NATIONAL RESEARCH
UNIVERSITY

Diffusion of Innovation

Social Network Analysis. MAGoLEGO course.
Lecture 7

Leonid Zhukov

lzhukov@hse.ru
www.leonidzhukov.net/hse/2019/sna

National Research University Higher School of Economics
School of Data Analysis and Artificial Intelligence, Department of Computer Science

Spreading process

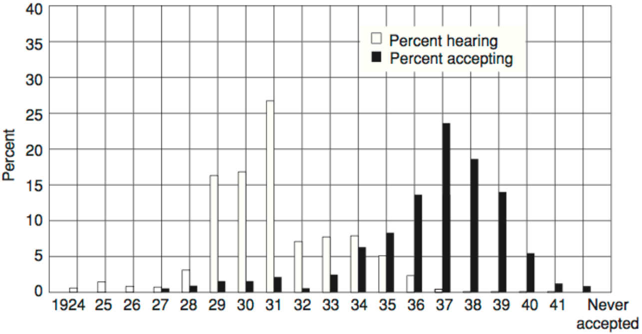
Propagation process:

- Information based models:
  - ideas, knowledge
  - virus and infection
  - rumors, news
- Decision based models:
  - adoption of innovation
  - joining political protest
  - purchase decision

Local individual decision rules will lead to very different global results.
"microscopic" changes $\rightarrow$ "macroscopic" results

Ryan-Gross study

Ryan-Gross study of hybrid seed corn delayed adoption (after first exposure)

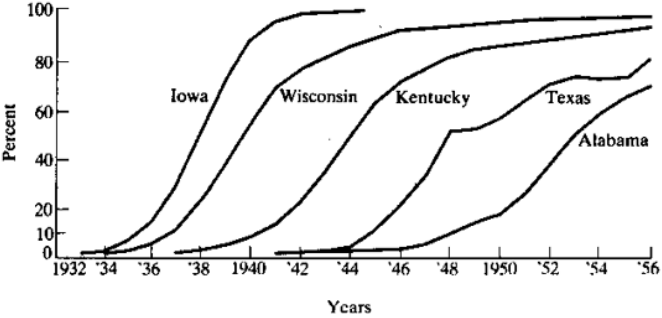


Information effect vs adopting of innovation

Ryan and Gross, 1943

Ryan-Gross study

Hybrid corn adoption

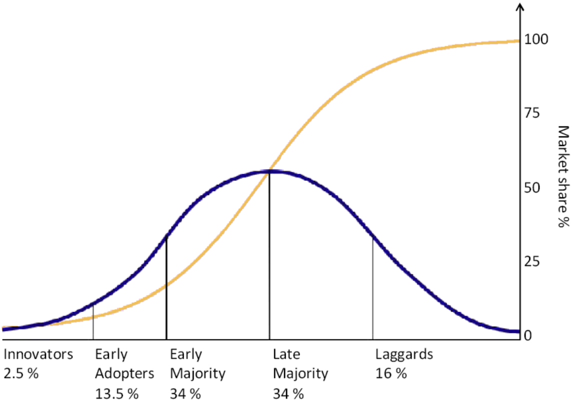


Percentage of total acreage planted

Griliches, 1957

Diffusion of innovation

Everett Rogers, "Diffusion of innovation" book, 1962



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

Diffusion of innovation

What influences potential adopters:

- relative advantage of the innovation
- compatibility with current ways of doing things
- complexity of the innovation
- triability - the ease of testing
- observability of results

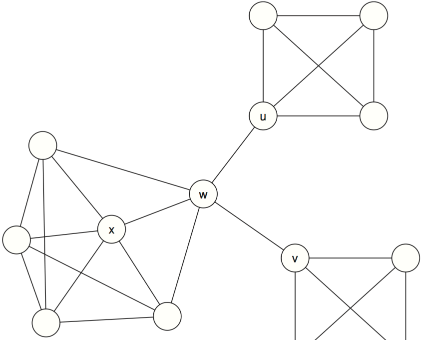
Some questions remain:

- how a new technology can take over?
- who different technologies coexist?
- what stops new technology propagation?

Everett Rogers,1962

Network structure

From the population level to local structure



Network coordination game

Local interaction game: Let *u* and *v* are players, and *A* and *b* are possible strategies

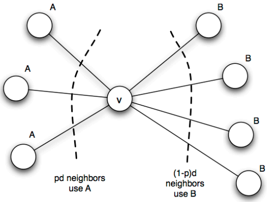
Payoffs

- 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 opposite behavior, each get payoff 0

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

Threshold model

Network coordination game, direct-benefit effect



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

Node *v* to make decision *A* or *B*, *p* - portion of type *A* neighbors to accept *A*:

$$a \cdot p \cdot d > b \cdot (1 - p) \cdot d$$

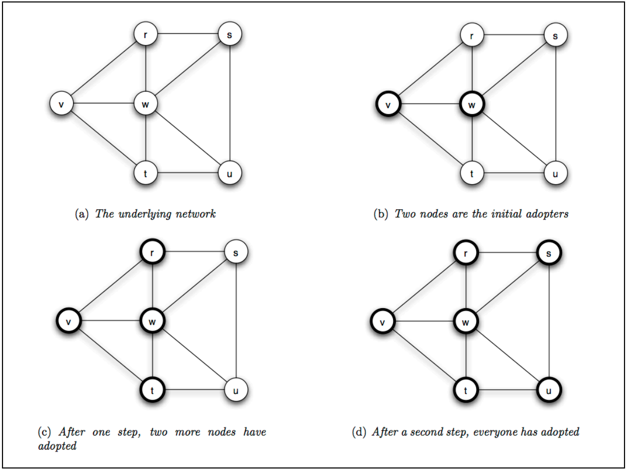
$$p \geq b/(a + b)$$

Threshold:

$$q = \frac{b}{a + b}$$

Cascades

Cascade - sequence of changes of behavior, "chain reaction"



(a) *The underlying network*

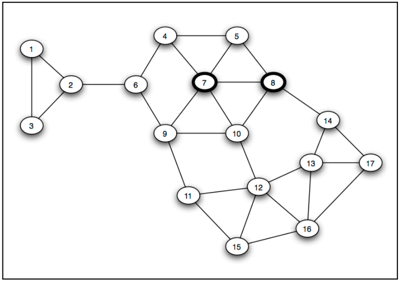
(b) *Two nodes are the initial adopters*

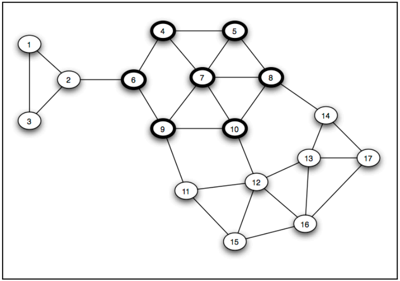
(c) *After one step, two more nodes have adopted*

(d) *After a second step, everyone has adopted*

Let $a = 3, b = 2$, threshold $q = 2/(2+3) = 2/5$

Cascade propagation

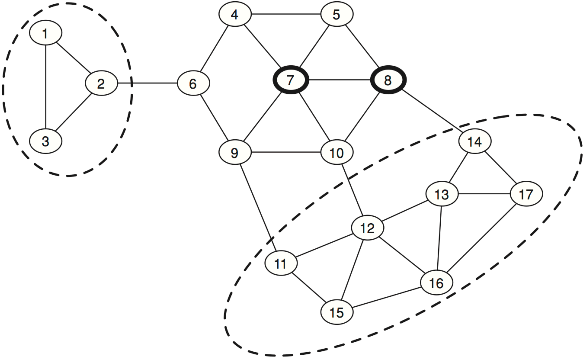




- Let $a = 3$, $b = 2$, threshold $q = 2/(2+3) = 2/5$
- Start from nodes 7,8: $1/3 < 2/5 < 1/2 < 2/3$
- Cascade size - number of nodes that changed the behavior
- Complete cascade when every node changes the behavior

Cascades and clusters

Group of nodes form a cluster of density ρ if every node in the set has at least fraction ρ of its neighbors in the set



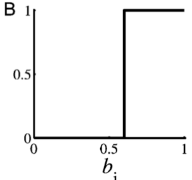
Both clusters of density $\rho = 2/3$. For cascade to get into cluster $q \leq 1 - \rho$.

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

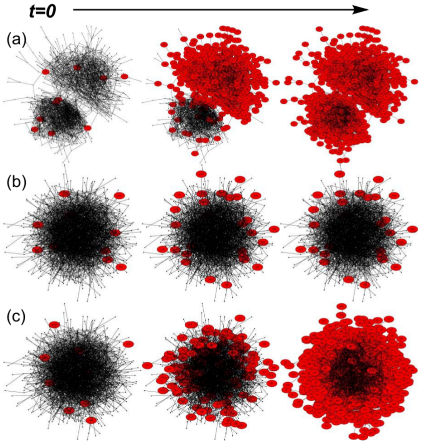
$$\sum_{active\ j \in N(i)} w_{ji} > \theta_i$$

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



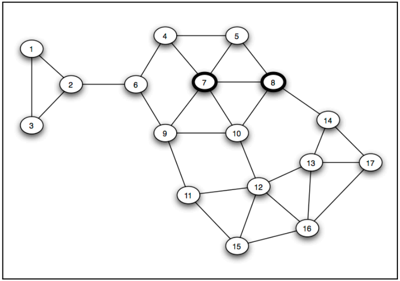
Cascades in random networks

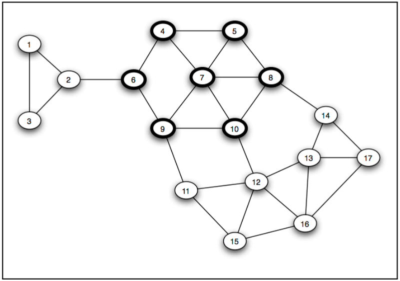
multiple seed nodes



(a) Empirical network; (b), (c) - randomized network

Influence maximization problem





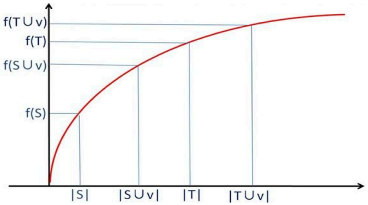
- Initial set of active nodes A_o
- Cascade size $\sigma(A_o)$ - expected number of active nodes when propagation stops
- Find k-set of nodes A_o that produces maximal cascade $\sigma(A_o)$
- k-set of "maximum influence" nodes
- NP-hard

D. Kempe, J. Kleinberg, E. Tardos, 2003, 2005

Influence maximization

Greedy maximization algorithm:
Given: Graph and set size k
Output: Maximum influence set A

1. Select a node v_1 that maximizes the influence $\sigma(v_1)$
2. Fix v_1 and find v_2 such that maximizes $\sigma(v_1, v_2)$
3. Repeat k times
4. Output maximum influence set: $A = \{v_1, v_2 ... v_k\}$



Approximation algorithm

Algorithm: Greedy optimization

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

Output: Maxumum influence set S

Set $S \leftarrow 0$

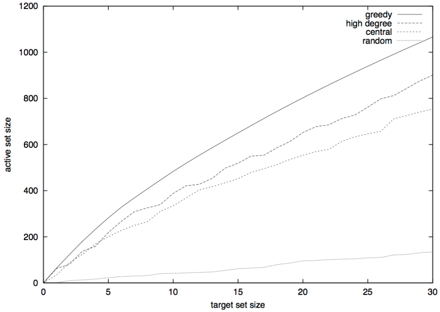
for $i = 1 : k$ **do**

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

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

Experimental results

Linear threshold model
network: collaboration graph 10,000 nodes, 53,000 edges



Greedy algorithm finds a set S such that its influence set $\sigma(S)$ is $\sigma(S) \geq (1 - \frac{1}{e})\sigma(S^*)$ from the true optimal (maximal) set $\sigma(S^*)$

D. Kempe, J. Kleinberg, E. Tardos, 2003

References

- Contagion, S. Morris, Review of Economic Studies, 67, p 57-78, 2000
- 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, 2005
- A Simple Model of Global Cascades on Random Networks. D. Watts, 2002.