6.896 Sublinear Time Algorithms	October 5, 2004
Lecture 7	
Lecturer: Ronitt Rubinfeld	*Scribe: John Danaher*

1 The adjacency list model

Given $0 < \epsilon < 1$ and a bound m on the number of edges. One can query either the degree of a vertex u or the ith neighbor of a vertex u in one step. Two graphs are ϵ-far apart if changing one into the other requires changing (adding or deleting) at least ϵm edges.

(This model is necessary for Problem Set 1, problem 7.)

2 Testing triangle-freedom

2.1 Regularity

Definition 1 *For $A, B \subseteq V$ such that $A \cap B = \emptyset$, and $|A| > 1, |B| > 1$, we define $e(A, B)$ to be the number of edges between A and B. We also define $d(A, B)$, the density of edges between A and B, to be $\frac{e(A,B)}{|A||B|}$.*

Definition 2 *The pair of sets of vertices A, B is γ-regular if for all $A' \subseteq A, B' \subseteq B$ such that $|A'| \geq \gamma |A|$ and $|B'| \geq B$, it is true that $|d(A, B) - d(A', B')| < \gamma$.*

Intuitively, this means that for any big-enough A' and B', the "inner density" between A' and B' is close to the "outer density" between A and B. This is a property which random graph has, so regularity says, in some sense, that the graph is like a random graph. But the next theorem gives a sense that regular graphs definitely have some properties we don't expect random graphs to have.

Lemma 3 (Szemeredi's Regularity Lemma) *For all m and for any $\epsilon > 0$, there exists some T independent of $|V|$ such that if $G = (V, E)$ with $|V| > T$ and $\mathcal{A}$ is an equipartition of V into m sets, then there is an equipartition $\mathcal{B}$ which is a refinement of $\mathcal{A}$ and is made up of k sets satisfying (1) $m \leq k \leq T$ and (2) at most $\epsilon \binom{k}{2}$ pairs of sets are not ϵ-regular.*

Note that an equipartion is defined as a partition into subsets whose sizes differ by at most one. Also note that if $m = 1$ then the lemma is trivially true. Finally, note that T will be really, really huge, (in fact, a tower of 2's of length polynomial in $1/\epsilon$), so V will also be really, really huge, making this really a theoretical result.

2.2 Triangles in a random tripartite graph

Consider a random tripartite graph of density at least η. (That is, a graph constructed by equipartioning the vertices into three sets V_1, V_2, V_3, such that each V_i is an independent set, and for each pair of vertices u, v in different V_i's, an edge (u, v) is inserted with probability η.) How many triangles do we expect to see in this graph? Let $\sigma_{u,v,w}$ be an indicator variable for the event that vertices $u \in V_1, v \in V_2$, and $w \in V_3$ form a triangle.

$$\mathrm{E}[\sigma_{u,v,w}] = \Pr[u, v, w \text{ is } \Delta] = \eta^3$$

$$\mathrm{E}\left[\sum_{u,v,w} \sigma_{u,v,w}\right] = \sum_{u,v,w} \mathrm{E}[\sigma_{u,v,w}] = (n/3)^3 \eta^3$$

So we expect a random graph to contain $(\eta n/3)^3$ triangles.

2.3 Triangles in a regular tripartite graph

If we instead start from a tripartite graph in which each pair V_i, V_j is regular with density η, then we still have $\Omega((\eta n)^3)$ triangles (that is, when η is not too close to 1).

Lemma 4 (Komlos Simonovits for special case of triangles) *For all $\eta > 0$, there exists some γ^Δ and δ^Δ such that if A, B, C are disjoint subsets of V and each pair is γ-regular with density $\geq \eta$, then G contains at least $\delta^\Delta |A||B||C|$ triangles with a vertex each from A, B, and C.*

For the case of triangles, some particular values that work are $\gamma^\Delta = \frac{1}{2}\eta$ and $\delta^\Delta = \frac{1}{8}(1-\eta)\eta^3$.
Proof Let A^* be the set of nodes in A with at least $(\eta - \gamma^\Delta)|B|$ neighbors in B and at least $(\eta - \gamma^\Delta)|C|$ neighbors in C.

Claim 5 $|A^*| \geq (1 - 2\gamma^\Delta)|A|$.

Proof Let A' be the nodes in A with less than $(\eta - \gamma^\Delta)|B|$ neighbors in B. By the hypothesis, $d(A, B) \geq \eta$, and we also have $d(A', B) < \frac{(\eta-\gamma^\Delta)|B||A'|}{|B||A'|} = \eta - \gamma^\Delta$. So

$$|d(A, B) - d(A', B)| > \gamma^\Delta$$

and since the pair A, B is regular, we have (using $B' = B$) that $|A'| < \gamma^\Delta |A|$.
Similarly, we let A'' be the nodes in A with less than $(\eta - \gamma^\Delta)|C|$ neighbors in C, and see that $|A''| < \gamma^\Delta |A|$. So

$$|A^*| = |A \setminus (A' \cup A'')| \geq |A| - 2\gamma^\Delta|A| = (1 - 2\gamma^\Delta)|A|$$

■

Now, for any $u \in A^*$, let B_u and C_u be the neighbors of u in B and C, respectively. (Note that $|B_u| \geq (\eta - \gamma^\Delta)|B|$ and $|C_u| \geq (\eta - \gamma^\Delta)|C|$.) How many distinct edges are there between B_u and C_u? Since $d(B, C) \geq \eta$ and the pair B, C is regular, $d(B_u, C_u) > \eta - \gamma^\Delta$. So

$$e(B_u, C_u) > (\eta - \gamma^\Delta)|B_u||C_u| \geq (\eta - \gamma^\Delta)^3|B||C|$$

And how many triangles contain nodes in $|A^*|$ when all $\geq (1 - 2\gamma^\Delta)|A|$ choices of u are taken into account? Substituting for γ^Δ and δ^Δ, the number of triangles is at least

$$(1 - 2\gamma^\Delta)|A|(\eta - \gamma^\Delta)^3|B||C| = (1 - \eta)(\eta - \frac{1}{2}\eta)^3|A||B|C| = \frac{1}{8}(1 - \eta)(\eta^3)|A||B||C| = \delta^\Delta|A||B||C|$$

■

2.4 Triangles in a general graph

Theorem 6 (Alon simplification and strengthening of AFICS for triangles) *For all ϵ, there exists δ such that if graph G with $|V| = n$ is ϵ-far from triangle-free in the adjacency matrix model, then G contains at least $\delta\binom{n}{3}$ distinct triangles.*

Corollary 7 *The following algorithm gives a 1-sided test for triangle-freedom.*

1. *Do the following $O(1/\delta^\Delta)$ times:*
   (a) *Randomly pick vertices pick v_1, v_2, v_3.*
   (b) *If v_1, v_2, v_3 form a triangle, output FAIL.*

2. *Output PASS.*

We will fully prove the theorem in the next lecture. A sketch of the proof is as follows:

- Use the regularity lemma to get a partition $V_1 \ldots V_k$ such that $5/\epsilon \le k \le T(5/\epsilon, \epsilon')$ (that is, $\epsilon n/5 \ge n/k \ge n/T$) for some $\epsilon' = \min\{\epsilon/5, \gamma^{\Delta}\}$. The partition should create at most $\le \epsilon' \binom{\epsilon}{2}$ pairs of sets which are not regular.
- Construct G' by deleting from G (i) edges internal to any V_i (ii) edges between nonregular pairs of V_i's (iii) edges between "low density" pairs.
- Claim that this doesn't delete too many edges (less than ϵn^2), so there must still be a triangle left.
- Since any triangle still remaining must be between three pairwise-regular high-density sets, the lemma shows that there must be a lot of triangles left, and therefore there were a lot of triangles to begin with.