Algorithms, Games, and Networks	April 9, 2013

Lecture 22

Lecturer: Ariel Procaccia — *Scribe: Noam Brown*

1 Overview

In the next few lectures we'll take a look at social networks. Specifically, we'll look at how ideas, behaviors, or anything else might spread through a connected population. One very clear example is a product like Facebook or Google+. Each process starts with a small group of early adopters, and then spreads through existing social connections to the larger population. For example, all Google employees may start with a Google+ account, but an outside person would only adopt the product if a sufficient number of friends are already using it. Other examples include religious beliefs, political ideas, or other new ideas and products.

In this lecture, we'll particularly look at what it takes for a new product or behavior to catch on and "infect" the entire population. We'll assume the product starts with a small group of early adopters, and attempts to spread to an entire infinite population.

2 Formalization

We define a social network as an undirected graph $G = (V, E)$, where V is countably infinite but the number of edges connected to any vertex $v \in V$ is finite. We treat the vertices as people who have a choice between an old behavior A, and a new behavior B. We also use a parameter $q \in (0, 1)$ which, intuitively, measures how attractive the new behavior is.

Consider nodes u and v, and their edge $(u, v) \in E$. We'll define their rewards as follows:

- If both choose A, then they receive q.
- If both adopt B, then they receive $1 - q$.
- Otherwise, both receive 0.

The overall payoff to v is the sum of payoffs from each neighbor.

We denote the degree of v as d_v, and the number of neighbors of v who adopt X by d_v^X. Consider the question, when should v switch from choosing A to choosing B? v would need the expected reward from B to be higher than from A. By our definitions, the expected reward for choosing A is qd_v^A, and the reward from adopting B is $(1-q)d_v^B$. Thus, v adopts B if $(1-q)d_v^B \geq qd_v^A$. After evaluating, this becomes $d_v^B \geq qd_v$. That is, v only adopts B if at least a q fraction of v's connections have already adopted B. Thus, q acts as a threshold.

3 Cascading Behavior

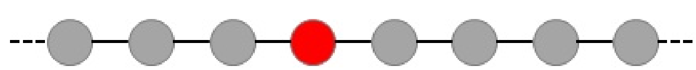
We now extend this model further by introducing time steps $t = 1, 2, ...$ At every time step, each node will evaluate its options (that is, choosing either A or B), and choose to adopt the higher-payoff behavior for the next time step. The choices and adoptions are all done simultaneously, and all players play a best response without considering how other players will adapt. Furthermore, a node is not stuck with B once it is adopted. At a future time step, a node can switch back to A if it wishes to.

We assume there is a finite set S of nodes who initially adopt B. After one round, we denote the set of nodes that are now adopting B as $h_q(S)$. After k rounds, the set of nodes adopting B is denoted as $h_q^k(S)$. We say v is converted by S if $\exists k$ s.t. $v \in h_q^k(S)$. That is, v is converted if it has ever adopted B, even if it later switches back to A. We say S is *contagious* if every node is converted (even if they are not all simultaneously adopting B).

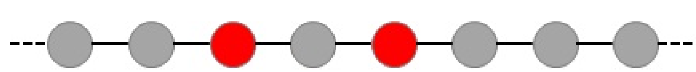
We're particularly interested in the case of a finite S. In some cases, a finite set S can be contagious in a graph G. In other cases, it can't. Clearly though, it is easier for a set S to be contagious when q is lower. We'll define the *contagion threshold* of G to be the maximum q such that a finite contagious set exists.

4 Examples

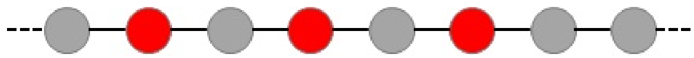
We now consider some particular examples. First, consider the graph shown below with infinitely many nodes arranged in a line, with each node connected to its two neighbors. What would the contagion threshold of this graph be? For simplicity, suppose the set S consists of a single node at time $t = 0$.



It's pretty easy to see the answer is $q = \frac{1}{2}$. Clearly, at time $t = 1$ the neighbors of S would become infected if and only if $q \leq \frac{1}{2}$. Furthermore, since the original node has 0/2 neighbors adopting B, it adopts A at time $t = 1$. Thus, at $t = 1$ we have

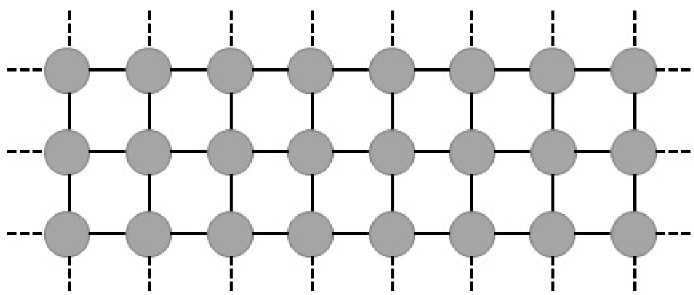


At $t = 2$, we then have



This trend continues so that every node is eventually converted (though never all of them at the same time).

We now consider a somewhat more complicated example. Suppose G is a square lattice, with each node connected to its 4 neighbors as shown below. What is the contagion threshold of G in this case?



Notice that any finite S can be bounded by a box. Clearly, in order for S to be contagious, it must eventually be possible to convert a node on the outside border of this box. But each node outside the box has at most 1 neighbor inside the box (out of 4 neighbors). Thus, the contagion threshold is at most $\frac{1}{4}$. It's also clear to see that $\frac{1}{4}$ is sufficient to convert the entire graph starting with any S.

5 Progressive Processes

So far we've only considered the *Nonprogressive Process* where nodes can switch from A to B or B to A. We now consider the *Progressive Process* where nodes can only switch from A to B, and can never go back to A. Everything else from our model remains the same. We denote the set of nodes adopting B in the progressive process with contagion threshold q by $\bar{h}_q(S)$. Intuitively, one would think that the contagion threshold of a graph would be higher in the progressive case. However, as we'll see now, the contagion threshold is the same in either case.

Theorem 1 *For any graph G, the contagion threshold is identical for progressive and non-progressive processes.*

Proof. We first prove the lemma: $\bar{h}_q^j(X) = h_q(\bar{h}_q^{j-1}(X)) \cup X$ using induction. For the base case, clearly when $j = 1$ we have $\bar{h}_q(X) = h_q(X) \cup X$. It can easily be seen that $\bar{h}_q^{j-1}(X) \supseteq h_q(\bar{h}_q^{j-1}(X)) \cup X$, so we need only show $\bar{h}_q^{j-1}(X) \subseteq h_q(\bar{h}_q^{j-1}(X)) \cup X$. By the induction assumption, we have $\bar{h}_q^{j-1}(X) = h_q(\bar{h}_q^{j-2}(X)) \cup X$. Clearly $\bar{h}_q^{j-2}(X) \subseteq \bar{h}_q^{j-1}(X)$, so we have $\bar{h}_q^{j-1}(X) \subseteq h_q(\bar{h}_q^{j-1}(X)) \cup X$.

We now move onto the main part of the proof. Obviously, if S is contagious for h_q then a finite contagious set exists for $\bar{h}_q$ (namely, the same set S). We'll now show the converse. Assume S is contagious for $\bar{h}_q$. Let $N(S) = \{u : v \in S, (u,v) \in E\}$ (that is, the set of all neighbors of S). Since S is contagious for $\bar{h}_q$, so clearly $\exists l$ such that $S \cup N(S) \subseteq \bar{h}_q^l(S)$. We'll show $\bar{h}_q^l(S)$ is contagious for h_q.

By our earlier lemma, for $j > l$, $\bar{h}_q^j(S) = h_q(\bar{h}_q^{j-1}(S)) \cup S$. But we can see that $S \subseteq h_q(\bar{h}_q^{j-1}(S))$, because $S \cup N(S) \subseteq \bar{h}_q^{j-1}(S)$, which means all the neighbors of every $v \in S$ are playing B, and so clearly v would continue to play B. Simplifying, we now have that for $j > l$, $\bar{h}_q^j(S) = h_q(\bar{h}_q^{j-1}(S))$. Therefore, we can see that $h_q^{j-1}(\bar{h}_q^l(S)) = \bar{h}_q^j(S)$. Recall that S is contagious for $\bar{h}_q$, so therefore $\bar{h}_q^l(S)$ is contagious for h_q. $\square$

6 Maximum Contagion Threshold

Let's now move on to a different question: Does there exist a graph with contagion threshold $> \frac{1}{2}$? It turns out there doesn't.

Theorem 2 *For any graph G, the contagion threshold $\leq \frac{1}{2}$.*

Here is a sketch of the proof. First, due to the previous theorem, note that we need only look at the progressive case.

Let $\delta(X) = \{(u,v) \in E : u \in X, v \in V \setminus X\}$. Consider the vertices $v \in \bar{h}_q^j(S) \setminus \bar{h}_q^{j-1}(S)$. For each v, because $q > \frac{1}{2}$, v has *strictly* more edges into $\bar{h}_q^{j-1}$ than outside of $\bar{h}_q^{j-1}$. Therefore, $\delta(\bar{h}_q^j(S)) < \delta(\bar{h}_q^{j-1}(S))$. Recall that S must be finite, and each vertex has finite edges. Thus, $\delta(\bar{h}_q^j(S))$ is finite $\forall j$, and shrinks at every time step. Therefore, the set must eventually cease expanding and remain forever finite. $\square$

7 Other Models

There are other models we could consider as well. For example, so far we have only looked at undirected graphs. A directed graph could model asymmetric influence. For these models, let $N(v) = \{u \in V : (u,v) \in E\}$. We'll assume a progresive contagion, and say a node is *active* if it adopts B, and *activated* if it switches from A to B this round.

In the Linear Threshold Model, we assign each edge $(u,v) \in E$ some nonnegative weight w_{uv}, or $w_{uv} = 0$ if no weight is assigned. Assume $\forall v \in V$, $\sum_u w_{uv} \leq 1$. Each $v \in V$ has a threshold θ_v. Then v becomes active if

$$\sum_{\text{active } u} w_{uv} \geq \theta_v$$

The Linear Model assumes additive influence, but we may not want to assume that. We could also use the General Threshold Model. In this model, v has a monotonic function $g_v(\cdot)$ defined on subsets $X \subseteq N(v)$. Then v becomes activated if the activated subset $X \subseteq N(v)$ satisfies $g_v(X) \geq \theta_v$.

Finally, we can consider the Cascade Model. In this model, when $\exists(u, v) \in E$ s.t. u is active and v is not, u has only one chance to activate v. Each v has an $incremental function$ $p_v(u, X)$, which equals the probability that u activates v when X have tried and failed.

There are a couple special cases in the Cascade Model. The first is diminishing returns, where $p_v(u, X) \geq p_v(u, Y)$ when $X \subseteq Y$. The second is called Independent cascade, where $p_v(u, X) = p_{uv}$.