A Note on Embedding Complete Graphs into Hypercubes

Michel Goemans* Michael Klugerman†
Alexander Russell‡ Ravi Sundaram§

October 19, 1995

Abstract

An embedding of K_n into a hypercube is a mapping of the n vertices of K_n to distinct vertices of the hypercube, and the associated cost is the sum over all pairs of (mapped) vertices of the Hamming distance between the vertices. Let $f(n)$ denote the minimum cost over all embeddings of K_n into a hypercube (of any dimension). In this note we prove that $f(n) = (n-1)^2$, unless $n = 4$ or $n = 8$, in which case $f(n) = (n-1)^2 - 1$. As an application, we use this theorem to derive an alternate proof of the fact that the Isolation Heuristic (and the accompanying variant) for the Multiway Cut problem of [DJP+ 92] is tight for all k.

1 Preliminaries

K_n denotes the complete graph on n vertices. The hypercube of dimension n has 2^n vertices, each vertex being labeled with a string of 0's and 1's of length n. The Hamming distance between two vertices of the hypercube is the number of positions in which the labels of the two vertices are different.

*Department of Mathematics, MIT. Email: goemans@math.mit.edu.
†Escher Group. Email: klugermn@theory.lcs.mit.edu.
‡Department of Mathematics, MIT. Email: acr@theory.lcs.mit.edu.
§Laboratory for Computer Science, MIT. E-mail: koods@theory.lcs.mit.edu.

A cut of a graph is a subset of vertices of the graph. The edges of a cut are the set of edges which go from vertices within the subset to vertices outside.

2 Result

Theorem 1 *Let E be any set of n* distinct *vertices $v_1, \ldots, v_n$ in a hypercube. Let $f(E) = \sum_{i<j} d(v_i, v_j)$ (where $d(.,.)$ is the Hamming distance function). Then*

$$f(n) = \min_E f(E) = \begin{cases} (n-1)^2 - 1 & \text{if } n = 4, 8 \\ (n-1)^2 & \text{otherwise.} \end{cases}$$

Proof: It is clear that the minimum in the case of $n = 4$ and $n = 8$ is achieved by letting E be the set of all vertices in a two-dimensional and three-dimensional hypercube, respectively. For all other n the minimum is achieved by taking a vertex and its $n-1$ neighbours in an $n-1$-dimensional hypercube.

It remains to show that we cannot do better. Consider any set with e nodes of even Hamming weight and o nodes of odd Hamming weight, $e + o = n$. To prove our bound, we consider the (Hamming) distances between even nodes, between odd nodes, and then between even and odd nodes.

It is clear that we have a lower bound of $(n-1)^2$ if either e or o is 1. Assume, then, without loss of generality that $o \geq e \geq 2$. At best, each pair of even nodes has distance 2. Similarly for the odd nodes. Consider the distances between even and odd nodes. For each pair of even nodes e_1 and e_2, there are at most 2 nodes o_1 and o_2 which are each at distance 1 from e_1 and e_2. The remaining $o - 2$ odd weight nodes are each an average distance of at least 2 from both e_1 and e_2. Based on these averages one arrives (as shown below) at a lower bound for the minimum total Hamming distance of

$$2\binom{e}{2} + 2\binom{o}{2} + \left(\frac{2(o-2)+2}{o}\right)eo = n^2 - n - 2e.$$

Note that $n^2 - n - 2e \geq n^2 - 2n$. Hence, if $e \neq o$, we arrive at a lower bound of $(n-1)^2$. Now, consider the remaining case when $e = o \neq 1$. When $e = 2, 4$ we have explicit constructions which match the lower bound of $n^2 - 2n$. It remains to consider the case $e = o > 3$. In order for the above argument concerning inter-node distances to be tight, the set must have special structure. Specifically, any pair of nodes having the same parity must have distance 2. This implies that the even weight nodes lie in a Hamming ball of radius 1. Similarly for the odd weight nodes. The tightness of the above argument also implies that for any two odd nodes, no even node has distance more than 3 from either of them.

We show that this cannot occur (demonstrating that the lower bound increases to $(n-1)^2$). Without loss of generality, suppose 4 of the odd nodes have addresses $0\ldots01, 0\ldots010, 0\ldots0100$ and $0\ldots01000$. Consider v, an even node with non-zero weight (there must be such a node). Considering the cases $weight(v) = 2$ and $weight(v) > 2$ we see that v is at distance at least 3 from some pair of these nodes. ■

3 Application

The original motivation for solving the problem of embedding complete graphs in hypercubes arose from the Mulitway or n-Way Cut problem. In the n-Way Cut we are given an edge-weighted graph and n distinguished vertices called terminals and asked for a minimum weight set of edges that separates every terminal pair. This problem is simply the min-cut max-flow problem when $n = 2$. In [DJP+ 92] it was shown that the problem becomes hard for $n = 3$. They also gave a simple approximation algorithm, the Isolation Heuristic, for arbitrary graphs that came within a factor of $2(1 - 1/n)$ of the optimal. They also gave a variant of the Isolation Heuristic which does does better for $n = 4$ and $n = 8$. They state in the paper, without proof, that similar approaches are bound to fail for all other values of n. Below we formalize what exactly the Isolation Heuristic and related variants are doing and prove that improvements cannot be obtained, except for $n = 4$ and $n = 8$.

The Isolation Heuristic (and its variant for the cases $n = 4$ and $n = 8$) can be thought of as essentially finding a minimum cost collection of cuts that separates all pairs of vertices in the unweighted complete graph, K_n, on n vertices. Here, the cost of a collection (of cuts) is the sum of the costs of the cuts in it, and the cost of a cut is just the number of edges in it.

Lemma 1 *If K_n has a cut collection of cost C separating all pairs of vertices then the n-Way Cut problem has an approximation within a factor of*

$2C/n(n-1)$.

Proof: The proof is a straightforward averaging argument. We associate each vertex of K_n one-to-one to a terminal of the graph in the n-Way Cut problem. To a particular cut of K_n we correspond the equivalent min-cost cut of the n-Way Cut problem graph. Consider all possible mappings of the vertices of K_n one-to-one to terminals of the graph. Since the average cost is within a factor of $2C/n(n-1)$ of the optimal to the n-Way Cut problem there exists a mapping which achieves this bound. Note that we do not give an effective way to compute this approximate solution. At this point we are concerned only with existence. ■

Lemma 2 *The minimum cost of any cut collection that separates all pairs of vertices of K_n is equal to $f(n)$.*

Proof: Given any cut collection $\mathcal{C} = \{c_1, c_2, \ldots, c_k\}$ we can create an embedding of equivalent cost in a hypercube of dimension k. We have one dimension per cut and a vertex of K_n gets mapped to that vertex of the hypercube with a 1 in the i'th position of the label iff the original vertex of of K_n is in the ith cut. It is easy to see that if E is the set of mapped vertices then $f(E)$ is equal to the cost of $\mathcal{C}$.

Similarly, given any embedding E in a hypercube of dimension k one can create a cut collection of equivalent cost by having one cut for each dimension and putting all those mapped vertices in the cut which have a 1 in the label at the dimension corresponding to the cut. ■

Corollary 1 *The best that the Isolation Heuristic and its variants can do is to get within a factor of $2(1-1/n)$, except when $n = 4$ or 8 in which case they can get within a factor of $2(1-1/n-1/2n(n-1))$, of the optimal to the n-Way Cut problem.*

Proof: Follows from Theorem 1 and Lemmas 2 and 3. ■

References

[DJP+ 92] E. Dahlhaus, D. S. Johnson, C. H. Papadimitriou, P.D. Seymour and M. Yannakakis. *The complexity of multiway cuts.* In Proc. 24th STOC, pp 241-251, 1992.