Paths of specified length in random k-partite graphs

C.R. Subramanian

Discrete Mathematics and Theoretical Computer Science **4**, 2001, 133–138

Paths of specified length in a random k-partite graph

C.R. Subramanian[†]

Department of Computer Science and Automation,
Indian Institute of Science, Bangalore-560012, INDIA.
email*: crs@csa.iisc.ernet.in*

received Aug 22, 2000, accepted Aug 10, 2001.

Fix positive integers k and l. Consider a random k-partite graph on n vertices obtained by partitioning the vertex set into $V_i, (i = 1, \ldots, k)$ each having size $\Omega(n)$ and choosing each possible edge with probability p. Consider any vertex x in any V_i and any vertex y. We show that the expected number of simple paths of even length l between x and y differ significantly depending on whether y belongs to the same V_i (as x does) or not. A similar phenomenon occurs when l is odd. This result holds even when k, l vary slowly with n. This fact has implications to coloring random graphs. The proof is based on establishing bijections between sets of paths.

Keywords: random graphs, paths, bijections

1 Motivation

This problem arose in the analysis of algorithms for coloring random k-colorable graphs [2, 3]. Consider a random graph drawn as explained in the abstract. To separate a color class, we fix a vertex x in the largest (or smallest) V_i and and compute the number of l-paths (paths of length l), $n(x,y,l)$, between x and an arbitrary vertex y. Depending on whether y belongs to the same class as x belongs to, the expectation of this quantity differs significantly. If we can show that $n(x,y,l)$ is close to its expected value almost surely, this gives us a way of separating the class containing x. Repeating this $k-2$ times, one gets a k-coloring. The expectation of $n(x,y,l)$ is $N(x,y,l)p^l$, where $N(x,y,l)$ is the total number of l-paths in the complete k-partite graph formed by V_is. The result stated in the abstract shows that the expectations differ significantly as required.

We do not discuss the algorithmic issues here since they have been outlined in [2]. We only prove the results stated in the abstract using only counting arguments. Even though the results are obviously true for bipartite graphs, for $k \geq 3$, it is not so straightforward. We believe the arguments used here would be of interest to know. The basic idea is to partition (for each pair of start-end vertices) the corresponding set of l-paths into groups (based on the color classes of intermediate vertices). Then, for two different

[†]Present address: The Institute of Mathematical Sciences, Taramani, Chennai - 600 113, INDIA. *email :* crs@imsc.ernet.in

pairs, we establish an (almost) bijection between the corresponding groups. For any such mapped pair of groups, we also establish an (almost) bijection between the l-paths in them. This establishes the required statement.

2 Paths of specified length

Definition 2.1 *By an l-path between two vertices x and y, we mean a* simple path *of length l between x and y. A simple path is one in which no vertex appears more than once. An l-path is represented as a $(l+1)$-tuple $\langle x, v_1, \ldots, v_{l-1}, y\rangle$ of vertices such that successive vertices in this sequence belong to different partite sets V_i.*

Notations : G is a *complete* k-partite graph on the partite sets $V_1, \ldots, V_k$ with each $|V_i| \geq n/C$ for some constant $C \geq k$. For each i, n_i denotes the size $|V_i|$. For each i, let $W_i \doteq V_i \cup \ldots \cup V_k$. For all $i (1 \leq i \leq k-1)$, for all $x \in V_i$, for all $y \in W_i$ such that $y \neq x$, let $N(x,y,l,i)$ denote the number of l-paths between x and y involving only vertices from W_i. Given a tuple σ with integral component values and an integer j, $c(\sigma, j)$ denotes the number of times j appears in σ.

We obtain the following results.

Theorem 2.1 *Assume that $n_1 \leq \ldots \leq n_k$. Let l be any fixed* even *integer ≥ 2. For all i, $1 \leq i \leq k-1$, for all $x, y \in V_i$, for all $z \in W_i - V_i$, we have*

$$\begin{aligned} N(x,z,l,i) &= \Theta(n^{l-1}) \text{ if } i \leq k-2 \\ N(x,y,l,i), N(x,y,l,i) - N(x,z,l,i) &= \Theta(n^{l-1}) \end{aligned}$$

Proof: Consider any $i (i = 1, \ldots, k-1)$ and any $x, y \in V_i$ and $z \in V_r, r > i$ and fix these parameters. We use the factorial functions defined as follows : $(n)_0 = 1$. $(n)_l = n(n-1)\ldots(n-l+1)$, $l \geq 1$. Let $P(x,y)$ denote the set of all l-paths between x and y involving only vertices from the $k-i+1$ partite sets V_j $(i \leq j \leq k)$. $P(x,z)$ is defined similarly. That is,

$P(x,y) = \{ \langle x, v_1, \ldots, v_{l-1}, y\rangle \mid$ the sequence is an l-path between x and $y\}$.

$P(x,z) = \{ \langle x, v_1, \ldots, v_{l-1}, z\rangle \mid$ the sequence is an l-path between x and $z\}$.

Clearly, we have $|P(x,y)| = O(n^{l-1})$ and $|P(x,z)| = O(n^{l-1})$. Also if $i = k-1$, then there are only two partite sets, namely, V_{k-1} and V_k and hence $P(x,z) = \emptyset$ and $N(x,z,l,i) = 0$. Define

$B_l^s = \{ \langle \sigma_1, \ldots, \sigma_{l-1}\rangle \mid \sigma_1 \neq i, \sigma_{l-1} \neq i, i \leq \sigma_j \leq k, \sigma_j \neq \sigma_{j+1} \forall j\}$

$B_l^d = \{ \langle \sigma_1, \ldots, \sigma_{l-1}\rangle \mid \sigma_1 \neq i, \sigma_{l-1} \neq r, i \leq \sigma_j \leq k, \sigma_j \neq \sigma_{j+1} \forall j\}$

In the above, the superscript s (or d) is a short notation for the word "same" (or "different"). We have $|B_l^s|, |B_l^d| \leq k^{l-1}$.

Now $f : P(x,y) \rightarrow B_l^s$ is a mapping which identifies each l-path $\langle x, v_1, \ldots, v_{l-1}, y\rangle$ with the unique $(l-1)$-tuple $\langle \sigma_1, \ldots, \sigma_{l-1}\rangle$ in B_l^s where if $v_m \in V_j$ then $\sigma_m = j$. Similarly, we can define a mapping $g : P(x,z) \rightarrow B_l^d$ which identifies each l-path in $P(x,z)$ with a unique $(l-1)$-tuple in B_l^d. We use the elements of B_l^s (or B_l^d) to partition the set $P(x,y)$ (or $P(x,z)$) as follows.

$P(x,y) = \bigcup_{\sigma \in B_l^s} P_\sigma$ where $P_\sigma = \{\tau \in P(x,y) \mid f(\tau) = \sigma\}$.

$P(x,z) = \bigcup_{\sigma \in B_l^d} P_\sigma$ where $P_\sigma = \{\tau \in P(x,z) \mid g(\tau) = \sigma\}$.

Now, for each $\sigma \in B_l^s \cup B_l^d$, $|P_\sigma| = (\prod_{i \leq j \leq k} (n_j)_{c(\sigma,j)}) = (\prod_{i \leq j \leq k} (n_j)^{c(\sigma,j)}) \cdot [1 - o(1)]$. As a result, for each $\sigma \in B_l^s \cup B_l^d$, $|P_\sigma| = \Theta(n^{l-1})$. The $[1 - o(1)]$ factor arises not only because of factorials, but also because x, y and z have to be excluded from consideration.

Also B_l^s is non-empty and it contains at least one element, namely, the tuple $\langle r, i, r, i, \ldots, r \rangle$. Hence $N(x,y,l,i) = |P(x,y)| = \Theta(n^{l-1})$. Also, if $i \leq k-2$, then there are at least 3 partite sets to be considered and hence B_l^d is non-empty. Hence $N(x,z,l,i) = |P(x,z)| = \Theta(n^{l-1})$ *if* $i \leq k-2$.
We need to prove that $|P(x,y)| - |P(x,z)| = \Theta(n^{l-1})$. In order to prove this, it is enough to prove that the following two assertions are true.

1. $|B_l^s| \geq |B_l^d| + 1$ and

2. There exists a *one-to-one* mapping $h : B_l^d \to B_l^s$ such that for each $\tau \in B_l^d$, we have $|P_{h(\tau)}| \geq |P_\tau|[1 - o(1)]$.

We prove that the two assertions are true as follows. Now, partition B_l^s, B_l^d into

$B_l^s = B_{l,1}^s \cup \ldots \cup B_{l,l-1}^s \cup B_{l,l}^s$

$B_l^d = B_{l,2}^d \cup \ldots \cup B_{l,l-1}^d \cup B_{l,l}^d$

where

$B_{l,l}^s = \{\sigma \in B_l^s \mid \sigma_{l-1} \neq i, \sigma_{l-1} \neq r\}$.

$B_{l,l}^d = \{\sigma \in B_l^d \mid \sigma_{l-1} \neq i, \sigma_{l-1} \neq r\}$.

$B_{l,j}^s = \{\sigma \in B_l^s \mid \sigma_{j-1} \neq i, \sigma_{j-1} \neq r,\ \sigma_m = i, r\ for\ m \geq j\}$, for $2 \leq j \leq l-1$.

$B_{l,j}^d = \{\sigma \in B_l^d \mid \sigma_{j-1} \neq i, \sigma_{j-1} \neq r,\ \sigma_m = i, r\ for\ m \geq j\}$, for $2 \leq j \leq l-1$.

$B_{l,1}^s = \{\langle r, i, r, i, \ldots, r \rangle\}$

Now $B_{l,1}^d$ cannot be defined similarly since l is even. It is easy to see that the definitions form a well-defined partition of B_l^s and B_l^d. In other words, for each $\sigma \in B_l^s$, there exists a unique value of j between 1 and l such that $\sigma \in B_{l,j}^s$. Similarly, for each $\tau \in B_l^d$, there exists a unique value of j between 2 and l such that $\tau \in B_{l,j}^d$.
Now we claim that for all j such that $2 \leq j \leq l$, $|B_{l,j}^s| = |B_{l,j}^d|$. For $j = l$, this follows from $B_{l,l}^s = B_{l,l}^d$. For $j < l$, consider the mapping $h_j : B_{l,j}^d \to B_{l,j}^s$ defined as follows. Let $\tau \in B_{l,j}^d$ be any tuple. Then, $h_j(\tau) = \sigma$ where σ is defined as

- For all m $(1 \leq m \leq j-1)$, $\sigma_m = \tau_m$.

- For all m such that $j \leq m \leq l-1$, $\sigma_m = i$ if $\tau_m = r$ and $\sigma_m = r$ if $\tau_m = i$.

Clearly $\sigma \in B^s_{l,j}$. Also it can be verified that h_j is a one-to-one and onto mapping. Since $B^s_{l,j}$ and $B^d_{l,j}$ are finite sets, it follows that $|B^s_{l,j}| = |B^d_{l,j}|$.
Thus, we have $|B^s_l| \geq |B^d_l| + 1$ and the first assertion is true.
To prove the second assertion, define the mapping $h : B^d_l \rightarrow B^s_l$ to be as follows. For each $\tau \in B^d_l$, define $h(\tau) = h_j(\tau)$ where j is such that $\tau \in B^d_{l,j}$. Clearly, h is a one-to-one mapping since each h_j is a one-to-one mapping.
We prove that for each $\tau \in B^d_l$, we have $|P_{h(\tau)}| \geq |P_\tau|[1 - o(1)]$. Let $\tau \in B^d_l$ be any tuple and let σ denote the tuple $h(\tau)$. We know $\tau \in B^d_{l,j}$ for some j, $2 \leq j \leq l$.
If $j = l$, then we have $\sigma = \tau$ and hence $|P_\sigma| \geq |P_\tau|[1 - o(1)]$.
If $j = l-2, l-4, \ldots, 2$, then clearly, $c(\tau, m) = c(\sigma, m)$ for all values of m $(i \leq m \leq k)$ and hence $|P_\sigma| \geq |P_\tau|[1 - o(1)]$.
If $j = l-1, l-3, \ldots, 3$, then clearly, $c(\tau, m) = c(\sigma, m)$ for all values of m $(i \leq m \leq k)$ such that $m \neq i$, $m \neq r$. Also, $c(\sigma, r) = c(\tau, r) + 1$ and $c(\tau, i) = c(\sigma, i) + 1$. Since $n_i \leq n_r$ $(r > i)$ by assumption, we have $|P_\sigma| \geq |P_\tau|[1 - o(1)]$.

Thus, we have

$$
\begin{aligned}
N(x,y,l,i) - N(x,z,l,i) &= |P(x,y)| - P(x,z)| \\
&= |\bigcup_{\sigma \in B^s_l} P_\sigma| - |\bigcup_{\tau \in B^d_l} P_\tau| \\
&= \sum_{j=l-1,\ldots,3} \left(\sum_{\sigma \in B^s_{l,j}} |P_\sigma| \right) + \sum_{j=l,l-2,\ldots,2} \left(\sum_{\sigma \in B^s_{l,j}} |P_\sigma| \right) + \sum_{\sigma = \langle r,i,\ldots,r \rangle} |P_\sigma| \\
&\quad - \sum_{j=l-1,\ldots,3} \left(\sum_{\tau \in B^d_{l,j}} |P_\tau| \right) - \sum_{j=l,l-2,\ldots,2} \left(\sum_{\tau \in B^d_{l,j}} |P_\tau| \right) \\
&\geq |P_\sigma| - o(|P_\sigma|) \text{ where } \sigma = \langle r, i, \ldots, r \rangle. \qquad (1)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
N(x,y,l,i) - N(x,z,l,i) &= \Theta((n_r)_{l/2}(n_i)_{l/2-1}) \\
&= \Theta(n^{l-1})
\end{aligned}
$$

Hence,

$$
\begin{aligned}
N(x,z,l,i) &= \Theta(n^{l-1}) \text{ if } i \leq k-2 \\
N(x,y,l,i), N(x,y,l,i) - N(x,z,l,i) &= \Theta(n^{l-1}).
\end{aligned}
$$

This completes the proof of the theorem. ∎

Using similar arguments, we can prove the following theorem also.

Theorem 2.2 *Assume that* $n_1 \geq \ldots \geq n_k$. *Let* l *be any fixed* odd *integer* ≥ 3. *For all* $i\,(1 \leq i \leq k-1)$, *for all* $x, y \in V_i$, *for all* $z \in W_i - V_i$, *we have*

$$\begin{aligned} N(x,y,l,i) &= \Theta(n^{l-1}) \text{ if } i \leq k-2 \\ N(x,z,l,i),\, N(x,z,l,i) - N(x,y,l,i) &= \Theta(n^{l-1}) \end{aligned}$$

3 Conclusions

1. The main result of the paper is that the number of l-paths joining a vertex x (in the largest or smallest V_i depending on the parity of l) and a vertex y differs significantly depending on where y comes from. A close look at the proof (particularly, derivation of (1)) shows that this holds even if we allow k, l and C to vary with n, provided $lk^l C^l = o(n)$.

2. It would be interesting to extend these results to structures other than simple paths. Such results can be applied to the design and analysis of efficient algorithms for random graphs (see [1] for a survey).

References

[1] A.M. Frieze and C. McDiarmid, "Algorithmic Theory of Random Graphs", *Random Structures and Algorithms*, 10:5-42, 1997.

[2] C.R. Subramanian, "Improved Algorithms for Coloring Random Graphs", Proceedings of the *Fifth International Symposium on Algorithms and Computation*, 1994, Springer-Verlag LNCS 834.

[3] C.R. Subramanian, "Algorithms for Coloring Random and Semi-Random Graphs", Phd thesis, Department of Computer Science and Automation, Indian Institute of Science, Bangalore, 1994.