6.896 Sublinear Time Algorithms	September 30, 2004
Lecture 6	
Lecturer: Ronitt Rubinfeld	*Scribe: Swastik Kopparty*

1 Continuing the Proof of Correctness of the Bipartiteness Tester

We assume that the graph is ϵ-far from bipartite and need to show that the algorithm fails with probability at least 2/3.

Recall that we partitioned the sample into 2 parts, U and S.

- $|U| = m_1 = \Theta(\log\left(\frac{1}{\epsilon}\right)/\epsilon)$
- $|S| = m_2 = Cm_1/\epsilon$

U induced a partition of all the vertices it covered into C_1, C_2. Also recall that U is *good* if less than $\epsilon n/4$ influential nodes are not adjacent to some vertex in U. Last time we showed that $\Pr_U[U$ good $] \geq 5/6$.

The other main result we carried over is that whenever U is good, then for *any* partition of the whole graph, there are at least $\epsilon n^2/3$ violating edges between C_1, C_2 (that is an edge between 2 vertices in the same C_i). Now we will finish off the proof by showing that this implies that if U is good, then with high probability $U \cup S$ will not be bipartite.

Lemma 1 *For any good U and any partition (U_1, U_2) of U, with probability $\geq 1 - \frac{2^{-|U|}}{6}$ (over the choice of S), the following statement holds: For all partitions (S_1, S_2) of S, $\exists$ some edge between vertices in $U \cup S$ violating $(U_1 \cup S_1, U_2 \cup S_2)$.*

Before proving this, we will prove a corollary showing its connection to the algorithm.

Corollary 2 *Pr_S[Every partition (U_1, U_2) of U and (S_1, S_2) of S has a violating edge w.r.t $(U_1 \cup S_1, U_2 \cup S_2)$] $\leq 1/6$.*

Remark If the algorithm picks U, S such that every partition of $U \cup S$ has a violating edge, then it must fail.

Proof $\Pr_S[\exists$ partition of (U_1, U_2) with no violating edge in some $(S_1, S_2)]$
$\leq 2^{|U|} \Pr_S[$ some fixed (U_1, U_2) has no violating edge in some $(S_1, S_2)]$
$\leq 2^{|U|} \frac{2^{-|U|}}{6} = 1/6$.
■

Now we prove the lemma by proving 2 claims.

Claim 3 *Pr_S[S has u, v from both C_1 or C_2 and (u, v) is an edge] $\geq 1 - 2^{-|U|}/6$.*

View $S = \{s_1, s_2, \ldots, s_{m_2}\}$ as $m_2/2$ pairs of nodes s_{2i-1}, s_{2i}. We will consider only the edge between vertices of a pair (and thus get independent edges).

$\forall$ pairs s_{2i-1}, s_{2i}, Pr[both in the same C_i] $\geq \epsilon/3$.

So the probability that none of the pairs are in both $\leq (1-\epsilon/3)^{m_2/2} \leq (1-\epsilon/3)^{Cm_1/2\epsilon} \leq e^{-Cm_1} = \frac{2^{-|U|}}{6}$, proving the claim.

Our other claim justifies our construction of C_1, C_2.

Claim 4 *If S picks u, v that are both in C_i for some i, then every partition of S has a violating edge w.r.t. (U_1, U_2).*

Without loss of generality, $u, v \in C_2$. Then take any other partition of S. If u, v are in the same partition, it has a violating edge. Otherwise u, v are in different partitions, but are both connected to U_1. So again we get a violating edge.

Now we can prove the main theorem.

Proof Given the the graph is ϵ-far from bipartite, we have to show that Pr[algorithm outputs pass] $\leq$ 1/3.

Pr[output pass] $\leq$ Pr[U bad OR U good but S does not contain a C_iC_i violating edge] $\leq 1/6+1/6 = 1/3$.

■

2 First order properties and a tester for "triangle free"

A *first order graph property* is a property that can be expressed in the following form:

$$\forall u_1, \dots u_k \exists v_1, \dots v_l \forall \dots R(u_1, \dots u_k, v_1, \dots, v_l, \dots)$$

where R is a relation that consists of AND, OR, NOT and $\sim$ (where $u \sim v$ represents "there is an edge between u and v").

We will give a tester for triangle freeness in the adjacency matrix model. Observe that the property "triangle-free" is first order:

$$\forall v_1, v_2, v_3[(v_1 \neq v_2 \wedge v_2 \neq v_3 \wedge v_3 = v_1) \Rightarrow \neg(v_1 \sim v_2 \wedge v_2 \sim v_3 \wedge v_3 \sim v_1)]$$

The existence of our tester is merely a special case of the following vastly general result of Alon, Fischer, Krivilevich and Szegedy:

Theorem 5 *Every first order graph property of the "$\exists\forall$" type can be tested for ϵ-closeness in the adjacency matrix model with a number of queries that depends only on ϵ.*

We will prove the result only for triangle free, that too in the next class.