-neighborhoods of the nodes. By Lemma 2.4, the size of such a cover satisfies $|P_2| \leq \sqrt{n(1+\ln n)}$. Let R be a routing scheme for a graph $G = (V,E)$ constructed by the PIR2 preprocessing algorithm with balls of size $t = \sqrt{n(1+\ln n)}$ and with P_1 and P_2 being the greedy covers constructed as above. Then, by Proposition 3.7, $\mathrm{Comp}_G(R) \leq t + |P_1| + |P_2| \leq 3\sqrt{n(1+\ln n)}$; $\mathrm{Memory}_G(R,v) = O(n^{1/2}\log^{3/2} n + \deg(v)\log n)$ by Proposition 3.9; $\mathrm{Memory}_G(R) = O(n^{3/2}\log^{3/2} n)$ by Proposition 3.10. □

In the rest of this section we present some additional analytic results on the efficiency of the PIR1 routing strategy. These results hold regardless of whether the cover P is constructed using the randomized or the greedy algorithm, and some of these results hold for *any* cover P of size $O(\sqrt{n \ln n})$. Moreover, most of the results apply also for the PIR2 strategy.

Let $PIR1[P]$ denote the PIR1 strategy with $t = \sqrt{n(1+\ln n)}$ and a cover P for the t-balls, and let $PIRS1[P]$ denote the set of routing schemes generated by $PIR1[P]$.

Let us first describe a property of the dilation incurred by the PIR1 strategy. For every graph G of diameter D, denote by $\overline{D}$ the average distance in G, i.e., $\overline{D} = \frac{1}{n(n-1)} \sum_{u \neq v} \mathrm{dist}_G(u,v)$.

Theorem 3.13 *Let G be a graph of diameter D and of average distance $\overline{D}$. Then, for every $R \in PIRS1$, and for all nodes $u, v \in V$, $\mathrm{AvStr}_G(R) \leq 2 + D - \overline{D} - \Theta(\sqrt{(\log n)/n})$.*

Proof. Consider a routing scheme $R \in PIRS1$ on G. Let U be the set of all pairs of nodes of G, $|U| = n(n-1)$. Partition U into $U = A \cup B$, where A is the set of the pairs at distance D and B is the set of nodes at distance $< D$.

$$\overline{D} \;\leq\; \frac{1}{|U|}\left(|A|D + |B|(D-1)\right) = D - \frac{|B|}{|U|}\,,$$

so $|B|/|U| \leq D - \overline{D}$. Also, by the dilation bound of Theorem 3.1, each pair $(u,v) \in A$ contributes at most 2 to the average stretch factor. As for pairs $(u,v) \in B$, note that for each such pair also $(v,u) \in B$, and by the analysis of Proposition 3.6, each such pair contributes for an average stretch factor of at most 3. Hence

$$\begin{aligned}
\mathrm{AvStr}_G(PIR1[P]) &\leq \frac{1}{|U|}\left(2|A| + 3|B|\right) \\
&\leq \frac{1}{|U|}\left(2|U| - 2|B| + 3|B|\right) \;=\; 2 + \frac{|B|}{|U|} \\
&\leq 2 + D - \overline{D}.
\end{aligned}$$

Also observe that at least $n(t-1)$ pairs of U route along shortest paths. Hence

$$\mathrm{AvStr}_G(PIR1[P]) \;\leq\; 2 + D - \overline{D} - \frac{n(t-1)}{n(n-1)}.$$

The proof follows from the fact that $t = \Theta(\sqrt{n \log n})$. □

It follows that the average stretch factor for graphs with $\overline{D} = D - o(1)$ is at most $2 + o(1)$. For instance, for the complete graph we get $\mathrm{AvStr}_{K_n}(PIR1[P]) \leq 2 - \Theta(\sqrt{(\log n)/n})$.

The next result shows that the bound of Theorems 3.1 and 3.11 on the average stretch factor is tight, as an average stretch factor of 3 can be reached asymptotically with the PIR1 strategy.

Theorem 3.14 *There exists an n-node graph G and an identity assignment* ID *for its nodes such that for every set of pivots P of G computed by the greedy cover algorithm (respectively, the randomized algorithm), $\mathrm{AvStr}_G(PIR1[P]) > 3 - o(1)$ (resp., with probability at least $1 - 1/\sqrt{n}$).*

Proof. Given two suitable integers $m, k \geq 1$, let us build an n-node graph $G = (V, E)$ defined as follows. Let $V = A \cup B$, where A is a clique of size mk, consisting of k subsets of size m, denoted $A_1, \ldots, A_k$. Likewise, B is composed of $2k$ cliques of size $t-1$, denoted $B_1, \ldots, B_k$ and $B'_1, \ldots, B'_k$. In addition, for every $i \in \{1, \ldots, k\}$, each node $u \in B_i$ is connected to all the nodes of A_i, and B'_i

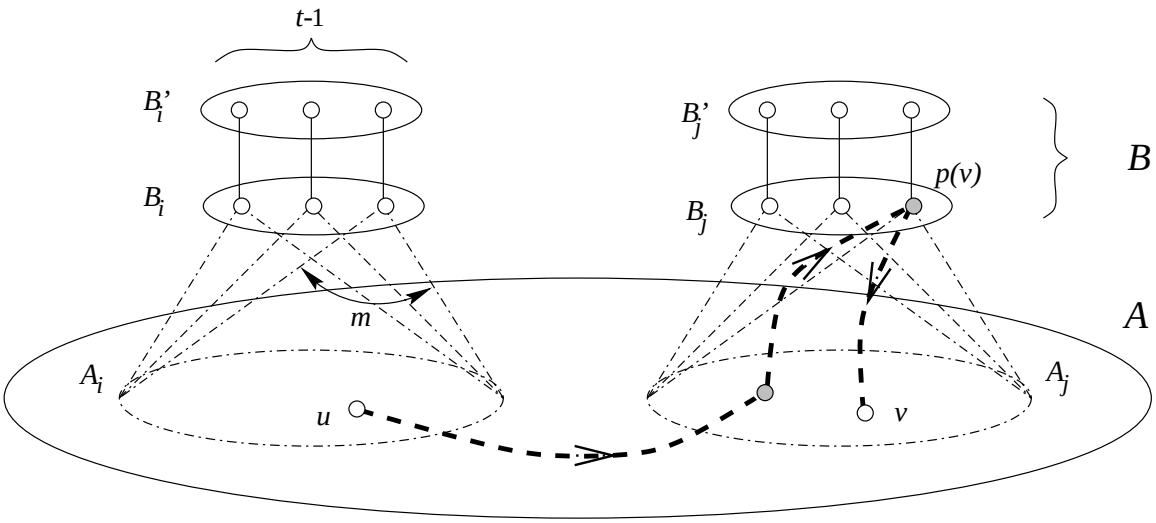

Figure 3: A example for proving an average stretch factor of $3 - o(1)$.

is connected to B_i by a matching. See Figure 3. Hence $n = |V| = mk + 2(t-1)k$ (where m and k must be selected so as to ensure that $t = \sqrt{n(1+\ln n)}$, as required from the size of the balls in the PIRS1[P] schemes.) See Figure 3.

Consider the following identity assignment ID for the nodes of G, selected in order to enforce a specific choice of the closest pivot. The identities for nodes of B are selected arbitrarily from the set $\{1,\ldots,|B|\}$, and the identities of the nodes of A are selected arbitrarily from the set $\{|B|+1,\ldots,n\}$.

Claim 3.15 *For every $u \in V$, $\deg(u) \geq t-1$. Moreover, if $u \in A_i$, then $\mathcal{B}_u = \{u\} \cup B_i$, and if $u \in B_i \cup B'_i$, then $\mathcal{B}_u \subseteq B_i \cup B'_i$.*

Proof. For every $u \in A$ and $v \in B$, $\deg(u) = |A| - 1 + t$, and $\deg(v) \geq t-1$. Hence, for every $u \in V$, $\mathcal{B}_u$ consists of u and the $t-1$ lowest-ID neighbors of u. Let us consider some $i \in \{1,\ldots,k\}$. For each $u \in A_i$, the $t-1$ lowest-ID neighbors of u constitute all the nodes of B_i, so $\mathcal{B}_u \setminus \{u\} = B_i$. For each $u \in B_i \cup B'_i$, u has exactly $t-1$ neighbors in $B_i \cup B'_i$, because the identities of the nodes of B are lower than those of the nodes of A. Hence $\mathcal{B}_u \subseteq B_i \cup B'_i$. □

Let P be the set of pivots. By Claim 3.15, for every $i \in \{1,\ldots,k\}$, and for every $u \in A_i \cup B_i \cup B'_i$, $\mathcal{B}_u \subseteq A_i \cup B_i \cup B'_i$. Hence whatever P is, there exists at least one $p \in P$ such that $p \in A_i \cup B_i \cup B'_i$. Let $\mathcal{E}_{OK}$ denote the event that for every $i \in \{1,\ldots,k\}$, there is a pivot $p \in P$ such that $p \in B_i$. Then,

Claim 3.16 (1) *If P is computed by the greedy cover algorithm, then $\mathbb{P}(\mathcal{E}_{OK}) = 1$;*
(2) *if P is computed by the randomized algorithm, then $\mathbb{P}(\mathcal{E}_{OK}) > 1 - 1/\sqrt{n}$.*

Proof. (1) Let p be the first pivot selected in $A_i \cup B_i \cup B'_i$. Note that if $p \in A_i$ then p covers only the ball around p itself. Also, if $p \in B'_i$ then p covers at most $\deg(p) = |B'_i| + 1 = t$ balls. On the other hand, if $p \in B_i$ then p covers at least $|A_i| + |B_i| + 1 = m + t$. It follows that for each i, the greedy algorithm will necessarily choose a pivot in B_i.

(2) Recall that the randomized algorithm randomly picks each node of V into the cover with probability $p = (c \ln n)/t$ for constant $c > 1$. For every $i \in \{1, \ldots, k\}$, let $X_i = |B_i \cap P|$ denote the random variable counting the number of pivots in B_i. For every i,

$$\mathbb{P}(X_i = 0) \;=\; (1-p)^{|B_i|} \;=\; (1-p)^{t-1},$$

so $\mathbb{P}(\exists i, X_i = 0) \leq k(1-p)^{t-1}$. It follows that

$$\mathbb{P}(\mathcal{E}_{OK}) = \mathbb{P}(\forall i, X_i > 0) \;\geq\; 1 - k(1-p)^{t-1}\ .$$

Note that $k = n/(m + 2(t-1)) \leq \sqrt{n}/2$. In order to bound the probability we use the following known fact.

Fact 3.17 *For all $0 < \alpha \leq 1/2$ and $\beta > 0$, $(1-\alpha)^\beta < e^{-\alpha\beta}$.*

Hence

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{OK}) \;&\geq\; 1 - k\,e^{-p(t-1)} \;\geq\; 1 - \frac{\sqrt{n}}{2}\,e^{-c(1-1/t)\ln n} \\ &>\; 1 - \frac{1}{2}\,n^{-(1/2-1/t)} \;>\; 1 - n^{-1/2}.\end{aligned}$$

□

By Claim 3.16, every clique B_i contains at least one pivot. Consider a routing scheme $R \in PIRS1[P]$ on G. We want to lower bound $\mathrm{AvStr}_G(R)$. Because $\mathrm{dist}(u,v) = 1$ for all $u, v \in A$, we have

$$\mathrm{AvStr}_G(R) \;>\; \frac{1}{n(n-1)} \sum_{u \neq v \in A} \mathrm{dist}_G(R, u, v).$$

For every node $u \in V$, let $p(u) \in P$ denote the closest pivot of u, according to the distance definition. For every $p \in P$, let T_p be the minimum spanning tree rooted in p, and used by R. Consider some pair $u, v \in A$, $u \neq v$. Note that if $\mathrm{dist}_G(u, p(v)) = 2$ and $\{u, v\}$ is not an edge of $T_{p(v)}$ then $\mathrm{dist}_G(R, u, v) = 3$. Given u, the number of nodes $v \in A$ such that $\{u, v\}$ is an edge of $T_{p(v)}$ is $k-1$, since all the nodes of a same component A_i share a unique pivot in B_i, hence use the same tree to route inside A_i. Moreover, $\mathrm{dist}_G(u, p(v)) = 1$ if and only if u and v are both in the same A_i. In total, the number of nodes v such that $\mathrm{dist}_G(R, u, v) \neq 3$ is at most $m + k - 1$, so

$$\sum_{v \in A \setminus \{u\}} \mathrm{dist}_G(R, u, v) \;\geq\; 3(|A| - 1 - (m + k - 1)) \;=\; 3mk - 3(m + k).$$

Summing over all $u \in A$, it follows that

$$\sum_{u \neq v \in A} \mathrm{dist}_G(R, u, v) \;\geq\; 3(mk)^2 - 3mk(m + k).$$

Letting $k = \sqrt{n}/\ln n$, we have $mk = n - 2(t-1)k = n - cn/\sqrt{\ln n}$ and $m + k < c'\sqrt{n}\ln n$, for large n and for suitable constants $c, c' > 0$. Hence

$$\begin{aligned}\mathrm{AvStr}_G(R) \;&>\; \frac{3(n - cn/\sqrt{\ln n})^2 - 3(n - cn/\sqrt{\ln n})c'\sqrt{n}\ln n}{n(n-1)} \\ &>\; \frac{3n^2 - 6cn/\sqrt{\ln n} - 3c'n^{3/2}\sqrt{\ln n}}{n^2} \;>\; 3 - \frac{6c}{n\sqrt{\ln n}} - \frac{3c'\sqrt{\ln n}}{\sqrt{n}} \\ &>\; 3 - O\left(\sqrt{\frac{\log n}{n}}\right)\ ,\end{aligned}$$

completing the proof. □

Next we present a lower bound for the average stretch factor which is independent on the choice of the pivots and the node identities. According to Theorem 3.13, this bound is tight.

Theorem 3.18 *For every set of pivots P, and for every $R \in PIRS1[P]$, $\mathrm{AvStr}_{K_n}(PIR1[P]) = 2 - \Theta\left(\sqrt{(\log n)/n}\right)$.*

Proof. Let P, P' be two sets of pivots on K_n, and let $R \in PIRS1[P]$ and $R' \in PIRS1[P']$. Observe that $\sum_{u \neq v} \mathrm{dist}_{K_n}(R, u, v) = \sum_{u \neq v} \mathrm{dist}_{K_n}(R', u, v)$. Indeed, whatever the set of pivots is, the route from a given node u to v would be of length 1 if $v \in \mathcal{B}_u$, and of length 2 otherwise, and $|\mathcal{B}_u|$ is independent of the set of pivots. Consequently,

$$\begin{aligned} \mathrm{AvStr}_{K_n}(PIR1[P]) &= \frac{1}{n(n-1)} \sum_{u \neq v} \mathrm{dist}_{K_n}(R, u, v) = \frac{n\left(\sqrt{n \ln n} - 1 + 2(n - \sqrt{n \ln n})\right)}{n(n-1)} \\ &= 2 - \Theta\left(\sqrt{\frac{\log n}{n}}\right) . \end{aligned}$$

□

Finally, we show that the behavior of the average stretch factor of the PIR1 strategy is asymptotically 1.875 for almost all unweighted graphs. Recall that $\mathcal{G}_{n,p}$ denotes the class of n-node random graphs, where p represents the probability to have an edge between any two nodes. Let $\Gamma_v = \Gamma_v(1)$. For random graphs selected from $\mathcal{G}_{n,1/2}$ we have the following bounds. (These simple variants of the bounds suits our needs; see [5] for sharper statements.)

Lemma 3.19 *With probability at least $1 - 1/n^3$, a random unweighted graph $G \in \mathcal{G}_{n,1/2}$ enjoys the following properties:*

1. *G is of diameter 2,*
2. *for every node $v \in V$, $\frac{n}{2} - 3\sqrt{n \ln n} \leq |\Gamma_v| \leq \frac{n}{2} + 3\sqrt{n \ln n}$,*
3. *for every two nodes $v, w \in V$, $\frac{n}{4} - 3\sqrt{n \ln n} \leq |\Gamma_v \cap \Gamma_w| \leq \frac{n}{4} + 3\sqrt{n \ln n}$.*

Proof. We prove only the third claim. Fix a pair of nodes $v, w \in V$. For every other node x, let $J(x)$ denote the event that x is adjacent jointly to both v and w,

$$J(x) = \text{“}x \in \Gamma_v \cap \Gamma_w\text{”}.$$

Note that this event occurs with probability $1/4$, and the events $J(x)$ and $J(x')$ are independent for every $x, x' \in V$. Let Z denote a random variable counting $|\Gamma_v \cap \Gamma_w|$. Then $Z = \sum_{x \neq v,w} z_x$, where z_x is the characteristic random variable of the event $J(x)$. Hence Z is the sum of $n-2$ mutually independent Bernoulli variables, and its expected value is $\mathbb{E}(Z) = (n-2)/4$, and hence applying Chernoff's bound (cf. [1]) we get,

$$\begin{aligned} \mathbb{P}\left(Z > \frac{n-2}{4} + 3\sqrt{n \ln n}\right) &< \mathbb{P}\left(Z > \left(1 + 7\sqrt{\frac{\ln n}{n}}\right) \mathbb{E}(Z)\right) \\ &< \exp\left(-\frac{49 \ln n \, \mathbb{E}(Z)}{3n}\right) < \frac{1}{n^3} \end{aligned}$$

for $n \geq 8$, and likewise

$$\mathbb{P}\left(Z < \frac{n-2}{4} - 3\sqrt{n \ln n}\right) \;<\; \frac{1}{n^3}\,,$$

and the claim follows. □

Theorem 3.20 *For almost every n-node graph G,*

$$1.875 - O\left(\sqrt{\frac{\log n}{n}}\right) \;\leq\; \mathrm{AvStr}_G(PIR1[P]) \;\leq\; 1.875 + O\left(\sqrt{\frac{\log n}{n}}\right).$$

Proof. To prove the theorem, it suffices to consider a graph $G \in \mathcal{G}_{n,1/2}$ and to show that with probability $1 - o(1)$, $\mathrm{AvStr}_G(R) = 1.875 \pm O(\sqrt{(\log n)/n})$, for $R \in PIRS1$. We will prove this by establishing the claim of the theorem for every random graph satisfying the properties asserted in Lemma 3.19, namely, for all but a $1/n^3$ fraction of the class of n-node graphs.

Suppose that the graph G at hand satisfies the properties of Lemma 3.19. Fix some node v in G. Note that under the above assumption, the nodes of $V \setminus \{v\}$ are partitioned into two sets, $A = \Gamma_v \setminus \{v\}$ and $B = V \setminus \Gamma_v$, where every node in B is adjacent to some node of A. Also note that by choice of the pivots, $p(v)$ is necessarily in A.

Let us calculate the average stretch over all the routes leading to v. For every node $x \in V$, let $s(x) = \mathrm{dist}_G(R, x, v) \;/\; \mathrm{dist}_G(x, v)$ denote the stretch of the route from x to v. By case analysis we show that $s(x)$ can assume only four possible values, namely, 1, 1.5, 2 or 3.

Optimal stretch ($s(x) = 1$) is achieved for nodes x in two sets, namely, $\mathcal{B}_v$ (where the shortest path and the actual route are both of length 1) and $B \cap \Gamma_{p(v)}$ (where the length of both paths is 2). By construction, $|\mathcal{B}_v| = \sqrt{n \ln n}$. By the third claim in Lemma 3.19,

$$\frac{n}{4} - 6\sqrt{n \ln n} \;\leq\; |B \cap \Gamma_{p(v)}| \;\leq\; \frac{n}{4} + 6\sqrt{n \ln n}.$$

The only nodes x for which $s(x) = 1.5$ are those nodes in $B \setminus \Gamma_{p(v)}$. By the second and third claims in Lemma 3.19 again,

$$\frac{n}{4} - 9\sqrt{n \ln n} \;\leq\; |B \setminus \Gamma_{p(v)}| \;\leq\; \frac{n}{4} + 9\sqrt{n \ln n}.$$

The nodes x for which $s(x) = 2$ are those in $(A \setminus \mathcal{B}_v) \cap \Gamma_{p(v)}$. By the third claim in Lemma 3.19 again,

$$\frac{n}{4} - 3\sqrt{n \ln n} \;\leq\; |(A \setminus \mathcal{B}_v) \cap \Gamma_{p(v)}| \;\leq\; \frac{n}{4} + 3\sqrt{n \ln n}.$$

Finally, the nodes x for which $s(x) = 3$ are those in $A \setminus (\mathcal{B}_v \cup \Gamma_{p(v)})$. By the third claim in Lemma 3.19 again,

$$\frac{n}{4} - 6\sqrt{n \ln n} \;\leq\; |A \setminus (\mathcal{B}_v \cup \Gamma_{p(v)})| \;\leq\; \frac{n}{4} + 6\sqrt{n \ln n}.$$

It follows that the average stretch from some $x \in V,\ x \neq v$ to v is at most

$$\begin{aligned}
&\frac{1}{n-1}\Big(1 \cdot (|\mathcal{B}_v| + |B \cap \Gamma_{p(v)}|) + 1.5 \cdot |B \setminus \Gamma_{p(v)}| + 2 \cdot |(A \setminus \mathcal{B}_v) \cap \Gamma_{p(v)}| + 3 \cdot |A \setminus (\mathcal{B}_v \cup \Gamma_{p(v)})|\Big) \\
\leq\; &\frac{1}{n-1}\left((1 + 1.5 + 2 + 3) \cdot \frac{n}{4} + O(\sqrt{n \ln n})\right) \;\leq\; 1.875 + O\left(\sqrt{\frac{\log n}{n}}\right)
\end{aligned}$$

and at least $1.875 - O(\sqrt{(\log n)/n})$, and hence the same bounds apply to the average over all destinations v, yielding

$$1.875 - O\left(\sqrt{\frac{\log n}{n}}\right) \;\leq\; \mathrm{AvStr}_G(R) \;\leq\; 1.875 + O\left(\sqrt{\frac{\log n}{n}}\right) .$$

□

4 Lower Bounds on Loop-Free Routing Schemes

It is relevant to ask if an amount of $O(\sqrt{n \log n})$ bits of memory in each node (on average) for some stretch factor s is the best we can hope for with loop-free routing schemes. It turns out that a lower bound of $\Omega(\sqrt{n})$ memory bits holds for any stretch factor. To prove it, we will take advantage of the fact that loop-free schemes route optimally on trees.

Lemma 4.1 *Let T be a tree, let $s \geq 1$ and let R be a loop-free routing scheme on T. Then* $\mathrm{Stretch}_T(R) \leq s$ *if and only if* $\mathrm{Stretch}_T(R) = 1$, *i.e., R is a shortest path routing scheme.*

Proof. A loop-free routing scheme routes along loop-free paths, which happen to be shortest paths on a tree. □

Applying Theorem 3.1, since the schemes of PIRS1 are loop-free and of bounded stretch factor, it follows that they are shortest path routing schemes for trees.

Theorem 4.2 *There exists an n-node graph G such that for every s, and every loop-free routing scheme R on G of stretch factor s that uses names taken from $\{1, \ldots, n\}$, there is a node x such that* $\mathrm{Memory}_G(R, x) \geq c\sqrt{n} - O(\log n)$, *where $c \approx 3.71$.*

Proof. Consider a tree T of n nodes and root x, and let $d = \deg(x)$. Let R be any loop-free routing scheme on T, and let $l_x \in \{1, \ldots, n\}$ be the label of x under R. By Lemma 4.1, the local routing function R_x routes along shortest paths. For every $i \in \{1, \ldots, d\}$, denote by n_i the number of nodes in the subtree of T rooted at the ith child of x. Without loss of generality assume $n_1 \leq \ldots \leq n_d$. Observe that knowing n, d and l_x suffices for computing the sequence $S = (n_1, \ldots, n_d)$ with R_x. Indeed, it is sufficient to enumerate the set of labels distinct of l_x, and compute the answer of R_x for each of them. The port number must be the same for all the nodes that belong to the same subtree. We remark that $\sum_{i=1}^{d} n_i = n - 1$.

Let $K(X)$ be the minimum number of bits required in order to code an object X. By the above discussion, $K(S) \leq K(R_x) + O(\log n)$, because S can be described knowing R_x, n, d and l_x (the last three integers can be stored using $O(\log n)$ bits). A *partition* of n is an integer sequence $U = (u_1, \ldots, u_k)$ such that $1 \leq u_1 \leq \ldots \leq u_k$ and $\sum_{i=1}^{k} u_i = n$. Note that for each partition U of $n - 1$ there exists a distinct n-node tree T_U, with root u of degree $\deg(u) = k$, whose children induce subtrees of size exactly u_i, for $i \in \{1, \ldots, k\}$. Hence for every loop-free routing R' on T_U, $K(R'_u) \geq \log_2(p_{n-1}) - O(\log n)$, where p_n denotes the number of partitions of n.

We now rely on the following well-known formula due to Harder and Ramanujan [21, Equation (4.2.7) page 44].

$$\ln p_n \sim c'\sqrt{n}, \quad \text{for } c' = \pi\sqrt{2/3}.$$

Thus for sufficiently large n, $K(R'_u) \geq c\sqrt{n} - O(\log n)$, where $c \approx 3.71$, completing the proof. □

Note that an upper bound of $c\sqrt{n}$ bits per node for trees has been established in [16, page 6] for shortest path loop-free routing scheme.

In contrast to the theorem, a universal strategy that uses $O(k\, n^{1/k} \log n)$ bits of memory per router and yields stretch factor $O(k^2\, 9^k)$, for every $k \geq 1$, is presented in [2]. Moreover, the generated schemes are name-independent. It follows that for any graph satisfying Theorem 4.2 there exists a routing scheme using names taken from $\{1, \ldots, n\}$ that guarantees $o(\sqrt{n})$ bits per router and a constant stretch factor (e.g., choose $k = 3$). However, that scheme cannot be loop-free, as the theorem indicates that a loop-free routing scheme with such an efficiency-space tradeoff does not exist for such graphs.

As done previously in Theorem 3.14, one may ask whether the stretch factor guaranteed by the PIR1 strategy is the best that can be achieved with memory requirements $O(n^{3/2} \log^{3/2} n)$. Actually, the smallest possible stretch factor one can hope for when considering a routing scheme with a memory requirements $o(n^2)$ is $s = 3$. Indeed, in [17] it is shown that there exist some graphs on which every routing scheme of stretch factor $s < 3$ requires at least $\Omega(n^2)$ bits.

5 Comparison with General Routing Schemes

This section compares the PIR1 strategy with other general strategies. Let HCP_k be the routing strategy developed in [3] called *Hierarchical Covering Pivot*.

Theorem 5.1 (Awerbuch et al. [2]) *For every n-node weighted graph G, and for every integer $k \geq 1$, and there exists $R \in \mathrm{HCP}_k$ on G such that* $\mathrm{Memory}_G(R) = O(k\, n^{1+1/k} \log^2 n)$, *and* $\mathrm{Stretch}_G(R) \leq 2^k - 1$. *Moreover R is constructible in polynomial time.*

The HCP_k strategy first selects a hierarchy of sets of pivots, $P_{k-1} \subset \ldots \subset P_1 \subset P_0 = V$, each one of size about $n^{1-i/k}$. Each pivot $v \in P_i$ is able to route on a shortest path to vertices in its neighborhood of size roughly $n^{(i+1)/k}$. When a vertex u wants to send a message to v, the message is transferred trough the chain of pivots associated with u from successively higher levels, until it reaches a pivot able to route directly to v. For concreteness, let us consider $k = 2$. Then the scheme generated by HCP_2 routes messages trough one level of pivots, P_1. ($P_0 = V$ is not considered here). Hence the route from u to v follows, successively, u, $p(u)$ (the pivot closest to u in P_1), then v. In comparison, the route generated by the PIR1 or PIR2 strategy would be u, $p(v)$, and v, allowing a loop-free routing, and bounding the compactness of its interval routing implementation.

The HCP_k strategy is name-independent. However, in the complexity memory analysis, it is assumed that the length of node identities is $O(\log n)$ bits (the pivots need to keep some translation tables storing the identities of some nodes). Note also that the space complexity analysis of the schemes can be slightly improved, to show that $\mathrm{Memory}_G(R) = O(k\, n^{1+1/k} \log^{2-1/k} n)$, for every $R \in \mathrm{HCP}_k$.

As done previously, given a set of pivots P, we denote by $\mathrm{HCP}_k[P]$ the subset of HCP_k that consists of only the covering pivot routing schemes having a pivot set P. It is easy to see that, for a given pivot set P of a graph G, $\mathrm{AvStr}_G(\mathrm{HCP}_{k+1}[P]) \geq \mathrm{AvStr}_G(\mathrm{HCP}_k[P])$. Moreover $\mathrm{Memory}(\mathrm{HCP}_1, G) = O(n^2 \log n)$ which is higher than the memory requirements of the PIR strategy. This is why, in the remainder of the paper, we are interested in the schemes of the strategy HCP_2. In [2], the set of schemes HCP_2 is called *Covering Pivot*. For simplicity we denote $\mathrm{CP} = \mathrm{HCP}_2$.

We first show that in terms of their average stretch factor, the schemes of PIR1 are asymptotically never worse than those of CP.

Theorem 5.2 *Let G be an n-node graph, let P be an arbitrary set of pivots of G, $R_{PIR1} \in PIRS1[P]$, and $R_{CP} \in \mathrm{CP}[P]$. Then* $\mathrm{AvStr}_G(R_{PIR1}) \leq \mathrm{AvStr}_G(R_{CP}) + 2\sqrt{(1+\ln n)/(n-1)}$.

Proof. Consider a set of pivots P for both schemes, and any two nodes u, v at distance d. Three cases may occur:
(1) $\mathcal{B}_u \cap \mathcal{B}_v = \emptyset$;
(2) $v \in \mathcal{B}_u$ and $u \in \mathcal{B}_v$; and
(3) $v \in \mathcal{B}_u$ and $u \notin \mathcal{B}_v$, or the reverse.

For the cases (1) and (2)

$$\mathrm{dist}(R_{PIR1}, u, v) + \mathrm{dist}(R_{PIR1}, v, u) \;=\; \mathrm{dist}(R_{CP}, u, v) + \mathrm{dist}(R_{CP}, v, u).$$

In case (3) $\mathrm{dist}(R_{PIR1}, u, v) = \mathrm{dist}(R_{CP}, u, v) = d$, whereas $\mathrm{dist}(R_{PIR1}, v, u) \leq 5d$ but $\mathrm{dist}(R_{CP}, v, u) \leq 3d$. So, R_{PIR1} may induce longer routing paths for the case (3). However, given a node u, the number of pairs (u, v) such that $v \in \mathcal{B}_u$ and $u \notin \mathcal{B}_v$ is bounded by $|\mathcal{B}_u| = \sqrt{n(1+\ln n)}$. Therefore the total number of such pairs in the whole graph is at most $n\sqrt{n(1+\ln n)}$, which is only a $\sqrt{(1+\ln n)/(n-1)}$ fraction of all the pairs. For each such pair the stretch can increase by at most 2, so $\mathrm{AvStr}_G(R_{PIR1}) - \mathrm{AvStr}_G(R_{CP}) \leq 2\sqrt{(1+\ln n)/(n-1)}$. □

As a corollary of Theorem 5.2, we have that $\mathrm{AvStr}_{K_n}(\mathrm{CP}) = 2 - O(\sqrt{(\log n)/n})$, and also that there exists some worst-case graph G with specific node labeling such that $\mathrm{AvStr}_G(\mathrm{CP}[P]) = 3 - o(1)$, for every pivot set P chosen randomly or computed with the greedy algorithm. Similarly, almost every graph G satisfies $\mathrm{AvStr}_G(\mathrm{CP}) \geq 1.875 - O(\sqrt{(\log n)/n})$.

Next we illustrate the advantage of the PIR1 strategy over the CP strategy w.r.t. average stretch factor.

Theorem 5.3 *Let $T = K_{1,n-1}$, and let r be its root. If* $\mathrm{ID}(r)$ *is not the lowest* ID, *then for every set of pivots P of T computed by the random or the greedy cover algorithm,* $\mathrm{AvStr}_T(\mathrm{CP}[P]) > 2 - o(1)$ *(w.h.p.), whereas* $\mathrm{AvStr}_T(PIR1[P]) = 1$.

Proof. Clearly $\mathrm{AvStr}_T(R) = 1$ for every $R \in PIRS1$, as PIR1 always routes optimally on trees. Suppose that the ID's assigned to the nodes of T are 1 through n, and that $\mathrm{ID}(r) \neq 1$. Let $\ell = \left\lfloor \sqrt{n(1+\ln n)} \right\rfloor - 2$. Note that in this case, the ball $\mathcal{B}_v$ built by the scheme around each leaf v

consists of the nodes $\{1,\ldots,x\} \cup \{r, v\}$, where

$$x = \begin{cases} \ell, & v, r > \ell, \\ \ell + 1, & v > \ell \text{ and } r \leq \ell \text{ or vice versa}, \\ \ell + 2, & v, r \leq \ell, \end{cases}$$

and similarly for the ball $\mathcal{B}_r$ around the root.

Let us first consider the greedy pivot selection algorithm. Assume that in the greedy algorithm, ties (between two candidates to join the pivot set) are broken in favor of the smallest-ID candidate. Note that when selecting the first pivot, all nodes in $\{1,\ldots,\ell\} \cup \{r\}$ are equal candidates, and hence the algorithm will pick 1 as the first (and only) pivot. Hence the resulting routing scheme $R \in \mathrm{CP}[P]$ will route messages between all but $O(n^{3/2}\sqrt{\log n})$ pairs of nodes (specifically, between any $v, w \in V$ such that $v, w > \ell$ and $v, w \neq r$) in 4 steps rather than 2. Thus $\mathrm{AvStr}_T(\mathrm{CP}[P]) > 2 - O(\sqrt{(\log n)/n})$.

As for the randomized cover algorithm, we are guaranteed that with high probability all balls are covered by at least one pivot, but the probability that r was chosen as pivot is roughly $\ln n \,/\, \sqrt{n}$, and moreover, the probability that r was the *smallest* ID pivot chosen is even smaller. Hence with probability at least $1 - \ln n \,/\, \sqrt{n}$, the analysis of the greedy strategy applies to the randomized strategy as well, and $\mathrm{AvStr}_T(\mathrm{CP}[P]) > 2 - O(\sqrt{(\log n)/n})$. □

6 Conclusion

In high-speed networks the bottleneck is the delay in the nodes and not the propagation time of the messages. As high-speed networks gain popularity, it is essential to design direct routing schemes, which could be implemented in hardware and at the same time, when dealing with large networks, it is important to decrease the amount of memory kept in the nodes for routing purposes. In this work we present for the first time a routing strategy that generates for every (weighted) network, a direct routing scheme with small stretch factor and modest memory requirements. On the other hand, we show that there exist n-node graphs on which any loop-free routing scheme uses $\Omega(\sqrt{n})$ bits of memory in at least one node, thus our routing strategy cannot be generalized to obtain a family of routing schemes with different values of stretch factor and memory such as the hierarchical routing schemes (e.g., [2]).

Following are some directions for further research.

- Is it possible to improve the worst case stretch factor using the same amount of memory? By [17], it cannot be improved below 3.
- What is the best trade-off between the memory requirements and the (average) stretch factor for 1-phase routing schemes?

Acknowledgments: We thank Shlomo Moran, Shmuel Zaks, and Jean-Michel Couvreur for helpful discussions.

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, John Wiley & Sons, 1992.

[2] B. Awerbuch, A. Bar-Noy, N. Linial, and D. Peleg, *Compact distributed data structures for adaptive routing*, in 21^{st} Symposium on Theory of Computing (STOC), vol. 2, 1989, pp. 230–240.

[3] ——, *Improved routing strategies with succint tables*, Journal of Algorithms, 11 (1990), pp. 307–341.

[4] B. Awerbuch and D. Peleg, *Sparse partitions*, in 31^{th} Symposium on Foundations of Computer Science (FOCS), 1990, pp. 503–513.

[5] B. Bollobás, *Random Graphs*, Academic Press, New York, 1975.

[6] T. Eilam, S. Moran, and S. Zaks, *A lower bound for linear interval routing*, in 10^{th} International Workshop on Distributed Algorithms (WDAG), vol. 1151 of Lecture Notes in Computer Science, Springer-Verlag, 1996.

[7] M. Flammini, G. Gambosi, U. Nanni, and R. B. Tan, *Multi-dimensional interval routing schemes*, in 9^{th} International Workshop on Distributed Algorithms (WDAG), vol. 972 of Lecture Notes in Computer Science, Springer-Verlag, 1995. To appear in TCSA.

[8] M. Flammini and E. Nardelli, *On the path length in interval routing schemes.* Manuscript, 1996.

[9] P. Fraigniaud and C. Gavoille, *Memory requirement for universal routing schemes*, in 14^{th} Annual ACM Symposium on Principles of Distributed Computing (PODC), Aug. 1995, pp. 223–230.

[10] G. N. Frederickson and R. Janardan, *Seperator-based strategies for efficient message routing*, in 27^{th} Symposium on Foundations of Computer Science (FOCS), May 1986, pp. 428–437.

[11] ——, *Designing networks with compact routing tables*, Algorithmica, 3 (1988), pp. 171–190.

[12] ——, *Efficient message routing in planar networks*, SIAM Journal on Computing, 18 (1989), pp. 843–857.

[13] ——, *Space-efficient message routing in c-decomposable networks*, SIAM Journal on Computing, 19 (1990), pp. 164–181.

[14] G. Gambosi and P. Vocca, *Topological routing schemes*, in 10^{th} International Workshop on Distributed Algorithms (WDAG), vol. 1151 of Lecture Notes in Computer Science, Springer-Verlag, Oct. 1996.

[15] C. Gavoille, *On the dilation of interval routing*, in 22^{nd} International Symposium on Mathematical Foundations of Computer Science (MFCS), vol. 1300 of Lecture Notes in Computer Science, Springer-Verlag, Aug. 1997, pp. 259–268.

[16] ——, *A survey on interval routing scheme*, Research Report RR-1182-97, LaBRI, University of Bordeaux, 351, cours de la Libération, 33405 Talence Cedex, France, Oct. 1997. Submitted for publication.

[17] C. Gavoille and M. Gengler, *Space-efficiency of routing schemes of stretch factor three*, in 4^{th} International Colloquium on Structural Information & Communication Complexity (SIROCCO), July 1997.

[18] C. Gavoille and E. Guévremont, *Worst case bounds for shortest path interval routing*, Research Report 95-02, LIP, École Normale Supérieure de Lyon, 69364 Lyon Cedex 07, France, Jan. 1995. To appear in Journal of Algorithms.

[19] C. Gavoille and D. Peleg, *The compactness of interval routing*, Research Report RR-1176-97, LaBRI, University of Bordeaux, 351, cours de la Libération, 33405 Talence Cedex, France, Sept. 1997. Submitted for publication.

[20] C. Gavoille and S. Pérennès, *Memory requirement for routing in distributed networks*, in 15^{th} Annual ACM Symposium on Principles of Distributed Computing (PODC), May 1996, pp. 125–133.

[21] M. J. Hall, *Combinatorial Theory (second edition)*, Wiley-Interscience Publication, 1986.

[22] Inmos, *The T9000 Transputer Products Overview Manual*, 1991.

[23] R. Královič, P. Ružička, and D. Štefankovič, *The complexity of shortest path and dilation bounded interval routing*, in 3^{rd} International Euro-Par Conference, vol. 1300 of Lectures Notes in Computer Science, Springer-Verlag, Aug. 1997, pp. 258–265.

[24] L. Lovász, *On the ratio of optimal integral and fractional covers*, Discrete Mathematics, 13 (1975), pp. 383–390.

[25] D. Peleg, *An overview of locality-sensitive distributed computing*, Unpublished Monograph, The Weizmann Institute, Rehovot, Israel, 1997.

[26] D. Peleg and E. Upfal, *A trade-off between space and efficiency for routing tables*, Journal of the ACM, 36 (1989), pp. 510–530.

[27] N. Santoro and R. Khatib, *Labelling and implicit routing in networks*, The Computer Journal, 28 (1985), pp. 5–8.

[28] S. S. H. Tse and F. C. M. Lau, *An optimal lower bound for interval routing in general networks*, in 4^{th} International Colloquium on Structural Information & Communication Complexity (SIROCCO), July 1997.

[29] J. van Leeuwen and R. B. Tan, *Routing with compact routing table*, in The Book of L, Springer-Verlag, 1986, pp. 259–273.