CS6845: Pseudorandomness 30 Jan 2018

9. Undirected reachability using random walks

9.1 Introduction

In this lecture, we will prove that we can solve undirected reachability with a randomized algorithm that uses only logarithmic space.

9.2 Undirected reachability with logarithmic space

Now we will see an algorithm to check if an undirected graph G has a path from a vertex s to t that uses only logarithmic space.

- **Input**: Graph G with degree d, nodes s and t.
- Perform a random walk of length $k = \text{poly}(n, d)$ starting from s.
- Accept if the random walk ends in t, else reject.

How do we prove the correctness of such an algorithm? We will first bound $\lambda(G)$ and use "spectral expansion $\Rightarrow$ mixing" lemma.

9.2.1 Bounding $\lambda(G)$

Let G be a d-regular connected non-bipartite graph with self-loops. We know that $\lambda(G)^2 = \max_{x \perp u, \|x\|=1} \|xA\|^2$. Rewriting this, $\langle xA, xA \rangle = xA^2x^T = \sum_{i,j}(A^2)_{i,j}x_ix_j$. Now, A^2 is the adjacency matrix of the multigraph G^2 which has e edges between two vertices u and v if there are e paths of length at most two between u and v in G. Each edge of the multigraph contributes $\frac{2}{d^2}x_ix_j$ to the sum $\sum_{i,j}(A^2)_{i,j}x_ix_j$. So we can rewrite that sum as $\sum_{(i,j)\in E} \frac{2}{d^2}x_ix_j$, where E is the multiset of edges of the graph G^2. Therefore, we have

$$\begin{aligned}
\langle xA, xA \rangle &= \sum_{(i,j)\in E} \frac{2}{d^2} x_i x_j \\
&= \frac{1}{d^2} \left(\sum_{(i,j)\in E} (x_i^2 + x_j^2) - \sum_{(i,j)\in E} (x_i - x_j)^2 \right) \\
&= \frac{1}{d^2} \left(d^2 - \sum_{(i,j)\in E} (x_i - x_j)^2 \right) \\
&= 1 - \frac{1}{d^2} \sum_{(i,j)\in E} (x_i - x_j)^2
\end{aligned}$$

Therefore, $1 - \lambda(G)^2 = \frac{1}{d^2} \min_{x \perp u, \|x\|=1} \sum_{(i,j)\in E} (x_i - x_j)^2$. This is non-zero since the graph G is not bipartite, is connected and has self-loops. Let $\hat{x}$ be the vector that achieves the minimum. Since $\hat{x} \perp u$, we have $\sum \hat{x}_i = 0$. Also, since $\|\hat{x}\| = 1$, we have $\sum \hat{x}_i^2 = 1$. Let u and v be indices such that $\hat{x}_u$ is the maximum and $\hat{x}_v$ the minimum. Either $\hat{x}_u \geq \frac{1}{\sqrt{n}}$ or $\hat{x}_v \leq \frac{-1}{\sqrt{n}}$. Let P be a shortest path between u and v in G^2. Such a path exists in G^2 since G has self-loops and is connected.

$$
\begin{aligned}
1 - \lambda(G)^2 &= \frac{1}{d^2} \sum_{(i,j)\in E} (\hat{x}_i - \hat{x}_j)^2 \\
&\geq \frac{1}{d^2} \sum_{(i,j)\in P} (\hat{x}_i - \hat{x}_j)^2 \\
&\geq \frac{1}{d^2|P|} \left(\sum_{(i,j)\in P} |\hat{x}_i - \hat{x}_j| \right)^2 \text{ (Cauchy-Schwartz)} \\
&\geq \frac{1}{d^2|P|} |\hat{x}_u - \hat{x}_v|^2 \geq \frac{1}{d^2|P|n} \geq \frac{1}{d^2 n^2}
\end{aligned}
$$

Therefore, $\lambda(G) \leq 1 - \frac{1}{d^2 n^2}$.

9.2.2 Proving the correctness of the algorithm

Suppose G is any graph. We replace each vertex with degree k with a k-cycle and add self-loops. We will assume that G is connected. If it is not, then two cases arise: If s and t are not in the same connected component we will always reject. On the other hand, if s and t are in the same connected component, then the bound from the previous calculation is applicable to this connected component. So, we can assume G is connected w.l.o.g. Now, we know that $\|\pi A^k - u\| \leq \lambda(G)^k$. Therefore, $\left|\pi A^k - \frac{1}{n}\right| \leq \left(1 - \frac{1}{d^2 n^2}\right)^k$. If $k = d^2 n^2 \log 2n$, then $\lambda(G)^k \geq 1/2n$. Therefore, $\pi A^k(t) \geq 1/2n$. So, the probability that a random walk of length $dn^2 \log 2n$ from s will not reach t is at most $1 - 1/2n$. To improve this error probability we repeat this algorithm by performing multiple independent random walks.