Social Network Analysis

CASCADE BEHAVIORS AND NETWORK EFFECTS

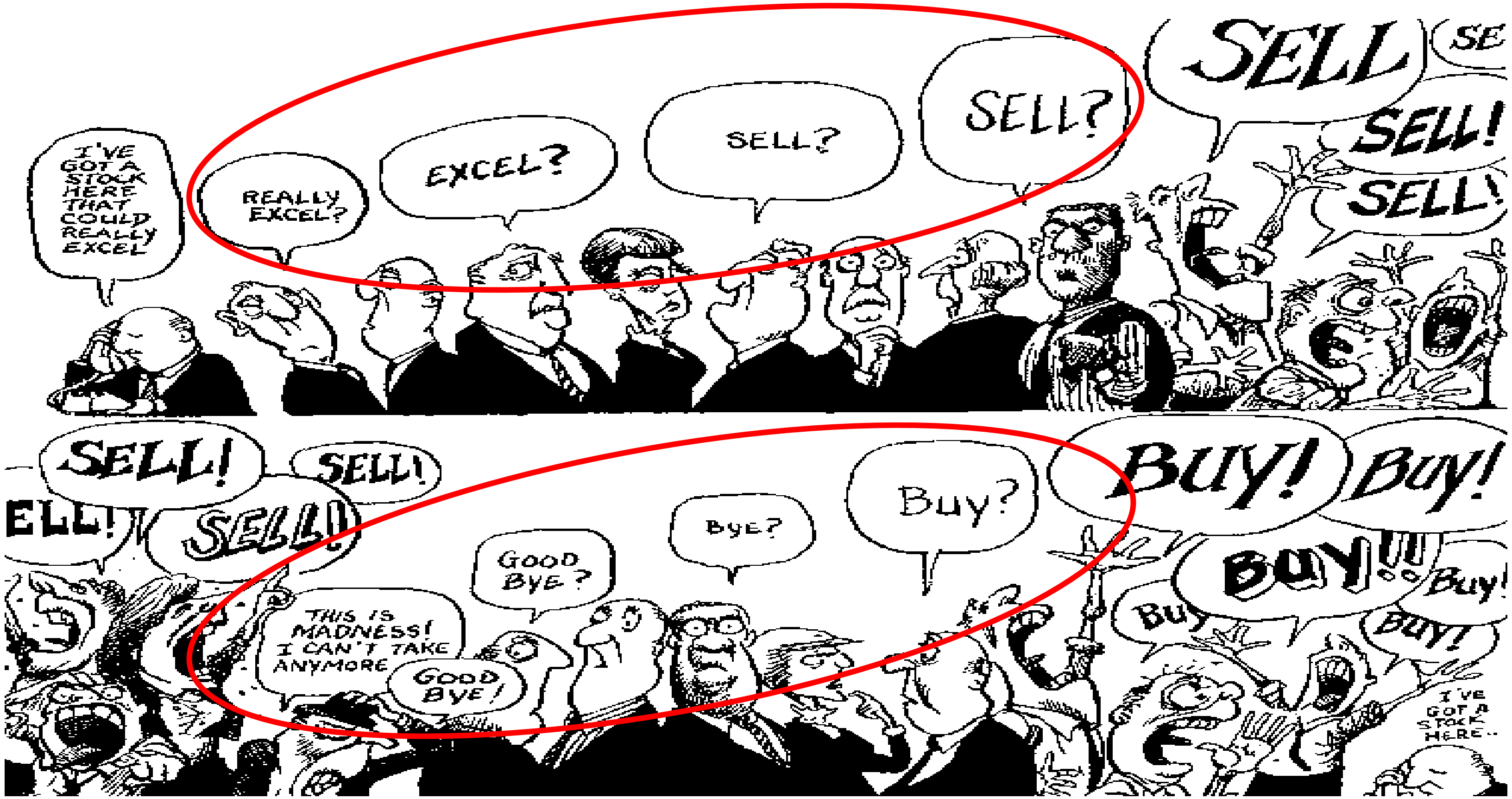
Information Diffusion

- Diffusion is the net movement of anything from a region of higher concentration to a region of lower concentration
- Driven by a gradient in concentration
- Information Diffusion is the process by which information is spread from one place to another through interactions
- Diffusion process involves three main elements:
  - Sender: An entity (or a group of entities) responsible for initiating the diffusion process
  - Receiver: An entity (or a group of entities) receives the diffusion information from the sender(s)
  - Medium: The channel through which the diffusion information is sent from the sender(s) to the receiver(s).

Cascade Behavior: Real-world Instances

- Healthcare
  - Disease Propagation
  - Epidemic Spreading
- Socio-political Cascades
  - Arab Spring Movement in 2010 – 2012
  - From small protests began in Leipzig to Fall of the Berlin Wall in 1989
  - #MeToo movement against sexual abuse and sexual harassment
- Financial Market Cascades
  - Market Bubble
  - A stock becomes overly popular among investors
  - Viral marketing
- Social network
  - Rumor spread
  - Belief Spread
  - Fake News virality

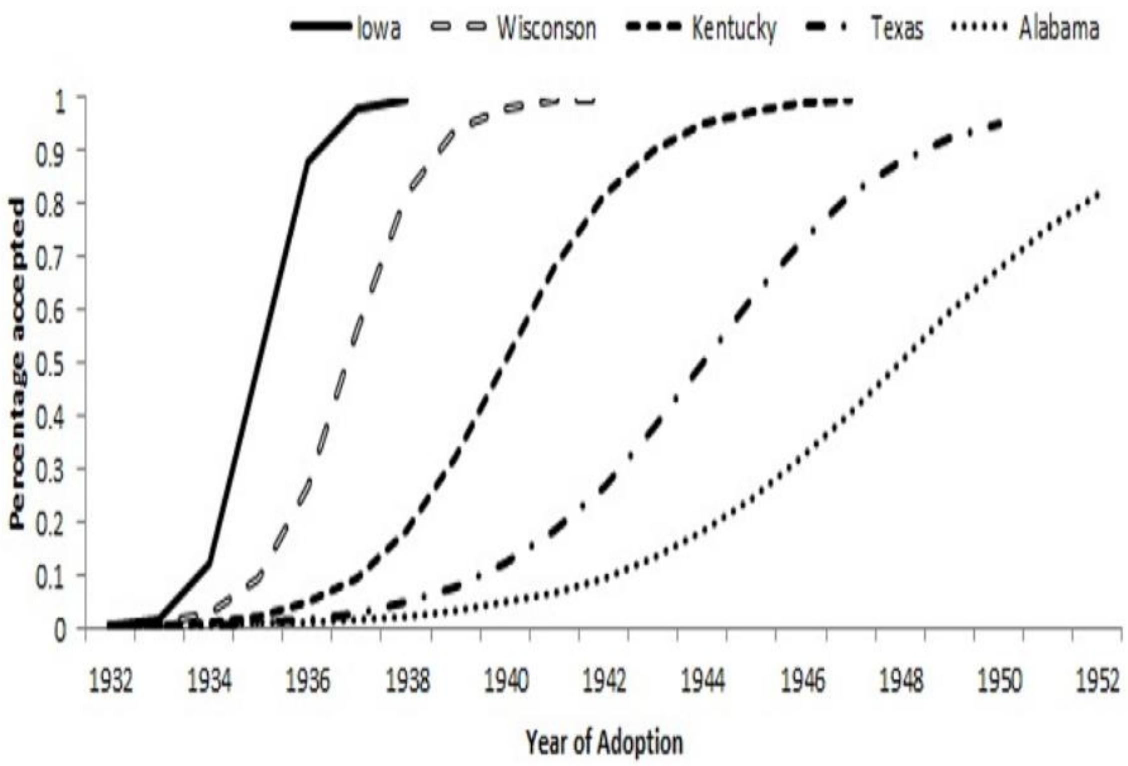
Information Diffusion: Information Cascade



https://blogs.cornell.edu/info2040/2011/11/17/ipos-and-information-cascades/

- An Information cascade is a phenomenon in which a number of people make the same decision in a sequential fashion.
- The phenomenon is found widely in behavioral economics and network theory
- Similar to, but not identical to herd behavior

Information Diffusion: Diffusion of Innovations



Adoption of hybrid seed corn by farmers in USA

http://homepage.cs.uiowa.edu/~sriram/196/spring12/lectureNotes/Lecture15.pdf

- Success/failure of an innovation is highly guided by the structure of the network formed by the initial adopters
- adoption of hybrid seed corn by farmers
- influenced by their neighbors in the community
- adoption of a new drug by the doctors
- Assurance from social peer connections

Information Diffusion: Herd Behavior

http://www.kiffingish.com/2013/12/false-sense-of-security.html

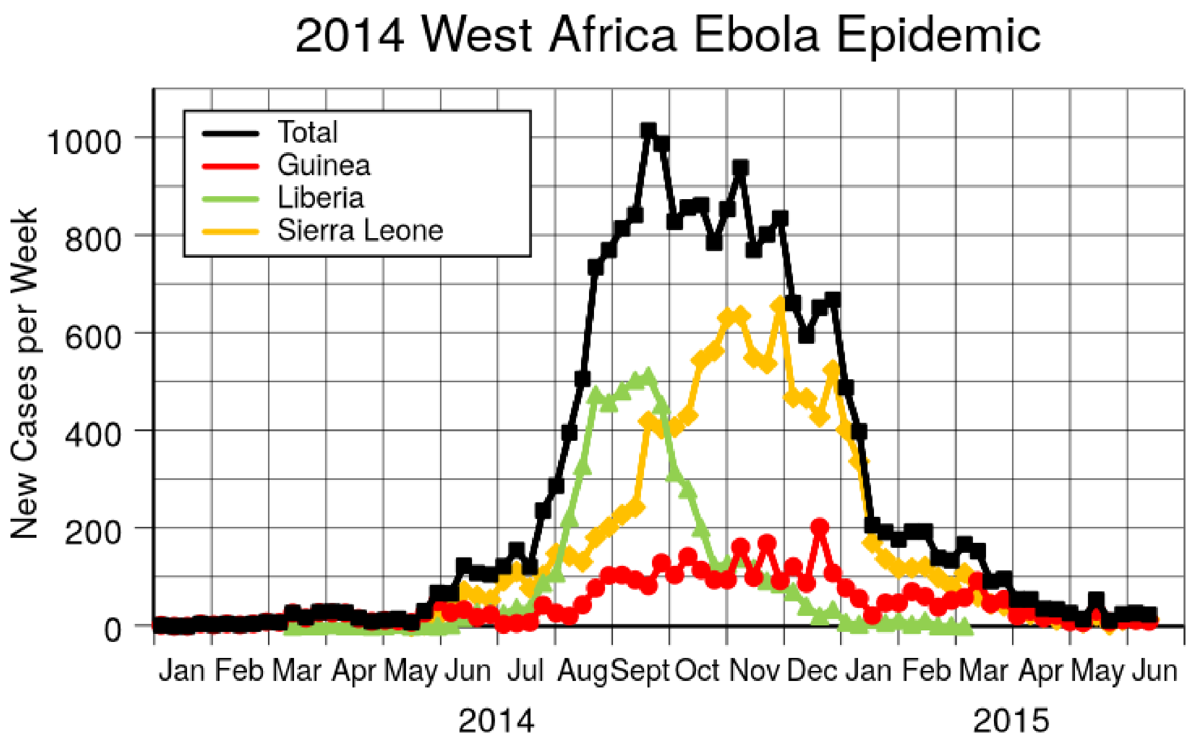
- The behavior of individuals in a group acting collectively without centralized direction
- Human based herd behaviour: demonstrations, riots, general strikes, religious gatherings, judgement and opinion-forming, etc.
- Often a useful tool in marketing; if used properly, can lead to increases in sales
- Herding behavior turns violent sometimes, particularly when confronted by an opposing ethnic or racial group

Information Diffusion: Echo Chambers

https://theconversation.com/the-problem-of-living-inside-echo-chambers-110486

- Situations in which beliefs are amplified or reinforced by communication and repetition inside a closed system
- A harmonious group of people amalgamate and develop tunnel vision
- Social communities become fragmented by echo chambers
- Causes powerful reinforcements of rumors and fake news due to the unchallenged trust in the evidence supplied by their peers

Information Diffusion: Epidemics

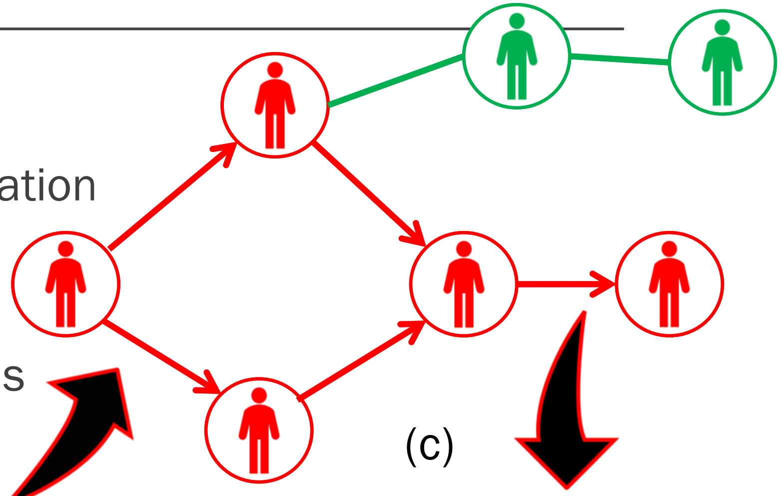


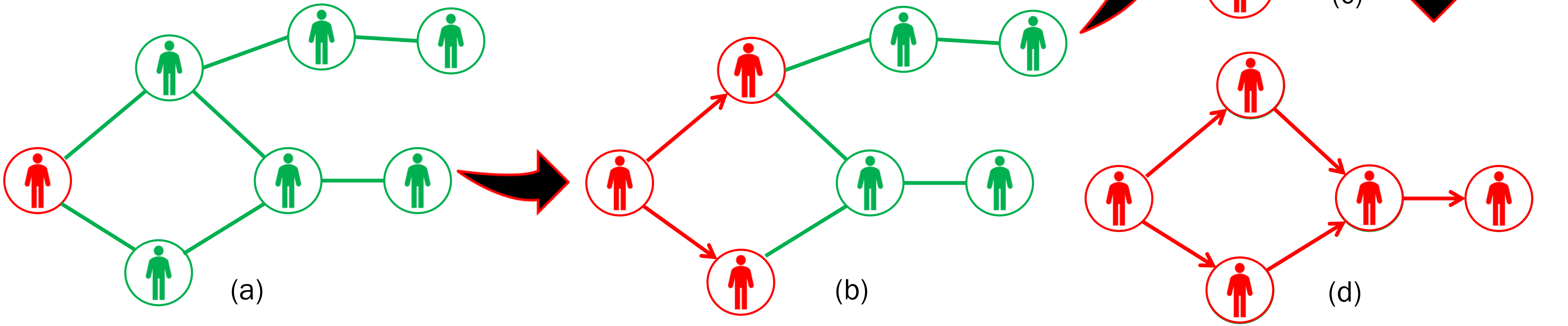
https://commons.wikimedia.org/wiki/File:2014_West_Africa_Ebola_Epidemic_-_New_Cases_per_Week.svg

- Rapid spread of disease to a large number of people in a given population within a short period of time
- Epidemic models are similar to diffusion of innovations models
- Only difference is: individuals do not decide whether to become infected or not

Information Diffusion: Terminologies

- A Contagion is an entity that spreads across a network
- Adoption refers to the event of infection or diffusion. Also known as activation
- Adopters represent the final set of infected nodes
- Final propagation tree obtained by the spread of the infection is known as cascade





Cascade Model: Decision-based Model

- Given a network, each node has the freedom to decide whether to adopt a contagion or not
- Originated from the idea of local interaction models described by Morris in 2000
- Decision at each node is influenced by the behavior of nodes in its neighborhood
- Nodes decide to adopt a new contagion driven by a direct benefit or payoff
- The payoff by adopting a contagion is directly proportional to the number of its neighbors that have adopted the same contagion
- Can be explained using a two-player coordination game
  - Given a number of strategies, the end goal of the players is to coordinate on the same strategy to maximize their payoffs

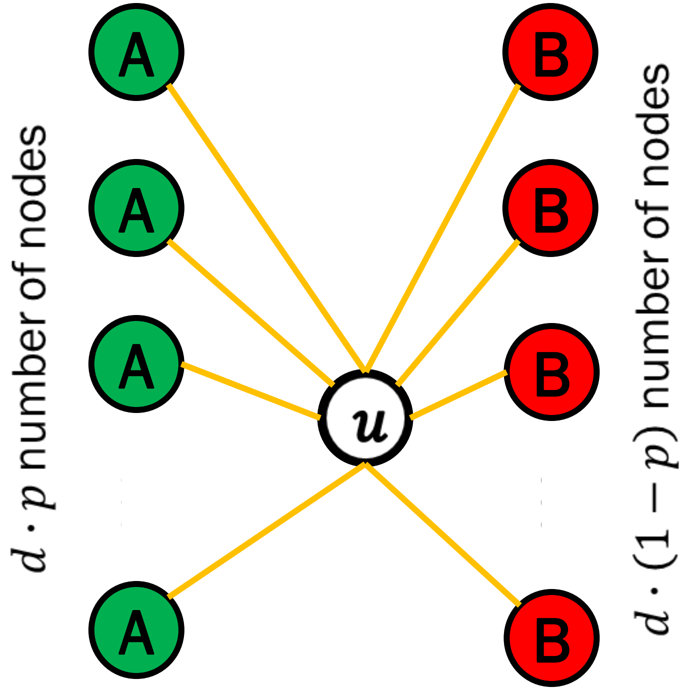
Decision-based Cascade Model: Two-player Coordination Game

$\boldsymbol{u}$'s decision	$\boldsymbol{v}$'s decision	Payoff
A	A	a*
B	B	b*
A	B	0
B	A	0

Payoff distribution for different adoption strategies
* a and b are positive constants

- A and B: two possible strategies that each node in network $G(V, E)$ could adopt
- Each node u will play its own independent game
- Final payoff is the sum of payoffs for all the games
- To calculate the required threshold at which a node u would decide to go with strategy A

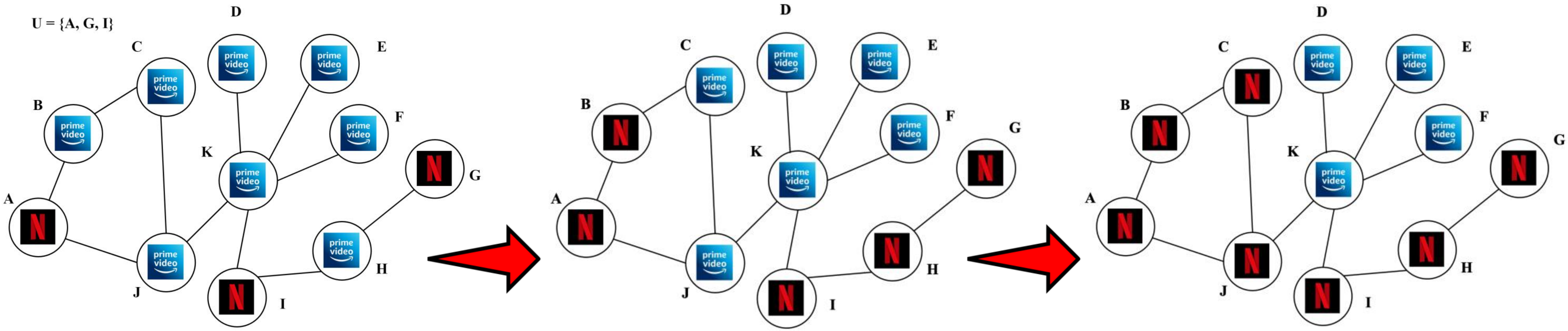
Decision-based Cascade Model: Two-player Coordination Game



- Node u has d neighbours
  - p fraction of neighbours adopt strategy A
  - Rest adopts strategy B
- Total payoff for node u if it goes with strategy A $= a \cdot d \cdot p$
- Total payoff for node u if it goes with strategy B $= b \cdot d \cdot (1-p)$
- Node u would adopt contagion A if

$$p \geq \frac{b}{a+b}$$

Decision-based Cascade Model: Illustration



The threshold for a switch from Amazon Prime Video to Netflix at a node is 0.50

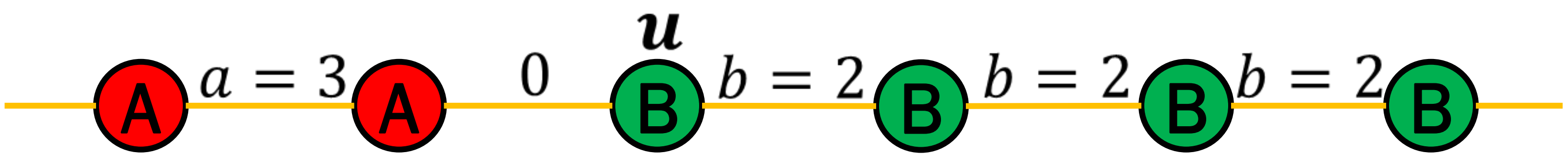
Multiple Choice Decision-based Cascade Model

- Allows a node to adopt more than one strategy/behavior
- In case a node prefers to go with both the strategies A and B, it would incur an additional cost c
- The revised payoff distribution:

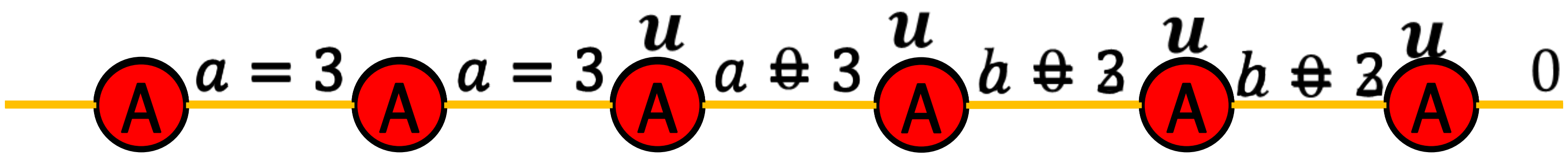
$\boldsymbol{u}$'s decision	$\boldsymbol{v}$'s decision	Payoff
AB	A	$a*$
AB	B	$b*$
AB	AB	$max(a, b)$

Payoff for a multiple choice decision model
* a and b are positive constants

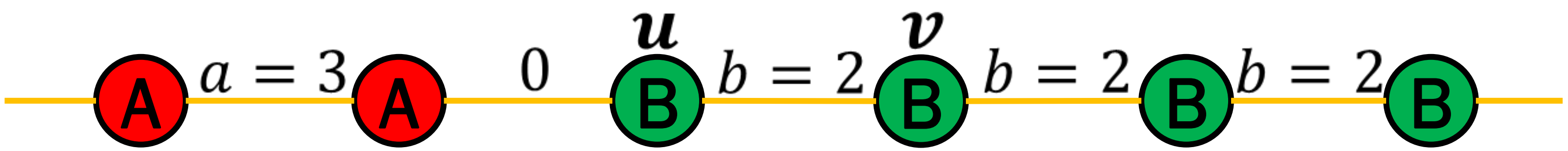
Cascades for Infinite Chain Networks: Single Choice



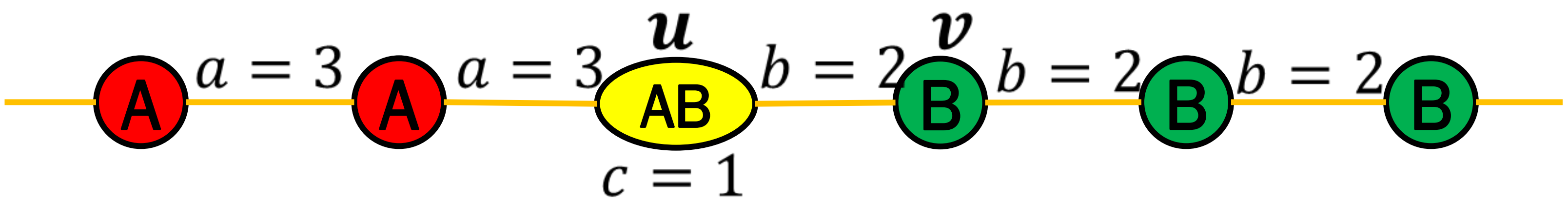
- Consider the case: $a = 3, b = 2$
- Two possible choice for node u
  - Stick with strategy B, total payoff: 0 + 2 = 2
  - Switch to strategy A, total payoff: 3 + 0 = 3
- So, node u would adopt strategy A
- And the cascade continues…



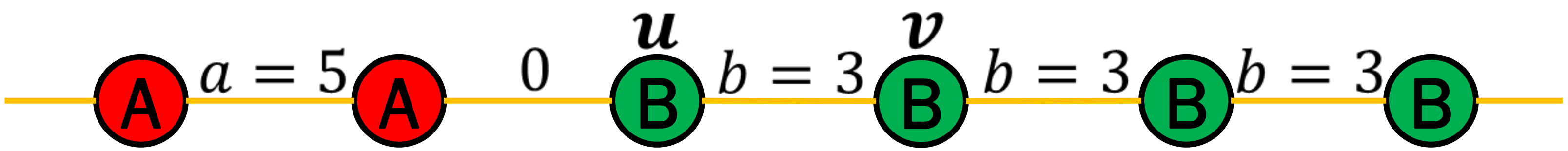
Cascades for Infinite Chain Networks: Multiple Choice: Case I



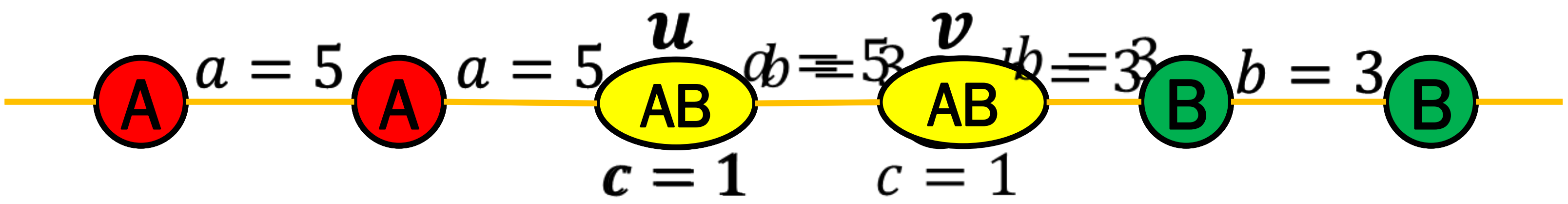
- Consider the case: $a = 3, b = 2, c = 1$
- Two possible choice for node u
  - Stick with strategy B, total payoff: 0 + 2 = 2
  - Switch to strategy A, total payoff: 3 + 0 = 3
  - Switch to strategy AB, total payoff: 3 + 2 – 1 = 4
- So, node u would adopt strategy AB
- And system is stable now!!



Cascades for Infinite Chain Networks: Multiple Choice: Case II

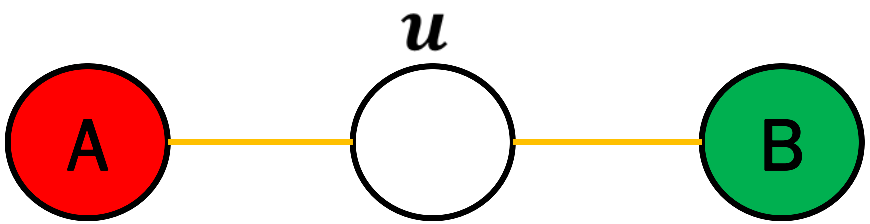


- Consider the case: $a = 5, b = 3, c = 1$
- Two possible choice for node u
  - Stick with strategy B, total payoff: 0 + 3 = 3
  - Switch to strategy A, total payoff: 5 + 0 = 5
  - Switch to strategy AB, total payoff: **5 + 3 – 1 = 7**
- So, node u would adopt strategy AB
- Two possible choice for node v
  - Stick with strategy B, total payoff: 3 + 3 = 6
  - Switch to strategy A, total payoff: 5 + 0 = 5
  - Switch to strategy AB, total payoff: **5 + 3 – 1 = 7**
- So, node v would adopt strategy AB
- And the cascade continues!!

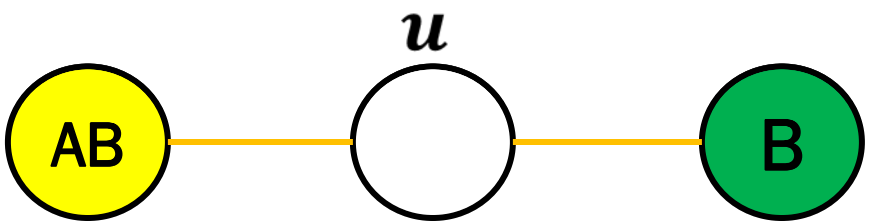


Cascade in Infinite Chain Networks: Generic Model

- Let us consider an infinite chain network with strategy set $\{A, B, AB\}$
- We consider the scenario: $a = a, b = 1, c = c$
- Two possible cases may arise:

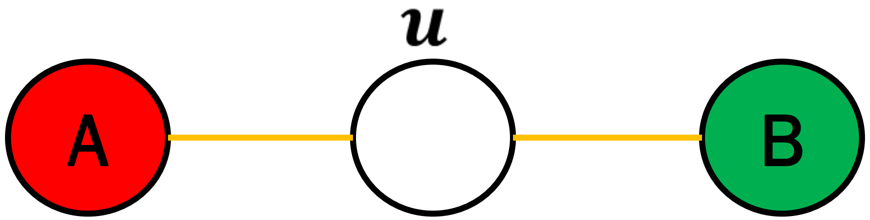


Case A



Case B

Generic Model: Case A



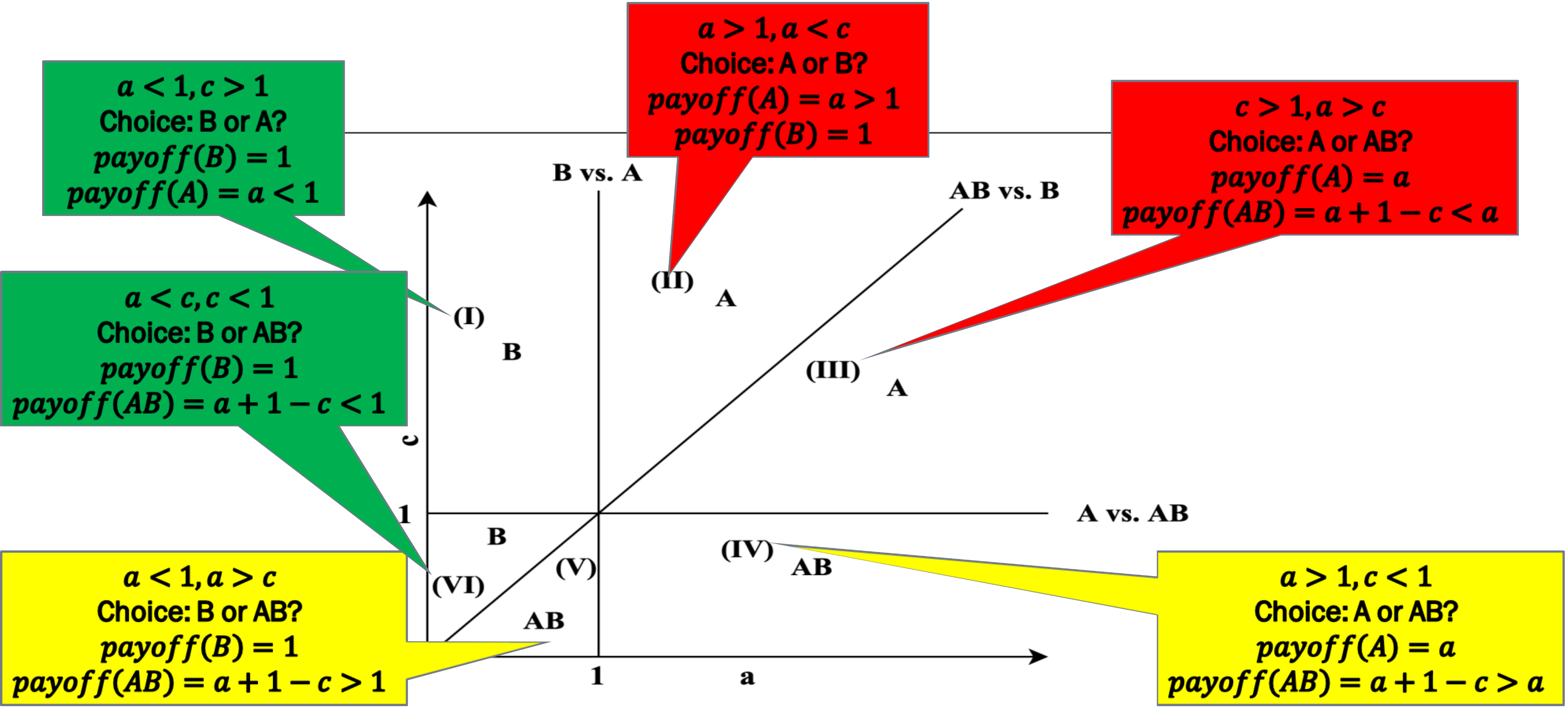
❑Three possible options for node u

1. Adopt Behavior A; Payoff $= a + 0 = a$
2. Adopt Behavior B; Payoff $= 0 + 1 = 1$
3. Adopt Behavior AB; Payoff $= a + 1 - c$

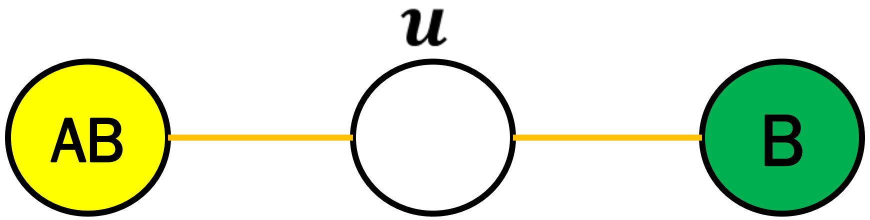
❑Breakpoint Equations:

a. B versus A: $a = 1, a < 1$: Prefer strategy B; $a > 1$: Prefer strategy A
b. AB versus B: $a = c, a < c$: Prefer strategy B; $a > c$: Prefer strategy AB
c. A versus AB: $\mathrm{c} = 1, c < 1$: Prefer strategy AB; $\mathrm{c} > 1$: Prefer strategy A

Generic Model: Case A



Generic Model: Case B



- Three possible options for node u
  1. Adopt Behavior A; Payoff $= a + 0 = a$
  2. Adopt Behavior B; Payoff $= 1 + 1 = 2$
  3. Adopt Behavior AB; Payoff $= a + 1 - c$, if $max(a, 1) = a$

- Breakpoint Equations:
  a. B versus A: $a = 2, a < 2$: Prefer strategy B; $a > 2$: Prefer strategy A
  b. AB versus B: $a - c = 1, a - c < 1$: Prefer strategy B; $a - c > 1$: Prefer strategy AB
  c. A versus AB: $\mathrm{c} = 1, c < 1$: Prefer strategy AB; $\mathrm{c} > 1$: Prefer strategy A

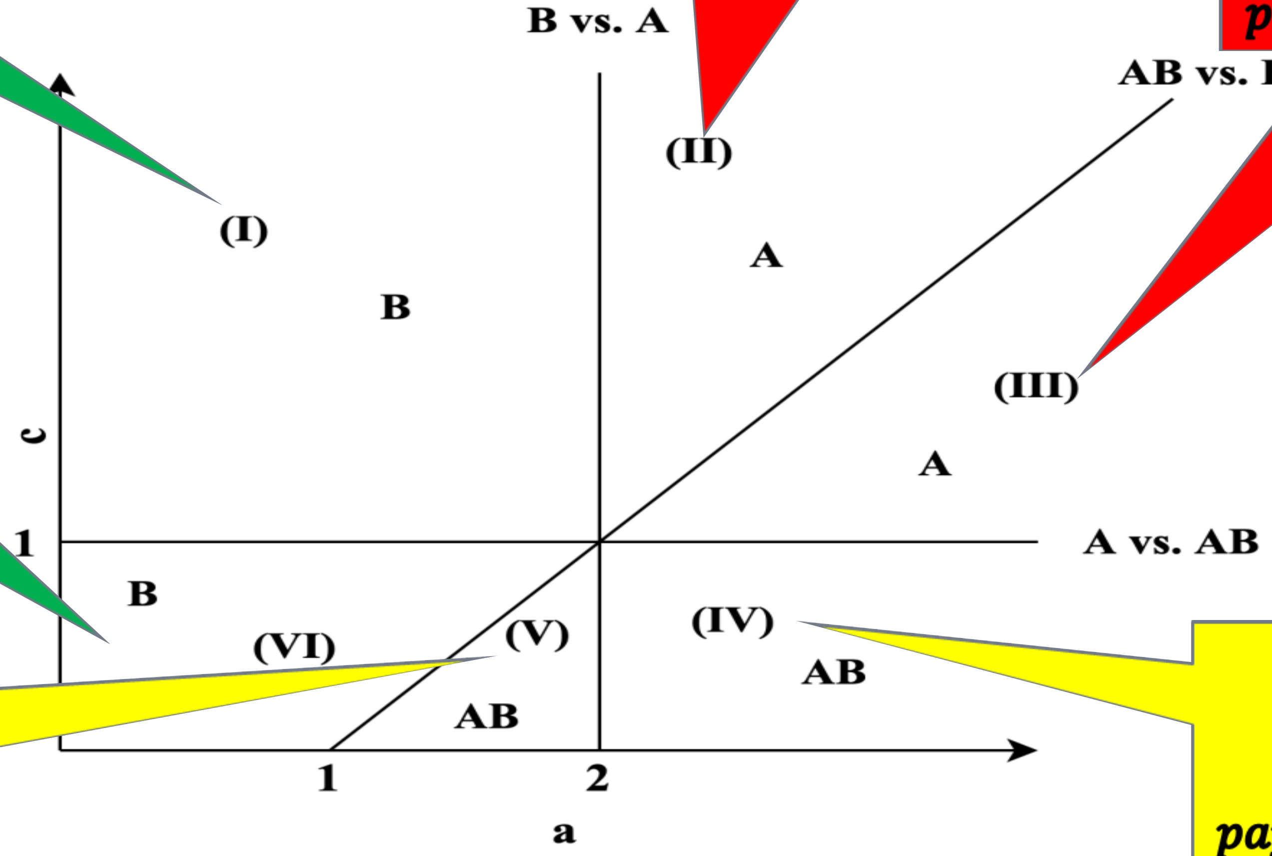
Generic Model: Case B

$a < 2, c > 1$
Choice: B or A?
$payoff(B) = 2$
$payoff(A) = a < 2$

$a > 2, a - c < 1$
Choice: A or B?
$payoff(A) = a > 2$
$payoff(B) = 2$

$c > 1, a - c > 1$
Choice: A or AB?
$payoff(A) = a$
$payoff(AB) = a + 1 - c < a$

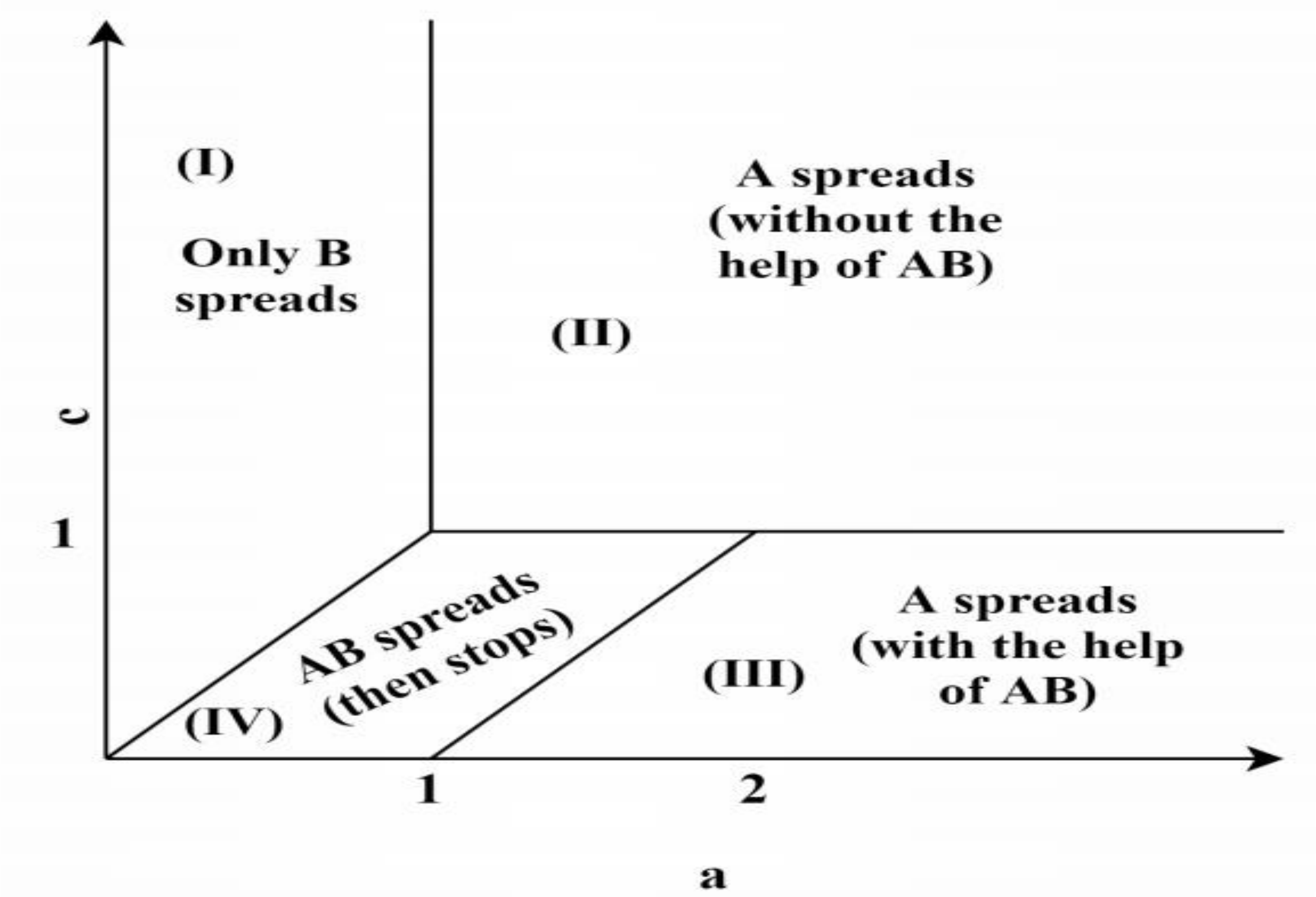
$a - c < 1, c < 1$
Choice: B or AB?
$payoff(B) = 2$
$payoff(AB) = a + 1 - c < 2$



$a < 2, a - c > 1$
Choice: B or AB?
$payoff(B) = 2$
$payoff(AB) = a + 1 - c > 2$

$a > 2, c < 1$
Choice: A or AB?
$payoff(A) = a$
$payoff(AB) = a + 1 - c > a$

Generic Model: Combined



Case Study: The "Indignados" Movement/15-M Movement

http://bit.ly/3sS4VVE

- A series of protests, demonstrations, and occupations against austerity policies in Spain
- Began around the local and regional elections of 2011 and 2012 (starting 15 May 2011)
- According to RTVE, the Spanish public broadcasting company, between 6.5 and 8 million Spaniards participated in these events
- Protesters/participants coordinated their movements using Twitter
- Dynamics of the participation of the citizens can be viewed as a cascade problem

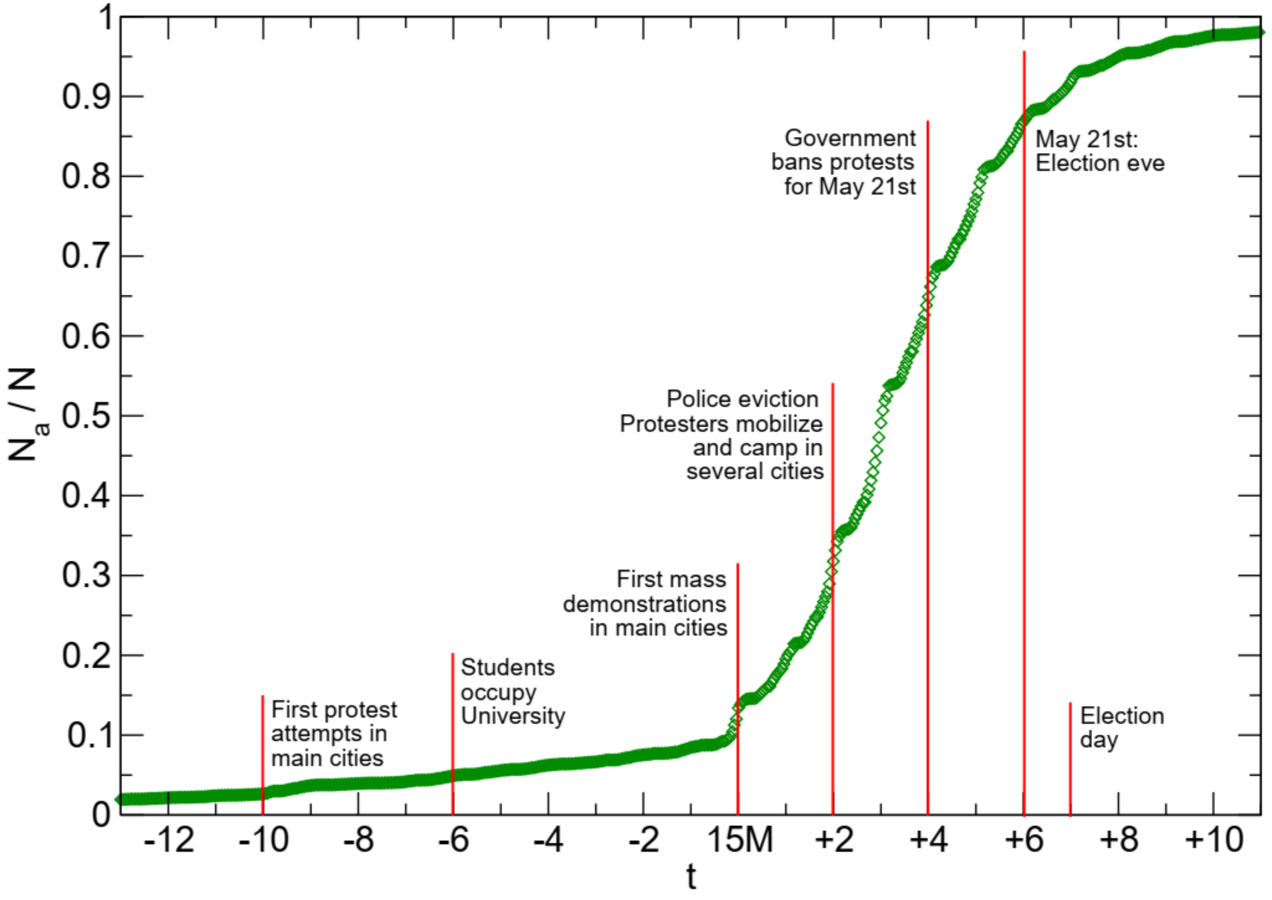
The "Indignados" Movement: Problem Formulation

- Answers to the following three questions are critical for the formulation
  - When does one generally start to tweet about the protests?
  - Does neighbor have any influence to the answer of the above question?
  - If yes, how much is the influence?
- González-Bailón et al. (2011) identified 70 hashtags used by the protesters
- Tweets containing these hashtags collected for 1 month
  - 581,750 tweets collected
  - 87,569 users were identified as relevant
- Condition for relevancy: any user who tweeted with the specified set of hashtags and has followers and followees who did the same

The "Indignados" Movement: Cascade Networks

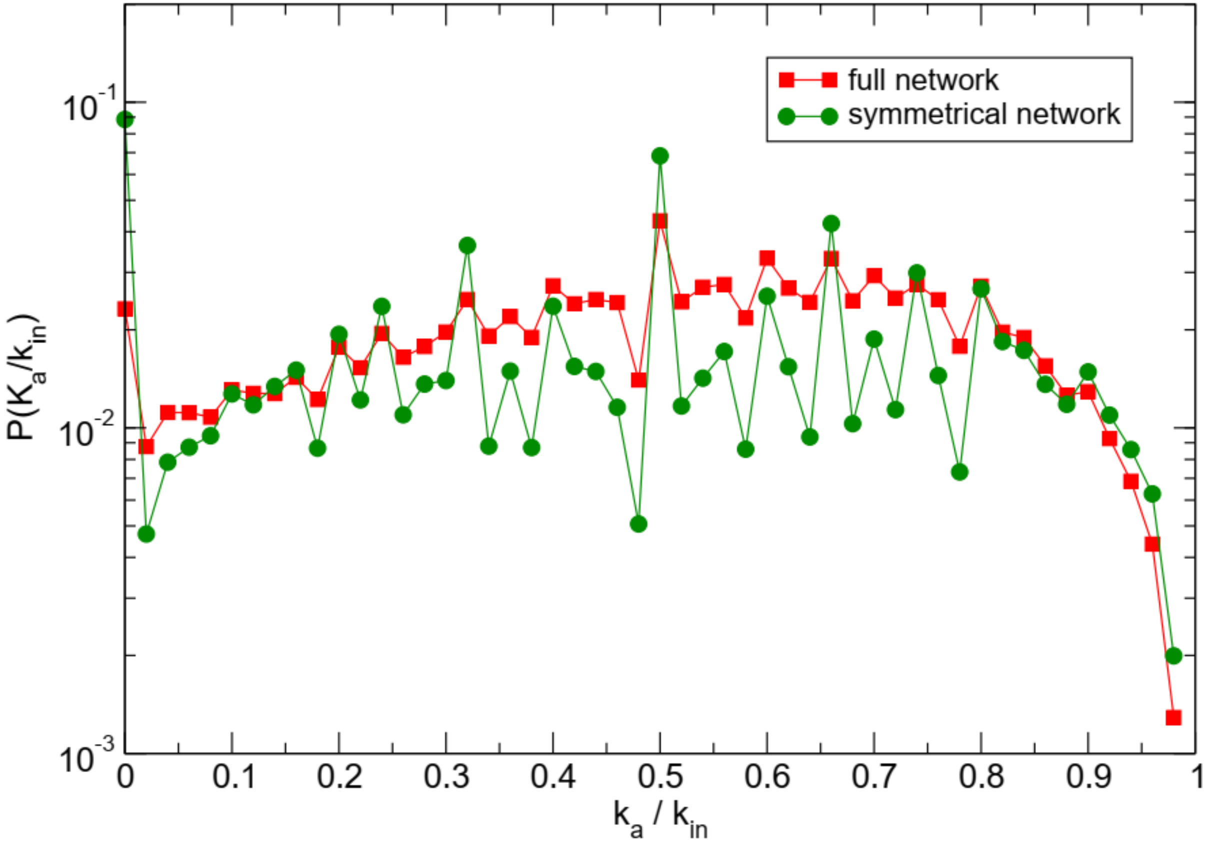
- Two kinds of networks in consideration:
  - Full Network: If a user follows another user, an edge will be formed between them connecting the two users in the network
  - Symmetric Network: If both the users follow each other, a single edge will be formed between the users in the network
- User activation time is refers to the time when the user starts tweeting about the protests
- K_{in} denotes the number of neighbors when a user becomes active
- K_a denotes the number of active neighbors when a user becomes active
- Activation Threshold $\left(\frac{K_a}{K_{in}}\right)$ refers to the fraction of active neighbors when a user becomes active
- Recruitment refers to the event wherein a user tweets about the ongoing protest

The "Indignados" Movement: Cascade Networks



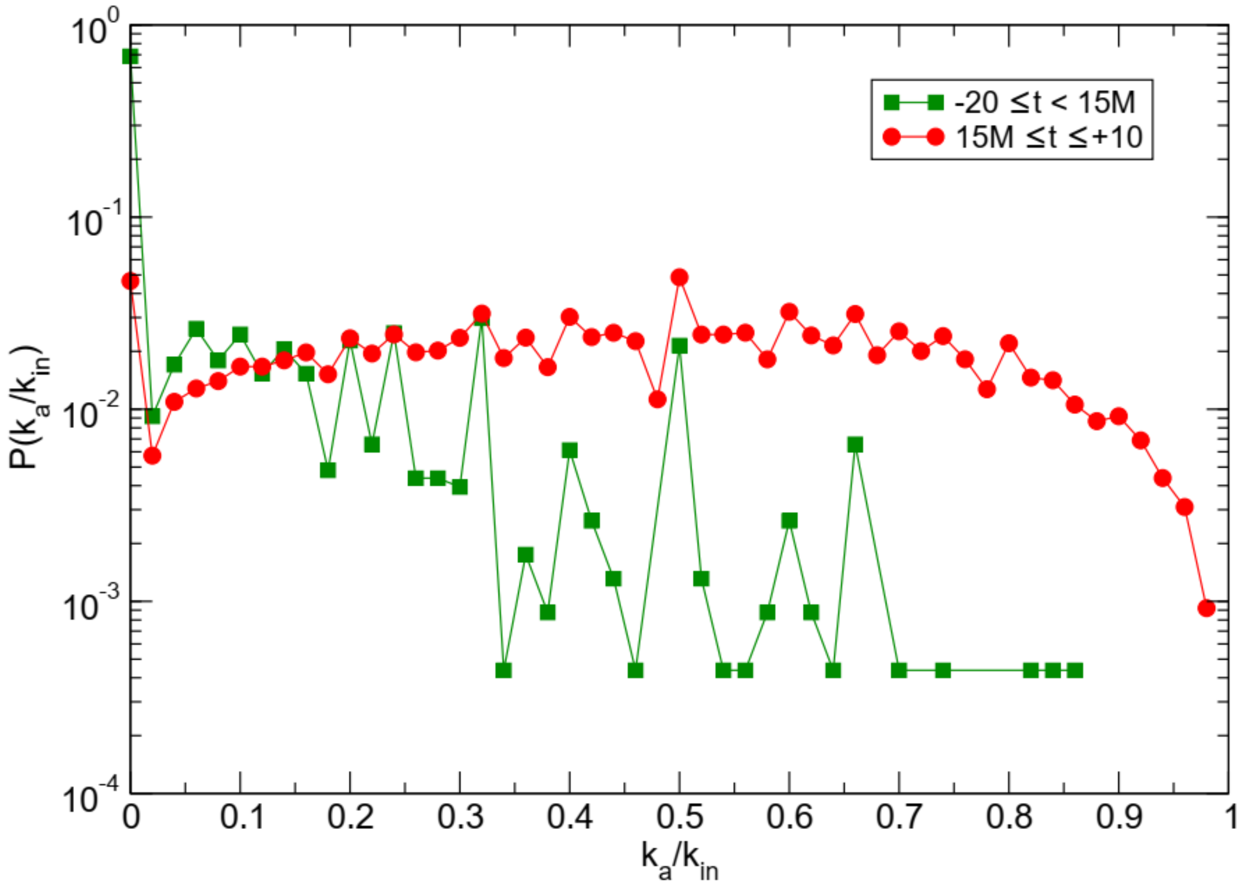
- Fraction of recruited users over time
- X-axis: the time spanning a month during the movement
- Y-axis: the number of active users normalized by the total users
- Curve reaches 0.98 by the end of the month

The “Indignados” Movement: Role of “Social Pressure” in Recruitment



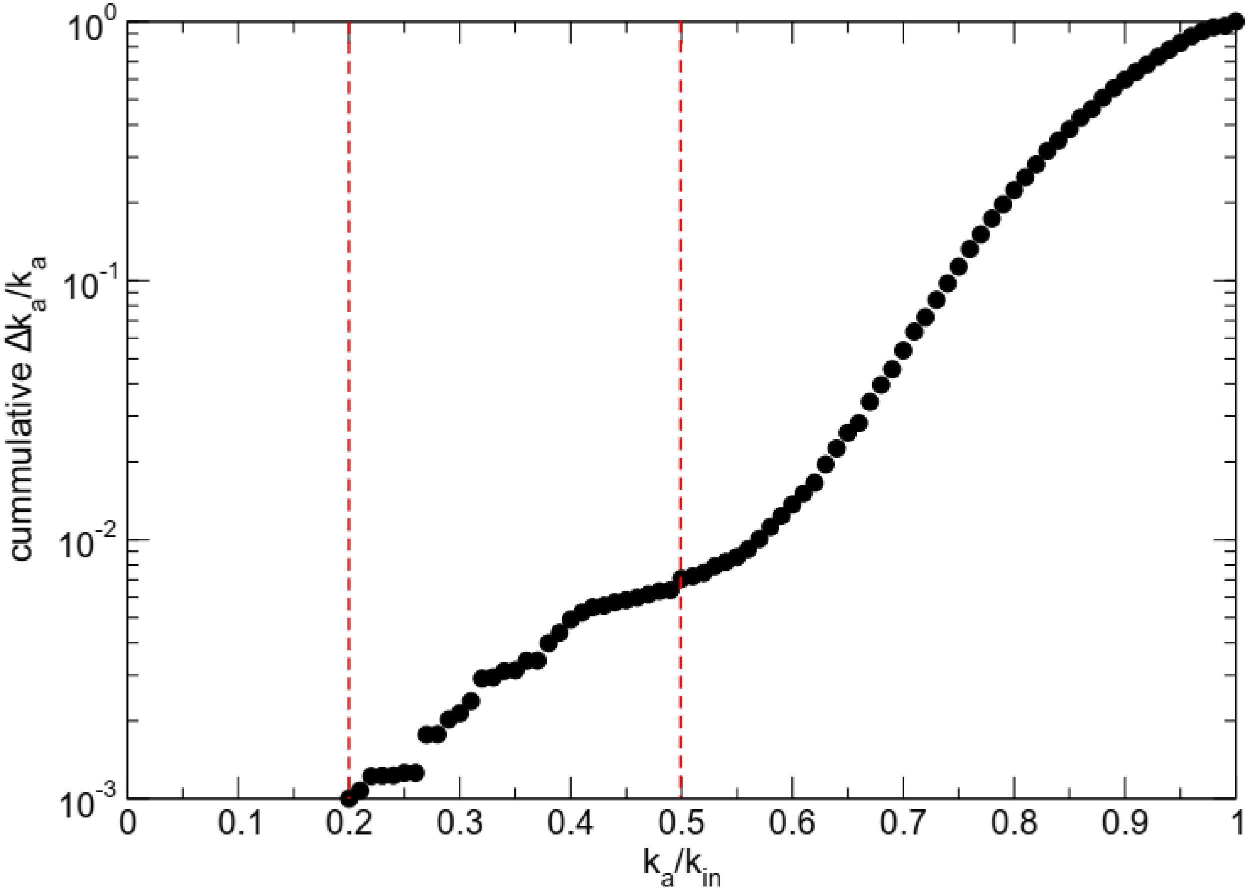
- ❑ $\frac{K_a}{K_{in}} \approx 0$: User joined the protest under little or zero social pressure
- ❑ $\frac{K_a}{K_{in}} \approx 1$: User joined the protest under acute or heavy social pressure
- ❑ A uniform distribution of activation threshold with two local peaks
  - at 0.0, representing the leaders or the self-active users of the protests
  - at 0.5, indicating that many users join the protests after half of their neighbors do
- ❑ Both the networks show nearly similar behaviour

The “Indignados” Movement: Role of “Social Pressure” in Recruitment



- The month-long timeline is divided into two time frames
  - 20-days timeframe before 15th May 2011
  - 10-days timeframe after 15th May 2011
- 15th May was the first day of the mass demonstrations
- Activation threshold distribution did not change much amongst the early activated, low-threshold users after 15th May
- Media coverage didn’t significantly influence the recruitment of early activated low-threshold users

The "Indignados" Movement: Role of "Social Pressure" in Recruitment



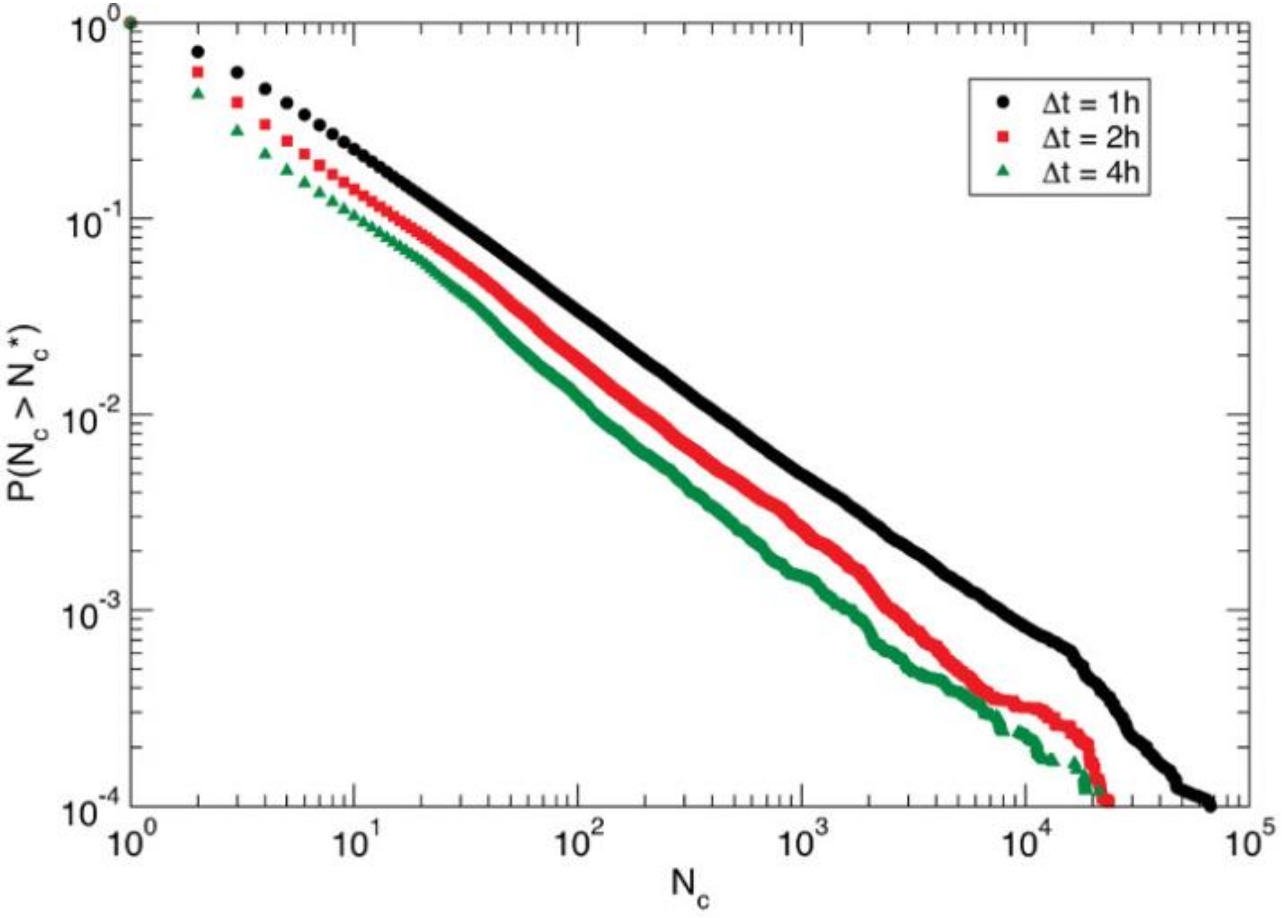
- If several neighbors of a user suddenly become active, then how likely is the user to become active?
- Burstiness is defined as the relative increase in the active neighbors

$$\frac{\Delta K_a}{K_a} = \frac{K_a^{t+1} - K_a^t}{K_a^t}$$

where t denotes a day

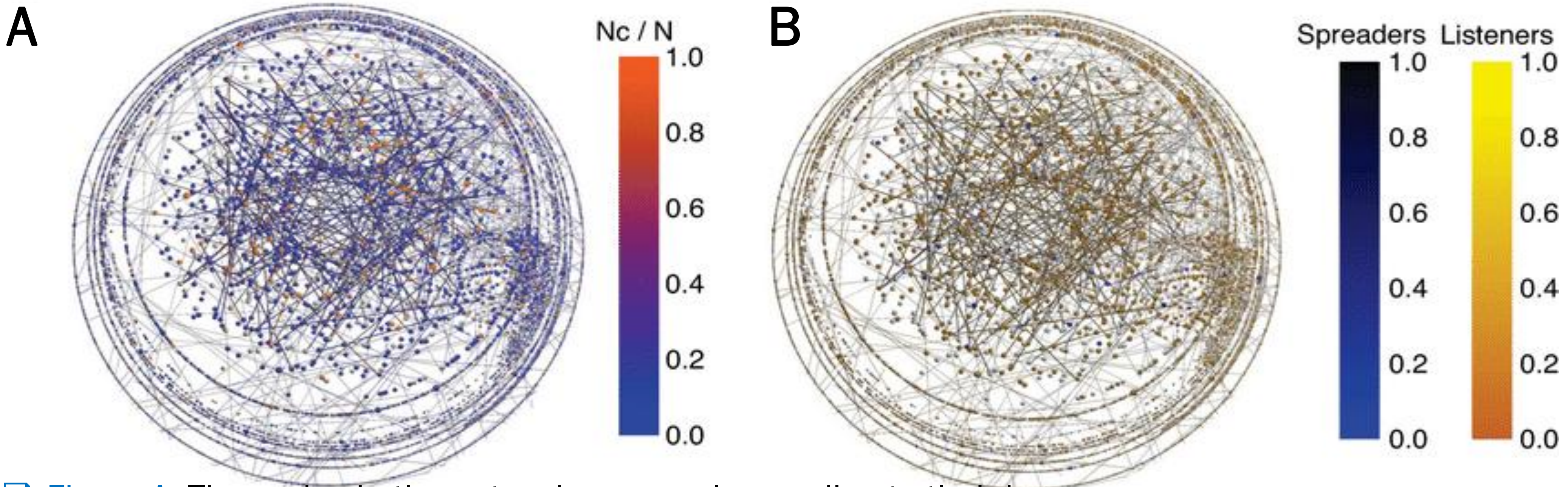
- Slope is steeper for high activation threshold users than the low activation threshold users
- High activation threshold users are more likely to join the protests if they see a sudden increase in participation amongst their neighbors

The “Indignados” Movement: Information Cascade



- ❑ If the messages of a user and her followers lie within the time difference Δt, then the user and the followers are said to form a cascade
- ❑ X-Axis denotes the size of cascades
- ❑ Y-Axis denotes the fraction of cascades having the corresponding size for different Δt
- ❑ Higher the size of the cascade, lower the number of such cascades

The "Indignados" Movement: Information Cascade



- Figure A: The nodes in the network arranged according to their k-core
- Cascades for higher K-core subgraphs (i.e., the central users) are responsible for starting successful cascades
- Figure B: Nodes in blue are users who participated in the diffusion of protest messages
- Figure B: Nodes in orange were exposed to the messages but did not send messages of their own

Decision-based Cascade Model: Limitations

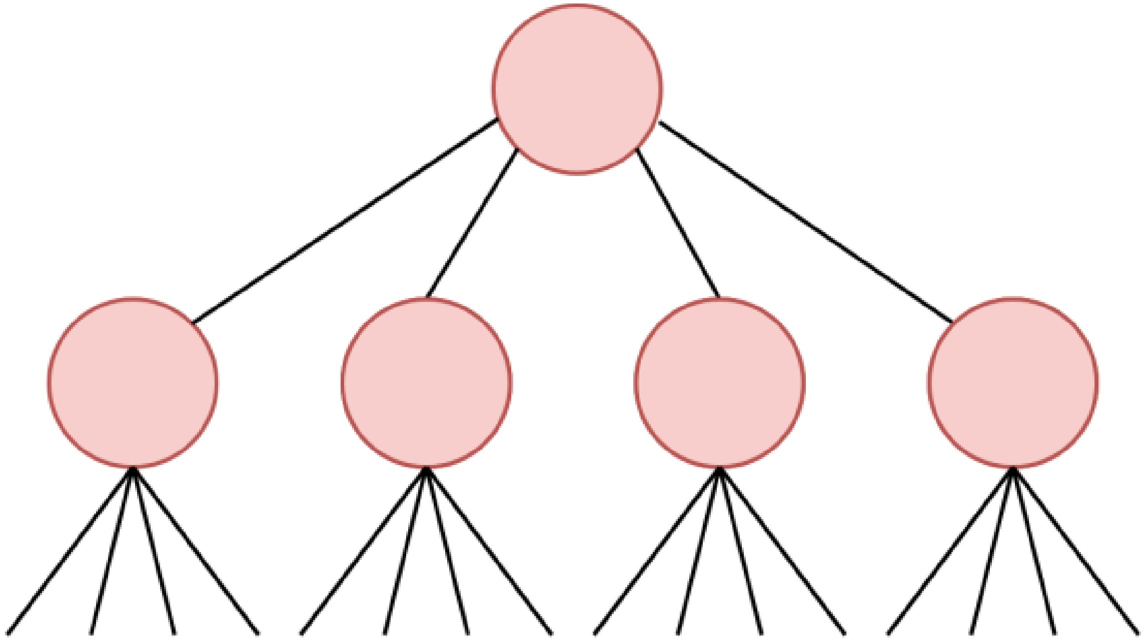
- Cascade will continue to grow only when its growth is associated with the highest reward amongst each of the nodes

- In many real-world scenarios, such hard decision making criteria or payoff functions are not available

- Infection spreading mechanism of a virus
  - We can model the spread of the virus as a cascade
  - Cascade growth is not in the hands of the node.

- Alternative Approach: Probabilistic Cascade Model

Probabilistic Cascade Model: Random Tree



A random tree with $\boldsymbol{d} = \mathbf{4}$

- Basic assumptions
  - Person at the root node of the random tree is always infected
  - Each person in the random tree meets d new people. So, the random tree is a d-nary tree
  - Each person, on meeting an infected person, has the probability of getting infected as q $(q > 0)$
- For the virus to stay active and keep on spreading (cascade)
  - probability that a node at a depth h will be infected should be a positive real number
  - Same must hold for all h

$$\lim_{h\to\infty} P[a\ node\ at\ depth\ h\ is\ infected] > 0$$

- The cascade would die out if

$$\lim_{h\to\infty} P[a\ node\ at\ depth\ h\ is\ infected] = 0$$

Probabilistic Cascade Model: Random Tree

- If p_h be the probability of a node being infected at level h, then

$$p_h = 1 - (1 - q \cdot p_{h-1})^d$$

- The recurrence relation can have the following functional form:

$$f(x) = 1 - (1 - qx)^d$$

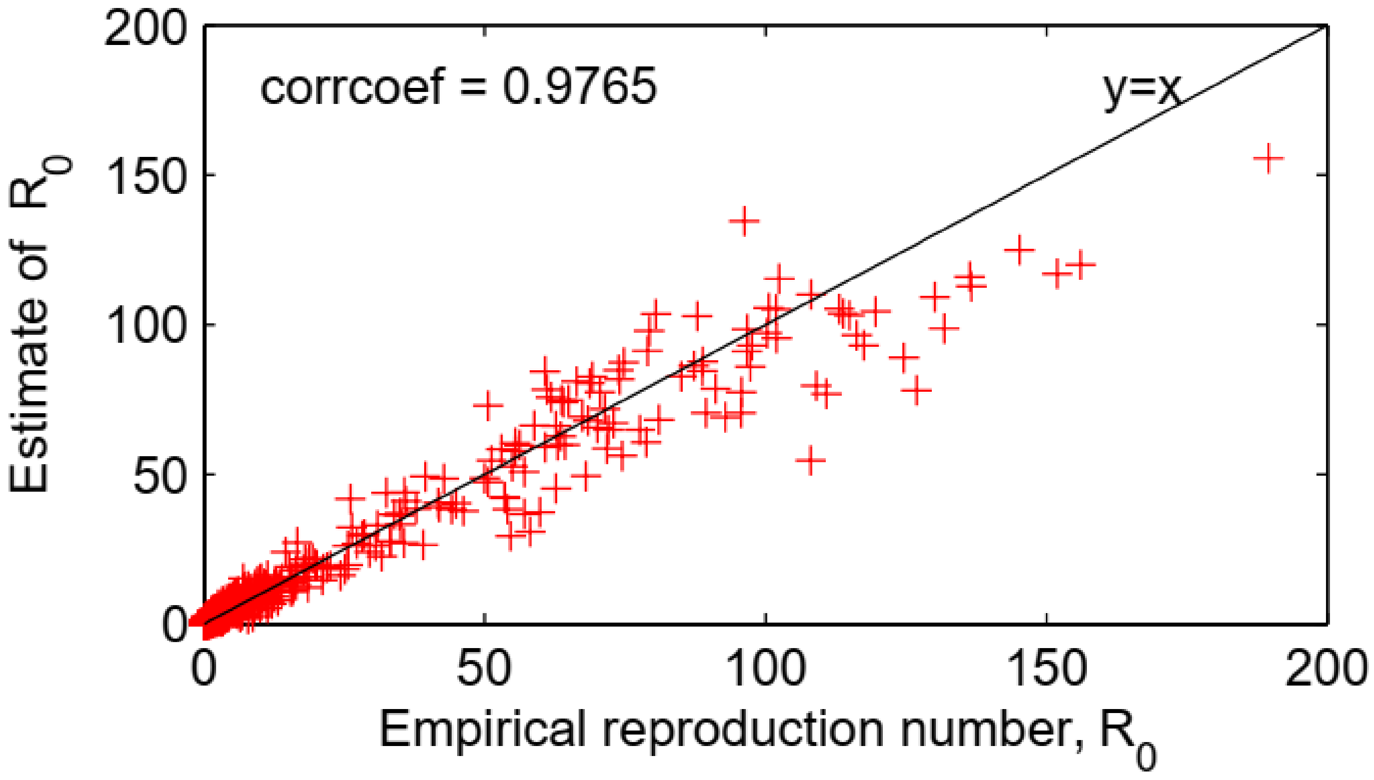
- The properties of f
  - $f(x)$ is monotonic function
  - $f'(x)$ is non-increasing
  - $f'(x)$ is monotonic, non-increasing in $[0,1]$
- $f(0) = 0$ and $f'(0) = q \cdot d$

Probabilistic Cascade Model: Random Tree

- Since $f'(x)$ is monotonic non-increasing, $f'(x) \leq q \cdot d$
- For epidemic to die out, $f(x) < x \Rightarrow q \cdot d < 1$
- The quantity $q \cdot d$ is called Reproductive number in the literature, denoted R_0
- If $R_0 \geq 1$, the epidemic grows in an exponential manner
- If $R_0 < 1$, the epidemic spread reduces constantly and eventually dies out

- two methods to contain the spread of the epidemic
  - reduce the value of d ⇛ keep the already-infected nodes in isolation
  - reduce the value of q ⇛ reduce transmission rate by promoting better hygiene practices

Probabilistic Cascade Model: The Reproductive Number



Cha et al. (2008)

- ❑ Experimental validation of R_0
- ❑ Dataset: Crawled from Flickr social network
  - ❑ ~2 million users interacting
  - ❑ ~11 million images
  - ❑ ~34 million 'like' interactions
- ❑ Cascade: if a user in her friend list 'liked' a picture before you, then the user's like is a part of the social cascade of the photo
- ❑ formula that is employed

$$R_0 = q \times d \times \frac{avg(d_i^2)}{avg(d_i)^2}$$

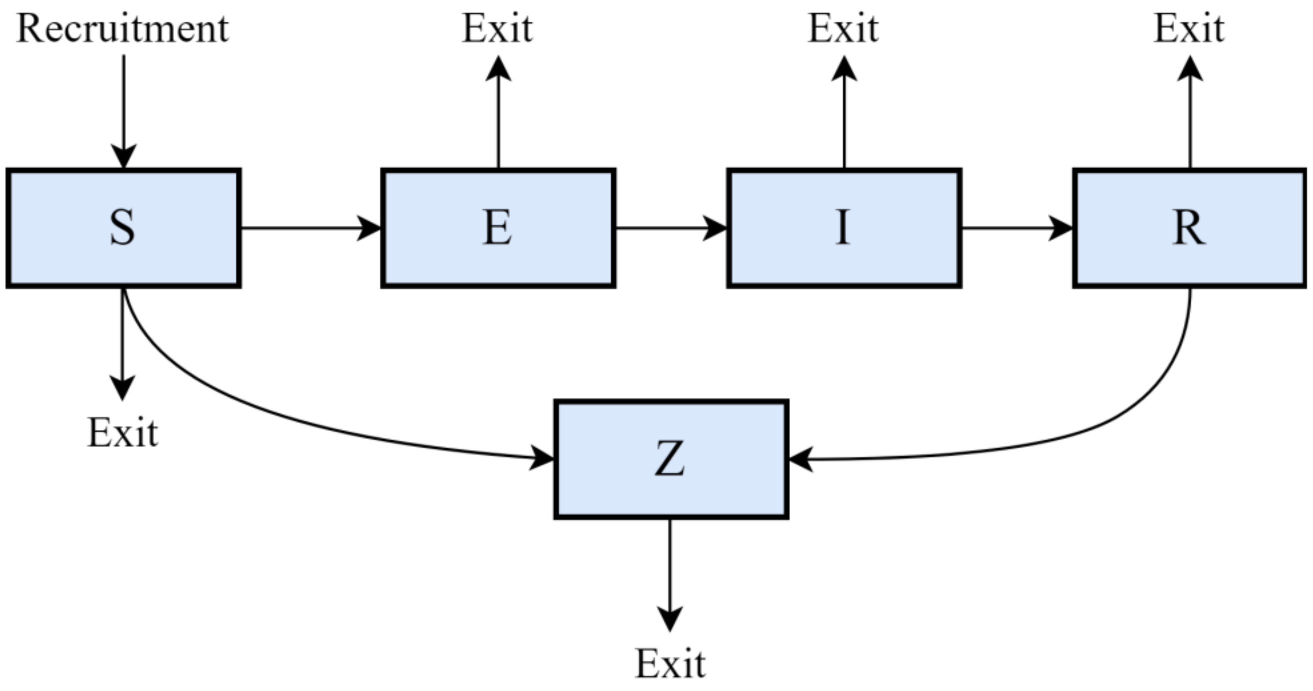
Compartmental Models of Epidemiology(Spreading models of viruses)

- ❑Origination of such models in the early 20th century
- ❑Seminal work by Kermack and McKendrick in 1927
- ❑Models are most often run with ordinary differential equations
- ❑Stochastic (random) framework, which are more realistic, are also possible
- ❑Two important parameters:
- ❑Birth Rate (β): probability with which a neighbor node attacks another node with the virus
- ❑Death Rate (δ): probability with which an infected node heals

Compartmental Models of Epidemiology: SEIR Model

- A generalized framework to model the spread of epidemics

- SEIR (or S+E+I+R) is an acronym of
  - Susceptible (S): those who may become infected
  - Exposed (E): those who are infected, but not yet capable of spreading the infection/idea
  - Infected (I): those who are capable of further propagating the infection/idea
  - Recovered (R): those who have recovered from or become immune to the infection/idea
  - immune(Z): Susceptible who no longer follow the infection/idea (Another possible state)



- Many possible variations of the model

Compartmental Models of Epidemiology: SIR Model (Chickenpox or plague virus model)

$$\text{Susceptible} \xrightarrow{\beta} \text{Infected} \xrightarrow{\delta} \text{Recovered}$$

- A node can go through only three stages: (i) Susceptible, (ii) Infected, and (iii) Recovered

- Rate of change of 'susceptible population' is:

$$\frac{dS}{dt} = -\beta \times S \times I$$

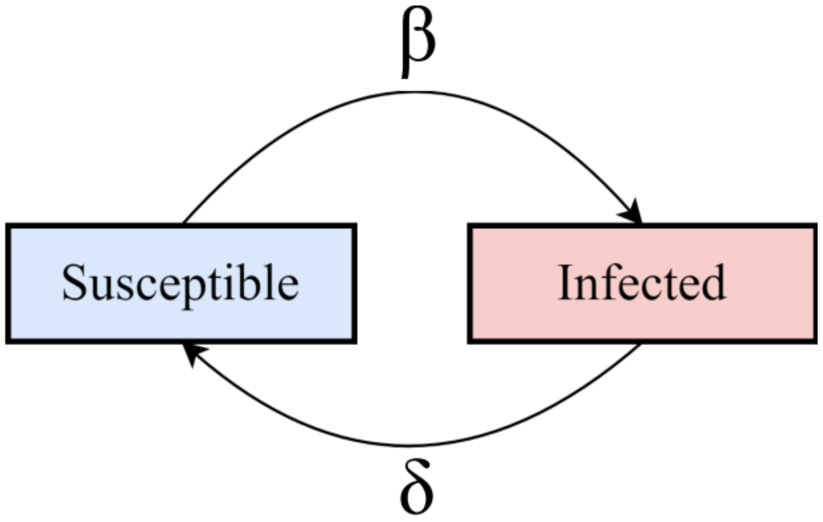
- Rate of change of 'recovered population' is:

$$\frac{dR}{dt} = \delta \times I$$

- Rate of change of 'infected population' is:

$$\frac{dI}{dt} = \beta SI - \delta I$$

Compartmental Models of Epidemiology: SIS Model(flu virus)



- a node can go through the phases of 'susceptible' to 'infected' to 'susceptible' again
- Common cold can recur with a high probability can be modelled by SIS

- Rate of change of 'susceptible population' is:

$$\frac{dS}{dt} = -\beta SI + \delta I$$

- Rate of change of 'infected population' is:

$$\frac{dI}{dt} = \beta SI - \delta I$$

- Strength of a virus $= \frac{\beta}{\delta}$

- Epidemic threshold, denoted by τ

Compartmental Models of Epidemiology: SIS Model

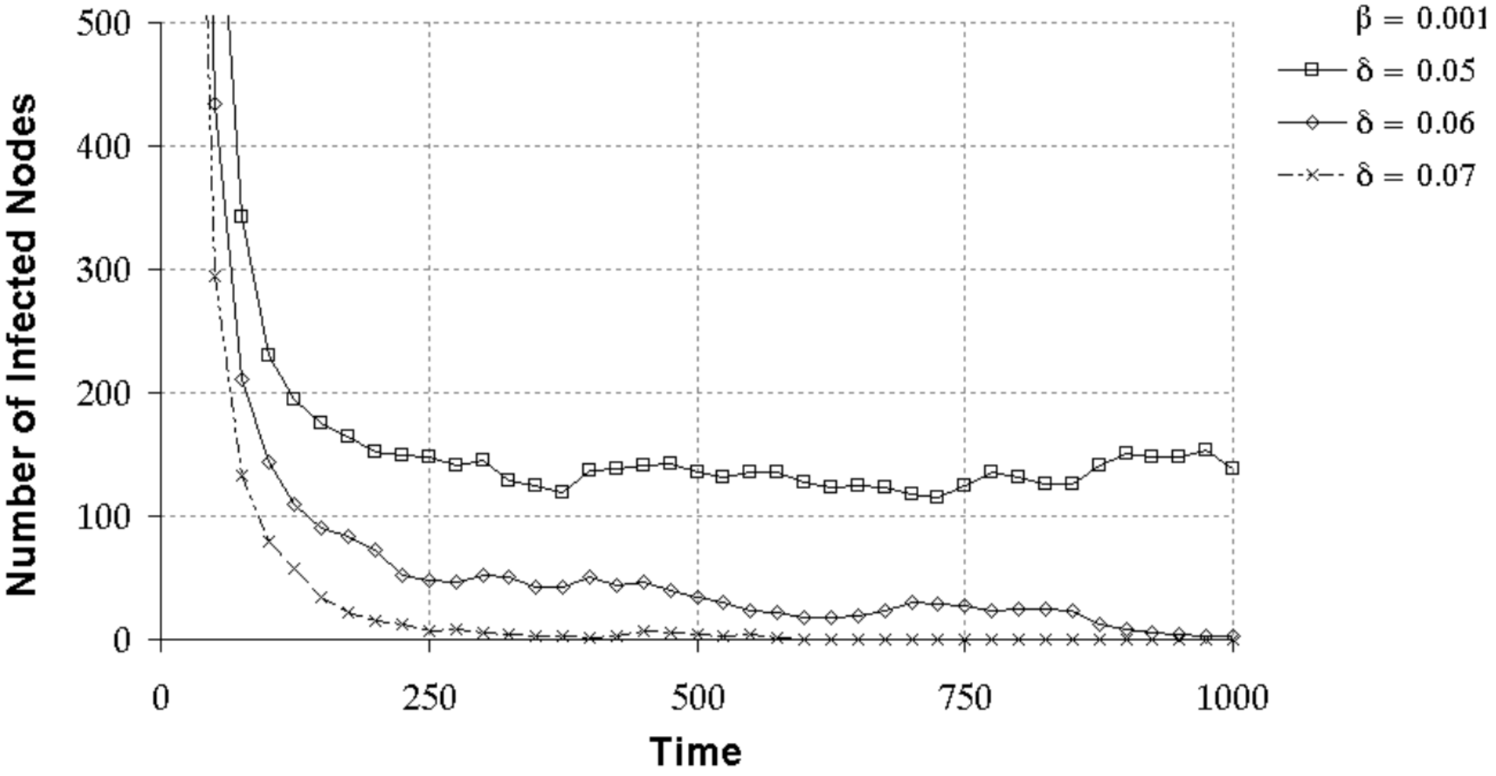
- The epidemic dies out if virus strength $< \tau$
- τ is nothing but the reciprocal of the largest eigenvalue of adjacency matrix representing the underlying network
- The epidemic dies out if

$$\frac{\beta}{\delta} < \tau = \frac{1}{\lambda_{1,A}}$$

where A: adjacency matrix of the underlying network

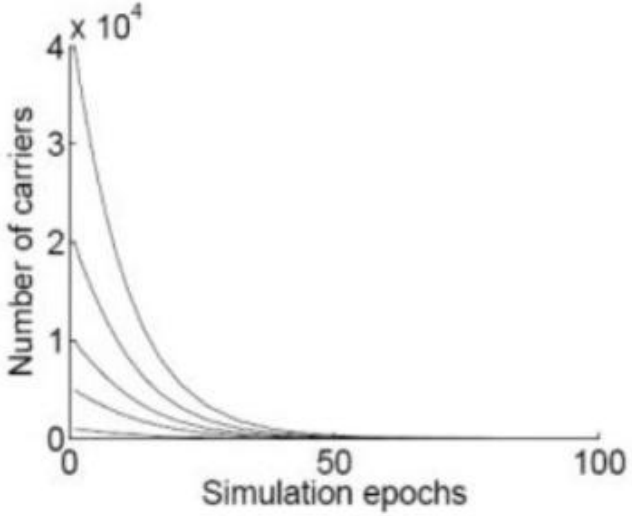
$\lambda_{1,A}$: largest eigen value of A

Compartmental Models of Epidemiology: SIS Model

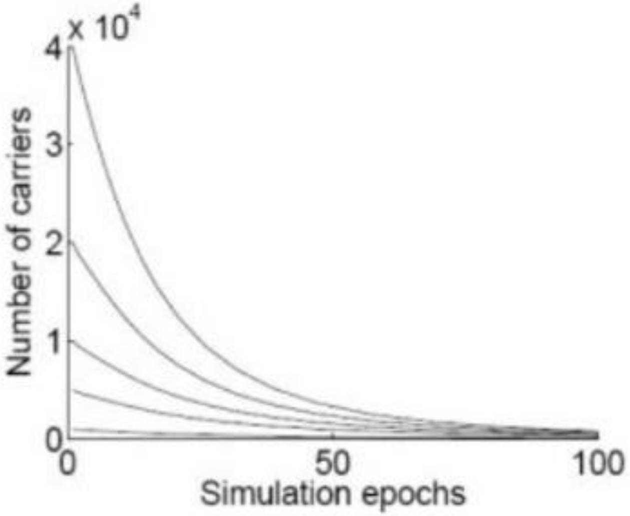


Reduction of Infected nodes for different β and δ

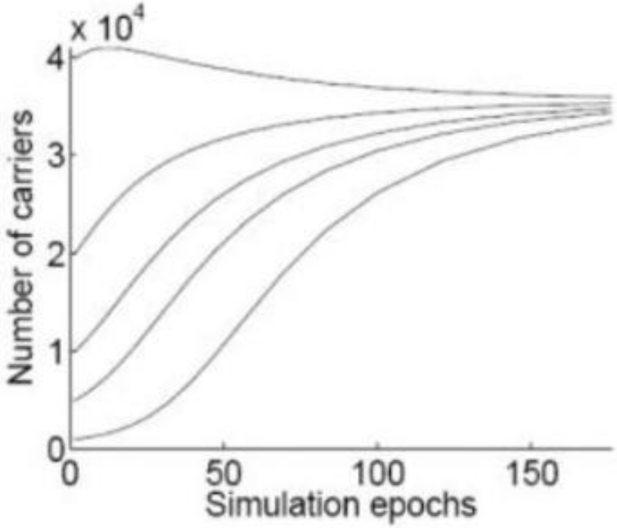
Compartmental Models of Epidemiology: SIS Model



(a) Below the threshold, s=0.912



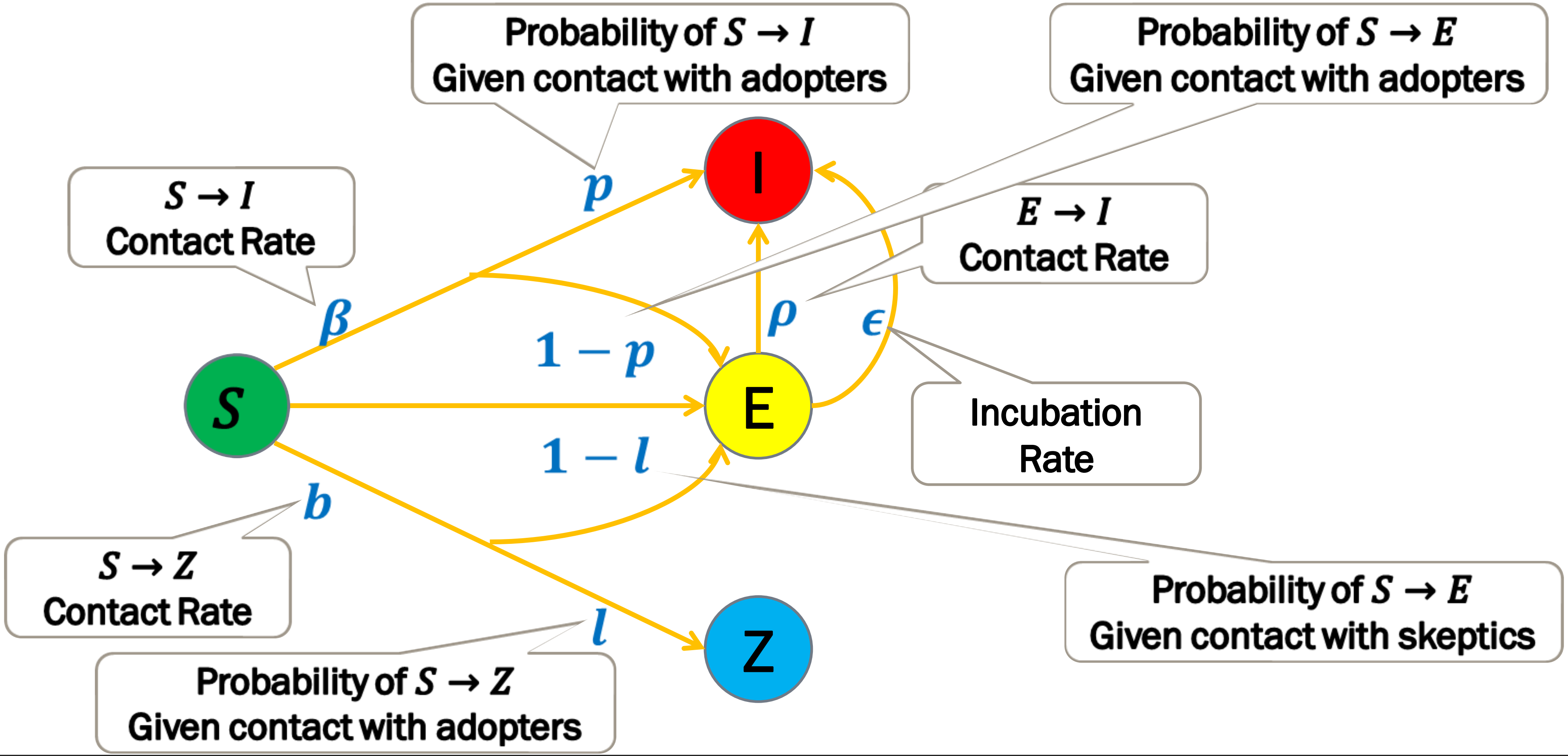
(b) At the threshold, s=1.003



(c) Above the threshold, s=1.1

Increasing or decreasing the number of initial carriers will make no difference in extending or reducing the duration of the epidemic

Analysing Rumour Spread: SEIZ Model



Analysing Rumour Spread: SEIZ Model

- A rumor in many ways is like a disease
  - Susceptible: those who are active on social media
  - Exposed: those who have seen/heard the rumor but did not believe it yet
  - Infected: those who believe the rumour
  - Skeptics: those who did not believe the rumor
- A new kind of metric defined for the event:

$$R_{SI} = \frac{(1-p)\beta + (1-l)b}{\rho + \epsilon}$$

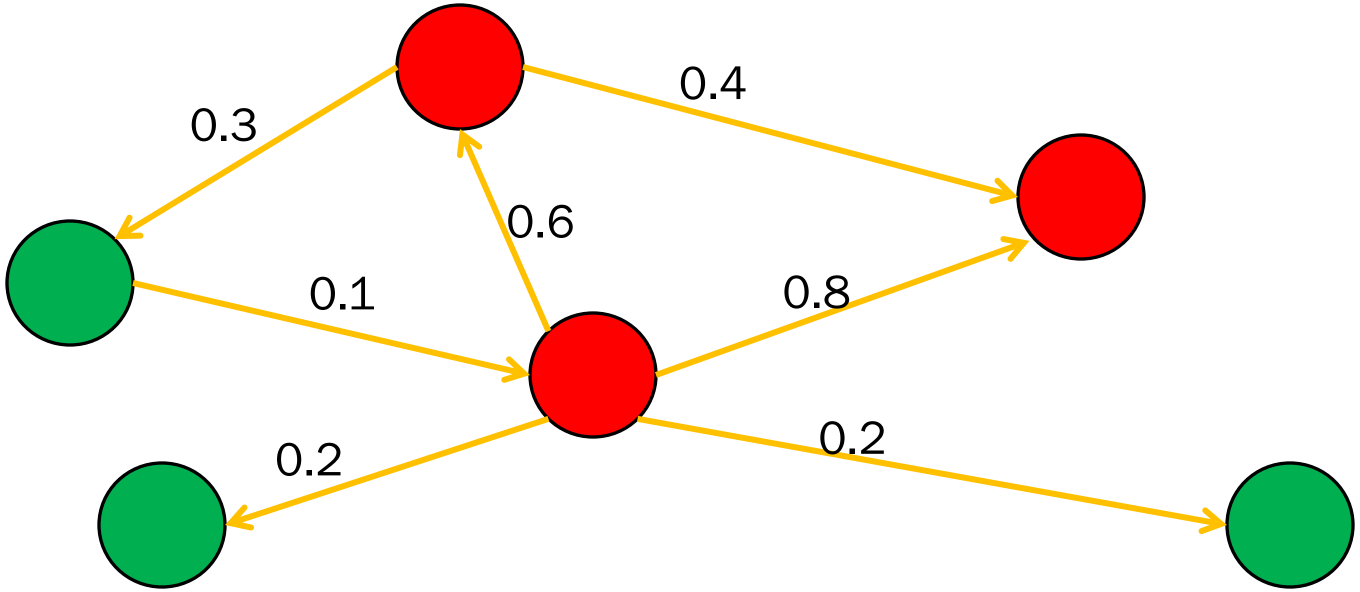
- A kind of a flux ratio between the ratio of effects entering a node which is being examined to those leaving that node
  - high for real-life events
  - low for rumors

Analysing Rumour Spread: SEIZ Model

- Mathematical representation of the model:
  - $\frac{dS}{dt} = -\beta S \frac{I}{N} - bS \frac{Z}{N}$
  - $\frac{dE}{dt} = (1-p)\beta S \frac{I}{N} + (1-l)bS \frac{Z}{N} - \rho E \frac{I}{N} - \epsilon E$
  - $\frac{dI}{dt} = p\beta S \frac{I}{N} + \rho E \frac{I}{N} + \epsilon E$
  - $\frac{dZ}{dt} = l\beta S \frac{Z}{N}$

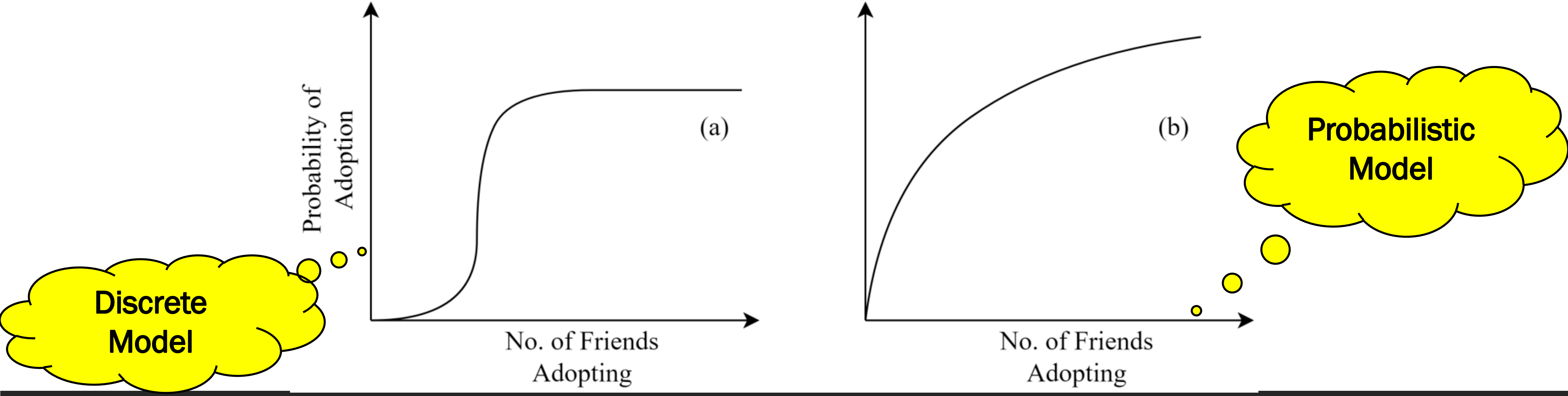
Independent Cascade Models

- ❑Spread of infection with uniform probability between any node pair may not be realistic!
- ❑Transmission of disease maybe more probable between certain pairs of nodes than other pairs
- ❑An edge between u and v having a probability p_{uv} of transmission between them

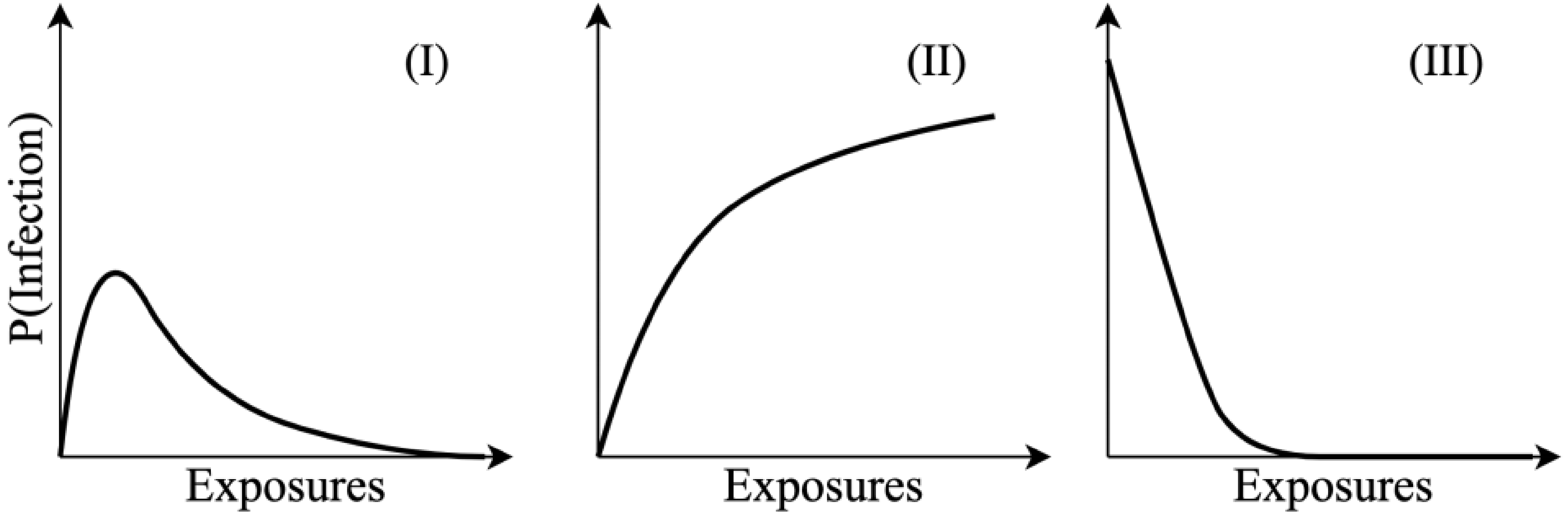


Independent Cascade Models

- ❑Exposure: event of a node being exposed to a contagious incident
- ❑Adoption: event of the node acting on the contagious incident
- ❑Hypothesis: probability of adoption is influenced by the number of neighbors who have adopted

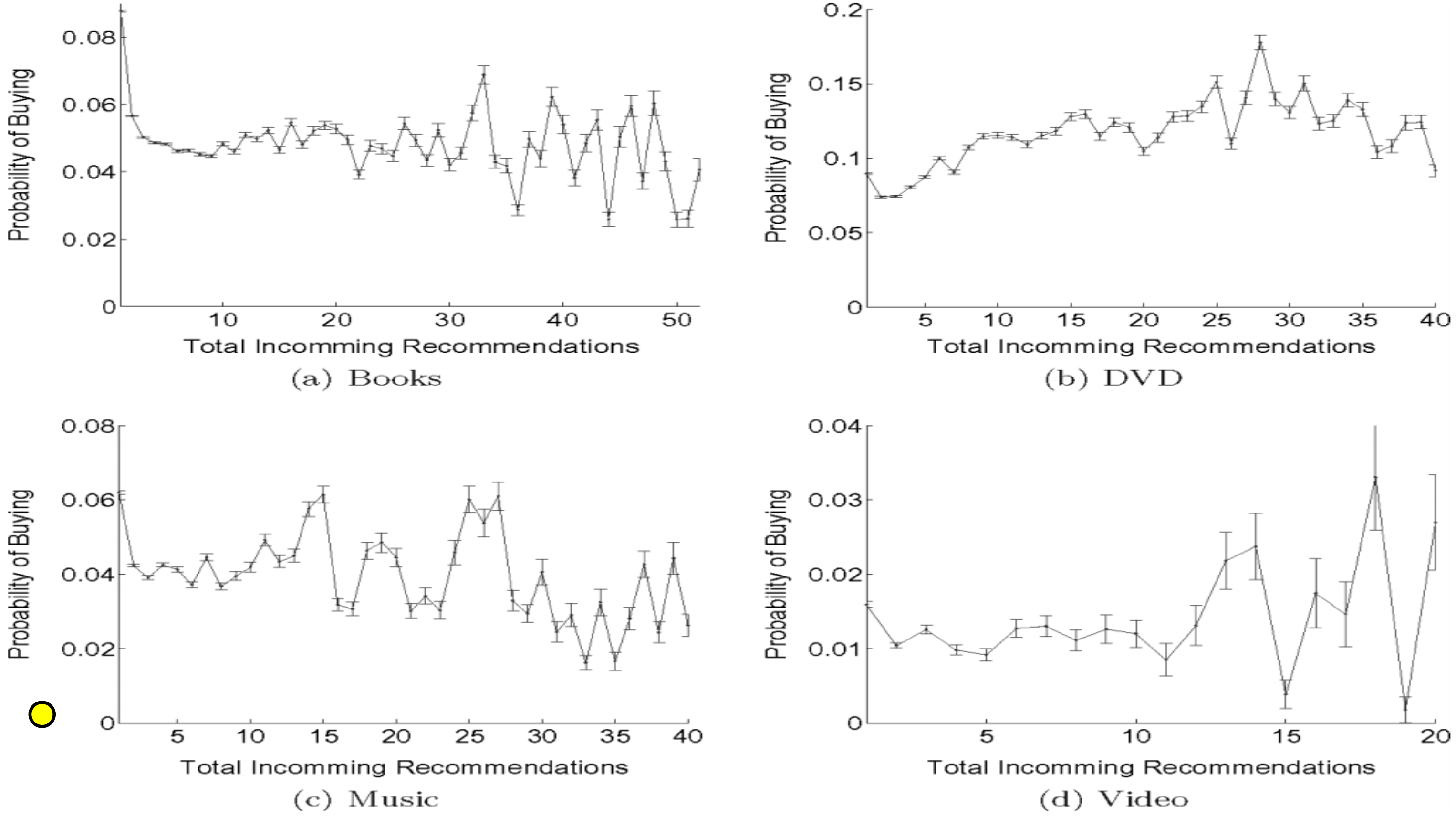


Independent Cascade Models: Exposure versus Adoption



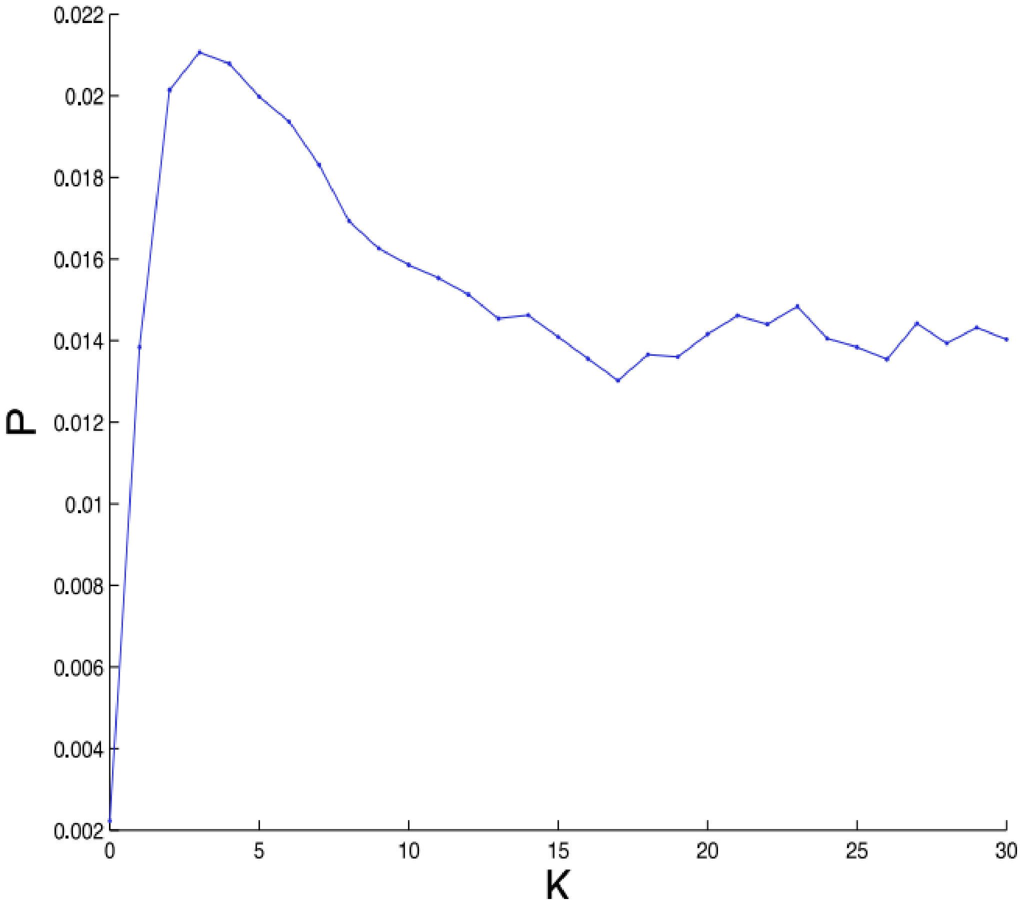
Independent Cascade Models: Exposure versus Adoption (Marketing)

Obtained from study of Referral programs building marketing habits



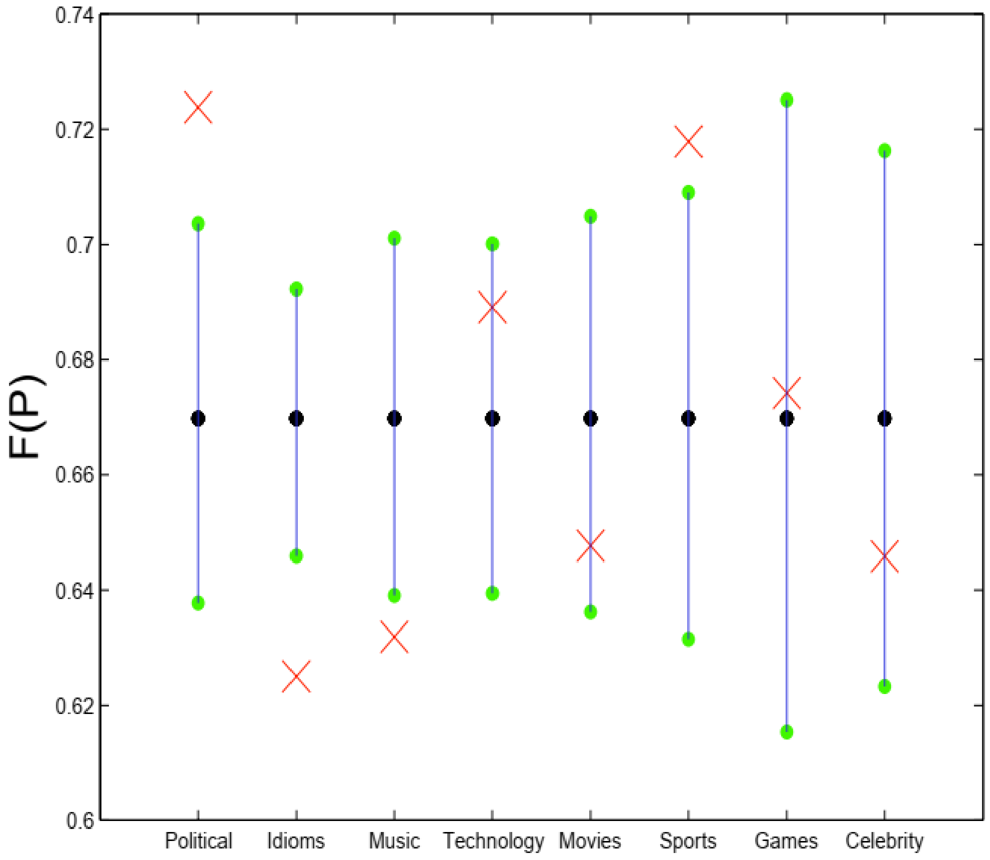
Independent Cascade Models: Exposure versus Adoption (Twitter)

- Exposure: seeing retweets of a particular tweet by different users
- Adoption: retweeting that particular tweet after seeing other retweets
- Important Curve Parameters
  - Width of the curve: The length of extension of the curve along the x-axis
  - Stickiness of cascade: It is the maximum height of the curve along y-axis
  - Persistence of cascade: This is defined as the ratio of the area under the curve and the rectangle defined by width and height of the curve



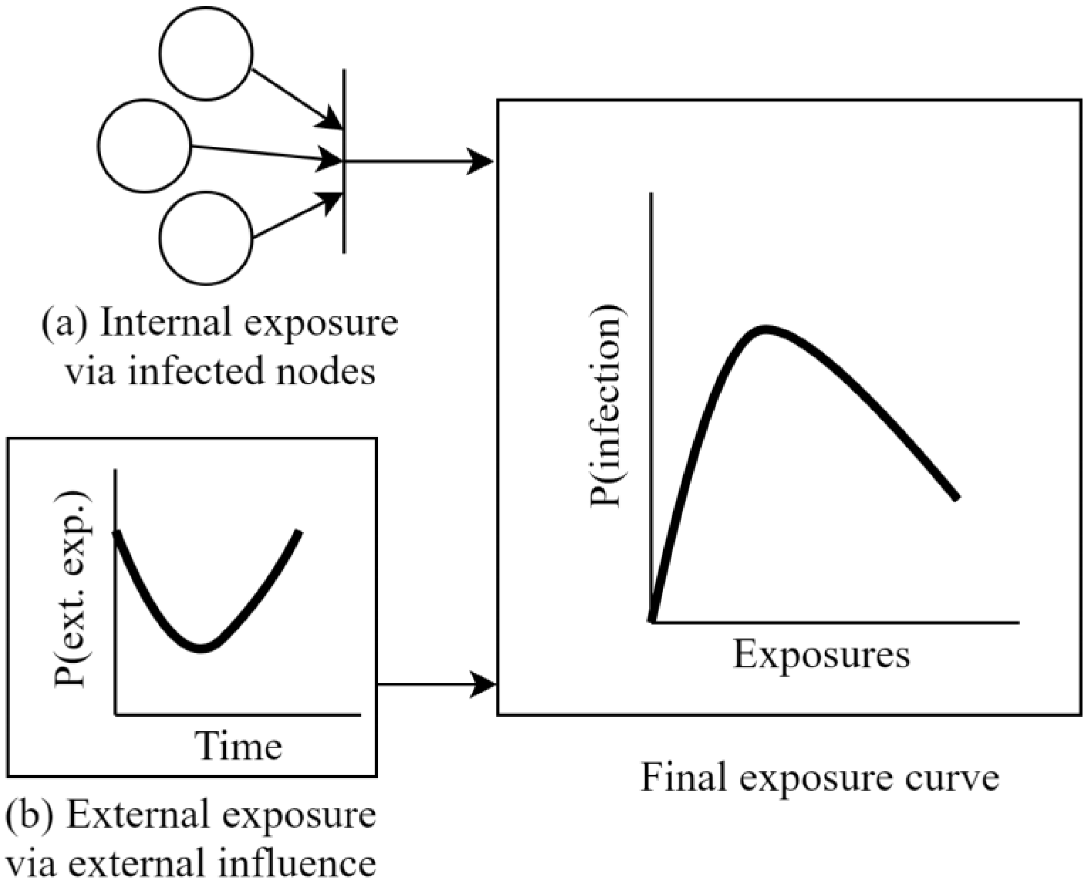
Independent Cascade Models: Exposure versus Adoption (Twitter)

- Comparing the exposure curves corresponding to different Twitter hashtags
  - Identified 8 topics, each consisting of more than or equal to 20 hashtags
  - A NULL set consisting of random subset of the collected hashtags as a base set
- $F(P)$ denotes the persistence levels
- Black dot denotes the persistence of the NULL subset
- Red cross denotes the average persistence of the topic



Independent Cascade Models: Exposure versus Adoption (Twitter)

- ❑ an exposure can beinjected externally also through media
- ❑ Available Input
  - ❑ a graph G
  - ❑ a set of node adoption times
- ❑ Study objective:
  - to model the external event
  - to find out the exposure curve



Cascade Prediction

- Can be divided into two main categories:
  - Classification problem – we predict if the diffused information/content would become popular in future
  - regression problem – we learn different numerical aspects of a cascade in future
    - final size,
    - growth
    - shape

Cascade Prediction: DeepCas

- proposed by Li et al. in 2017
- An end-to-end deep learning approach to predict the size of a cascade in future
- Uses a representation of the cascade network itself
- Cascade Graph: Let C be a set of cascades which originate in a network $G(V, E)$ at time t_0. The graph $g_c^t = (V_c^t, E_c^t)$ denotes the cascade graph for a cascade $c \in C$ obtained after time t from t_0
- Size Increment of Cascade: Δs_c represents an increase in the size of cascade c in a Δt time interval, $\Delta s_c = |V_c^{t+\Delta t}| - |V_c^t|$
- DeepCas presents a neural network framework such that given a cascade graph g_c, it predicts the increment Δs_c

DeepCas: Flow Diagram

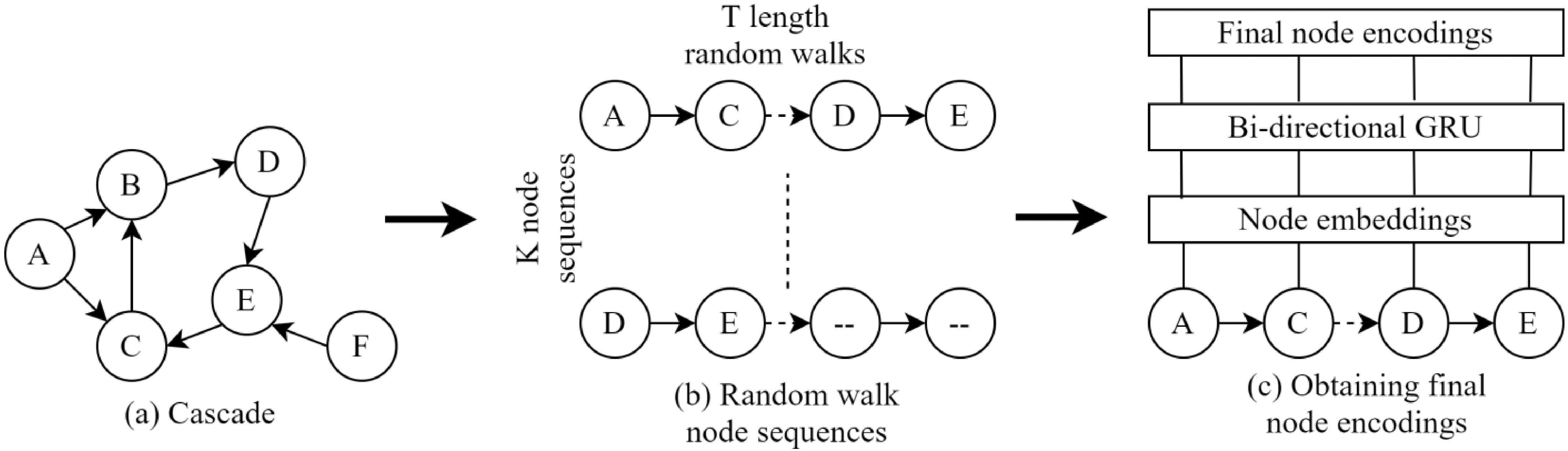
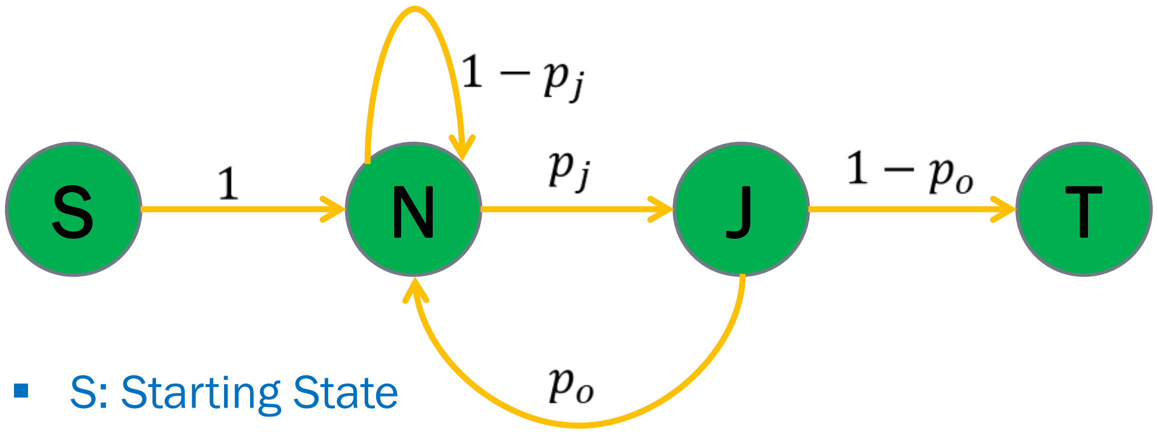


Figure 7.34: Flow diagram for DeepCas.

DeepCas: Markov Chain Model for Random Walk

- ❑ Node sequences are considered to incorporate the local and global network structure characteristics
- ❑ To sample node sequences, random walks on g_c using the Markov chain model is used



- S: Starting State
- N: Neighbor Node State
- J: Random Node Jump State;
- T: Terminal State

❑ Random walker is in state N at node v, then transition probability to select a neighbor node u to jump to:

$$p(u \in N_c(v)|v) = \frac{sc_t(u) + \alpha}{\sum_{s \in N_c(v)}(sc_t(s) + \alpha)}$$

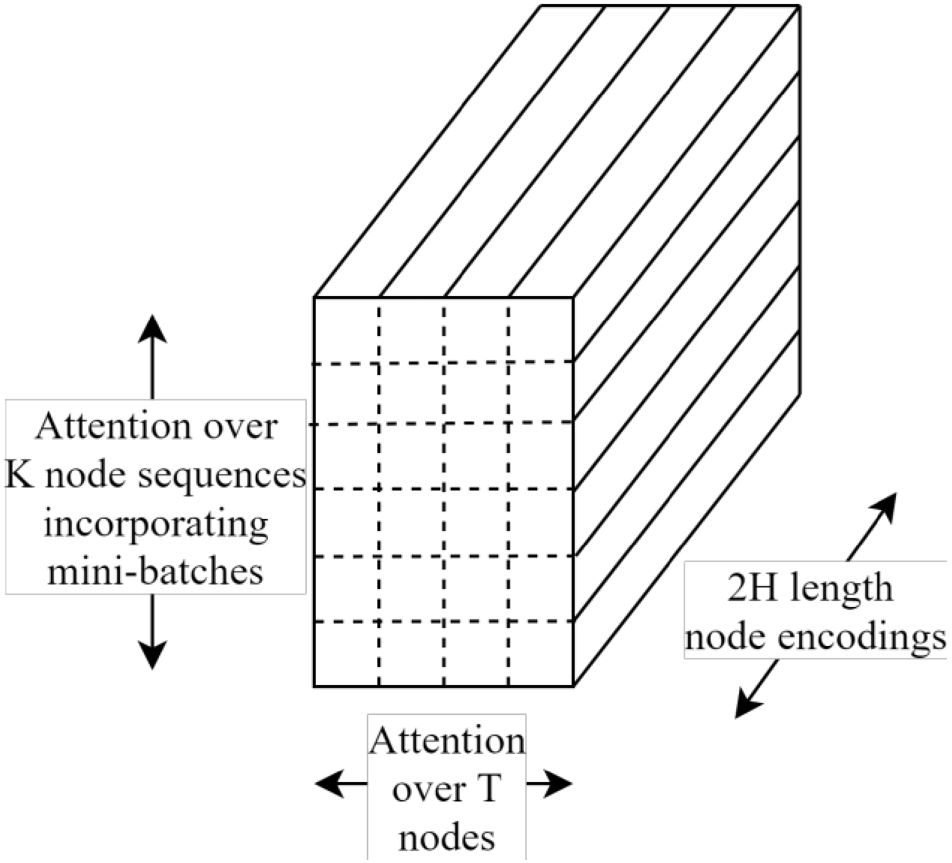
❑ random walker is at the random node jump state J, then transition probability of random jump:

$$p(u) = \frac{sc_t(u) + \alpha}{\sum_{s \in V_c}(sc_t(s) + \alpha)}$$

DeepCas: Node Embedding

- For determining length of node sequences (T) and number of node sequences (K), the probabilities p_j and p_o are learned using the neural network architecture
- The method use a node embedding to represent each node
  - N^{node}: number of nodes in a node sequence
  - $A \in \mathbb{R}^{H \times N^{node}}$: a given Embedding matrix
  - $q \in \mathbb{R}^{N^{node}}$: a vector
  - Then the vector embedding representation $x \in \mathbb{R}^{H}$ for each node is obtained as: $x = Aq$
- The authors used a bi-directional Gated Recurrent Unit (GRU) to obtain an encoding of the node sequences
- $\overleftrightarrow{h_i^k} \in \mathbb{R}^{2H}$: an encoding of the i^{th} node in the k^{th} node sequence resulted from concatenating the hidden vectors obtained from the forward and backward GRU

DeepCas: Final Cascade Graph Representation



- Use of attention mechanism to learn T and K
- Final representation of g_c:

$$h(g_c) = \sum_{k=1}^{K} \sum_{i=1}^{T} \left((1 - a_c)^{\lfloor k/B \rfloor} a_c \right) \lambda_i \overleftrightarrow{h_i^k}$$

- Objective function to minimize:

$$\mathcal{O} = \frac{1}{|C|} \sum_{c} (f(g_c) - \Delta s_c)^2$$

Cascade Prediction: DeepHawkes

- An end-to-end deep learning approach to predict the popularity of a piece of information diffused in an online social network via retweet prediction
- Model proposed by Cao et al. in 2017
- $C^i = \{(u_j^i, v_j^i, t_j^i)\}$: cascade used to describe message $m_i \in M$
- (u_j^i, v_j^i, t_j^i): j^{th} retweet such that v_j^i retweets message m_i of u_j^i wherein the time gap between the original post and the j^{th} retweet is t_j^i
- R_t^i: The popularity of m_i till time t
  - Total number of retweets within that time frame

DeepHawkes: Predict the Popularity

$$p_t^i = \sum_{j:t_j^i<t} \mu_j^i \varphi(t - t_j^i)$$

- p_t^i: the rate at which new retweets arrive for a message tweet m^i at time t
- t_j^i: time gap between the original post, and the j^{th} retweet
- μ_j^i: users who will be influenced by the j^{th} retweet
- $\varphi(\cdot)$: time decay function

❑ Factors influencing the DeepHawkes formulation:
  - ❑ Influence of users: messages retweeted by influential users tend to get further more retweets over time
  - ❑ Self-exciting mechanism: every retweet of a message is responsible for attracting more retweets thereby increasing its popularity over time
  - ❑ Time decay effect: influential exposure of retweets degrades as time passes

END