Counting the Number of Hamilton Cycles in Random Digraphs

Alan Frieze*and Stephen Suen
Department of Mathematics,
Carnegie Mellon University,
Pittsburgh, U.S.A.

February 23, 2002

Abstract

We show that there exists a *a fully polynomial randomized approximation scheme* for counting the number of Hamilton cycles in almost all directed graphs.

1 INTRODUCTION

In this note we consider the problem of counting the number of Hamilton cycles $h = h(D)$ in a digraph D. More precisely we consider the possibility of computing an estimate $\hat{h} = \hat{h}(D)$ which satisfies

$$(1 - \epsilon)\hat{h} \leq h \leq (1 + \epsilon)\hat{h}, \tag{1}$$

*Supported by NSF grant CCR-8900112

where ϵ is the required accuracy.

We consider randomised algorithms and so we only require (1) to hold with probability $\geq 1-\delta$, where δ is given as input. A randomised algorithm is called a Fully Polynomial Time Approximation Scheme (FPRAS) if it runs in time which is polynomial in $|V(D)|, 1/\epsilon$ and $\log 1/\delta$. Since it is NP-Complete to determine whether or not D has a Hamilton cycle, we do not expect there to be an FPRAS which works for all inputs D, unless RP=NP.

In this paper we consider a scheme which is likely to be efficient when the input is the random graph $D_{n,m}$ which has vertex set $\{1, 2, \ldots, n\}$ and m random directed edges. We will also require that $m^3/n^2 \to \infty$ with n.

So now let H_n denote the (random) number of Hamilton cycles in $D_{n,m}$. Associated with a digraph D is a bipartite graph $B(D)$ which has vertex sets $V(D), V'(D) = \{v' : v \in V(D)\}$ and an edge (u, v') iff D has a directed edge (u, v).

It is well known that each perfect matching M of $B(D)$ corresponds to a unique set of vertex disjoint cycles $C(M)$ which cover all vertices of D. Let M_n denote the number of perfect matchings in $B_{n,m} = B(D_{n,m})$. Our algorithm is very simple:

1. Estimate M_n to within a factor $1 \pm \epsilon/3$.

2. Estimate H_n/M_n to within a factor $1 \pm \epsilon/3$.

3. Multiply these estimates together.

Jerrum and Sinclair [3] have proved the existence of an FPRAS which works for almost all $B_{n,m}$ and this can be applied to carry out Step 1.

For Step 2 we generate (near) random perfect matchings M in $B_{n,m}$ and count the proportion of times that $C(M)$ has one cycle i.e. *is* a Hamilton cycle. (The generation of near random perfect matchings is part of the scheme in [3].) It is a standard observation in this area of computation that if H_n/M_n is bounded below by $1/p(n)$ for some polynomial $p(n)$ then $O(p(n)\epsilon^{-2}\log 1/\delta)$ trials are sufficient to estimate H_n/M_n with the required accuracy.

Now

$$\mathbf{E}(M_n) = \mu_{n,m} = n!\frac{\binom{n^2-n}{m-n}}{\binom{n^2}{m}}$$

and Jerrum [2] has recently shown that, for example,

$$\mathbf{Pr}[M_n \geq 2\mu_{m,n}] = O\left(\frac{n^2}{m^3}\right). \tag{2}$$

On the other hand

$$\mathbf{E}(H_n) = (n-1)!\frac{\binom{n^2-n}{m-n}}{\binom{n^2}{m}}$$

and the main result of this paper comes from using the second moment method to prove that

$$\mathbf{Pr}\left[H_n \leq \frac{(n-1)!}{2}\frac{\binom{n^2-n}{m-n}}{\binom{n^2}{m}}\right] = O\left(\frac{n^2}{m^3}\right). \tag{3}$$

Thus

$$\mathbf{Pr}(M_n \geq 4nH_n) = O\left(\frac{n^2}{m^3}\right)$$

and our scheme will work with high probability when $m^3/n^2 \to \infty$ with n.

It can be seen from our expression for $E(H_n)$ that we have assumed that $D_{n,m}$ can contain loops. If we do not allow loops then our result still holds.

To see this we simply add a (random) number of loops and apply the previous analysis. The number of Hamilton cycles stays the same and the number of perfect matchings cannot decrease and so we see that we do no worse in the loop-free case.

2 SECOND MOMENT CALCULATION

We will show here that when $m^2/n^3 \to \infty$ as $n \to \infty$,

$$\mathbf{E}[H_n^2] = (1 + O(n^3/m^2))(\mathbf{E}H_n)^2. \tag{4}$$

Note that

$$\begin{aligned}
\mathbf{E}[H_n] &= (n-1)!\binom{n^2-n}{m-n}\binom{n^2}{m}^{-1} \\
&= (n-1)!\frac{m(m-1)\dots(m-n+1)}{n^2(n^2-1)\dots(n^2-n+1)} \\
&= (n-1)!\left(\frac{m}{n^2}\right)^n \exp\left(-\frac{n(n-1)}{2m} + \frac{n(n-1)}{2n^2} + \mathrm{O}\left(\frac{n^3}{m^2} + \frac{1}{n}\right)\right) \\
&= (n-1)!\left(\frac{m}{n^2}\right)^n \exp\left(-\frac{n^2}{2m} + \frac{1}{2} + \mathrm{O}\left(\frac{n}{m} + \frac{n^3}{m^2} + \frac{1}{n}\right)\right),
\end{aligned}$$

and since $m^2/n^3 \to \infty$ as $n \to \infty$, we have

$$\mathbf{E}[H_n] = (n-1)!\left(\frac{m}{n^2}\right)^n \exp\left(-\frac{n^2}{2m} + \frac{1}{2}\right)\left(1 + \mathrm{O}\left(\frac{n^3}{m^2}\right)\right). \tag{5}$$

We next would like to estimate $\mathbf{E}[H_n^2]$. Given a Hamilton cycle H in the complete digraph DK_n with n vertices, where $n \geq 2$, we first obtain a formula for the number $f_n(t)$ of Hamilton cycles H' in DK_n such that H and H' have exactly t edges in common. Fix t ($t \leq n-2$) edges on H. Suppose that these t edges form k paths with $t+k$ vertices on these paths. Note that a Hamilton cycle H' in DK_n containing these t edges can be uniquely identified by a Hamilton cycle on DK_{n-t} with k distinguished vertices representing the k paths. Therefore, given t specified edges on H, the number of Hamilton cycles H' in DK_n that have exactly the t specified edges in common with H is equal to $f_{n-t}(0)$. That is, for $t = 0, 1, \ldots, n-2$,

$$f_n(t) = \binom{n}{t} f_{n-t}(0).$$

Note that $f_n(n-1) = 0$ and $f_n(n) = 1$.

We next proceed to find the number $f_n(0)$ (where $n \geq 2$) of Hamilton cycles in DK_n that has no edges in common with H. Assume now that the Hamilton cycle H has edges $e_1, e_2, \ldots, e_n$. For $i = 1, 2, \ldots, n$, let A_i be the set of all Hamilton cycles in DK_n containing e_i. Note that $|A_1 \cap A_2 \cap \ldots \cap A_k|$ is the number of Hamilton cycles in DK_n containing edges $e_1, e_2, \ldots, e_k$ and by previous reasoning, is equal to $(n-k-1)!$ when $k \leq n-2$. Note that when $k = n-1$ or n, $|A_1 \cap A_2 \cap \ldots \cap A_k|$ is equal to 1. Thus using the principle of inclusion-exclusion, we have for $n \geq 2$,

$$\begin{aligned} f_n(0) &= \sum_{k=0}^{n} \binom{n}{k} (-1)^k |A_1 \cap A_2 \cap \ldots \cap A_k| \\ &= \sum_{k=0}^{n-1} \binom{n}{k} (-1)^k (n-k-1)! + (-1)^n. \end{aligned}$$

Since $f_1(0) = 0$ we may take $f_n(t) = \binom{n}{t} f_{n-t}(0)$ when $t \leq n-1$ and $f_n(n) = 1$. Writing $g_n(t) = \binom{n^2-2n+t}{m-2n+t} / \binom{n^2}{m}$, it follows that

$$\mathbf{E}[H_n^2] = (n-1)! \sum_{t=0}^{n} f_n(t) g_n(t)$$

$$
\begin{aligned}
&= (n-1)! \sum_{t=0}^{n-1} \binom{n}{t} f_{n-t} g_n(t) + (n-1)! g_n(n) \\
&= (n-1)!^2 \sum_{t=0}^{n-1} \sum_{k=0}^{n-t-1} \frac{(-1)^k n g_n(t)}{t! k! (n-t-k)} + (n-1)!^2 \sum_{t=0}^{n} \frac{(-1)^{n-t} n g_n(t)}{t!(n-t)!} \\
&= (n-1)!^2 \left(S'_n + S''_n\right), \qquad (6)
\end{aligned}
$$

where

$$
S'_n = \sum_{t=0}^{n-1} \sum_{k=0}^{n-t-1} \frac{(-1)^k n g_n(t)}{t! k! (n-t-k)} \quad \text{and} \quad S''_n = \sum_{t=0}^{n} \frac{(-1)^{n-t} n g_n(t)}{t!(n-t)!}.
$$

Hence, in view of equations(5), (6) and the fact that $m^2/n^3 \to \infty$ as $n \to \infty$, it will suffice to show that as $n \to \infty$,

$$
S'_n + S''_n = (1 + \mathrm{o}(1)) \left(\frac{m}{n^2}\right)^{2n} \exp\left(-\frac{n^2}{m} + 1\right).
$$

Now

$$
S''_n = \frac{(-1)^n}{(n-1)!\binom{n^2}{m}} \sum_{t=0}^{n} (-1)^t \binom{n}{t} \binom{n^2-2n+t}{n^2-m},
$$

and it is not difficult to show that (see Graham, Knuth and Patashnik [1](5.24))

$$
\sum_{t=0}^{n} (-1)^t \binom{n}{t} \binom{n^2-2n+t}{n^2-m} = (-1)^n \binom{n^2-2n}{n^2-m-n},
$$

and that since $n^2 \geq m \geq n$,

$$
\frac{\binom{n^2-2n}{n^2-m-n}}{\binom{n^2}{m}} \leq \left(\frac{m}{n^2}\right)^n.
$$

Thus

$$
S''_n \leq \frac{1}{(n-1)!} \left(\frac{m}{n^2}\right)^n. \qquad (7)
$$

To estimate S'_n, we note that

$$\begin{aligned}
S'_n &= \sum_{t=0}^{n-1}\sum_{k=0}^{n-t-1} \frac{(-1)^k n g_n(t)}{t!k!(n-t-k)} \\
&= \sum_{t=0}^{n-1}\sum_{k=0}^{n-t-1} \frac{(-1)^k g_n(t)}{t!k!} + \sum_{t=0}^{n-2}\sum_{k=0}^{n-t-2} \frac{(-1)^k g_n(t+1)}{t!k!(n-t-k-1)} \\
&\quad + \sum_{t=0}^{n-2}\sum_{k=0}^{n-t-2} \frac{(-1)^k g_n(t)}{t!k!(n-t-k-1)} \\
&= \sum_{t=0}^{n-1}\sum_{k=0}^{n-t-1} \frac{(-1)^k g_n(t)}{t!k!} \\
&\quad + \sum_{t=0}^{n-2}\sum_{k=0}^{n-t-2} \frac{(-1)^k g_n(t)}{t!k!(n-t-k-1)} \left(\frac{n^2-2n+t+1}{m-2n+t+1} - 1\right) \\
&= T'_n + T''_n, \text{ say} \qquad (8)
\end{aligned}$$

where

$$T'_n = \sum_{t=0}^{n-1}\sum_{k=0}^{n-t-1} \frac{(-1)^k g_n(t)}{t!k!} \text{ and } T''_n = \sum_{t=0}^{n-2}\sum_{k=0}^{n-t-2} \frac{(-1)^k g_n(t) h_n(t)}{t!k!(n-t-k-1)},$$

and

$$h_n(t) = \frac{n^2-2n+t+1}{m-2n+t+1} - 1.$$

We next choose τ as the greatest even integer not greater than $n/4$ and we write

$$T'_n = \sum_{t=0}^{\tau}\sum_{k=0}^{\tau} \frac{(-1)^k g_n(t)}{t!k!} + \sum_{t=\tau+1}^{n-1}\sum_{k=0}^{\tau} \frac{(-1)^k g_n(t)}{t!k!} + \sum_{t=0}^{\tau}\sum_{k=\tau+1}^{n-t-1} \frac{(-1)^k g_n(t)}{t!k!}.$$

Note that

$$\sum_{k=0}^{\tau} \frac{(-1)^k}{k!} \leq \exp(-1),$$

and that

$$g_n(t)/t! \text{ is non-increasing wrt } t \text{ for } t \in \{\tau, \ldots, n-1\}.$$

Hence,

$$T_n' \leq e^{-1} \sum_{t=0}^{\tau} \frac{g_n(t)}{t!} + n^2 \frac{g_n(\tau)}{\tau!} + \frac{n}{\tau!} \sum_{t=0}^{\tau} \frac{g_n(t)}{t!}. \tag{9}$$

For T_n'', note first that for $t \leq n/4$,

$$\sum_{k=0}^{\tau} \frac{(-1)^k}{k!(n-t-k)} \leq \frac{1}{(n-t)},$$

and that

$$h_n(t) = \frac{n^2 - 2n + t + 1}{m - 2n + t + 1} - 1 = \mathrm{O}\left(\frac{n^2}{m}\right).$$

So it follows similarly as above that

$$\begin{aligned} T_n'' &\leq \sum_{t=0}^{\tau} \frac{g_n(t)}{t!(n-t)} + \mathrm{O}\left(\frac{n^2}{m}\right) \frac{n^2 g_n(\tau)}{\tau!} + \frac{1}{\tau!} \mathrm{O}\left(\frac{n^2}{m}\right) \sum_{t=0}^{\tau} \frac{n g_n(t)}{t!} \\ &\leq \frac{2}{n} \sum_{t=0}^{\tau} \frac{g_n(t)}{t!} + \mathrm{O}\left(\frac{n^4}{m}\right) \frac{g_n(\tau)}{\tau!} + \mathrm{O}\left(\frac{n^3}{m\tau!}\right) \sum_{t=0}^{\tau} \frac{g_n(t)}{t!}. \end{aligned} \tag{10}$$

We therefore have from equations (8), (9) and (10) that

$$S_n' \leq \left(1 + \mathrm{O}\left(n^{-1}\right)\right) e^{-1} \sum_{t=0}^{\tau} \frac{g_n(t)}{t!} + \mathrm{O}\left(\frac{n^4 g_n(\tau)}{m\tau!}\right). \tag{11}$$

We next note that writing $M = M(n) = m^{-1} - n^{-2}$,

$$\begin{aligned} g_n(t) &= \binom{n^2 - 2n + t}{m - 2n + t} \binom{n^2}{m}^{-1} \\ &= \left(\frac{m}{n^2}\right)^{2n-t} \prod_{i=1}^{2n-t-1} \left(\frac{1 - i/m}{1 - i/n^2}\right) \\ &\leq \left(\frac{m}{n^2}\right)^{2n-t} \exp\left(-\frac{1}{2}(2n - t - 1)^2 M\right). \end{aligned} \tag{12}$$

Hence,

$$\sum_{t=0}^{\tau} \frac{g_n(t)}{t!} \leq \sum_{t=0}^{\tau} \frac{1}{t!} \left(\frac{m}{n^2}\right)^{2n-t} \exp\left(-\frac{1}{2}(2n - t - 1)^2 M\right)$$

$$
\begin{aligned}
&\leq \sum_{t=0}^{\tau} \frac{1}{t!} \left(\frac{m}{n^2}\right)^{2n-t} \exp\Big(-(2n^2 - 2nt - 2n)M\Big) \\
&\leq \left(\frac{m}{n^2}\right)^{2n} \exp(-2n(n-1)M) \sum_{t=0}^{\infty} \frac{1}{t!} \left(\frac{n^2 e^{2nM}}{m}\right) \\
&= \left(\frac{m}{n^2}\right)^{2n} \exp\left(-\frac{n^2}{m} + 2\right) \left(1 + \mathrm{O}\left(\frac{n^3}{m^2}\right)\right)
\end{aligned}
$$

and using inequality(12) again,

$$
\frac{g_n(\tau)}{\tau!} \leq \left(\frac{m}{n^2}\right)^{2n} \left(\frac{n^2 e}{m\tau}\right)^{\tau}.
$$

With the above estimates, we have from equations (7) and (11) that

$$
S_n' + S_n'' \leq \left(\frac{m}{n^2}\right)^{2n} \exp\left(-\frac{n^2}{m} + 1\right) \left(1 + \mathrm{O}\left(\frac{n^3}{m^2}\right)\right).
$$

(4) now follows from (5) and (6).

Final remarks: a similar scheme seems likely to work for undirected graphs, but the calculations are more arduous and will be left for a later paper.

References

[1] R.Graham, D.Knuth and O.Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, Addison-Wesley, new York, 1989.

[2] M.Jerrum, *An analysis of a Monte carlo algorithm for estimating the permanent*, to appear.

[3] M.Jerrum and A.Sinclair, *Approximating the permanent*, SIAM Journal on Computing 18 (1989) 1149-1178.