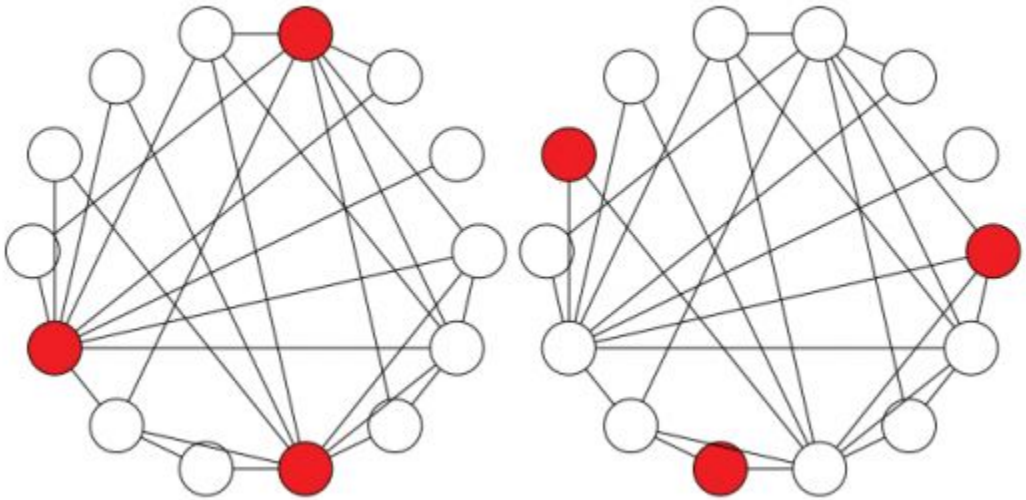
Discrete Ideas

Information Diffusion, Epidemics and Emergence of Knowledge in Networks

S. R. Sudarshan Iyengar
CSE department
IIT Ropar

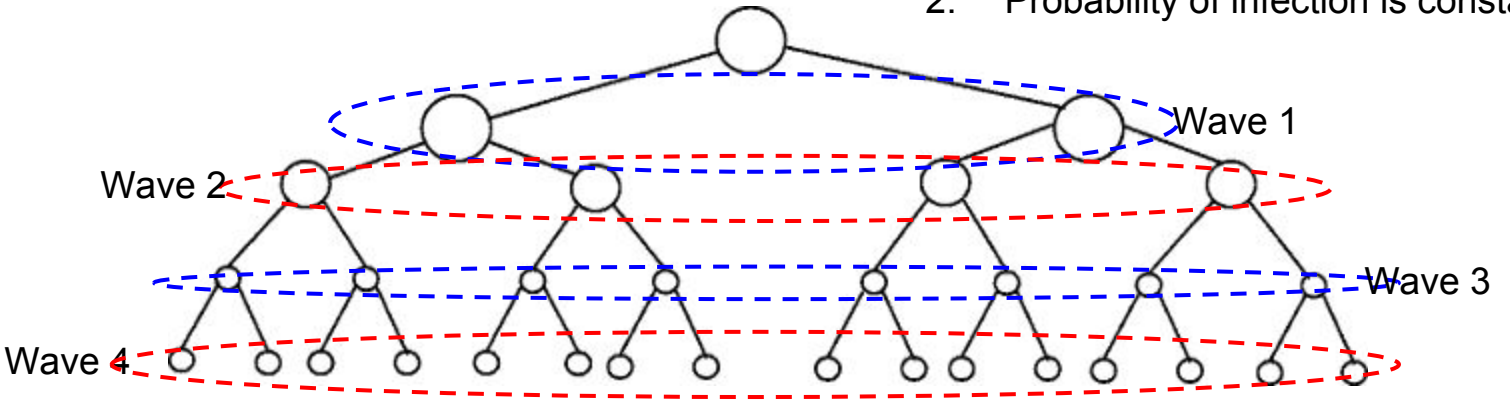
The Majority Illusion



The majority illusion in Social Networks, Kristina et al., PloS one, Vol 11, Issue 2, 17/Feb/2016

The Branching Model

1. Each node has k neighbours.
2. Probability of infection is constant = p



1. Expected number of neighbours each person infects = k . p
2. The process stops when none of the nodes in a wave is infected.
3. The value k.p has a cutting edge significance. What happens when k.p < 0 , when k.p>0 and when k.p ~ 0 ?

Emergence of Knowledge:

Epidemic or Information Diffusion?

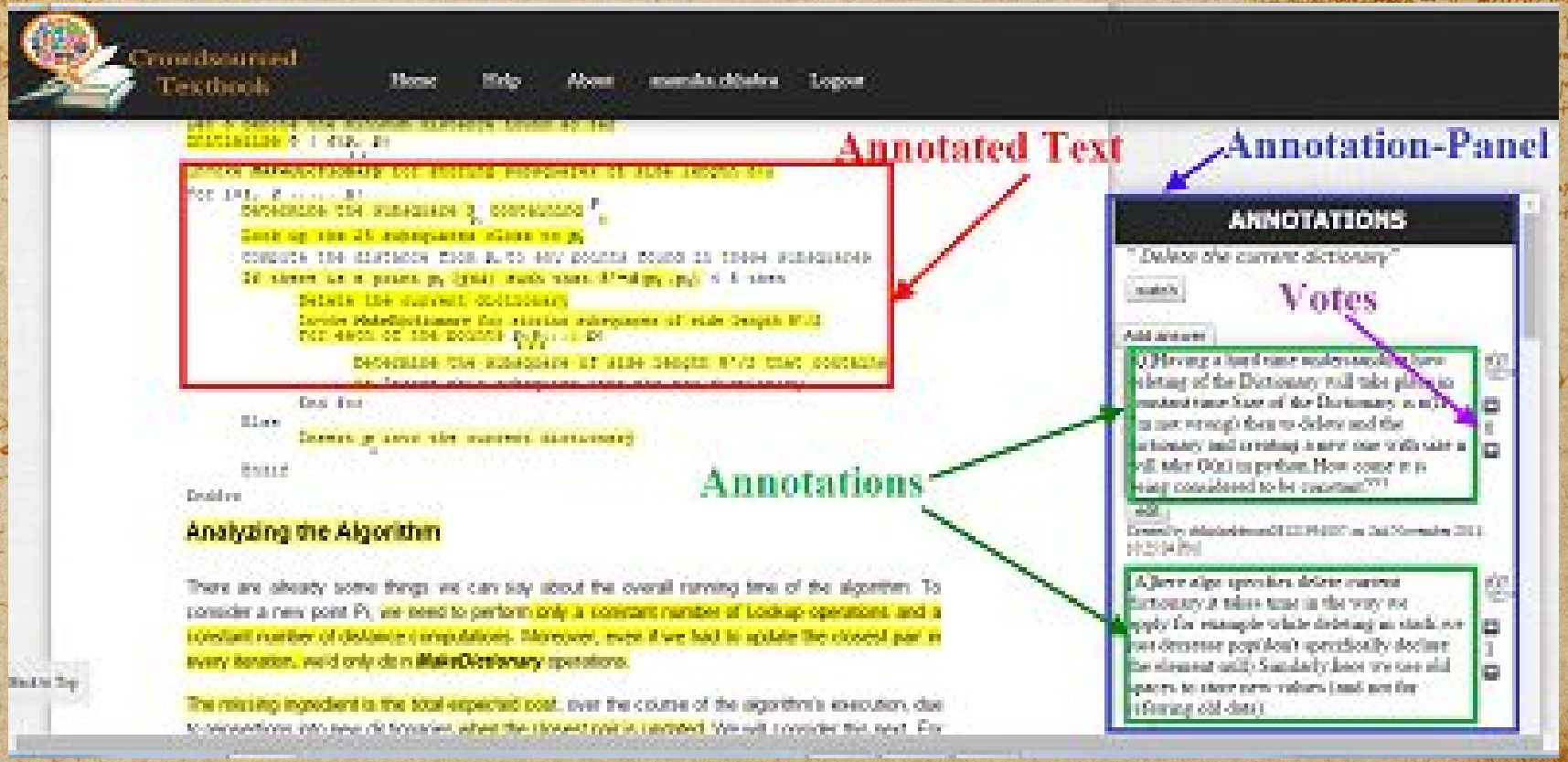

Knowledge Building In The Past

Crowdsourcing= Crowd + Outsourcing

WIKIPEDIA
The Free Encyclopedia

Quora

stackoverflow

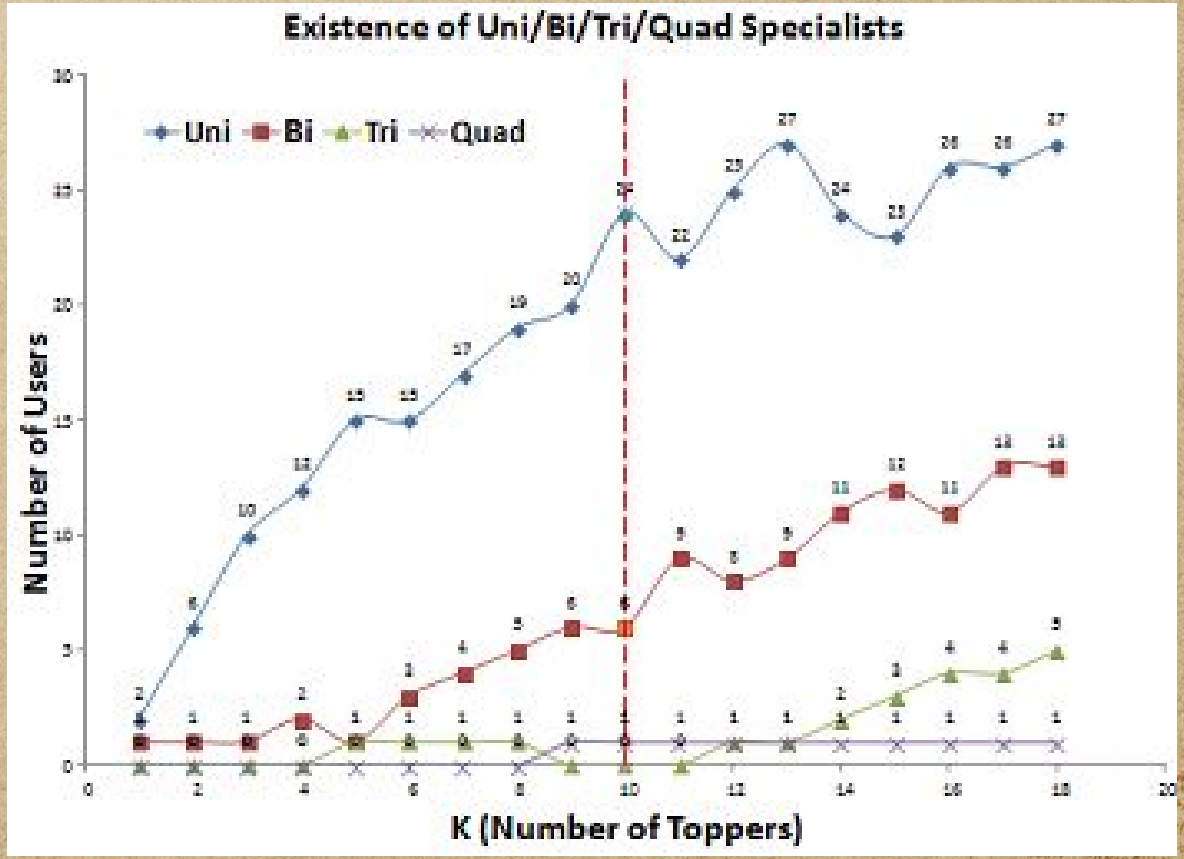


Crowdsourced Textbook

Presence Of A Variety Of Expertise

- Insights
- Questions
- Answers
- Pointers

Unispecialists...



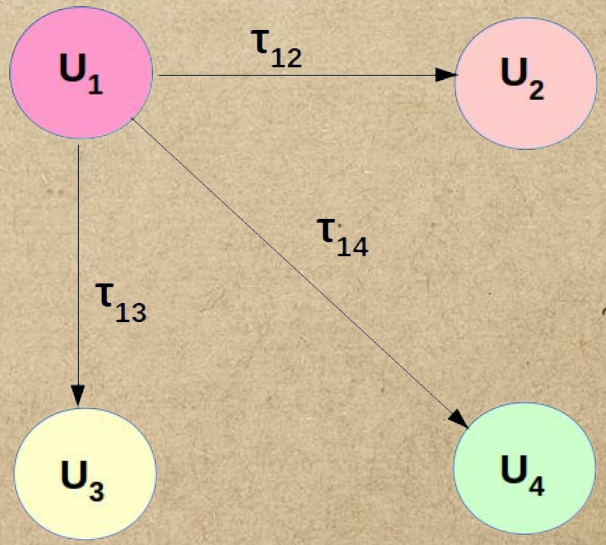
Ecosystem

There are categories of users in any Collaborative Knowledge Building System, which leads to an ECOSYSTEM.

★ Why do groups perform better than individuals?

★ Does the composition of the group matter?

The Apple Hitting Phenomenon...



The Triggering Phenomenon

A Two User Scenario

Elementary Theories On Triggering

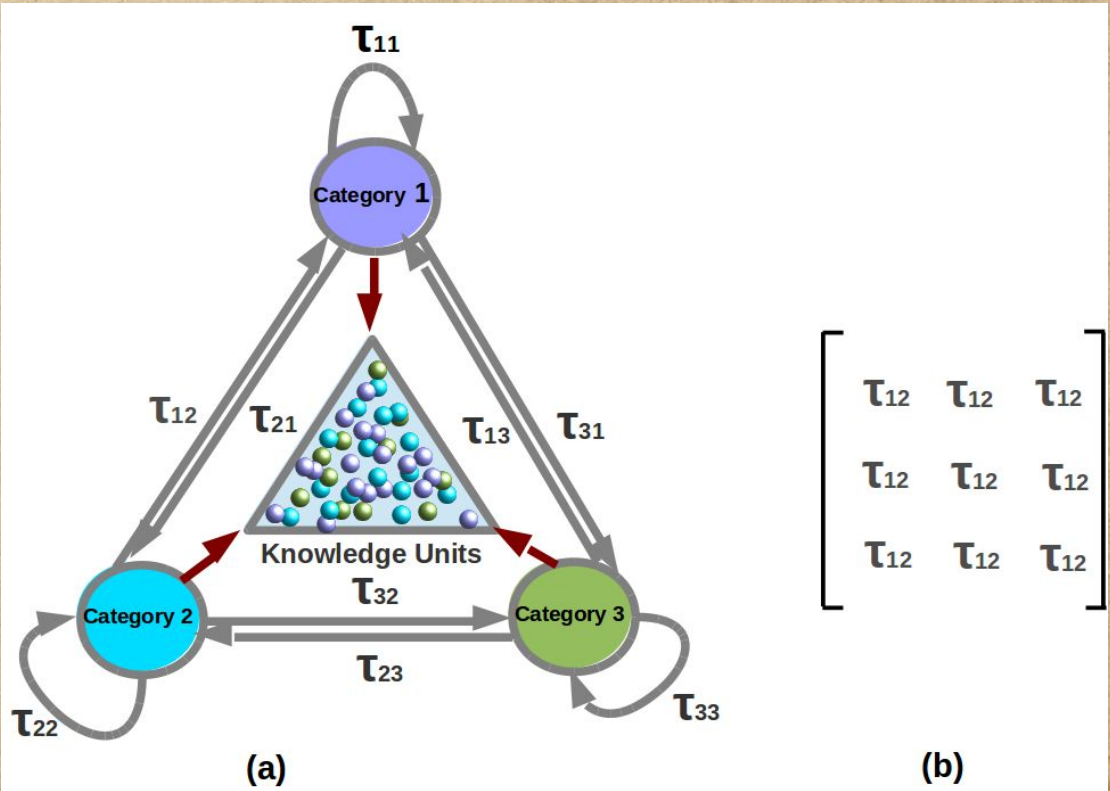
- Luhmann's Theory [1]
  - Social System and Cognitive System
  - Irritations
  - Structural Coupling
- Piaget's Model of Equilibration [2]
  - Incongruency leads to Cognitive Conflicts
  - Equilibration
- Information Processing Theory [3]
  - Knowledge Frames related by various conditions

[1] Seidl, David. "Luhmanns theory of autopoietic social systems." Ludwig-Maximilians-Universitt Mnchen-Munich School of Management (2004).

[2] Piaget, Jean. Piagets theory. Springer Berlin Heidelberg, 1976.

[3] Minsky, Marvin. "Frame-system theory." Thinking: Readings in cognitive science (1977): 355-376.

The Triggering Phenomenon in categories



KU'S AT A GIVEN TIME

KUs of a category that get added at time t are dependent on:

- ★ The number of KUs of all the categories that get added to the system at time (t− 1).
- ★ The triggering factors from all other categories to the considered category.
- ★ The number of users in the considered category.
- ★ The internal knowledge of the users of the considered category.

The Problem Statement

Given the:

- *Number of users,*
- *Categories and*
- *The triggering among the categories,*

Find the ideal number of users for each category to build maximum amount of knowledge

Sources of Knowledge Units in a Knowledge Building System

- *Internal Knowledge*
- *Triggered Knowledge*

Terminology

n_i = the number of users in category '**i**' , such that

$$\sum_{i=1}^{m} n_i = n$$

m = the number of categories,

r_i = the internal knowledge contribution of a user of category '**i**'.

$r_i(t)$ = the internal knowledge contribution of a user of category '**i**' at time '**t**', such that,

$$\sum_{t=0}^{\infty} r_i(t) = r_i$$

τ_{ij} = the number of knowledge units (KUs) of type '**i**' that get triggered due to one KU of type '**j**'.

KU'S AT A GIVEN TIME CONTD..

KUs of a category '**i**' that get added at time '**t**' is given by:

$$k_i(t) = n_i(\tau_{i1}k_1(t-1) + \tau_{i2}k_2(t-1) + \cdots + \tau_{im}k_m(t-1)) + n_i r_i(t)$$

i.e.,

$$k_i(t) = n_i\left(\sum_{j=1}^{m} \tau_{ij}k_j(t-1) + r_i(t)\right)$$

The Matrices Used...

K(t) = The column vector consisting of the knowledge generated by various categories at time '**t**'.

T = The triggering matrix

N = Matrix storing the number of users in each category

R(t) = the column matrix storing the function by which each category users add their internal knowledge to the system.

$$K(t) = \begin{bmatrix} k_1(t) \\ k_2(t) \\ \vdots \\ k_m(t) \end{bmatrix}$$

$$T = \begin{bmatrix} \tau_{11} & \tau_{12} & \cdots & \tau_{1m} \\ \tau_{21} & \tau_{22} & \cdots & \tau_{2m} \\ \vdots & & & \vdots \\ \tau_{m1} & \cdots & \cdots & \tau_{mm} \end{bmatrix}$$

$$N = \begin{pmatrix} n_1 & & & \\ & n_2 & & 0 \\ & & \ddots & \\ & 0 & & n_{m-1} \\ & & & & n_m \end{pmatrix}$$

$$R(t) = \begin{bmatrix} r_1(t) \\ r_2(t) \\ \vdots \\ r_m(t) \end{bmatrix}$$

KU'S AT A GIVEN TIME CONTD..

The knowledge generated by various categories at time 't' can be represented as:

$$K(t) = N(TK(t-1) + R(t))$$

Total Knowledge at time 't'

Lemma *The total knowledge generated in the system at time t is given by:*

$$K(t) = \sum_{i=0}^{t} (NT)^i NR(t-i)$$

Net Knowledge in the system

Theorem *Given the matrices N, T and R, the net knowledge in the system at the end of the Knowledge building process is given by:*

$$K = (I - NT)^{-1}NR$$

when

$$\rho(NT) < 1$$

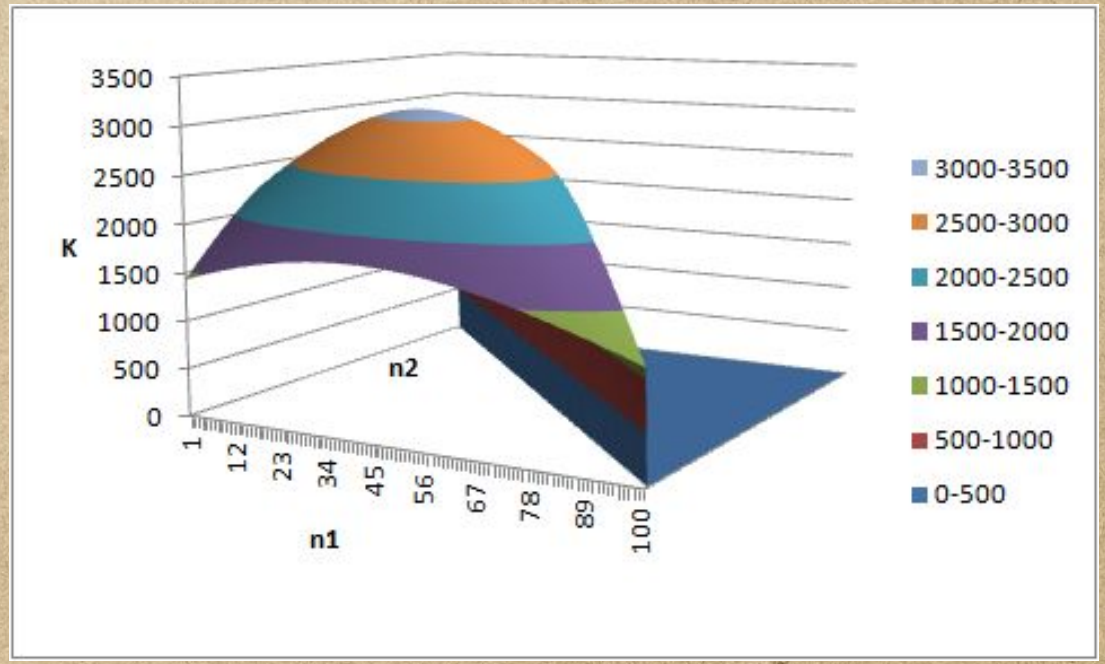
where I is an Identity Matrix of the order $m \times m$ and ρ is the spectral radius.

$$R = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}$$

$$K = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_m \end{bmatrix}$$

Ideal Distribution Of Users

The Surface Plot



Hill Climbing Algorithm

Algorithm 1 Hill Climbing Algorithm For Perfect Skill set Distribution

1. Pick a random distribution $D = [n_1, n_2, \ldots, n_m]$ of users across different categories, such that $\sum_{i=1}^{m} n_i = n$
2. Consider all the $2 * \binom{m}{2} = m(m-1)$ neighboring distributions(N) of D, defined as: $N = [\,]$
   for $i = 1 : m$
   for $j = 1 : m$
   If $i \neq j$:
   $N.append((D[1], D[2], \ldots, D[i] - 1, \ldots, D[j] + 1, \ldots, D[m]))$
3. Calculate $K_{N[j]} \forall\ 1 \leqslant j \leqslant m(m-1)$ based on the following formula:

$$K = (I - \delta)R$$

   where $K_{N[j]}$ represents the net knowledge in the system with the distribution $N[j]$ Choose i *such that* $K_{N[i]} = \max_j K_{N[j]}$
4. If $K_{N[i]} > K_D$ then $D = N[i]$ and we repeat steps 2 to 4
   else:
   return D //The perfect skill set distribution

Conclusion

- ★ Right mix of users leads to acceleration in the knowledge building process.
- ★ It should not necessarily be a group of experts, even the presence of non-experts helps.

“The sum of some is greater than parts!”

Thanks!

sudarshan@iitrpr.ac.in