Social and Information Networks

University of Toronto CSC303
Winter/Spring 2019

Week 8: March 2-6 (2020)

Announcements

- Regrading requests for Assignment 1 will be accepted up to Monday, March 9 at 3PM.
- Five questions have been posted for the second asasignment. There will be a couple more questons which will be added in the next couple of days. . Assignment 2 is due March 16 at 2:59 PM.
- Midterm March 4 and March 6. The test will cover everything in the first six week. Wednesday, March 4 part of the test is in the usual tutorial rooms. Friday, March 6 part of the text will be in GB248 and for those in the other tutorial, the test will take place in Haultain Building, room 403.
- Comments on the critical review assignment.
  - Due date: March 30
  - You need to find a conference or journal article that has appeared in the last 3 years; to be precise lets say, has appeared since January 1, 2017.
  - The article can be about any topic in the course.
  - You are to provide a critical review of the article as if you were on program committee or a reviewer for a journal. I have elaborated on this assignment in the Monday, February 26 lecture. Please send me your suggestikon for the paper you will critically review.

This weeks agenda

- Chapter 19: Influence spread in a social network
- Choosing a set of initial adopters

Chapter 19: Influence spread in a social network

- We begin a study of the spread/diffusion of products/influence in an existing social network (Chapter 19). This is in contrast to the population wide influence spread that we are passing over in Chapters 16 and 17. Chapter 18 (on power laws) also dealt with population wide influence phenomena. .

- The goal (as thorughout the course) is to qualitatively understand a process or observed phenomena in a highly stylized (but hopefully still interesting) setting.

- We will (as usual) be interested in what kind of general conclusions can be inferred from such an understanding?

The chapters preceding chapter 19

- In Chapters 16 (herding effects), 17 (direct benefit effects), and 18 (rich get richer models) we did not have a social network per se.
- These chapters dealt with population wide effects. Although :
  - One can construe Chapter 16 as taking place in a network where the i^{th} individual is connected to all $i-1$ previous individuals.
  - Chapter 17 can be construed as taking place in the complete graph network. Information about the entire population impacts decisions.
  - In Chapter 18 we studied studied a random process (e.g., link creation) by which an information network can grow. We also studied in this chapter, an example (music downloading in the Salganik et al exoeriment) where we can idientify how the presense of polulation wide information will influence an outcome. Like Chapter 16, we can think of this as taking placd in a social network where the i^{th} person knows some global information about the preceding $i-1$ individuals.
- But basically these are population wide effects absent from an existing social network where influence spreads without any global information.

Note: It is interesting to contrast the herding effect in chapter 16 with the impact of influence in the Salganik et al experiment in Chapter 18.

Social network effects

- Now we wish to consider an existing social network where edges (ties) between individuals represent some sort of friendship/relationship.
- This takes us back to concepts introduced in Chapters 3 and 4.
- There we saw the contrast between
  - selection (we tend to be friends with people of similar backgrounds, geography, interests)
  - social influence (we join clubs, are influenced) by our friends/relations.

Models of influence spread/diffusion

- One of the most important themes of the text (and CSC 303) is that we construct models to gain insight.
  - Our models are often (maybe always) very simplified given the complexity of real social and economic networks.
  - There is always a tradeoff between the adherence to reality and our ability to analyze and gain insight.
- How we model diffusion in a social network will clearly depend on what product, idea, membership, etc. we are considering.
- There are many assumptions as to how products, ideas, influence are spread in a social network and what are the set of individual alternatives.
- The main emphasis in Chapter 19 is on a very simple process of diffusion where each person has 2 alternative decisions:
  1. stay with a current "product" B
  2. or switch to a (new) product A.

A simple model of diffusion in a social network

- Let's assume that we are making decisions based on the direct benefit of being coordinated with our friends beyond any intinsic value associated with the decision (e.g. when the decision is the purchase of an item).

- A standard example is what laptop or cell phone we decide to buy to the extent that we are mostly influenced by our friends rather than by general population wide usage. What influences you most? Friends or general population information?

A simple model of diffusion in a social network

- Let's assume that we are making decisions based on the direct benefit of being coordinated with our friends beyond any intinsic value associated with the decision (e.g. when the decision is the purchase of an item).

- A standard example is what laptop or cell phone we decide to buy to the extent that we are mostly influenced by our friends rather than by general population wide usage. What influences you most? Friends or general population information?
  - Choosing between two weekly television shows that occur at the same time or who to vote for are other examples.

- In fact, the model given in this chapter dictates that certain decisions (i.e. to change from B to A) are irreversible.
  - The text calls this a "progressive process" in the sense that it progresses in only one direction. Any good examples of truly (or *essentially*) irreversible decisions?

A simple model of diffusion in a social network

- Let's assume that we are making decisions based on the direct benefit of being coordinated with our friends beyond any intinsic value associated with the decision (e.g. when the decision is the purchase of an item).

- A standard example is what laptop or cell phone we decide to buy to the extent that we are mostly influenced by our friends rather than by general population wide usage. What influences you most? Friends or general population information?
  - Choosing between two weekly television shows that occur at the same time or who to vote for are other examples.

- In fact, the model given in this chapter dictates that certain decisions (i.e. to change from B to A) are irreversible.
  - The text calls this a "progressive process" in the sense that it progresses in only one direction. Any good examples of truly (or *essentially*) irreversible decisions?
  - For example, the decision to get a tatoo.

A threshold model for spread

- We assume that some number of individuals are enticed (at some time $t = 0$) to adopt a new product A.
- Outside of these "initial adopters", we assume all other individuals in the network are initially using a different product B (or equivalently this is the first product in a given market).
- This is not really a competitive influence model as B is not really competing. (More comments later.)
- The first model we consider for diffusion is that every node v has a threshold q (in absolute or relative terms) for how many of its neighbors must have adopted product A before v adopts A.

Threshold model (continued)

- For simplicity the text initially assumes that every node v (i.e. individual) in the network has the same threshold but then later explains how to deal with individual thresholds.
- If at some time t, the threshold for a node v has been achieved, then by time time $t+1$, v will adopt product A.
- If the threshold has not been reached then v decides not to adopt A at this time.

Note

Although it is not explicitly stated, the initial adopters never reverse their adoption.

- Given these model assumptions, adopting A is irreversible for all nodes in the network.

Determining a (relative) threshold

- One way (some might say is usually the best way) to reason about a plausible threshold for a node is to view one's decision in economic terms.

- Specifically for every edge (v, w) in the network suppose
  - There is payoff a to v and w if both v and w have adopted product A.
  - There is payoff b to v and w if both v and w have adopted product B.
  - A zero payoff when v and w do not currently utilize the same product.

- This determines a simple coordination game.

$$
\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{w} \\
 & & A & B \\
\cline{3-4}
\multirow{2}{*}{v} & A & a, a & 0, 0 \\
\cline{3-4}
 & B & 0, 0 & b, b \\
\cline{3-4}
\end{array}
$$

Figure: $A - B$ coordination [Fig 19.1, E&K]

Coordination game induces threshold

- Suppose node v has not yet adopted A at time t, but a fraction p of the $d(v)$ neighbors of v have already adopted A, then:
  - By switching, the payoff to v is $p \times d(v) \times a$.
  - By staying with B, v has payoff $(1-p) \times d(v) \times b$.
- Thus node v will switch to A if

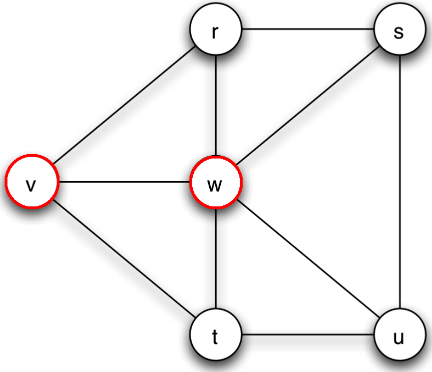
$$p \times d(v) \times a \geq (1-p) \times d(v) \times b$$

  (for simplicity say v switches when payoffs are equal).

- This is then equivalent to saying that v will switch whenever p is at least $\frac{b}{a+b} = q$ which is then the relative threshold.
- That is, whenever there is at least a (threshold) fraction q of the neighbours of node v that have adopted A, then v will also adopt A.

The process unfolds (example: $a = 3$ and $b = 2$)

[Fig 19.3, E&K]

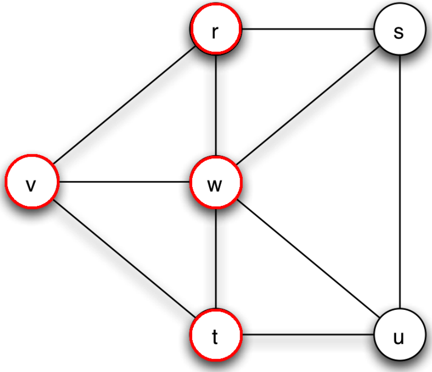


$t = 0$

- A node adopts A if and only if the threshold $q = \frac{b}{a+b} = 2/5$ is reached.
- Two nodes v and w are initial adopters.

The process unfolds (example: $a = 3$ and $b = 2$)

[Fig 19.3, E&K]

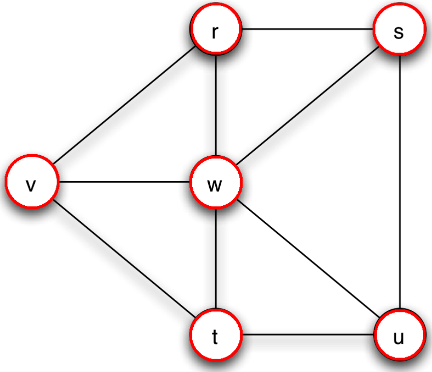


$t = 1$

- A node adopts A if and only if the threshold $q = \frac{b}{a+b} = 2/5$ is reached.
- Two nodes v and w are initial adopters.

The process unfolds (example: $a = 3$ and $b = 2$)

[Fig 19.3, E&K]

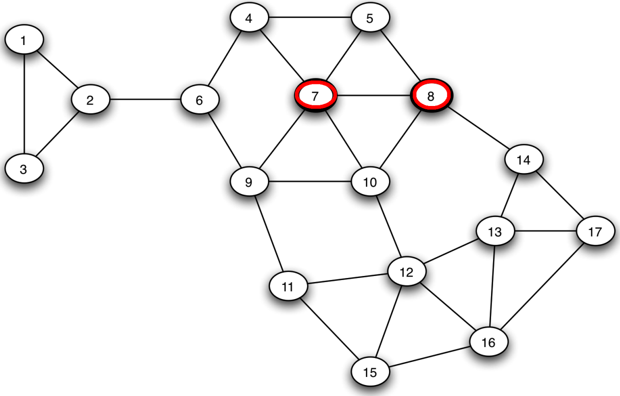


$t = 2$

- A node adopts A if and only if the threshold $q = \frac{b}{a+b} = 2/5$ is reached.
- Two nodes v and w are initial adopters.

Complete cascades vs tightly-knit communities (example: $a = 3$, $b = 2$, $q = 2/5$)

- The previous example showed a complete cascade where all nodes eventually adopt A.
- In the next example, "tightly-knit communities" block the spread.

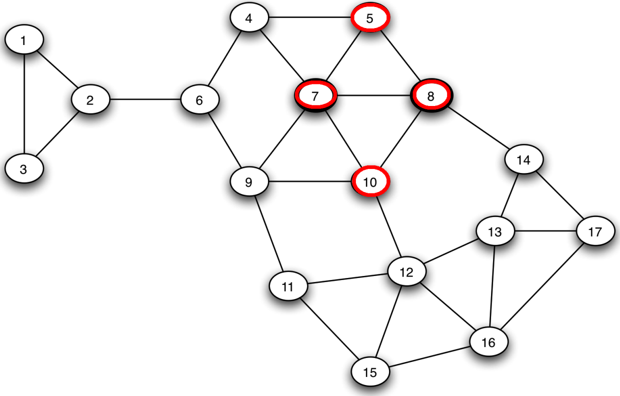


[Fig 19.4, E&K]

$t = 0$

Complete cascades vs tightly-knit communities (example: $a = 3$, $b = 2$, $q = 2/5$)

- The previous example showed a complete cascade where all nodes eventually adopt A.
- In the next example, "tightly-knit communities" block the spread.

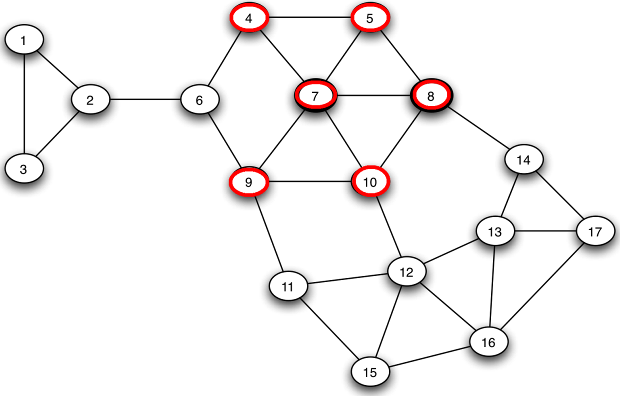


$t = 1$

[Fig 19.4, E&K]

Complete cascades vs tightly-knit communities (example: $a = 3$, $b = 2$, $q = 2/5$)

- The previous example showed a complete cascade where all nodes eventually adopt A.
- In the next example, "tightly-knit communities" block the spread.

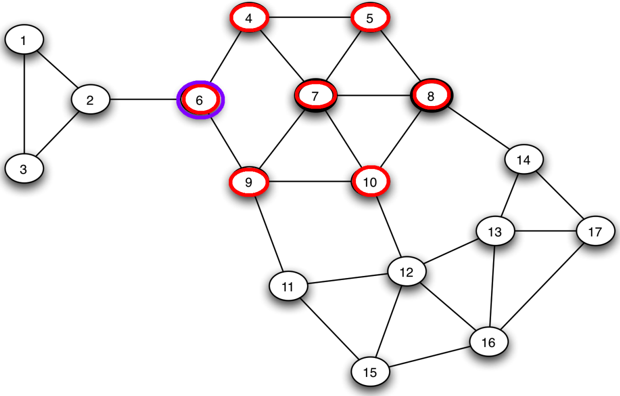


$t = 2$

[Fig 19.4, E&K]

Complete cascades vs tightly-knit communities (example: $a = 3$, $b = 2$, $q = 2/5$)

- The previous example showed a complete cascade where all nodes eventually adopt A.
- In the next example, "tightly-knit communities" block the spread.



$t = 3$

[Fig 19.4, E&K]

Factors determining the rate and extent of diffusion in a social network

1. The structure of the network.

2. The relative payoffs vs costs for adopting a new product.
   - We haven't spoken of costs yet but we usually do have a cost for adopting a new product.
   - We can introduce such a cost into the model by saying that v will not adopt the new A unless

$$p \times d(v) \times a \geq (1-p) \times d(v) \times b + cost$$

   - We could also add intrinsic values for A and B to both sides of the above inequality to determine the threshold for v adopting A.

3. The choice of initial adopters.
   - This raises an interesting computational question as to how to select the most influential nodes (within some budgetary constraint).

Defining a tightly-knit community

- We want to show that not only do tightly-knit communities cause a cascade to be blocked but moreover this is the only thing that can stop a cascade.
- To do so, we need a more precise definition.

Definition

A non-empty subset S of nodes is a blocking cluster of density p if every node $v \in S$ has at least a fraction p of its edges go to nodes in S.

Aside

- Clustering is a pervasive concept in many fields and contexts (beyond networks).
- It is an intuitive concept that can be defined in many ways.
- There does not appear to be any one definition that is always (or even usually) most preferred.

Clusters at different levels of granularity

- The given definition of a blocking cluster does not imply a unique way of clustering the nodes.
- Indeed if S and T are both clusters of density p, then the union of S and T is a cluster of density p.
  - Note: this is not generally true of the intersection of S and T.
- This clustering definition also implies that the set of all nodes is a cluster of density 1.

Clusters vs complete cascades

- Suppose we have a network threshold spread model with threshold q, an initial set of A adopters I and $V' = V - I$ is the set of nodes that are not initial adopters.

- Then we have the following (provable) intuitive result that characterizes when complete clusters will or will not form:

  - If V' contains a cluster C of density greater than $1 - q$, then the initial adopters will not cause a complete cascade. Furthermore, no node in C will adopt A.
  - If in a network with threshold q and an initial set I of adopters does not cause a complete cascade, then the non initial adopters nodes $V' = V - I$ must contain a cluster of density greater than $1 - q$.

When nodes have different thresholds

- As remarked before the assumption that all nodes have the same threshold is not essential.
- Consider a node v. Suppose now that for every adjacent edge (v, w), node v has payoff $a(v)$ (resp. $b(v)$) if both v and w have adopted product A (resp. B) and a zero payoff if v and w currently utilize different products.
- If node v has not yet adopted A at time t, but a fraction p of the $d(v)$ neighbours of v have already adopted A, then:
  - By switching, v has payoff $p \times d(v) \times a(v)$.
  - By staying with B, v has payoff $(1 - p) \times d(v) \times b(v)$.
- Thus node v will switch to A if

$$p \times d(v) \times a(v) \geq (1 - p) \times d(v) \times b(v).$$

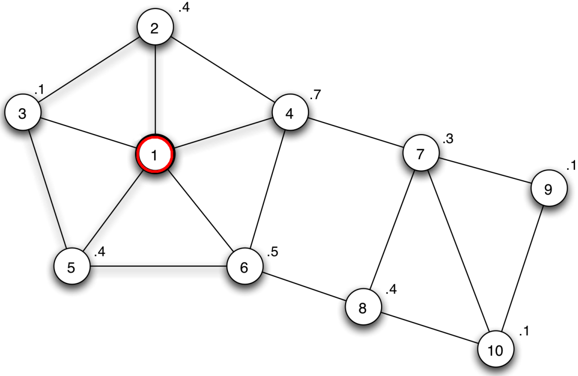
- This is then equivalent to saying that v will switch whenever

$$p \geq \frac{b(v)}{a(v) + b(v)} = q(v)$$

which is then the threshold for node v.

Redefining blocking clusters

- A blocking cluster is now a set of nodes C such that every node $v \in C$ has more than a fraction $1 - q(v)$ of its adjacent nodes in C.
- It follows (as in the case of homogenous threshold nodes) that a given set of adopters I in a network will not cause a complete cascade iff $V - I$ contains a blocking cluster C.

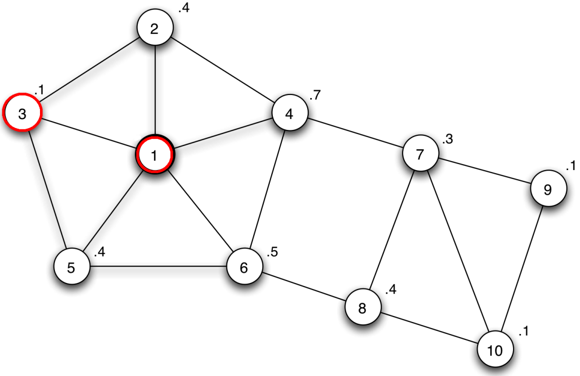


$t = 0$

[Fig 19.13, E&K]

Redefining blocking clusters

- A blocking cluster is now a set of nodes C such that every node $v \in C$ has more than a fraction $1 - q(v)$ of its adjacent nodes in C.
- It follows (as in the case of homogenous threshold nodes) that a given set of adopters I in a network will not cause a complete cascade iff $V - I$ contains a blocking cluster C.

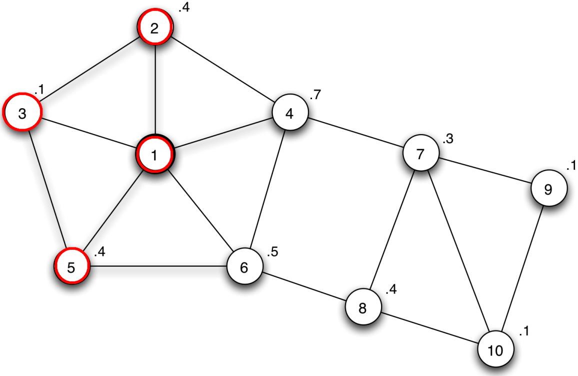


$t = 1$

[Fig 19.13, E&K]

Redefining blocking clusters

- A blocking cluster is now a set of nodes C such that every node $v \in C$ has more than a fraction $1 - q(v)$ of its adjacent nodes in C.
- It follows (as in the case of homogenous threshold nodes) that a given set of adopters I in a network will not cause a complete cascade iff $V - I$ contains a blocking cluster C.

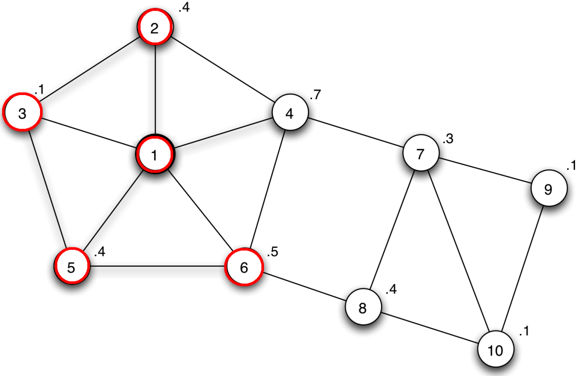


[Fig 19.13, E&K]

$t = 2$

Redefining blocking clusters

- A blocking cluster is now a set of nodes C such that every node $v \in C$ has more than a fraction $1 - q(v)$ of its adjacent nodes in C.
- It follows (as in the case of homogenous threshold nodes) that a given set of adopters I in a network will not cause a complete cascade iff $V - I$ contains a blocking cluster C.

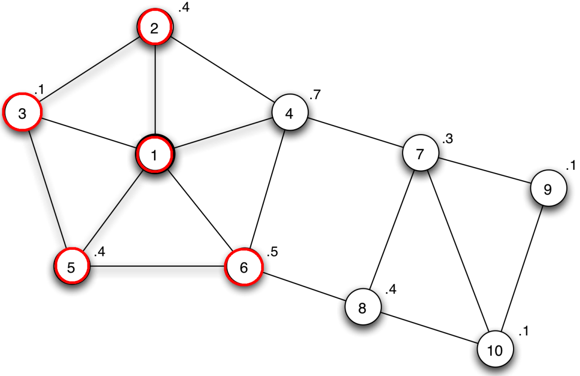


$t = 3$

[Fig 19.13, E&K]

Redefining blocking clusters

- A blocking cluster is now a set of nodes C such that every node $v \in C$ has more than a fraction $1 - q(v)$ of its adjacent nodes in C.
- It follows (as in the case of homogenous threshold nodes) that a given set of adopters I in a network will not cause a complete cascade iff $V - I$ contains a blocking cluster C.

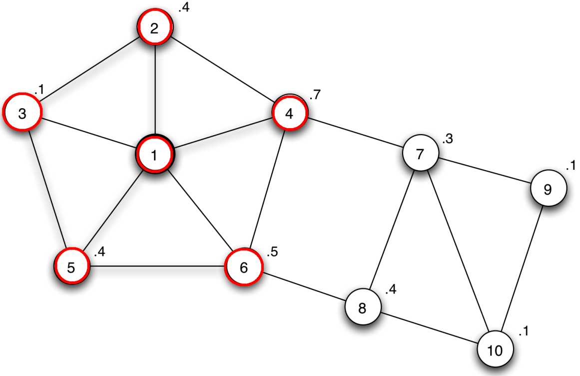


$t = 4$

[Fig 19.13, E&K]

Redefining blocking clusters

- A blocking cluster is now a set of nodes C such that every node $v \in C$ has more than a fraction $1 - q(v)$ of its adjacent nodes in C.
- It follows (as in the case of homogenous threshold nodes) that a given set of adopters I in a network will not cause a complete cascade iff $V - I$ contains a blocking cluster C.



[Fig 19.13, E&K]

$t = 5$

Choosing influential adopters

- Suppose we wish to spread a new technology and to do so we have money to influence some "small" set of initial adopters (e.g. by giving away the product or even paying people to adopt it).

- Even in this simple model of (non-competitive) influence spread, and even if we have complete knowledge of the social network, it is not at all clear how to chose an initial set of adopters so as to achieve the largest spread.

- Furthermore the spread process could be much more sophisticated.
  - For example, adoption by a node might be a more random process (say adopting with some probability relative to the nodes threshold) and maybe the influence of neighbors first increases and then decreases over time.

Choosing influential adopters continued

- Suppose we have funds/ability to influence k nodes to become initial adopters.
  - We can try all possible subsets of the entire $n = |V|$ nodes and for each such subset simulate the spread process.
  - But clearly as k gets larger, this "brute force" becomes prohibitive for large (and not even massive) networks.
- It turns out that the problem of the optimum set of initial adopters in many settings is an NP-hard problem.

Can we determine a "good" set of initial adopters?

- For even simple models of information spread as being discussed here, complexity theory (the P vs NP conjecture) argues that we cannot efficiently choose the best set of initial adopters. There is a class of networks for which (assuming the $P \neq NP$ conjecture) it is not possible to obtain an approximation within a factor n^c for any $c < 1$.

- Instead we will identify properties of a spread process that will allow a good approximation: a good set of initial adopters that will do "almost as well" as the best set.

Note: What follows is a discussion as to how to choose a set of initial adopters by a relatively efficient approximation algorithm when making some assumptions on the spread process. However, as we discussed we would need much more efficient methods for massive networks.

Influence maximization models; monotone submodular set functions

- Some spread models have the following nice properties.

Let $f(S)$ be size (or more generally a real value benefit since some nodes may be more valuable) of the final set S of adopters satisfying:

1. Monotonicity: $f(S) \leq f(T)$ if S is a subset of T
2. Submodularity: $f(S+v) - f(S) \geq f(T+v) - f(T)$ if S is a subset of T

- We also usally assume that $f(\varnothing) = 0$. Such normalized, monotone, submodular functions arise in many applications.
- The simple threshold examples considered thus far are monotone processes but are not submodular in general. Are these contrived worst case network examples?
- But some variants of the threshold model and related models do satisfy these properties. We consider two such stochastic models.