Modelling Knowledge

Bielefeld/Berlin-networks-group:

Andreas Krüger
Dima Volchenkov
Philippe Blanchard
Rainer Siegmund-Schultze
Sascha Delitzscher
Tyll Krüger

Brussels 22.11.2007
NEMO User Group

Diffusion of knowledge

Generalized Epidemic Process (GEP):

- classical epidemics
- threshold epidemics
- mean-field infection
- forgetting or activation
  i.e. passive vs. active knowledge

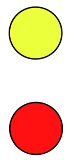
- Initial infection:
  „seed group“ of interconnected nodes

Bimodal Network with $N = N_{orgs} + N_{projs}$ nodes

$x \sim y \iff$ **edge** between x and y

$$\begin{aligned} d(x) &= \text{degree of x} \\ &= \text{number of projects in which } \textbf{organisation } x \text{ participates} \\ &\quad \textit{or} \text{ number of participating organisations in } \textbf{project } x \end{aligned}$$

$$\omega(x) = \omega_t(x) = \begin{cases} 0 & \textit{unaware} \\ 1 & \textit{knowing} \end{cases} \quad \text{at time } t$$



first model:
one type of knowledge

project x can be unaware/knowing and
organisation x can be unaware/knowing

first model:
no distinction between
organisations & projects

global observable

$$b = b_t = \frac{1}{N} \sum_{x=1}^{N} \omega(x) \quad \text{total knowledge prevalence (at time } t\text{)}$$

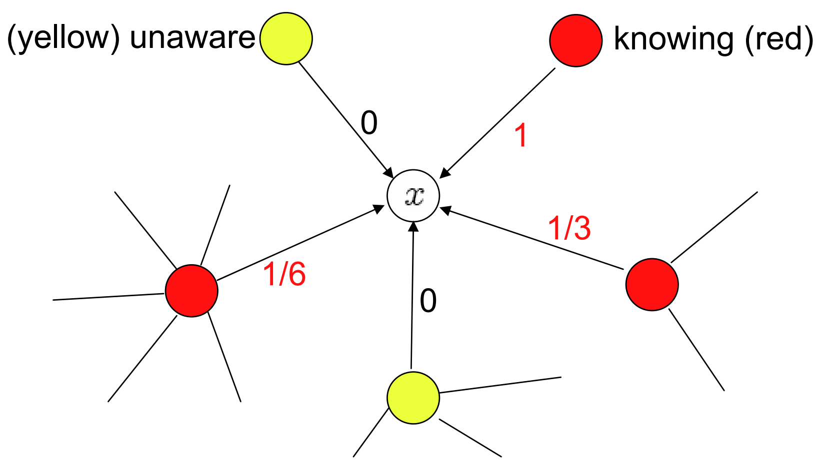
local observables

$$\Omega_t(x) = \sum_{x \sim y} \omega(y) \quad \textbf{number of knowing neighbours} \text{ of } x$$ = 3

$$\Phi_t(x) = \sum_{x \sim y} \frac{1}{d(y)} \omega(y) \quad \text{local } \textbf{knowledge inflow}$$ =1+1/3+1/6 =1.5

Inner structure of projects is *not* FullGraph, but now we account for that:

1/degree weighing of the knowing neighbours



(ε) epsilon-process ~classical epidemics

- Local infection by *knowing neighbours*
- The epsilon-process has a very low probability ε, but:
- The more neighbours knowing, the higher the probability to get knowing:

$$\mathbf{P}_{0\rightarrow 1} \approx \epsilon \cdot \Phi(x)$$

$$\Phi_t(x) = \sum_{x\sim y} \frac{1}{d(y)} \omega(y)$$

But this rather weak epsilon process only happens below a threshold … $1 \leq \Omega(x) < \Delta$

(α) alpha-process: delta-threshold infection

<u>If</u> the *number of knowing neighbours* exceeds a threshold Δ $\Omega(x) \geq \Delta$

suddenly there is a *higher* probability α to get knowing

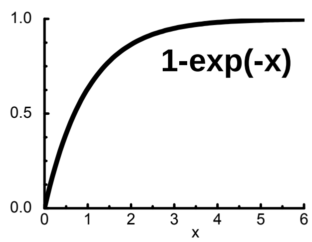
α >> ε

$$\mathbf{P}_{0\rightarrow 1} \approx \alpha \cdot (1 - e^{-\Phi(x)})$$

Degree weighed inflow

$$\Phi_t(x) = \sum_{x\sim y} \frac{1}{d(y)} \omega(y) \quad \text{shift into } [0...1]:$$

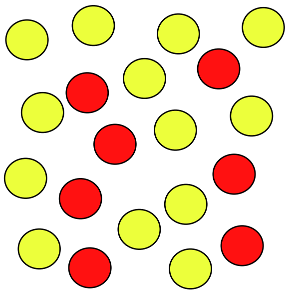
$$(1 - e^{-\Phi_t(x)}) = \begin{cases} \sim 0 & \text{for } \Phi_t(x) \text{ small} \\ \sim 1 & \text{for } \Phi_t(x) \text{ large} \end{cases}$$



(β) beta-process: mean-field influence also infects

- i.e. mass media, intuition about the state of the whole system, journals, …= „mean-field“.
- Proportional to square of relative prevalence

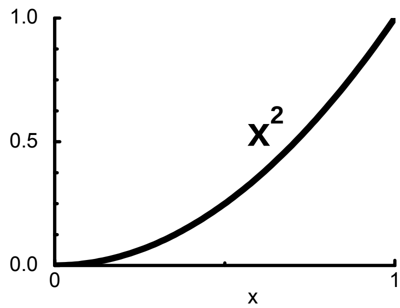
$$\mathbf{P}_{0\to 1} = \beta(b_t)^2$$



$$b_t = \frac{1}{N}\sum_{x=1}^{N}\omega(x)$$

total knowledge prevalence

7 / 20 = 0.35



(γ) gamma-process: forgetting *passive* knowledge

- The less-knowing my neighbours, the higher my γ-process-forgetting

Ratio of **unaware** neighbours

$$\mathbf{P}_{1\to 0} = \gamma\left(1 - \frac{\Omega_t(x)}{d(x)}\right)$$

Ratio of **knowing** neighbours

But I can only forget PASSIVE knowledge. ACTIVE knowledge stays with me..

(ζ) zeta-process: activation of passive knowledge

- Each time step there is a (constant) probability ζ to get from „passive“ to „active“ knowledge

$$P_{1 \to A} = \zeta$$

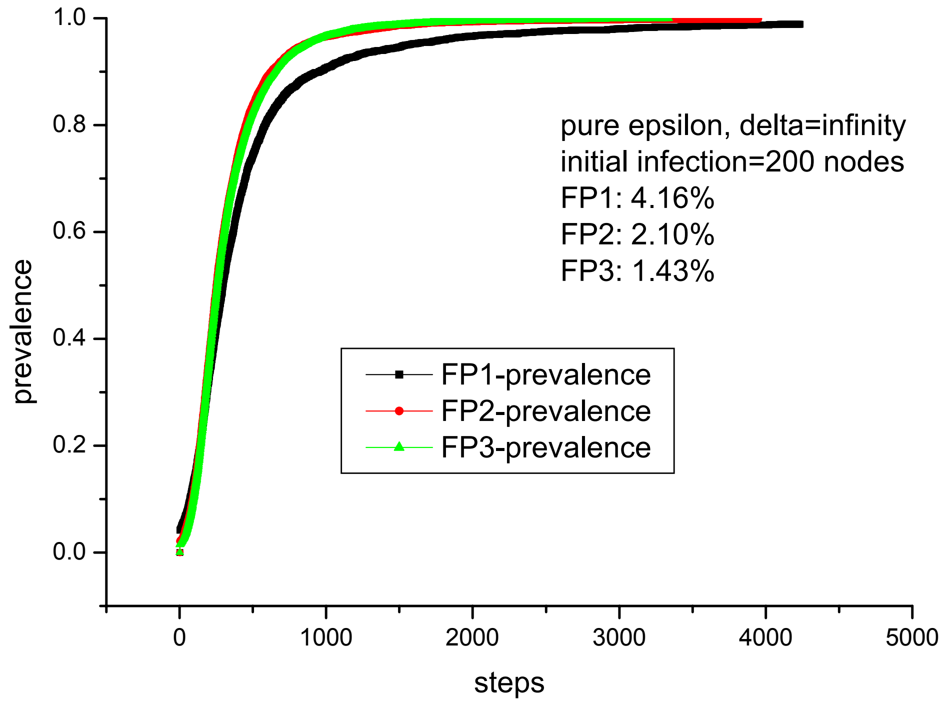
- Only passive knowledge can be forgotten.
  Once activated, the node stays knowing forever.
- Possible extensions:
  - Active knowledge „counts“ more than passive knowledge (e.g. A=3)
  - When several competing knowledge dimensions: *Active* knowledge of everything is *not* possible

Planned *next* extensions:

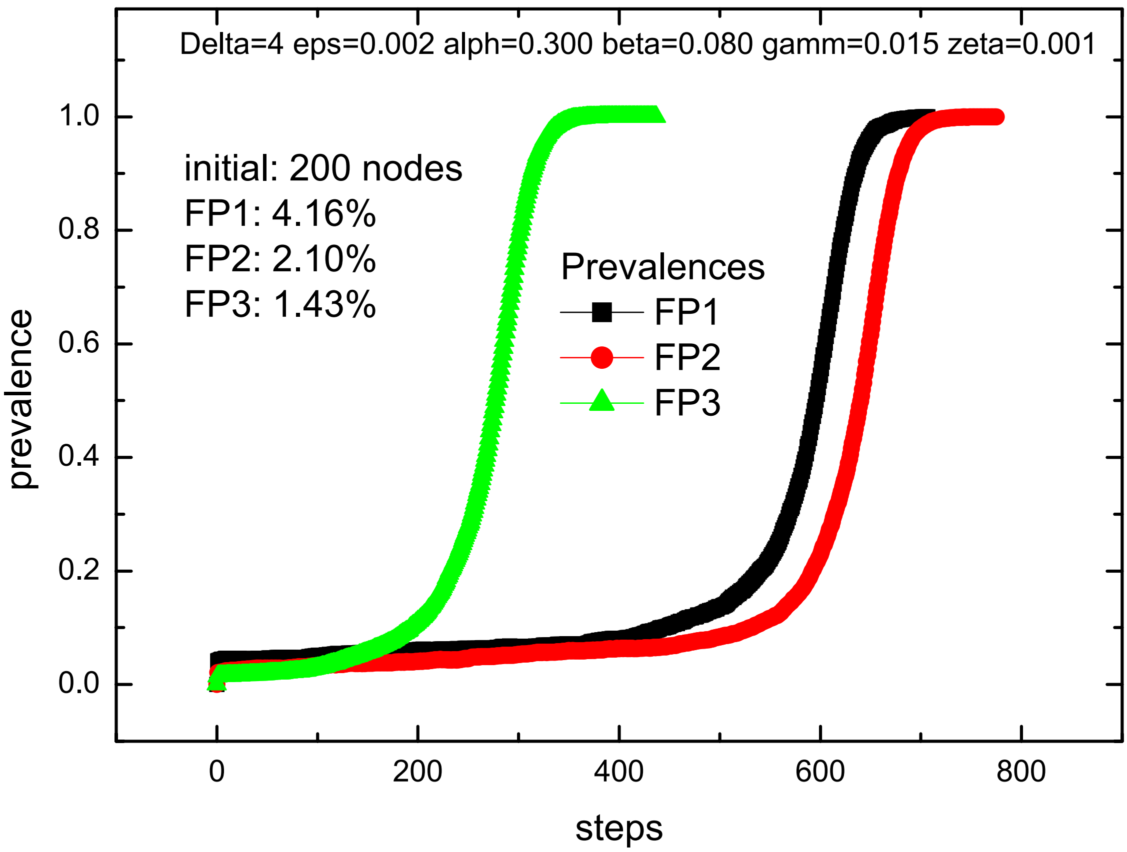
- infectious time is only short after infection
- competing knowledge types:
  - first steps into high-dimensional knowledge representation
  - no *active* knowledge of all types possible
  - Majority rules for local and mean-field processes

Pure Epsilon process epsilon=0.04

alpha=0 beta=0 gamma=0 zeta=0 delta=infinity



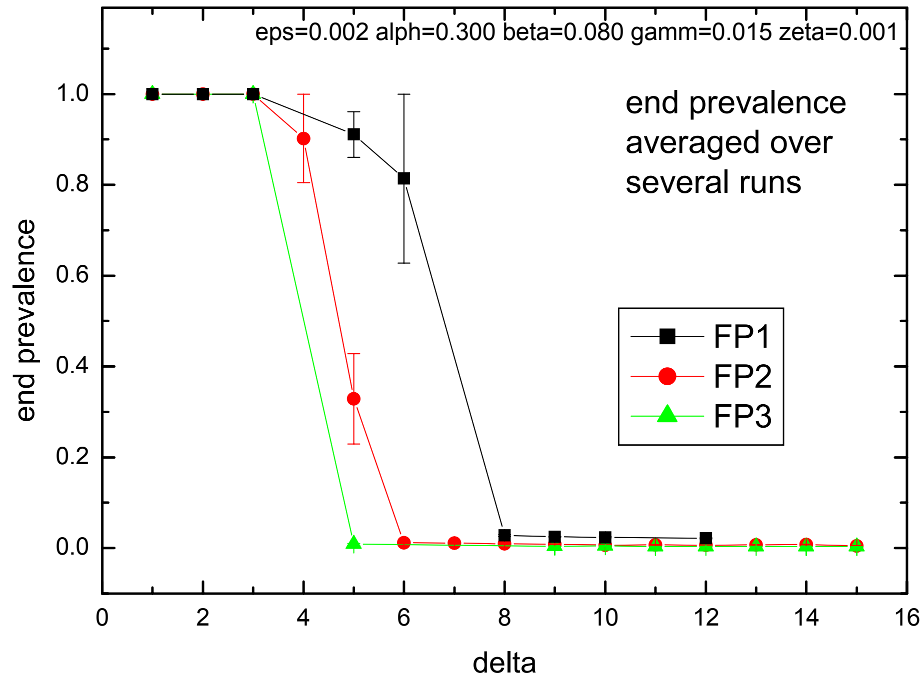
One infection run: **Prevalence** FP1 FP2 FP3 - comparison



FP1, 2, 3

Variation of delta-threshold

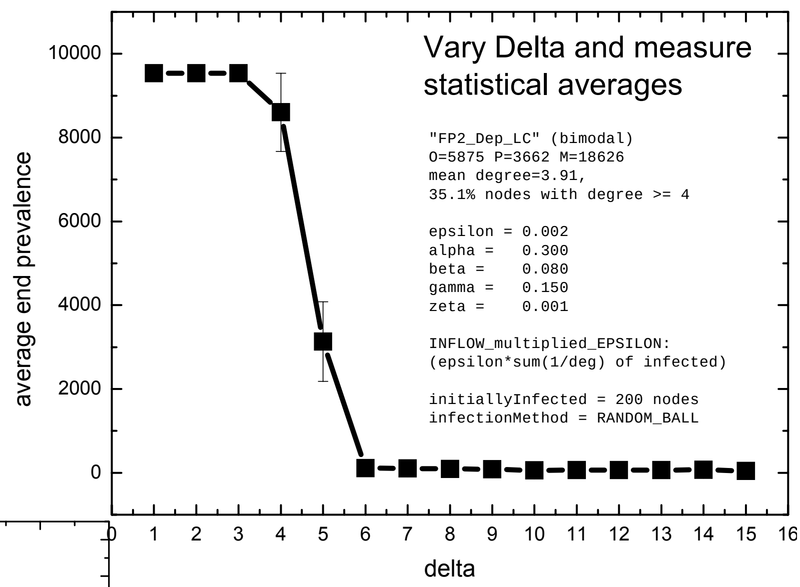
Many runs, averages of **end** results

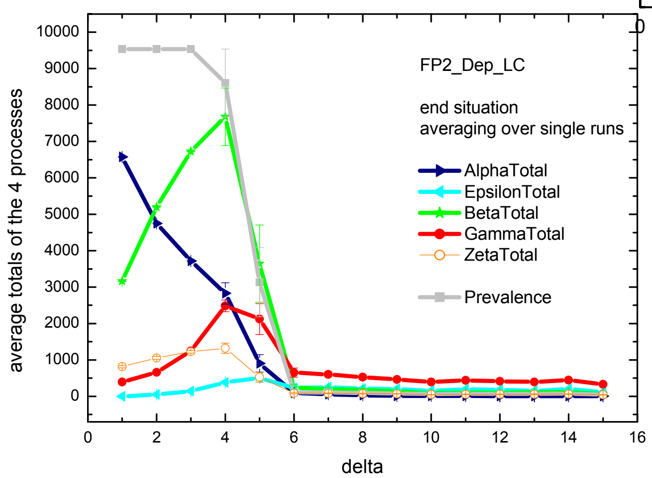


FP2

Variation of delta

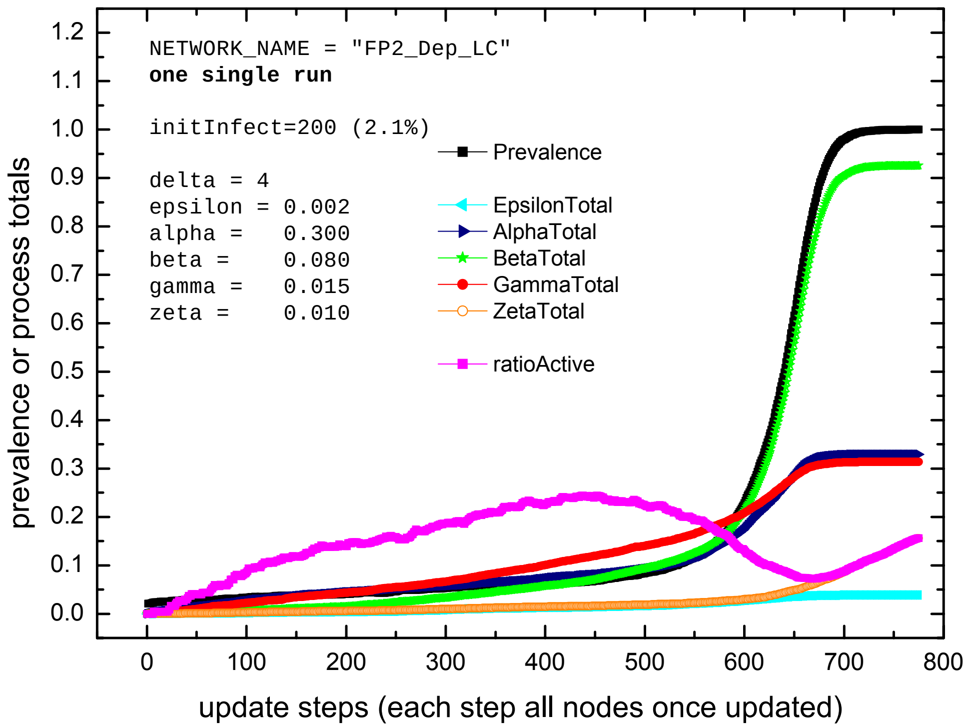
Many runs,
averages of **end** results

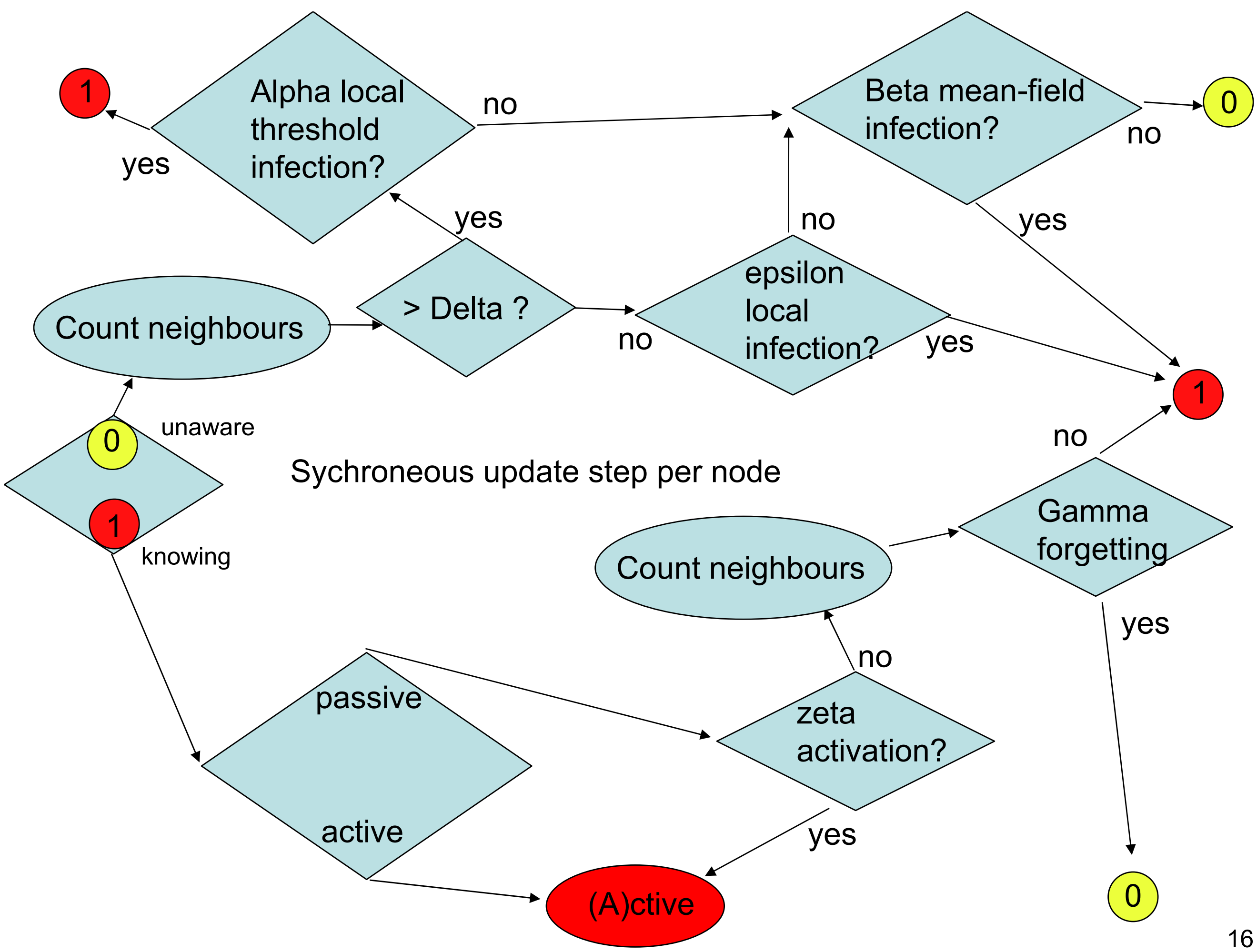




One infection run, on FP2 initial infection: 200 nodes

Each process can be studied





First simple model for best-partner choice

x=(x_1,…, x_n)
y=(y_1,…, y_n)

$|x_i - y_i| \geq \delta_{comp}$

Complementary Knowledge Subspace

$D_{comp} = \{i : |x_i - y_i| \geq \delta_{comp}\}$

$d_{comp} = \#D_{comp}$

complementary knowledge number

$|x_i - y_i| \leq \delta_{shared}$

Shared Knowledge Subspace

$D_{shared} = \{i : |x_i - y_i| \leq \delta_{shared}\}$

$d_{shared} = \#D_{shared}$

shared knowledge number

The probability $P(x \sim y)$ that x cooperates with y and produces new knowledge is a monotone rising function of the size of common speech, measured by d_{shared}:

$$P(x \sim y) \propto d_{shared}$$

The value of this new knowledge component is a monotone rising function of the number of topics in which the cooperators complement each other. This number is measured by d_{comp}:

$$x_{n+1} \propto d_{comp}$$