Title

Category-based routing in social networks: Membership dimension and the small-world phenomenon

Authors

Eppstein, David
Goodrich, Michael T
Löffler, Maarten
et al.

Publication Date

2013-11-01

DOI

10.1016/j.tcs.2013.04.027

Peer reviewed

Category-Based Routing in Social Networks: Membership Dimension and the Small-World Phenomenon

David Eppstein* — eppstein@uci.edu
Michael T. Goodrich* — goodrich@uci.edu
Maarten Löffler* — mloffler@uci.edu
Darren Strash* — dstrash@uci.edu
Lowell Trott* — ltrott@uci.edu

*Deptartment of Computer Science, University of California, Irvine, USA

***Abstract*—A classic experiment by Milgram shows that individuals can route messages along short paths in social networks, given only simple categorical information about recipients (such as "he is a prominent lawyer in Boston" or "she is a Freshman sociology major at Harvard"). That is, these networks have very short paths between pairs of nodes (the so-called *small-world phenomenon*); moreover, participants are able to route messages along these paths even though each person is only aware of a small part of the network topology. Some sociologists conjecture that participants in such scenarios use a *greedy routing* strategy in which they forward messages to acquaintances that have more categories in common with the recipient than they do, and similar strategies have recently been proposed for routing messages in dynamic ad-hoc networks of mobile devices. In this paper, we introduce a network property called *membership dimension*, which characterizes the cognitive load required to maintain relationships between participants and categories in a social network. We show that any connected network has a system of categories that will support greedy routing, but that these categories can be made to have small membership dimension if and only if the underlying network exhibits the small-world phenomenon.**

I. Introduction

In a pioneering experiment in the 1960's, Stanley Milgram and colleagues [14], [20], [24] studied message routing in real-world social networks. 296 randomly chosen people in Nebraska and Kansas were asked to route a letter to a lawyer in Boston by forwarding it to an acquaintance, who would receive the same instructions. Messages that reached their destinations typically passed between at most six acquaintances;[1] The observation that acquaintance graphs have such short paths has come to be called the *small-world phenomenon* [11], [25].

Even more surprising than the existence of these short paths is that participants are able to efficiently route messages using only local information and simple facts about targets, such as ethnicity, occupation, name, and location.

As a way to model the methods used by humans to route such messages, sociologists have studied the importance of *categories*, that is, various groups to which people belong, in the small-world phenomenon. In the early 1970's, Hunter and Shotland [8] found that messages routed between people in the same university category (such as student, faculty, etc.) had shorter paths than messages routed across categories. Killworth and Bernard [10] performed experiments in the late 1970's that they called *reverse small-world experiments* in which each participant was presented with a list of messages for hundreds of targets, identified by the categories of town, occupation, ethnic background, and gender, and asked to whom they would send each of these messages. The study concluded that the choices people make in selecting routes are overwhelmingly categorical in nature. In the late 1980's, Bernard *et al.* [3] extended this work to identify which of twenty categories are most important for message routing to people from various cultures. More recently, Watts *et al.* [26] present a hierarchical model for categorical organization in social networks for the sake of message routing. They propose groups as the leaves of rooted trees, with internal nodes defining groups-of-groups, and so on. They define an ultrametric on sets of such overlapping hierarchies and conjecture that people use the minimum distance in one of their trees to make message routing decisions. That is, they argue that individuals can understand their "social distance" to a target as the minimum distance between them and the target in one of their categories. Such a determination requires some global knowledge about the structures of the various group hierarchies.

Although this previous work shows the importance of categories and of hierarchies of categories in explaining the small world phenomenon, it does not explain where the categories come from or what properties they need to have in order to allow greedy routing to work. Hence, this prior work leaves open the following questions:

- Which social networks support systems of categories that allow participants to route messages using the simple greedy rule of sending a message to an acquaintance who has more categories in common with the target?
- How complicated a system of categories is needed for this purpose, and what properties of the underlying network can be used to characterize the complexity of the category system?

[1] This observation has also led to the concept of "six degrees of separation" between all people on earth and the trivia game, "Six Degrees of Kevin Bacon," where players take turns trying to link performers to the actor Kevin Bacon via at most six movie collaborations.

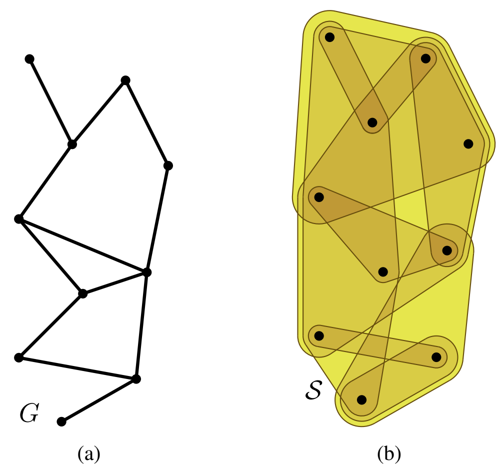

Figure 1. A set of elements U (drawn arbitrarily as points in the plane). (a) The graph G on U. (b) The categories $\mathcal{S}$ on U. In this example, the membership dimension is 4, because no element is contained in more than 4 groups.

Our goal in this paper, therefore, is to address these questions by studying the existence of mathematical and algorithmic frameworks that demonstrate the feasibility of local, greedy, category-based routing in social networks.

A. *Our Results*

Inspired by the work of Watts *et al.* [26], we view a social network as an undirected graph $G = (U, E)$, whose vertices represent people and whose edges represent relationships, taken together with a collection, $\mathcal{S} \subset 2^U$, of categories defined on the vertices in G. Figure 1 shows an example. In addition, given a network $G = (U, E)$ and category system $\mathcal{S}$, we define the *membership dimension* of $\mathcal{S}$ to be

$$\max_{u \in U} |\{C \in \mathcal{S} \colon\ u \in C\}|,$$

that is, the maximum number of groups to which any one person in the network belongs. The membership dimension characterizes the *cognitive load* of performing routing tasks in the given system of categories—if the membership dimension is small, each actor in the network only needs to know a proportionately small amount of information about his or her own categories, his or her neighbors' categories, and the categories of each message's eventual destination. Thus, we would expect real-world social networks to have small membership dimension.

In this paper, we provide a constructive proof that a category system can support greedy routing. Our results are not intended to model the actual formation of social categories, and we take no position on whether categories are formed from the network, the network is formed from categories, or both form together. Rather, our intention is to show the close relation between two natural parameters of a social network, its path length and its membership dimension. In particular:

- We show that the membership dimension of $(G, \mathcal{S})$ must be at least the diameter of G, $\operatorname{diam}(G)$, for a local, greedy, category-based routing strategy to work.
- We show that every connected graph $G = (U, E)$, has a collection $\mathcal{S}$ of categories such that local, greedy, category-based routing always works, with membership dimension $O((\operatorname{diam}(G) + \log |U|)^2)$.

Since Milgram's work [14], [20], [24], social scientists have believed that real-world social networks have diameters bounded by constants or slowly growing functions of the network size. Under a weak form of this assumption, that the diameter is $O(\log |U|)$, our results provide a natural model for how participants in a social network could efficiently route messages using a local, greedy, category-based routing strategy while remembering an amount of information that is only polylogarithmic in the size of the network.

B. *Previous Related Work*

Geometric greedy routing [6], [15] uses geographic location rather than categorical data to route messages. In this method, vertices have coordinates in a geometric metric space and messages are routed to any neighbor that is closer to the target's coordinates. Greedy routing may not succeed in certain geometric networks, so a number of techniques have been developed to assist such greedy routing schemes when they fail [4], [9], [16]. Introduced by Rao et al. [23], virtual coordinates can overcome the shortcomings of real-world coordinates and allow simple greedy forwarding to function without the assistance of fallback algorithms. This approach has been explored by other researchers [1], [12], [17], [22], who study various network properties that allow for greedy routing to succeed. Several researchers also study the existence of *succinct* greedy-routing strategies [5], [7], [18], [21], where the number of bits needed to represent the coordinates of each vertex is polylogarithmic in the size of the network; this notion of succinctness for geometric greedy routing is closely analogous to our definition of the membership dimension for categorical greedy routing.

Recent work by Mei *et al.* [19], studies category-based greedy routing as a heuristic for performing routing in dynamic delay-tolerant networks. Mei *et al.* assume that the network nodes have been organized into pre-defined categories based on the users' interests. Experiments suggest that using these categories for greedy routing is superior to routing heuristics based on location or simple random choices. One can interpret the categorical greedy routing techniques of Mei *et al.* and of this paper as being geometric routing schemes using virtual coordinates, where the coordinates represent category memberships. In this interpretation, the membership dimension of an embedding corresponds to the number of nonzero coordinates of each node, and our

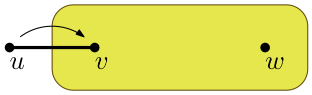

Figure 2. Illustration of the routing rule.

results show that such greedy routing schemes can be done succinctly in graphs with small diameter.

Similarly to the work of this paper, Kleinberg [11] studies the small-world phenomenon from an algorithmic perspective. However, his approach is orthogonal to ours: He focuses on location rather than categorical information as the critical factor for the ability to find short routes efficiently, and constructs a random network based on that information, whereas our approach takes the network as a given and studies the kinds of categorical structures needed to support category-based greedy routing.

In addition, it is worth noting that small world networks exhibit scale-free properties.

II. Routing based on Categorical Information

In this section, we introduce a mathematical model of categorical greedy routing, and provide basic definitions and properties that guarantee the success of this strategy.

A. Basic definitions

Abstracting away the social context, let U be the universe of n people defining the potential sources, targets, and intermediates for message routes, and let $G = (U, E)$ be an undirected graph whose m edges represent pairs of people who can communicate. For any two elements $s, t \in U$, let $\mathrm{sp}(s, t)$ be the length of the shortest path in G from s to t. The diameter $\mathrm{diam}(G) = \max_{s,t \in U} \mathrm{sp}(s, t)$ is the maximum length of any shortest path. For $s \in U$, define the *neighborhood* of s to be the set of neighbors $N(s) = \{u \in U \mid \{s, u\} \in E\}$ of s in G.

Now let $\mathcal{S} \subset 2^U$ be a set of subsets of U, which represent the abstract categories that elements of U belong to. For a given $u \in U$, we define $\mathrm{cat}(u) \subset \mathcal{S}$ to be the set of groups to which u belongs: $\mathrm{cat}(u) = \{C \in \mathcal{S} \mid u \in C\}$.

Definition 1 (membership dimension): The *membership dimension* of $\mathcal{S}$ is the maximum number of elements of $\mathcal{S}$ that any element of U is contained in, that is,

$$\mathrm{memdim}(\mathcal{S}) = \max_{u \in U} |\mathrm{cat}(u)|.$$

As discussed, there is evidence that in real world social networks and group structures $(G, \mathcal{S})$, both $\mathrm{diam}(G)$ and $\mathrm{memdim}(\mathcal{S})$ are significantly smaller than $|U|$.

B. The routing strategy

We now describe a simple greedy category-based strategy to route a message from one node to another. We clarify the distance function immediately following the rule definition.

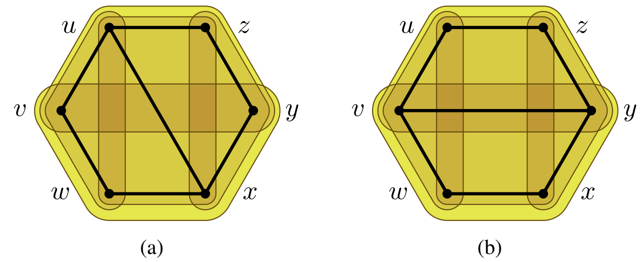

Figure 3. Two networks with the same elements and categories. (a) An example that is internally connected, but not shattered: no category contains y and a neighbor of v but not v itself. (b) An example that is shattered, but not internally connected: the induced graph of category $\{u, w, x, z\}$ is not connected.

Definition 2 (greedy routing rule): If a node u receives a message M intended for a destination $w \neq u$, then u should forward M to a neighbor $v \in N(u)$ that is closer to w than u is, that is, for which $d(v, w) < d(u, w)$.

The category-based distance function used by this rule is $d(s, t) = |\mathrm{cat}(t) \setminus \mathrm{cat}(s)|$, which measures the number of categories of the target that the current node does *not* share.[2] This number decreases as the number of shared groups of $\mathcal{S}$ between the current node and the target increases. We refer to the greedy routing strategy that uses this distance function as ROUTING (see Figure 2).

For category systems with low membership dimension, this strategy is easy to evaluate using only local knowledge about the categories of each neighbor of the current node and the categories of the target node.

C. Successful routing

We now investigate conditions under which ROUTING can successfully route messages between all pairs of nodes in a network. We identify several properties of a graph G and associated group structure $\mathcal{S}$ that directly influence the feasibility of routing. For routing to succeed, G must be connected. It seems natural to consider a stronger property:

Definition 3 (internally connected): $(G, \mathcal{S})$ is *internally connected* if for each $C \in \mathcal{S}$, G restricted to C is connected.

Figure 3(a) shows an example of an internally connected pair $(G, \mathcal{S})$. This is a very natural property for sociological groups to exhibit. People belonging to the same group will have greater cohesiveness, and if a group is not internally connected then it may be redefined to be the set of groups defined by its connected components.

Definition 4 (shattered): A pair $(G, \mathcal{S})$ is *shattered* if, for all $s, t \in U$, $s \neq t$, there is a neighbor $u \in N(s)$ and a set $C \in \mathcal{S}$ such that C contains u and t, but not s.

Figure 3(b) shows an example of a shattered pair. Note that in this definition, u and t could be the same node. This property falls out naturally from the instructions given in the

[2]Note that d is not a metric, since it is not necessarily symmetric.

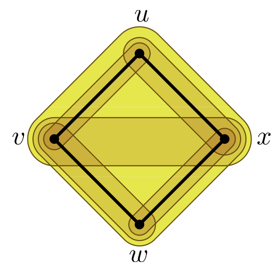

Figure 4. ROUTING does not work in this graph, even though it is internally connected and shattered. Routing from v to x fails: v, u, and w are all at distance 2 from x, so v has no neighbor that is closer than it to x.

real-world routing experiments of Milgram and others. In order for someone to advance a letter toward a target, there must be an acquaintance that shares additional interests with the target. Indeed, we now show that the shattered property is necessary for ROUTING to work.

Lemma 1: If $(G, \mathcal{S})$ is not shattered, ROUTING fails.

Proof: Since $(G, \mathcal{S})$ is not shattered, there exists a pair of vertices s and t, where s's neighbors are not in sets with t that do not contain s. Therefore, s's neighbors cannot share strictly more sets with t as s does, and ROUTING will fail to route from s to t. ■

If G is a tree, then these two properties together are sufficient for the routing strategy to always work:

Lemma 2: If G is a tree, and $(G, \mathcal{S})$ is internally connected and shattered, then ROUTING is guaranteed to work.

Proof: Let s and t be vertices in G. Since G is a tree, there is one simple path from s to t. Let (u, v) be an edge on the path from s to t. First, we claim that every set in $\mathcal{S}$ that contains both u and t also contains v. This follows from $(G, \mathcal{S})$ being internally connected: any set $C \in \mathcal{S}$ with $u, t \in C$ must also contain v, since v is on the only path between u and t. Therefore, v is contained in at least as many sets in $\mathcal{S}$ with t as u is. However, by the shattered property, v is in a set in $\mathcal{S}$ with t that does not contain u. Therefore v is in strictly more sets with t than u is. This property holds for every simple path; hence, ROUTING always works. ■

Although sufficient for routing in trees, the internally connected and shattered properties are not sufficient for ROUTING to work on arbitrary connected graphs. Figure 4 shows a counter-example—ROUTING is unable to route a message from the leftmost to the rightmost node, since there is no neighbor whose distance to the target is smaller.

III. Existence of Categories

In this section, we consider the following question: Is it possible to construct the family $\mathcal{S}$ so that ROUTING always works and $\mathcal{S}$ has low membership dimension?

We show that such a construction is always possible if we are given a connected graph as input. We also show that it is impossible to construct an $\mathcal{S}$ such that ROUTING will work if the graph is not known in advance.

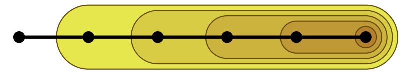

Figure 5. The sets B_v for each vertex v in the path. The sets A_v are constructed symmetrically.

A. *Constructing $\mathcal{S}$ given G*

Given a connected graph $G = (U, E)$ as input, we would like to construct a family $\mathcal{S} \subset 2^U$ so that ROUTING works, and the membership dimension of S is small. We concentrate foremost on constructions of category collections that are internally connected and shattered, because of the social significance of these properties. Nevertheless, even without these properties, we have the following lower bound.

Lemma 3: Let G and $\mathcal{S}$ be a graph and a category system, respectively, such that ROUTING works for G and $\mathcal{S}$. Then $\mathrm{memdim}(\mathcal{S}) \geq \mathrm{diam}(G)$.

Proof: By definition of the diameter, there are two vertices $s, t \in U$ such that $sp(s,t) = \mathrm{diam}(G)$. Let P be the path that ROUTING follows from s to t, and note that the length of P must be at least $\mathrm{diam}(G)$. An edge (u, v) can only be on P if $d(v,t) < d(u,t)$. Since $d(\cdot,\cdot)$ can only take integer values, $d(u,t) \geq d(v,t) + 1$. Therefore, $d(s,t) \geq |P|$. By definition, $d(s,t) = |\,\mathrm{cat}(t) \setminus \mathrm{cat}(s)|$, and $\mathrm{memdim}(\mathcal{S})$ is the maximum of $\mathrm{cat}(\cdot)$ over all elements; hence $\mathrm{memdim}(\mathcal{S}) \geq |\,\mathrm{cat}(t)| \geq |\,\mathrm{cat}(t) \setminus \mathrm{cat}(s)| = d(s,t) \geq |P| \geq \mathrm{diam}(G)$, as claimed. ■

For paths, this bound is tight:

Lemma 4: If G is a path, then there exists an $\mathcal{S}$ s.t. $(G, \mathcal{S})$ is shattered and internally connected with $\mathrm{memdim}(\mathcal{S}) = \mathrm{diam}(G)$.

Proof: Arbitrarily pick one of the two end vertices of G and let us refer to the vertices in G by their distance, 0 to $n-1$, from this vertex. For each vertex i, form two sets A_i and B_i, where $A_i = \{0, \ldots, i-1\}$ and $B_i = \{i+1, \ldots, n-1\}$, and let $\mathcal{S} = \bigcup_{v \in U} \{A_v, B_v\}$. Figure 5 illustrates this construction. Each set in $\mathcal{S}$ consists of a path of vertices and therefore $\mathcal{S}$ is internally connected. $\mathcal{S}$ is also shattered, since for all s and t, s has a neighbor that shares either A_s or B_s with t, but s is not in these sets. Considering $\mathrm{memdim}(\mathcal{S})$, note that each vertex i is contained in sets A_j for $0 \leq j < i$ and B_k for $k < i \leq n-1$. Therefore, each vertex is in exactly $n-1$ sets, which is $\mathrm{diam}(G)$. ■

It follows from Lemmas 2 and 4 that, if G is a path, one can construct $\mathcal{S}$ with $\mathrm{memdim}(\mathcal{S}) = \mathrm{diam}(G)$, so that ROUTING works in G.

There are other graphs for which it is relatively easy to set up a category set that is shattered and internally connected in a way that supports the ROUTING algorithm. For example, in a tree of height 1 (i.e., a star graph), with root r, we could create for each leaf of the tree two categories, one containing the leaf itself and one containing both the leaf and the root. Every path in this tree supports ROUTING. However, the

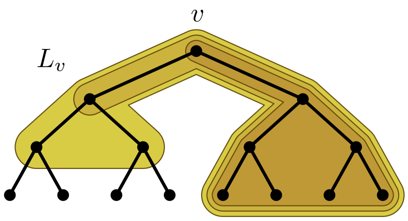

Figure 6. The collection of sets L_v for an example subtree at v.

membership dimension of this category system is high, since the root belongs to a linear number of categories. So even in this simple example, supporting ROUTING and achieving low membership dimension is a challenge. Moreover, this challenge becomes even more difficult already for a tree of height 2, since navigating from any leaf, x, to another leaf, y, requires that the parent of x belong to more categories with y than x—and this must be true for *every* other leaf, y. Thus, it is perhaps somewhat surprising that we can construct a set of categories, $\mathcal{S}$, for an arbitrary binary tree that causes this network to be shattered and internally connected (so the ROUTING strategy works, by Lemma 2) and such that $\mathcal{S}$ has small membership dimension.

Lemma 5: If G is a binary tree, then there exists an $\mathcal{S}$ s.t. $(G, \mathcal{S})$ is shattered and internally connected with $\mathrm{memdim}(\mathcal{S}) = O(\mathrm{diam}^2(G))$.

Proof: We show how to construct $\mathcal{S}$ from G. Arbitrarily pick a vertex $r \in U$ of degree at most 2 and root the binary tree at r, so each vertex v has left and right children, $\mathrm{left}(v)$ and $\mathrm{right}(v)$, and let $\mathrm{height}(v)$ be the length of the longest simple path from v to any descendant of v. For each vertex v, we create a set S_v, containing v's descendants (which includes v). We further construct two families, L_v and R_v, using helper sets $L_{v,i}$ and $R_{v,i}$. Let $L_{v,i}$ (resp., R_{v_i}) consist of v, the vertices in v's left (right) subtree down to depth i, and all vertices in v's right (left) subtree. Then define

$$L_v = \{L_{v,i} \mid \mathrm{depth}(v) \leq i \leq \mathrm{depth}(v) + \mathrm{height}(\mathrm{left}(v))\}.$$

Figure 6 illustrates this. The family R_v is defined symmetrically. Our $\mathcal{S}$ is then defined as

$$\mathcal{S} = \bigcup_{v \in U} \{S_v\} \cup L_v \cup R_v.$$

Each set in $\mathcal{S}$ is a connected subgraph of G, so $\mathcal{S}$ is internally connected. As the following argument shows, $\mathcal{S}$ is shattered: If s is an ancestor of t, then s's child u on the path to t is in set S_u which contains u and t but not s. Otherwise, let v be the lowest common ancestor of s and t, and assume without loss of generality that s in v's left subtree; then $L_{v,\mathrm{depth}(s)-1}$ contains s's parent and t but not s.

We now analyze the membership dimension of this construction. Let v be a vertex, and let $\mathrm{ancestors}(v)$ be the set of v's ancestors. For $u \in \mathrm{ancestors}(v)$, $v \in S_u$, and v belongs to $O(\mathrm{height}(u))$ sets of L_u and R_u. Then v belongs to $O\left(\sum_{u \in \mathrm{ancestors}(v)} \mathrm{height}(u)\right)$ sets, which is $O(\mathrm{diam}^2(G))$ for any v. ■

We now extend this result to arbitrary trees by applying *weight-balanced binary trees* [2], [13].

Definition 5 (weight balanced binary tree): A weight balanced binary tree is a binary tree that stores weighted items in its leaves. If item i has weight w_i, and all items have a combined weight of W then item i is stored at depth $O(\log(W/w_i))$.

Lemma 6: Let T be an n-node rooted tree with height h. We can embed T into a binary tree such that the ancestor–descendant relationship is preserved, and the resulting tree has height $O(h + \log n)$.

Proof: Let n_u be the number of descendants of vertex u in T. For each vertex u in T that has more than two children, we expand the subtree consisting of u and u's children into a binary tree as follows. Construct a weight balanced binary tree B on the children of u, where the weight of a child v is n_v. We let u be the root of B. Each child v of u in the original tree is then a leaf at depth $\log(n_u/n_v)$ in B. Performing this construction for each vertex u in the tree expands T into a binary tree with the ancestor–descendant relationship preserved from T.

Furthermore, each path from root to leaf in T is only expanded by $\log(n)$ nodes, which we can see as follows. Each parent-to-child edge (u, v) in T is replaced by a path of length $O(\log(n_u/n_v))$. Therefore for each path P from root r to leaf l in T, our construction expands P by length $O(\sum_{(u,v)\in P} \log(n_u/n_v))$, which is a sum telescoping to $O(\log(n_r/n_l)) = O(\log n)$. Therefore, the height of the new binary tree is $O(h + \log n)$. ■

Combining this lemma with Lemma 2, we get the following theorem.

Theorem 1: Given a tree T, it is possible to construct a family $\mathcal{S}$ of subsets such that ROUTING works for T and $\mathrm{memdim}(\mathcal{S}) = O((\mathrm{diam}(T) + \log n)^2)$.

Proof: Arbitrarily root T and embed T in a binary tree B using the method in Lemma 6. Then B has height $O(\mathrm{diam}(T) + \log n)$, and diameter $\mathrm{diam}(B) = O(\mathrm{diam}(T) + \log n)$. Applying the construction from Lemma 5 to B gives us a family $\mathcal{S}_B$ with $\mathrm{memdim}(\mathcal{S}_B) = O((\mathrm{diam}(T) + \log n)^2)$. We then construct a family $\mathcal{S}_T$, by removing vertices that are in B but not T from the sets in $\mathcal{S}_B$. By construction, $(T, \mathcal{S}_T)$ is shattered and internally connected, and $\mathrm{memdim}(\mathcal{S}_T) \leq \mathrm{memdim}(\mathcal{S}_B) = O((\mathrm{diam}(T) + \log n)^2)$. By Lemma 2, ROUTING works on T with category sets from $\mathcal{S}_T$. ■

We can further extend this theorem to arbitrary connected graphs, which is the main upper bound result of this paper.

Theorem 2: If G is connected, there exists $\mathcal{S}$ s.t. ROUTING works and $\mathrm{memdim}(\mathcal{S}) = O((\mathrm{diam}(G) + \log(n))^2)$.

Proof: Compute a low-diameter spanning tree T of G. This step can easily be done using breadth-first search, producing a tree with diameter at most $2\,\mathrm{diam}(G)$. We then use the construction from Theorem 1 on T. For greedy routing to work in a graph G, note that it is sufficient to show that it works in a spanning tree of G. Therefore, since ROUTING works in T, ROUTING also works in G. ■

IV. Conclusion and Open Problems

We have presented a construction of groups S on a connected graph G that allows a simple greedy routing algorithm, utilizing a notion of distance on group membership, to guarantee delivery between nodes in G. Such a construction will have membership dimension $O((\mathrm{diam}(G) + \log n)^2)$, demonstrating a small cognitive load for the members of G.

There are several directions for future work. For example, while we have shown that the membership dimension must be minimally the diameter of G, it remains to be shown if the membership dimension must be the square of the diameter plus a logarithmic factor for arbitrary graphs. We conjecture that the square term is not strictly needed in the membership dimension in order for ROUTING to work. Our group construction is performed for a general graph by selecting a low diameter spanning tree and using the presented tree construction, so it may be possible that there is a group construction that has lower membership dimension and more efficient routing if it is constructed directly in G.

In this paper all categories are given equal weight with respect to routing tasks and that participants use a simple greedy routing algorithm based solely on increasing the number of categories in common with the target. Future work could include study of a category-based routing strategy that allows participants to weight various categories higher than others, as in the work of Bernard *et al.* [3].

References

[1] P. Angelini, F. Frati, and L. Grilli. An algorithm to construct greedy drawings of triangulations. In *Graph Drawing*, LNCS 5417, pages 26–37. Springer, 2009.

[2] S. W. Bent, D. D. Sleator, and R. E. Tarjan. Biased search trees. *SIAM J. Computing*, 14(3):545–568, 1985.

[3] H. R. Bernard, P. D. Killworth, M. J. Evans, C. McCarty, and G. A. Shelley. Studying social relations cross-culturally. *Ethnology*, 27(2):155–179, 1988.

[4] P. Bose, P. Morin, I. Stojmenovic, and J. Urrutia. Routing with guaranteed delivery in ad hoc wireless networks. *Wireless Networks*, 7:609–616, 2001.

[5] D. Eppstein and M. T. Goodrich. Succinct greedy graph drawing in the hyperbolic plane. In *Graph Drawing*, LNCS 5417, pages 14–25. Springer, 2009.

[6] G. G. Finn. Routing and addressing problems in large metropolitan-scale internetworks. Research report, ISI, 1987.

[7] M. Goodrich and D. Strash. Succinct greedy geometric routing in the Euclidean plane. In *Algorithms and Computation*, LNCS 5878, pages 781–791. Springer, 2009.

[8] J. E. Hunter and R. L. Shotland. Treating data collected by the "small world" method as a Markov process. *Social Forces*, 52(3):321–332, 1974.

[9] B. Karp and H. T. Kung. GPSR: greedy perimeter stateless routing for wireless networks. In *6th ACM MobiCom*, pages 243–254, 2000.

[10] P. Killworth and H. Bernard. Reverse small world experiment. *Social Networks*, 159(1), 1978.

[11] J. Kleinberg. The small-world phenomenon: an algorithm perspective. In *32nd ACM STOC*, pages 163–170, 2000.

[12] R. Kleinberg. Geographic routing using hyperbolic space. In *26th IEEE INFOCOM*, pages 1902–1909, 2007.

[13] D. E. Knuth. Optimum binary search trees. *Acta Informatica*, 1:14–25, 1971.

[14] C. Korte and S. Milgram. Acquaintance networks between racial groups: Application of the small world method. *J. Personality and Social Psychology*, 15(2):101–108, 1970.

[15] E. Kranakis, H. Singh, and J. Urrutia. Compass routing on geometric networks. In *11th CCCG*, pages 51–54, 1999.

[16] F. Kuhn, R. Wattenhofer, Y. Zhang, and A. Zollinger. Geometric ad-hoc routing: of theory and practice. In *22nd ACM PODC*, pages 63–72, 2003.

[17] T. Leighton and A. Moitra. Some results on greedy embeddings in metric spaces. *Discrete Comput. Geom.*, 44:686–705, 2010.

[18] P. Maymounkov. Greedy embeddings, trees, and Euclidean vs. Lobachevsky geometry, 2006. http://pdos.csail.mit.edu/~petar/papers/maymounkov-greedy-prelim.pdf.

[19] A. Mei, G. Morabito, P. Santi, and J. Stefa. Social-aware stateless forwarding in pocket switched networks. In *30th IEEE INFOCOM*, 2011.

[20] S. Milgram. The small world problem. *Psychology Today*, 1(May):61–67, 1967.

[21] R. B. Muhammad. A distributed geometric routing algorithm for ad hoc wireless networks. In *3rd ITNG*, pages 961–963, 2007.

[22] C. H. Papadimitriou and D. Ratajczak. On a conjecture related to geometric routing. *Theor. Comput. Sci.*, 344:3–14, November 2005.

[23] A. Rao, S. Ratnasamy, C. Papadimitriou, S. Shenker, and I. Stoica. Geographic routing without location information. In *9th ACM MobiCom*, pages 96–108, 2003.

[24] J. Travers and S. Milgram. An experimental study of the small world problem. *Sociometry*, 32(4):425–443, 1969.

[25] D. J. Watts. Networks, dynamics, and the small-world phenomenon. *The American Journal of Sociology*, 105(2):493–527, 1999.

[26] D. J. Watts, P. S. Dodds, and M. E. J. Newman. Identity and search in social networks. *Science*, 296:1302–1305, 2002.