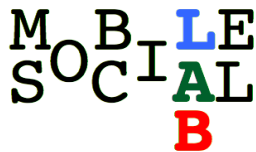


Lecture 20: Influence (3/4)

“How are our neighbors influencing us?
How can my start-up become viral?”

COMS 4995-1: Introduction to Social Networks
Tuesday, November 27th



Outline

- * Life under the influence
- * 1978: The global view, 'Network effect'
- * 2000: The local view, 'Neighbors effect'
- * 2003: The algorithmic view, 'Exploiting Influence'

Outline

- * Life under the influence
- * 1978: The global view, 'Network effect'
- * 2000: The local view, 'Neighbors effect'
- * 2003: The algorithmic view, 'Exploiting Influence'

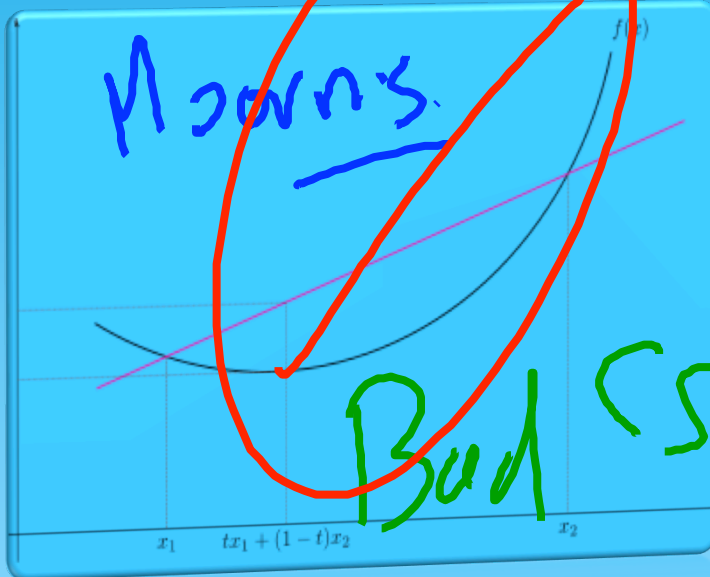
A general algorithmic problem

- * How to find the best initial seeding set S_0 ?
 - Maximizing the total spread, with a fixed size
- * A more general model of neighbor influence
 - Assumes threshold t_v uniform in $[0;1]$ and v becomes active as soon as $t_v \leq g_v(X)$
 - as u becomes active, activates neighbor v with prob. $p_v(u,X)$, where $X = \{\text{nodes in } N(v) \text{ previously active}\}$
 - Special cases: Granovetter, Morris, Independent
 - If p order independent, the two models equivalent

Critical Mass vs. Dim. Return

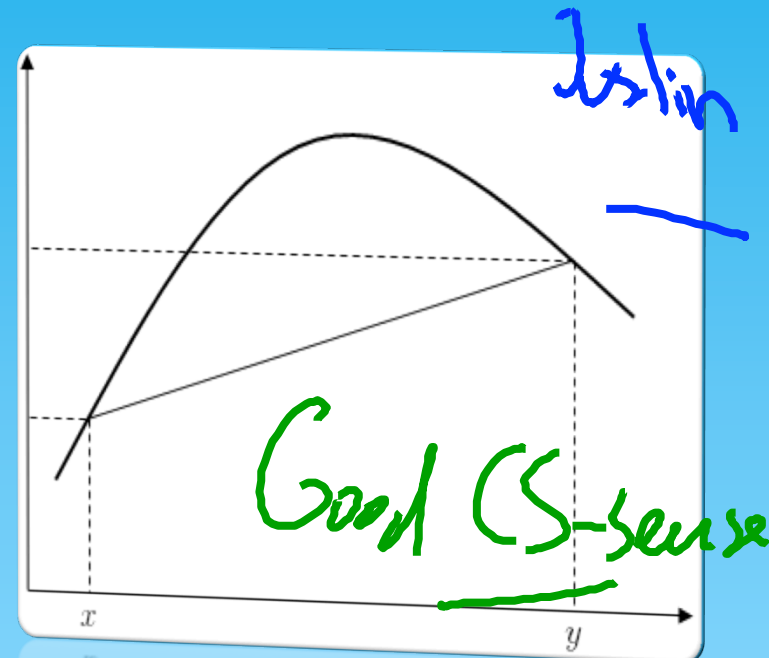
$p_v(u,X)$ “increases” with X

- Infection gets easier



$p_v(u,X)$ “decreases” with X

- infection gets harder



Maximizing spread of influence

- * Let us denote the objective function by $f(S)$
 - $f(S) = E[\# \text{ of active nodes}]$ at the end of the process
 - Assuming process starts with S
- * Thm: Computing $\max\{f(S) \mid |S|=k\}$ is NP hard
 - Even computing a $n^{1-\epsilon}$ approximation is hard ($\epsilon > 0$)
 - Proof: If g_v shows “critical mass” then relates to a set covering problem
 - Can we avoid “critical mass” and compute it?

Maximizing spread of influence

- * Thm: Whenever p_v show diminishing return
 - There exists a simple polynomial algorithm computing S such that $f(S) \geq (1-1/e)f(S^*)$ where S^* is the optimal subset of size k
 - Algorithm follows greedy “one node at a time” rule
Do k times: $S \leftarrow S \cup \operatorname{argmax}_v \{ f(S \cup \{v\}) - f(S) \}$

Maximizing the spread of influence through a social network,
D. Kempe, J. Kleinberg, E. Tardos, ACM KDD (2003)

Proof

* Three steps:

1. Show that the result holds if f is submodular, i.e.

$$S \subseteq T \text{ Implies } f(S \cup \{v\}) - f(S) \geq f(T \cup \{v\}) - f(T),$$

2. Show f is submodular under this condition on p_v
3. Finally, prove that each step is polynomial more involved (will be admitted here)