CS599: Structure and Dynamics of Networked Information (Spring 2005)
04/06/2005: The Morris Contagion Model
Scribes: Iftikhar Burhanuddin and Ashish Vaswani

Last time, we introduced a simple model of contagion proposed by Morris [1] Recall that this model assumes an infinite graph where all the nodes (which represent infividuals) have finite degrees. A node can be either *active* or *inactive* (also, *infected* or *uninfected*). All vertices have the same threshold $p \in [0, 1]$: they become active if at least a p-fraction of their neighbors is active.

A set $X \subset V$ is called *contagious*, if starting from X, all nodes are activated eventually. The *contagion threshold* $t(G)$ is the supremum of all p such that there exists a finite contagious set for threshold p. The *community threshold* $c(G)$ is the infimum over all α such that every co-finite[1] set of nodes contains an infinite $(1 - \alpha)$ community in the sense of previous lectures. It turns out that the contagion threshold is equal to the community threshold.

Theorem 1 *$c(G) = t(G)$.*

Proof. The normal way to prove an equality is to prove both inequalities $c(G) \leq t(G)$ and $c(G) \geq t(G)$. As both quantities are defined in terms of infimum/supremum, this direct approach will not work as easily. Instead, we will show for each p that $c(G) < p$ if and only if $t(G) < p$.

For any node set $S \subseteq V(G)$ and node threshold $p \in [0, 1]$, we let $f_p(S)$ be the set of nodes active after one step if exactly the set S is active previously, and all nodes have threshold p. Define $f_p^k(S) := f_p(f_p^{k-1}(S))$ inductively to be the set of nodes active after k steps (with $f_p^0(S) := S$, of course).

1. Assume that $c(G) < p$. By definition of the community threshold, this means that every cofinite set S contains an infinite $(1 - p)$ community.

   Let X be an arbitrary finite set. Then, $V \setminus X$ is co-finite, and thus contains an infinite $(1 - p)$-community C. Because $C \cap X = \emptyset$, induction on the number of steps k shows that no node from C is ever active. Hence, X cannot be contagious, and because X was an arbitrary finite set, we have proved that $t(G) < p$.

2. Assume that $t(G) < p$. Let S be a co-finite set, so $V \setminus S$ is finite. Therefore, by assumption, it is not contagious for parameter p. Let $Y := \cup_{k \geq 0} f_p^k(V \setminus S)$. Then, $\overline{Y}$ is a $(1-p)$-community, as each node of $\overline{Y}$ must have strictly less than an p fraction of its edges to nodes from Y. Moreover, $\overline{Y}$ is infinite, for otherwise, $\overline{Y} \cup \overline{S}$ would be a finite contagious set, a contradiction. Hence, every co-finite set S contains such an infinite p-community $\overline{Y}$.

■

Theorem 2 *There is no graph G with $t(G) > \frac{1}{2}$.*

Proof. For a set S, let $\phi(S) := |\delta(S)|$ be the number of edges leaving S. For $i \geq 0$, let S_i denote the set of infected nodes after i rounds. Observe that if at any point $S_i = S_{i+1}$, then clearly the infection process has terminated, as no new nodes can ever be activated.

Let $p > \frac{1}{2}$ be arbitrary, and S_0 a finite set starting out infected. Consider the sequence $\phi(S_i)$, $i \geq 0$, the number of edges leaving the infected sets in each iteration. When node v moves from $\overline{S}_i$ to S_{i+1} (i.e., v gets infected in the i^{th} step), node v has strictly more edges into S_i than into $\overline{S_i}$, because $p > \frac{1}{2}$. In the

---

[1] A cofinite set is a set whose complement is finite.

$(i+1)^{\text{st}}$ iteration, all the edges from v to S_i are not cut any more, whereas those from v to $\overline{S_{i+1}}$ are cut. Summing over all v, ϕ strictly decreases in each iteration in which nodes become infected. Because S_0 is a finite set, and each node has finite degree, $\phi(S_0)$ too must be finite. Therefore, ϕ can only decrease a finite number of times, and $S_k = S_{k+1}$ for some k. As only finitely many nodes become infected each round, only finitely many nodes will become infected eventually. As this holds for any finite set S_0, and any $p > \frac{1}{2}$, the contagion threshold $t(G)$ cannot exceed $\frac{1}{2}$. ■

In view of the above fact, a natural question to ask is which graphs have $t(G)$ close to $\frac{1}{2}$. We begin by making some definitions.

A *labeling* of a graph G is a bijection $\lambda : \mathbb{N} \to V$. A labeling λ for G is *p-inductive* with parameter $k_0 \in \mathbb{N}$ if for all $k \geq k_0$, each node $\lambda(k)$ has at least a p-fraction of edges to $\lambda(0), \ldots, \lambda(k-1)$. We define $\ell(G)$ to be the supremum over all p such that G has a p-inductive labeling. This new measure is again equal to the contagion and community thresholds.

Theorem 3 *$\ell(G) = t(G)$.*

Proof. Start with nodes $\lambda(0), \ldots, \lambda(k_0)$ infected. Then, by definition of p-inductiveness, all subsequent nodes will become infected at threshold p. Similarly, the order in which nodes become infected can be used as a labeling, which is easily seen to be p-inductive. ■

A *BFS-labeling* is a labeling λ for graph G consistent with the BFS traversal of G from some finite node set X. It is *δ-smooth* for parameters p and k_0 if all nodes $\lambda(k)$ for $k \geq k_0$ have between a $(p-\delta)$-fraction and a p-fraction of edges to $\lambda(0), \ldots, \lambda(k-1)$. Also, recall that a graph G has *subexponential growth* if for all $c > 1$ and all finite sets X, the ball sizes are $|B_r(X)| = o(c^r)$. (Where $B_r(X) := \{u \mid d(u, X) \leq r\}$ is the set of all nodes at distance at most r from X.)

Theorem 4 *If a graph G has subexponential growth and a δ-smooth BFS-labeling then $t(G) \geq \frac{1}{2} - \delta$*

Notice that this is not an "if and only if" statement. For instance, consider the grid with sufficiently large neighborhood circles, and add one random edge per node. For purposes of infection, this graph acts very much like the grid (as most edges are grid edges). However, as we saw in past lectures, this graph has exponential growth. Indeed, one might consider a definition of "small world" graphs to have both exponential growth and $t(G)$ close to $\frac{1}{2}$.

Proof. We will show that the labeling must have $p = \frac{1}{2}$. Then, because the labeling is $\frac{1}{2} - \delta$ inductive, Theorem 3 implies the result. We prove the statement by contra-position: if $p < \frac{1}{2}$, then the graph must have exponential growth. Consider a BFS consistent with λ. For each v, let $f(v)$ be the number of edges between v and nodes u with lower labels, and $g(v)$ the number of edges between v and nodes u with higher labels.

By definition, $f(v) \leq p(f(v) + g(v))$, so $g(v) \geq \frac{1-p}{p} f(v)$. Let $L_r = \{u \mid d(X, v) = r\}$ be the r^{th} layer of the BFS from X. We denote the edges within L_r by I_r, and the edges from L_{r-1} to L_r by F_r. Then, $\sum_{v \in L_r} f(v) = |I_r| + |F_r|$ (counting each edge towards the value of the higher endpoint), and $\sum_{v \in L_r} g(v) = |I_r| + |F_{r+1}|$ (counting edges for lower endpoints). Combining this with the inequality between $f(v)$ and $g(v)$, we obtain that $|I_r| + |F_{r+1}| \geq \frac{1-p}{p} \cdot (|I_r| + |F_r|)$, which we can rearrange to obtain

$$|F_{r+1}| \quad \geq \quad \tfrac{1-2p}{p} \cdot |I_r| + \tfrac{1-p}{p} \cdot |F_r| \quad \geq \tfrac{1-p}{p} \cdot |F_r|,$$

because $\frac{1-2p}{p} \geq 0$. Therefore, the sizes of F_r grow exponentially, and since the nodes have finite degrees, the number of vertices in the r^{th} layer must also grow exponentially. ■

References

[1] S. Morris. Contagion. *Review of Economic Studies*, 67:57–78, 2000.