On Spanning Trees and Cuts

Persi Diaconis
Statistics Department, Stanford University,
Stanford, CA 94305, USA
diaconis@math.stanford.edu,
Tomás Feder
268 Waverley St., Palo Alto, CA 94301, USA
tomas@theory.stanford.edu,
Amin Saberi
Management Science and Engineering Department, Stanford University,
Stanford, CA 94305, USA
saberi@stanford.edu

December 7, 2009

1 Dominating Set

A dominating set in a graph G is a set S such that every vertex in G is either in S or a neighbor of a vertex in S. Let G be a graph with n vertices and degree measure d. Suppose we repeatedly select a vertex from G uniformly at random, until the selected vertices form a dominating set. Let T be the number of vertices that must be selected until the selected vertices form a dominating set.

Theorem 1.1 *If the least degree is d, then $E(T) = \Omega(n/(d+1))$.*

Proof. Suppose we select $k = \lfloor n/(d+1) \rfloor$ vertices. The probability that a vertex of least degree d is not dominated is $(1-(d+1)/n)^k = 1/e$, so with constant probability we will need at least k vertices, so the expected number of vertices needed is at least $E(T) = \Omega(k)$. □

Theorem 1.2 *A split graph with $d \geq 1$ vertices forming a clique and the remaining $n-d$ forming an independent set joined to the clique of size d by a complete bipartite graph has average degree $\Theta(d)$ and $E(T) = \Theta(n/(d+1))$, matching the lower bound in Theorem 1.1.*

Proof. The average degree is $(d(n-d)+(n-1)d)/d = \Theta(d)$. The selection of vertices ends when we select either all $n-d$ independent vertices or one of the d vertices in the clique. The probability of selecting a vertex in the clique of size d is d/n, so the expected time for selecting such a vertex is $\Theta(n/d)$, which is less than the expected time to select all $n-d$ independent vertices. □

Theorem 1.3 *If all degrees are at least d, $E(T) = O((n/(d+1)) \log n)$.*

Proof. Suppose all degrees are at least d. Then the probability that a vertex will not be dominated in $k = c(n/(d+1)) \log n$ steps is at most $(1-n/(d+1))^k = 1/n^c$, so the probability over all vertices is at most $1/n^{c-1}$. Therefore $E(T) = O(k)$. □

Theorem 1.4 *A random graph with each edge chosen with probability d/n for $d \geq n^c$ with $0 < c < 1$ constant has degrees $\Theta(d)$ and $E(T) = \Theta((n/(d+1)) \log n)$, matching the upper bound in Theorem 1.3.*

Proof. Suppose $(n/(d+1))i$ vertices have been chosen. The probability that a vertex of degree d will not be covered is e^{-i}, and the graph will stay essentially of uniform degree, so we expect n/e^{-i} vertices to need to be covered, so we must choose $i = \Omega(\log n)$ to cover all the vertices. □

Theorem 1.5 *If all degrees are at most d, $E(T) = \Omega((n/(d+1)) \log(n/(d+1)))$ improves the tight lower bound Theorem 1.1, and matches the upper bound in Theorem 1.3 for $d \leq n^c$ with $0 < c < 1$ constant.*

Proof. Suppose k vertices are dominated. Then the number of edges that could be used in one direction for domination of a new vertex is $d(n-k)$, so we expect to dominate at most $D = (d+1)(n-k)/n$ new vertices. It will thus take $\Omega(n/(n-k))$ chosen vertices to achieve $d+1$ new vertices dominated in expectation. Considering the $\log(n/d)$ values i with $(d+1)2^i \leq n-k < (d+1)2^{i+1}$ each thus take chosen vertices in expectation, up to constant factors, the result follows. □

Theorem 1.6 *If G is a d-regular graph with $n = k(d+1)$ vertices consisting of k disjoint cliques of size $d+1$, then $E(T) = \Theta((n/(d+1)) log(n/(d+1)))$, matching the lower bound in Theorem 1.5.*

Proof. We must reduce the probability that a clique is not chosen to at most $1/k = (d+1)/n$, and a clique is not chosen at each iteration with probability $1-(d+1)/n$, so the result follows by noticing that $(1-(d+1)/n)^i = (d+1)/n$ solves to $i = (n/(d+1)) \log(n/(d+1))$. □

2 Vertex Cover

A hypergraph G is given by a vertex set V with $|V| = n$ and a collection S of subsets of V with $|S| = m$. A vertex cover in G is a collection U of vertices of V such that each set of S has at least one vertex in U. We consider the problem of repeatedly selecting vertices in G at random until a vertex cover is obtained, and denote by T the number of steps this takes. We wish to estimate $E(T)$.

Theorem 2.1 *If every set of S has size bounded by some constant c, and the minimum vertex cover has size v, then $E(T) = \Theta(n \log v)$.*

Proof. Let W be a maximal collection of disjoint sets in S. Then U must choose at least one vertex in each set of W, and a solution U can be found by choosing all the vertices of all sets of W. Thus $|W| \leq v \leq c|W|$. Since the probability of choosing any element of a set in W is between $1/n$ and c/n, the probability of choosing some such element is $\Theta(v/n)$, so in expectation we will choose such an element every $\Theta(n/v)$ choices, and to cover all $\Theta(v)$ such sets we need $\Theta(v \log v)$ such choices in expectation by the coupon collecting problem. Thus in total we get expected number of samples $\Theta((n/v)(v \log v)) = \Theta(n \log v)$. □

Theorem 2.2 *If the sets of S have least size s, then $E(T) = \Omega(n/s)$ and $E(T) = O((n/s)\log m)$.*

Proof. A set is covered with probability s/n, so it is not covered in n/s with probability $(1 - n/s)^{n/s} \geq 1/e$, so $E(T) = \Omega(n/s)$. If we choose $c(n/s)\log m$ elements, the probability that any of the m sets is not covered is $(1 - n/s)^{c(n/s)\log m} \leq 1/m^c$, so with probability at most $1/m^{c-1}$ not all sets will be covered, proving $E(T) = O((n/s)\log m)$. $\square$

Theorem 2.3 *Suppose $n = \binom{m}{k}$, and each element of V chooses k sets of the m sets of S in each of the possible ways. Then $E(T) = \Theta(\log m/\log(1/(1-k/m))$, in particular, $E(T) = \Theta((m/k)\log m)$ if $k \leq n/2$.*

Proof. Write $k = (1-\alpha)m$. The probability that a set will not be covered by a choice is α, so the number of sets left to be covered is reduced to $m\alpha^t$ after t choices in expectation, so we must choose t so that $m\alpha^t < 1$, or $(1/\alpha)^t > m$, $t = Theta(\log m/\log(1/\alpha))$ as desired. $\square$