Title: 2,3,…,k: From approximating the number of edges to approximating the number of k-cliques (with a sublinear number of queries)

Abstract:

In this talk I will present an algorithms for approximating the number of k-cliques in a graph when given query access to the graph. This problem was previously studied for the cases of k=2 (edges) and k=3 (triangles). We give an algorithm that works for any k >= 3, and is actually conceptually simpler than the k=3 algorithm. We consider the standard query model for general graphs via (1) degree queries, (2) neighbor queries and (3) pair queries. Let n denote the number of vertices in the graph, m the number of edges, and C_k the number of k-cliques. We design an algorithm that outputs a (1+\epsilon)-approximation (with high probability) for C_k, whose expected query complexity and running time are $O\left(\frac{n}{C_k^{1/k}} + \frac{m^{k/2}}{C_k}\right)\text{poly}(\log n, 1/\epsilon, k)$.

Hence, the complexity of the algorithm is sublinear in the size of the graph for $C_k = \Omega(m^{k/2-1})$. Furthermore, we prove a lower bound showing that the query complexity of our algorithm is essentially optimal (up to the dependence on $\log n, 1/\epsilon$, and k).

This is joint work with Talya Eden and C. Seshadhri.