Social diffusion

Leonid E. Zhukov

School of Applied Mathematics and Information Science
National Research University Higher School of Economics

21.04.2014

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

Models of "diffusion"

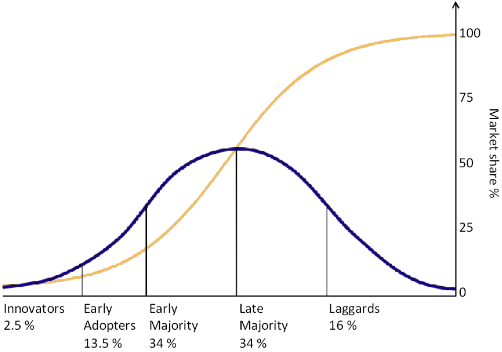
Two classes of models

- Viral propagation:
  - virus and infection, rumors, news
  - SI, SIS, SIR
- Decision based models:
  - adoption of innovation, joing a political party
  - Threshold models

*We will talk about social diffusion , not a physical diffusion process

Diffusion of innovation

Everett Rogers (sociologist) , "Diffusion of innovation" book, 1962
Theory that trys to explain how, why, at what rate new ideas, innovations, spread around



Bass Diffusion model

Frank Bass, 1969, "A new product growth model for consumer durables"

- Growth model of how new products get adopted
- Two types of agents, two key parameters:
  - p - innovation or spontanious adoption rate (coefficient of innovation)
  - q - rate of immitation (coefficient of immitation)
- Let $F(t)$ fraction of agents adopted by time t

$$F(t+1) = F(t) + p(1 - F(t))\delta t + q(1 - F(t))F(t)\delta t$$

$$\frac{dF(t)}{dt} = (p + qF(t))(1 - F(t))$$

Bass Diffusion model

- if only innovators, $q = 0$, exponential function:

$$\frac{dF(t)}{dt} = p(1 - F(t))$$

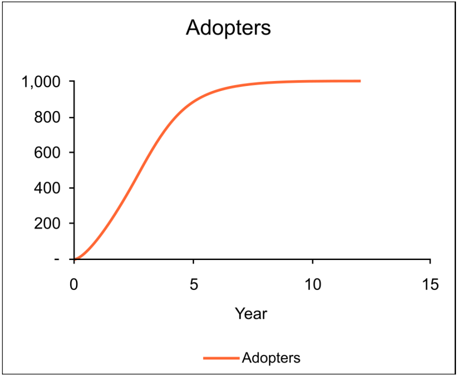
- if only immitators, $p = 0$, logistic function:

$$\frac{dF(t)}{dt} = qF(t)(1 - F(t))$$

Bass Diffusion model

Soltion of Bass model - S-curve. When $F(0) = 0$

$$F(t) = \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p} e^{-(p+q)t}}$$



Emprical $p \sim 0.01 - 0.03$, $q \sim 0.3 - 0.5$, when t in years

Threshold models

Threshold models of Collective Behavior. Mark Granovettor, 1978

- Adoption of innovation, voting, applause, leaving social occasion, riots
- Group of people, each to make a decision
- Binary mutually exclusive decision: adopt/reject, stay/go, join/not join
- Every person has own preference, decision threshold
- Costs and benefits depends on how many others make which choice
- Dynamical proces with equilibrium outcome (final proportion of making each decision)
- Example: insitgator + crowd
  $5, 5, 5, 5, 5, 5, 5, 5, 5$
  $1, 2, 3, 4, 5, 6, 7, 8, 9$ (domino effect)
  $2, 3, 4, 5, 6, 7, 8, 9$

Threshold models

- Let $i's$ threshold level $\theta(i)$, x - number of participans
- if $x \geq \theta(i)$ - join, $x < \theta(i)$ - not joint
- Let $f(x)$ - number of people with thresholod level $\theta = x$
  $F(x)$ - number of people with $\theta \leq x$ (cumulative function)

$$F(x) = \sum_{x'}^{x} f(x')$$

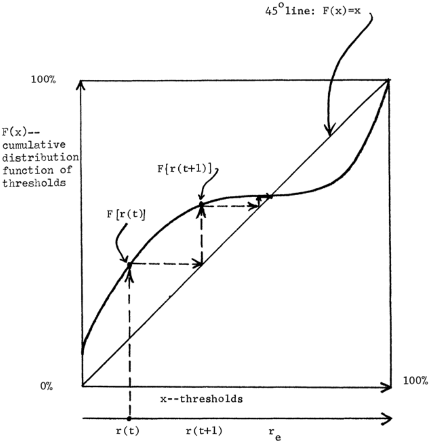
- Inititial state x_0 -already joined
- First time step:
  there are $F(x_0)$ people with threshold $\theta \leq x$ ready to join

$$x_1 = F(x_0)$$
$$x_2 = F(x_1)$$
$$x_{t+1} = F(x_t)$$

- Fixed point of the dynamical model

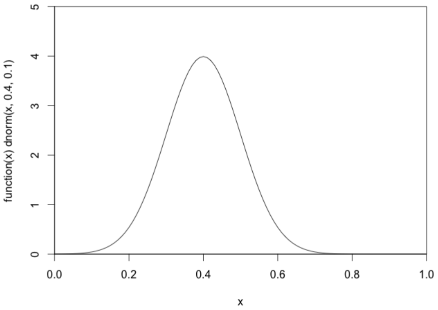
$$x^* = F(x^*)$$

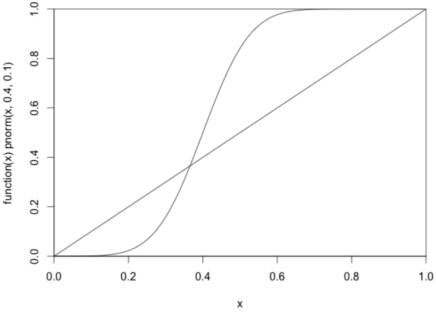
Granovetter model



$y = x$
$y = F(x)$

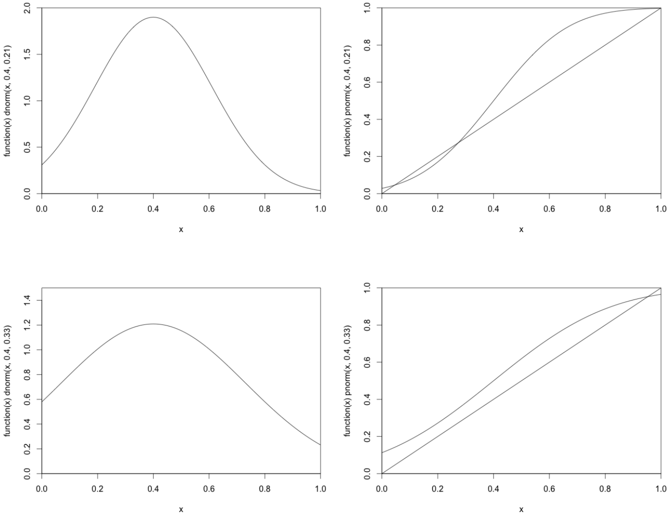
Granovetter model





$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x-\mu}{\sigma\sqrt{2}}} \qquad F(x) = \frac{1}{2}(1 + erf(\frac{x-\mu}{\sigma\sqrt{2}}))$$

Granovetter model



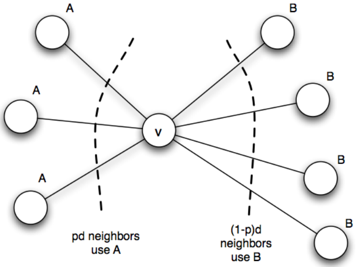
Network coordination game

Let u and v are players, and A and b are possible strategies
Payoffs

- if u and v both adopt behavior A, each get payoff $a > 0$
- if u and v both adopt behavior B, each get payoff $b > 0$
- if u and v adopt oppsoite behavior, each get payoff 0

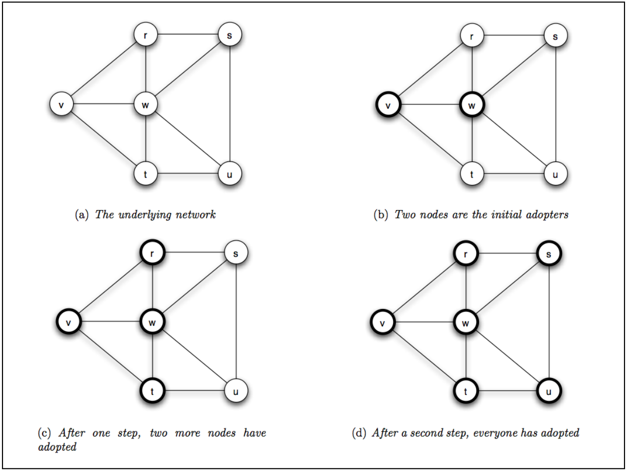
		w: A	w: B
v	A	a, a	$0, 0$
	B	$0, 0$	b, b

Network model



Node v to make decision A or B, p - portion of type A neighbors
To accept A:
$a \cdot p \cdot d > b \cdot (1 - p) \cdot d$
$p > b/(a + b)$

Cascades



(a) *The underlying network* (b) *Two nodes are the initial adopters*

(c) *After one step, two more nodes have adopted* (d) *After a second step, everyone has adopted*

$a = 3, b = 2$, threshold $p > 2/5$

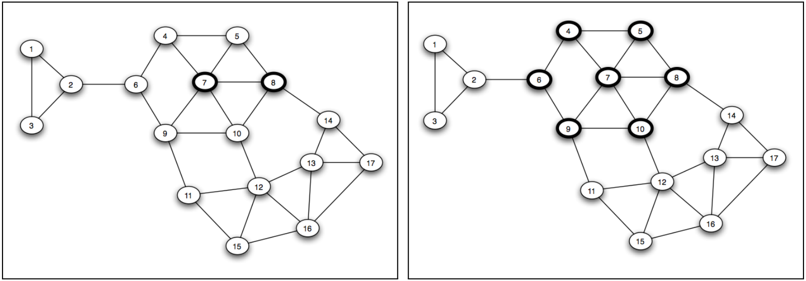
Linear threshold model

- Influence comes only from NN $N(i)$ nodes, w_{ij} influence $i \to j$
- Require $\sum_{j \in N(i)} w_{ji} \leq 1$
- Each node has a random acceptance threshold from $\theta_i \in [0, 1]$
- Activation: fraction of active nodes exceeds threshold

$$\sum_{active\ j \in N(i)} w_{ji} > \theta_i$$

- Initial set of active nodes A_o, iterative process with discrete time steps
- Progressive process, only nonactive $\to$ *active*

Maximal Cascades



- Inital set of active nodes A_o
- Cascade size $\sigma(A_o)$ - number of active nodes when propagation stops
- Find k-set of nodes A_o that produces maximal cascade $\sigma(A_o)$
- k-set of "maximum influence"nodes
- NP-hard

Submodular functions

- Set function f is submodular, if for sets S, T and $S \subseteq T$, $\forall v \notin T$

$$f(S \cup \{v\}) - f(S) \geq f(T \cup \{v\}) - f(T)$$

- Function of diminishing returns
- Function f is monotone, $f(S \cup \{v\}) \geq f(S)$

Theorem

Let F be a monotone submodular function and let S^* be the k-element set achieving maximal f. Let S be a k-element set obtained by repeatedly, for k-iterations, including an element producing the largest marginal increase in f.

$$f(S) \geq (1 - \frac{1}{e}) f(S^*)$$

Nemhauser, Wolsey, and Fisher, 1978

Influence maximization

- $\sigma()$- submodular function (D. Kempe, J. Kleinberg, E. Tardos, 1993)

$$\sigma(S) \geq (1 - \frac{1}{e})\sigma(S^*)$$

- Greedy algorithm for maximum influence set finds a set S such that its influence set $\sigma(S)$ is within $1/e = 0.367$ from the optimal (maximal) set $\sigma(S^*)$, $\sigma(S) \geq 0.629\sigma(S^*)$

Influence maximization

Approximation algorithm

Algorithm: Greedy optimization

Input: Graph $G(V, E)$, k

Output: Maxumum influence set S

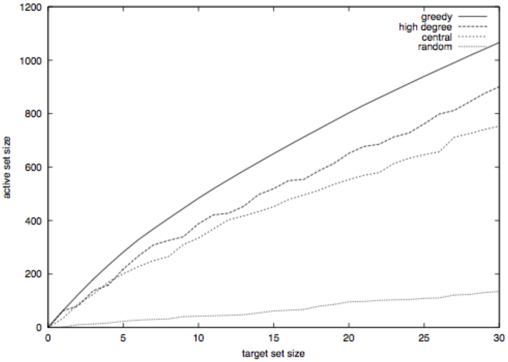
Set $S \leftarrow 0$

for $i = 1 : k$ **do**

  select $v = \arg\max_{u \in V \setminus S}(\sigma(S \cup \{u\}) - \sigma(S))$

  $S \leftarrow S \cup \{v\}$

Linear threshold model



network: collaboration graph
10,000 nodes, 53,000 edges

References

- Threshold Models of Collective Behavior Mark S. Granovetter, American Journal of Sociology 83(6):1420-1443, 1978.
- Maximizing the Spread of Influence through a Social Network, D. Kempe, J. Kleinberg, E. Tardos, 2003
- Influential Nodes in a Diffusion Model for Social Networks, D. Kempe, J. Kleinberg, E. Tardos