COMS E6998-2: Analysis of Social Information Networks — Feb 11, 2011

Lecture 4 – Influence

Instructor: *Augustin Chaintreau* — Scribes: *Tingting Xu*

1 Life under the influence

Epidemic is one of the influences under which we live our lives. It occurs when a certain disease spreads from people to people(the population here is usually in significant size) accidentally. It may be restricted to one locale, in which case called endemic. Or it may be pandemic which could even shape our history and present. Other influences have effect on development or behavior of someone which typically opposed to the individual free choice. In much broader sense, influence here means diffusion, spread or cascades which not necessarily dramatic, negative, or accidental.

Example 1. *Asch's experiment tells us people frequently followed the majority judgment, even when the majority was wrong. We live under group-pressure and tend to keep conformity to the majority. However, from the experiment graph we observed that the curve(i.e. Errors percentage) decreases when the number of opponents reached a specific magnitude and continually grew. The reason may be that followed people have more time to think.*

Example 2. *In the spread of happiness experiment, we can see that there are lots of factors which could affect the magnitude of the influence. And the influence of peers seems to follow the nature of the ties. Call up what we have learnt in the previous lecture, strong ties(e.g. nearby mutual friend or next door neighbour) may be have more significant influence on the subject than weak ties(e.g. distant friend). Similar results can be found in the spread of obesity and quit smoking.*

Example 3. *The spread of an innovation experiment demonstrates combinational results from the above two experiments. It shows that influence is a significant topic which can help us to figure out key social-science questions including how can we induce people to make behavior changes that are known to have large and positive health benefits.*

2 Diffusion as collective behavior

In this model we assume that people primarily decides to take an action as a monotone function of its adoption. Every node v has threshold t_v. If a fraction $x \geq t_v$ of nodes adopts, then v adopts. Here,

$$x = \frac{|\{v \in V | adopted\}|}{|\{v \in V\}|}$$

The example here may be adopting a smart phone. When the smart phone just comes into the market, there are not so much related applications there and you may not think about adopting it. With the user size growing significant, network effect occurs(i.e. user size and amount of applications increase with each other). And you will adopt it with a much higher probability.

Observation 4. *Let $F(x) = Fraction\ of\ nodes\ whose\ threshold \leq x$. This means that,*

$$F(x) = \frac{|\{v \in V | t_v \leq x\}|}{|\{v \in V\}|} \leq x$$

This is called Cumulative Density Function of threshold. Evolution follows F(F(...F(x))).

Proof. Suppose in i^{th} iteration, the fraction x_i of nodes adopts, then $F(x_i) = Fraction\ of\ nodes\ whose\ threshold \leq x_i$, then the fraction $F(x_i)$ will all adopt in the $(i+1)^{th}$ iteration, $x_i + 1 = F(x_i)$, hence, $F(x_i + 1) = F(F(x_i))$. □

Function F(x) has the following characterizations:

$$\begin{array}{lllllll} F(x) > x: & expansion & region & (adopted & fraction & increases) & \\ F(x) < x: & contraction & region & (adopted & fraction & decreases) & \\ F(x) = x: & equilibrium & (stable) & or & tipping & point & (unstable) \end{array}$$

On tipping point, a slight shift will cause dramatic change overall.

3 Contagion on arbitrary graphs

In reality, individuals know and care only about a local view of their network(i.e. fraction of neighbors). In this model, we suppose an individual primarily decides to take an action as a monotone function of his or her neighbors adoption. Every node v has threshold t_v. If a fraction $x \geq t_v$ of his or her neighbors activate, v activates.

Claim 5. *Assume S_0 is the initial set of nodes adopting the behavior. Then we have S_1 nodes with fraction $f \geq t_v$ of neighbors in S_0, S_2 nodes with fraction $f \geq t_v$ of neighbors in $S_0 \bigcup S_1$ and so on. Also assume all nodes in G have $t_v = p$, starting from the set S_0, the graph G(V,E) is p-converted if $V = S_0 \bigcup S_1 \bigcup ....$*

This claim means that if a Graph is p-converted, with contagion threshold p, starting from the initial set S_0, adoption will finally be realized in the whole graph. Since we have forbidden the set of adopters from shrinking we know that once $S_i = S_{i-1}$ for some i the set of adopters remains unchanged thereafter. This means that for the trend to spread to the entire graph we necessarily have that S_i is larger than S_{i-1} for all i. Note that this is not in general a sufficient condition, as the graph is infinite the sequence S_i may continue to grow forever without covering all the nodes.

Definition 6. *The contagion threshold cont(G): $\{$ p | there exists finite S_0 such that G is p-converted $\}$.*

The bigger the value of cont(G) is, the easier the contagion will happen in the graph. However we will see that the biggest value of cont(G) is 1/2. Please see proof later.

3.1 Relation Contagion - Density

A subset S is a p-dense cluster if a node in S has a fraction at least p of its neighbors in S.

Definition 7. *clu(G) = $\{ p$ such that $|\forall|A| < \infty$, the set V-A contains a p-dense cluster $\}$.*

Intuitively, if we forget about A, $clu(g)$ is simply the largest density of a cluster found in the graph. The presence of A here simply represent the fact that we are not interested in cluster density which are fragile (may become smaller if a few nodes are removed) and that we wish to characterize density that are robust.
It is obvious that cont(G) and clu(G) $\in [0, 1]$. Since the bigger the value of cont(G) is, the easier the contagion will happen in the graph, we know that dense cluster will block innovation. Essentially, a cascade comes to a stop when it runs into a dense cluster. Furthermore, this is the only thing that causes cascades to stop. From the theorem above we can get the following claim:

Theorem 8. $clu(G) = 1 - cont(G)$

The proof of the theorem comes from the following two facts:

Claim 9. *Consider a set of initial adopters set S_0 of behavior A, with a threshold of q(i.e. cont(G)) for nodes in the remaining network to adopt behavior A. (i) If the remaining network contains a cluster of density greater than 1-q(i.e. if we have $clu(G) > 1 - cont(G)$ here), then the set of initial adopters will NOT cause a complete cascade. (ii) Moreover, whenever a set of initial adopts does NOT cause a complete cascade with threshold q(i.e. cont(G)),the remaining network must contain a cluster of density greater than 1-q(i.e. here we must have $clu(G) > 1 - cont(G)$).*

Proof. In part (i), if the remaining network $V - S_0$ contains a cluster of density greater than 1-q, this means that the maximum fraction of active neighbors of nodes in the remaining network will be less than q. Then these nodes will not be affected according to the threshold rule, so that the cascade cannot influence them and they will never be infected. In part (ii), whenever a set of initial adopts does NOT cause a complete cascade with threshold q, this means that nodes in the remaining network have fraction less than q of active neighbors. Hence the fraction of their neighbors in the remaining network is greater than 1-q. Then since this is true for any starting set $S_0 = A$, we proved the claim that the remaining network must contain a cluster of density greater than 1-q. □

3.2 More results on contagion

Theorem 10. $cont(G) \leq 1/2$

Proof. Assume $cont(G) > 1/2$. We let S_0 be the initial set of adopters and S_i be all the adopters at the i^{th} step. Also, let B(S_i) denote the edges with one end in S_i and the other end not in S_i.We know that S_i is larger than S_{i-1} for all i. Then for any vertex v in $S_i - S_{i-1}$ we know that v has to have more neighbors in S_i than outside S_i because having $cont(G) > 1/2$ forces the adopted neighbors of a vertex to outnumber other neighbors in order to adopt the trend. Looking at this over all such v, we can see that B(S_i) must be smaller than B(S_{i-1}). Indeed, we can write $B(S_{i-1})$ as $\cup_{v \in S_i, v \notin S_{i-1}} \{(u, v) \in E | u \in S_{i-1}\}$, which is a disjoint union of set. For each of this set, we know that there exists a strictly smaller set of edges that contain all edges that connect v to a node not in S_{i-1}. The set of edges that connect v to a node not in S_i is even smaller (as S_i contains S_{i-1}) and these sets form a partition of $B(S_i)$ proving that it contains strictly less edge than $B(S_i)$. If we look at the sequence of numbers $\{|B(S_1)|, |B(S_2)|, |B(S_3)|, ...\}$, we know that they are a strictly decreasing sequence of natural numbers. Since there is a smallest natural number, namely 0, this isn't possible! This means that $|S_i|$ cannot be always larger than $|S_{i-1}|$, and therefore the set of adopters must stop growing at some point.

□

Now we know at least one graph(i.e. nodes on a line) with cont(G) = 1/2. This threshold can only be approaching with uniform interaction and low neighborhood growth. Another situation here is that nodes renew their decisions at each step. This doesn't affect the contagion threshold as nothing stranger than oscillations similar to the one in the second activity occur.

4 A general algorithmic problem

The real problem is that how to find an initial fixed size S_0(i.e. say k-element) which can maximize the total spread. We introduce a more general model of neighbor influence here. It may be described in two ways:

- Define for any node $u \in V$ a function g_u taking value in $[0; 1]$ and which is defined on all subset of neighbors of u (*i.e.*, for any $S \subseteq N(U)$ we define a value $g_u(S) \in [0; 1]$).

  Node's behavior is then characterized as follows. First, we assume that a set S_0 of nodes initially adopt the service, and that for any node $v \in V$ there exists a threshold θ_v which chosen once for all in $[0; 1]$ according to a uniform distribution.

  Then, for any time slot t, if during this time slot t, the set of neighbors of v which adopt the service is S, v will adopt the service if and only if we have $\theta_v \leq g_v(S)$.

- In another model, we assume that for any u and v such that (u, v) is an edge in E there exists a function $p_v(u, .)$ which takes value in $[0; 1]$ and is defined on all subset of neighbors of v that do not contain u (*i.e.*, for any $S \subseteq N(v)$ such that $u \notin S$, we define a value $p_v(u, S)$).

  The behavior of the nodes is then described as follows. A subset S_0 of nodes initially adopt the service or the innovation at time $t = 0$.

  Then for any $t = 0, 1, \ldots,$ whenever a node u adopts the service for the first time during t, for any node v that is a neighbor of u it makes a single attempt during this time slot to influence v. Note that if v has already been influenced and use the service nothing will happen. Otherwise it indicates that all previous attempt to influence v has failed. We denote by S_v the set of all nodes who attempted to influence v before u. What happens then is the following. With probability $p_v(u, S_v)$ (chosen independently from the past) the attempt is successful and v starts using the service at time $t + 1$. Otherwise, hence with a probability $p_v(u, S_v)$ the attempt does not succeed and hence u is added to S_v.

Special cases here are Granovetter or Morris fixed threshold model, Independent Cascade Model. When p is an order independent function, the two models are equivalent. (see Exercice)

4.1 Critical Mass vs. Diminishing Return

Critical Mass and Diminishing Return are two properties that may be satisfied or not by the above model. Both are found in practice and lead to very different qualitative behavior. For simplicity we will describe them in function of the model formulated with function p.

Observation 11. *In critical mass case, $p_v(u, X)$ increases with X. Infection gets easier to catch once more of your friends adopt. An example here may be adopting a smart phone. You don't really care if* *a few of your friends own an iPhone. However, when the number of your friends who owns an iPhone reaches some point, you change your mind. This explains why the curve grows slowly at the beginning and then changes dramatically. This model could be applied to diffusion of innovations, diseases, rioting and so on. This model may exhibit tipping point a slight shift in the group's level of threshold and adoption may lead to a large shift in overall behavior.*

Observation 12. *The diminishing return case is one where intuitively it is easier to infect when the behavior is not widespread. In this case, $p_v(u, X)$ decreases with X, and in other words the marginal chance to successfully influence decrease as the number of people who have already attempt increase. Infection gets relatively harder to pass as it becomes widespread. This model dominates in membership, growth in large social networks. The examples where it was observed is how the probability of joining a community depends on the number of k friends that are already members, or the probability of joining a conference when k co-authors are already members of the conference.*

4.2 Maximizing spread of influence

Assuming process starts with set S, the objective function f(S) = E[number of active nodes] at the end of the process.

Theorem 13. *Computing max $\{f(S)||S| = k\}$ is NP hard.*

In fact, even computing a $n^{1-\varepsilon}$ approximation is hard ($\varepsilon > 0$). This comes from the fact that, if g_v shows "critical mass" then computing the above relates to a set covering problem. In order for the maximization of influence to be tractable, an additional condition is hence required.

Theorem 14. *Whenever p_v exhibit the diminishing return property*
There exists a simple polynomial algorithm computing S such that $f(S) \geq (1 - 1/e)f(S^*)$
Algorithm follows greedy "one node at a time" rule. Do k times: $S < -S \bigcup argmax_v \{f(S \bigcup \{v\}) - f(S)\}$

This algorithm starts from a set S as $\varnothing$, and iteratively find $\{v\}$ to make $f\left(S \bigcup \{v\} - f\left(S\right)\right) > 0$(i.e. v is the element gives the maximum marginal gain). It repeat this k times. Note that it might be that the element giving maximum marginal gain is not actually critical (because adding two elements would do overall much better than having this one and another later), but the diminishing return property allow to show that the efficiency loss remains minimum.
Sketch of the analysis: We wish to find a k-element set S for which f(S) is maximized. The optimal solution can be efficiently approximated to within a factor of $(1 - 1/e - \varepsilon)$, where e is the base of the natural logarithm and ε is any positive real number. Thus, this is a performance guarantee slightly better than 63 percent.

5 Summary

Influence is prevalent. It takes many forms and can be seen in group-pressure, conformity, spread of happiness, adoption of innovations and so on. It is usually impacted by topology (i.e. large cluster density blocks innovation) and local dynamics(i.e. critical mass vs. diminishing return).