Lecture 35: Basic probability theory

Nitin Saxena ⋆

IIT Kanpur

1 Super-concentrators

This is a graph theory term inspired from telecommunications. We will see a construction from [1].

A *super-concentrator* is a directed acyclic graph $G = (V, E)$, with n special input nodes $I \subset V$ and n output nodes $O \subset V$. It satisfies the property that for any $1 \leq k \leq n$, *any* k subset of I is connected with *any* other k subset of O with k disjoint paths.

Note 1. This captures the physical situation where k pairs of people want to talk to each other simultaneously on the telephone.

This is a very strong connectivity property and can be used to design robust networks. It is very easy to design a super-concentrator with $O(n^2)$ edges (Exercise). We will see how linear size (number of edges) super-concentrators can be constructed. We will do it recursively. But we first need "concentrators".

For $n_1 > n_2 > u$, an (n_1, n_2, u) *concentrator* has n_1 input nodes and n_2 output nodes. For any $1 \leq k \leq u$ and an input k-subset S, *there exists* an output k-subset T which is connected to S using k disjoint paths.

Exercise 1. In concentrators, number of input and output nodes could be different and there is an upper limit on size k subset which we are allowed to choose. What is the other essential difference between concentrator and super-concentrator?

Quantification $\forall S \forall T$ has been weakened to $\forall S \exists T$.

For any j, we will first construct $(6j, 4j, 3j)$ concentrator using probabilistic methods. Again, the important thing is that we can construct such concentrators in linear size. Though, since concentrators have weaker properties, it is less surprising than the linear construction of super-concentrators.

The concentrator is going to be a directed *bipartite* graph with $6j$ input nodes as one part and $4j$ output nodes as the other part. There will be no extra vertices. We will show that if we pick a linear number $(36j)$ of edges randomly, with non-zero probability, it will be a concentrator.

Suppose every input node has out-degree 6 and every output node has in-degree 9 in our random graph. Every edge is labelled from one side by the vertex and a number between 1 and 6, from other side by an output vertex and a number between 1 and 9.

The random way to pick a graph is, for the first vertex v of input nodes and label 1 pick a partner ($4j \times 9 = 36j$ choices). Then, pick the partner for $(v, 2)$ and so on. Clearly there are $(36j)!$ ways of doing that.

Say a *non-concentrator* is a graph where there exists a subset S of size $\leq 3j$ of input nodes whose neighborhood is $< |S|$. We need to find the number of non-concentrators in the random process described above. We will show that it is less than $(36j)!$.

Exercise 2. Prove that if a graph is not a non-concentrator then it is a concentrator.

$k \leq 3j$ disjoint paths in this case means a matching between the k input and k output nodes. Apply Hall's marriage theorem.

Note 2. We have taken a stricter definition of non-concentrator than required. It is to help us in calculations.

Theorem 1. *The number of non-concentrators is less than* $(36j)!$.

⋆ Edited from Rajat Mittal's notes.

Proof. A non-concentrator has a subset S of input nodes with size $k \leq 3j$, s.t., $|N(S)| < |S|$. If this is the case for an S, then $N(S)$ is a subset of some $(k-1)$-subset of output nodes T.

For a particular k, there are $\binom{6j}{k}$ ways to choose S and less than $\binom{4j}{k}$ ways to choose T. S has $6k$ outgoing edges and T has less than $9k$ incoming edges. Fixing S and T, there are at most $\binom{9k}{6k} \cdot (6k)! \cdot (36j - 6k)!$ ways for $N(S)$ to fall in T.

We need to show that the following upper bound on the number of non-concentrators,

$$\sum_{k=0}^{3j} \binom{6j}{k}\binom{4j}{k}\binom{9k}{6k} \cdot (6k)!(36j-6k)! \; < \; (36j)!\,.$$

It reduces to showing,

$$\sum_{k=0}^{3j} \binom{6j}{k}\binom{4j}{k}\binom{9k}{6k} \Big/ \binom{36j}{6k} \; \ll \; 1\,.$$

Exercise 3. Prove this using binomial identities.

Calculation shows that the summands first decrease and then increase, as k grows. For both $k = 1$ and $k = 3j$, the summands are smaller than $1/j^4$. Thus, the overall sum is smaller than $3/j^3$.

□

Given a concentrator (bipartite), the task to construct a super-concentrator (non-bipartite) is not hard. We will use two concentrators and a single smaller super-concentrator:

From $6j$ input nodes I we will use a concentrator to connect them to new $4j$ nodes I'. Then, put a super-concentrator from these $4j$ nodes to next new $4j$ nodes O'. The final step is to connect the final $4j$ nodes to $6j$ output nodes O using the reverse concentrator. Additionally, create a *matching* $\mathcal{M}$ from vertices I to O.

Exercise 4. For getting a super-concentrator, prove that it is enough to show:

for all $1 \leq k \leq 3j$, any input k-subset S is connected to any output k-subset T with k disjoint paths.

The case of $3j < k \leq 6j$: Consider the vertices in S that are *unmatched* in T by $\mathcal{M}$. These will be at most $3j$ many. Thus, for these we use the paths through I', and for the rest we use $\mathcal{M}$.

Exercise 5. Show that for $k \leq 3j$, any input k-subset is connected to any output k-subset with k disjoint paths using concentrators.

This should convince you that we have a super-concentrator.

Exercise 6. Show that it has a linear, in n, number of edges.

Corollary 1. *There is a randomized algorithm to design a $20j$-vertex super-concentrator using $O(j)$ edges.*

The construction of concentrators can be made deterministic. This is done using *expander graphs*. Interested students are advised to read more about expander graphs.

References

1. U. Schöning. Gems of Theoretical Computer Science. *Springer-Verlag*, 1998.