

Uniform distribution of last new vertex in random walk on a graph determines (almost) the graph

Tim Dwyer

March 7, 2013

Abstract

Cover tours are a natural object of study in the investigation of random walks on graphs. The general analysis of cover tours seems rather complicated, e.g. what is the expected time to actually cover a graph? This question seems pretty difficult generally.

Rather than attack this problem directly, Lovász and Winkler [1] settled a question about what types of graphs can have a uniform distribution of the last new vertex for their final stop on a random walk. In particular, what precisely are the graphs for which a cover tour starting at a vertex has the property that each other vertex is equally likely to be the last new vertex, i.e. the last vertex to have been hit. It turns out that this property is held only by the complete graph and the cycle.

1 A few Definitions

First we give a few definitions.

Definition 1. *Given a simple connected graph G , a **random walk** on G is a stochastic process whose state space is simply the vertices of G . Presently we are only considering the random walks where the transition probability from a vertex v to a vertex w is precisely the reciprocal of the number of edges incident to v , i.e. the degree of v (hereafter denoted $\deg v$).*

For simplicity of notation we also want to introduce a few events to streamline the statement of results later.

Definition 2. *For a nonempty connected graph G we let $L(x, y)$ be the event that for a random walk starting at vertex x , y is the last new vertex visited. Note that we will consider the vertex x to "have been visited" at the moment the walk is started, that is at time $t = 0$, so by this convention $\mathbb{P}(L(x, x)) = 0$.*

Definition 3. For a nonempty connected graph G we let $H(i, j, k)$ denote the event that a random walk starting at the vertex i arrives at the vertex j before it arrives at the vertex k .

2 A preliminary result

Lovász and Winkler [1] refer to the following result as folklore, and give, to my mind, a very simple proof due to Folklore. We will repeat the proof here if only because it is quick.

Theorem 2.1 (Theorem 1 in LW). *If G is a cycle (so in particular G has at least three vertices) then for any triple of distinct vertices v_1, v_2, w in G we have*

$$\mathbb{P}(L(w, v_1)) = \mathbb{P}(L(w, v_2)).$$

Proof. Suppose that G is a cycle on n -vertices, and we label the vertices $0, 1, \dots, n-1$ with the understanding that if we refer to a vertex $k-1$ for arbitrary the if $k=0$ we mean the vertex $n-1$ (this is the convention of naming the vertices mod n and allows us to write formulas without having to give multiple cases, as such it is very convenient).

Consider the event $H(0, 2, 1)$, and we ask how is this different from the event $H(j-1, j+1, j)$ (for an arbitrary vertex j)? Well G is a vertex transitive graph, that is for each v, w there is a graph automorphism $\varphi : G \rightarrow G$ such that $\varphi(v) = w$. This is simple to see, and actually from the convention of labeling the vertices mod n , we write it explicitly $v+i \mapsto w+i$ for all $i \in \{0, \dots, n-1\}$. So we must have that $\mathbb{P}(H(0, 2, 1)) = \mathbb{P}(H(j-1, j+1, j))$. Similarly we can consider the map $v+i \mapsto w-i$ to see that $\mathbb{P}(H(0, 2, 1)) = \mathbb{P}(H(j+1, j-1, j))$

But now we ask ourselves, what does it mean for a random walk starting at vertex i to have j as its last new vertex? Well, then it must actually hit vertex $j-1$ before $j+1$ or $j+1$ before $j-1$ since these are the only vertices that allow access to j . So given that such a walk hits $j-1$ before $j+1$ and have hit both of these vertices before the walk ever arrives at j .

From this discussion we conclude that

$$\mathbb{P}(L(i, j)) = \mathbb{P}(L(j-1, j+1, j)).$$

Thinking for a moment about the previous statement, it says that the probability that random walk starting at a vertex i and has j as its last new vertex has no dependence in i . This is an interesting observation and is, I would think, more than a bit counterintuitive. But even more

than that, we've already observed the quantity on the right has absolutely no dependence on j , it is just a constant and we actually have

$$P(L(i, j)) = \mathbb{P}(H(0, 2, 1)).$$

This immediately implies the result, since for any vertices w, v_1, v_2 we have

$$P(L(w, v_1)) = P(H(0, 2, 1)) = P(L(w, v_2)).$$

□

This is an interesting property, it seems natural to ask what other graphs might share this property. It is not hard to see that the complete graph shares this property, due to the fact that a walk starting at a given vertex sees all other vertices as the same (a property not intuitively shared by the cycle).

3 A few more definitions

Before we actually get to the business of proving the stated result, we'll need another preliminary theorem, and again for the sake of streamlining that theorem, we'll introduce a few more definitions.

Definition 4. *A graph G is called **k-connected** if for any collection of vertices v_1, \dots, v_{k-1} , the induced graph, \hat{G} , obtained by deleting the vertices v_1, \dots, v_{k-1} is connected. In particular we are concerned with **2-connected** graphs, i.e. graphs such that whenever a single vertex is deleted, the remaining graph is connected.*

The next definition is not, to my knowledge, in anyway standard, but it seems very natural.

Definition 5. *A graph G is called **2-disconnected** if for any pair of nonadjacent vertices v_1, v_2 the induced graph, \hat{G} , obtained by deleting the vertices v_1, v_2 is not connected, i.e. there are some vertices in \hat{G} not connected by a path in \hat{G} .*

Definition 6. *For a random walk on a graph G beginning at a vertex x , we define $L(x, v, u)$ to be the event that the next to last vertex visited is v and the last vertex is u .*

4 Another preliminary result

Theorem 4.1 (Theorem 2 in LW). *If G is a connected graph and u, v are nonadjacent vertices then there is a vertex x which is adjacent to u such that*

$$\mathbb{P}(L(x, v)) \leq \mathbb{P}(L(u, v)).$$

Further, if the graph \hat{G} induced by deleting the vertices u, v is connected then this inequality is strict.

Proof. Let $d = \deg u$, and let x_1, \dots, x_d be the neighbors of u in G . then for a random walk in G starting at u we can see that

$$L(u, v) = \bigcup_{i=1}^d (L(x_i, v) \cup L(x_i, v, u))$$

and that this union is disjoint. So by observing that

$$\mathbb{P}(L(u, v)) = \sum_{i=1}^d \frac{1}{d} \mathbb{P}((L(x_i, v) \cup L(x_i, v, u))),$$

we have that

$$\mathbb{P}(L(u, v)) = \frac{1}{d} \sum_{i=1}^d (\mathbb{P}(L(x_i, v)) + \mathbb{P}(L(x_i, v, u))), \quad (1)$$

since the probability of a disjoint union is just the sum of the probabilities. But (1) shows that $\mathbb{P}(L(u, v)) \geq \frac{1}{d} \sum_{i=1}^d (\mathbb{P}(L(x_i, v)))$. Indeed, supposing that the inequality as stated is false, i.e. that for all $i = 1, 2, \dots, d$ we have $\mathbb{P}(L(x_i, v)) > \mathbb{P}(L(u, v))$. Then we would have

$$\mathbb{P}(L(u, v)) \geq \frac{1}{d} \sum_{i=1}^d (\mathbb{P}(L(x_i, v))) > d \frac{1}{d} \mathbb{P}(L(u, v)) = \mathbb{P}(L(u, v))$$

which is certainly false, so the inequality holds.

For the second conclusion, if \hat{G} , as defined, is connected then for some i we must have $\mathbb{P}(L(x_i, v, u)) > 0$. This is because \hat{G} being connected would mean that a random walk could, starting at some vertex x_i could hit all vertices other than u, v . So in this case

$$\mathbb{P}(L(u, v)) > \frac{1}{d} \sum_{i=1}^d (\mathbb{P}(L(x_i, v)))$$

and so by identical reasoning, $\mathbb{P}(L(u, v)) > \mathbb{P}(L(x_i, v))$. □

5 The Big Result

So we've now seen that the cycle and the complete graph on n vertices share this property of the distribution of the last new vertex being uniform. Now with all the preliminaries out of the way, we're in a position to answer the question, are there possibly any other graphs with this property? The answer is a resounding no.

Lemma 5.1. *If G is a connected graph which is 2-connected and 2-disconnected then G is isomorphic to a cycle or the complete graph.*

The proof of Lemma 5.2 is a straightforward graph-theoretic argument, and so is omitted.

Theorem 5.2 (Theorem 3 in LW). *Let G be a graph such that for any vertices w, v_1, v_2 such that $\mathbb{P}(L(w, v_1)) = \mathbb{P}(L(w, v_2))$. Then either G is a cycle or a complete graph (up to isomorphism).*

Proof. Suppose G is a graph with the stated property. First we observe that G must in fact be 2-connected. Suppose it was not, then there would be a vertex v such that the graph, \hat{G} induced by deleting v is not connected. But then $\mathbb{P}(L(x, y)) = 0$ for any vertex x since there is no way that y could be the last new vertex in a random walk.

Second we claim that G must be 2-disconnected. Indeed suppose otherwise, then there exist a pair of vertices u, v such that the induced graph \hat{G} is connected. But then by Theorem 4.1 for some neighbor x of u we have $\mathbb{P}(L(x, v)) < \mathbb{P}(L(u, v))$. Though this may not seem immediately to be a contradiction (in the sense that it doesn't explicitly contradict the distribution property as stated), the distribution property implies that $\mathbb{P}(L(u, v)) = \frac{1}{n-1} = \mathbb{P}(L(x, v))$ which certainly gives a contradiction. So G must be a 2-connected, 2-disconnected graph, so by the lemma we have that G is either a cycle or a complete graph. \square

References

- [1] L. Lovász and P. Winkler (1993), A Note on the Last New Vertex visited by a Random Walk, *Journal of Graph Theory* **17**, 593–596.