GEP = Generalized Epidemic Process

Corruption as a GEP

Philippe Blanchard Andreas Krueger

Tyll Krueger Peter Martin

arxiv:physics/0505031

My doctorate thesis consists of:

- Introduction into Complex Networks & SNA (incl. "ethics" chapter)

1. The Network of EC-funded R&D projects (CORDIS database)
2. "Corruption" – simulation of a contagion process on networks
3. CAMBO clustering – finding clusters in networks by matrix reordering

→ www.AndreasKrueger.de/networks → dissertation/disputation

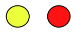
Corruption?
Imagine any contagion process with

1. Neighbour infection
   - **Threshold** contagion, i.e. local infection **only if** „level of corruption of my neighbours exceeds Δ"
   - plus small infection probability if less than Δ
2. Mean field infection
3. Mean field desinfection

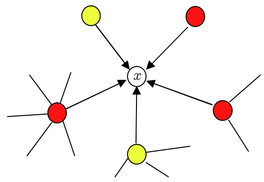
e.g.:

- opinions, fashions, …
- waves of scientific hypes, discussed topics…
- processes of innovation diffusion, **knowledge diffusion**
- ***Corruption…***

Corruption state variables

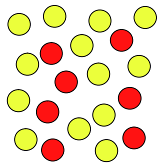


ω(x, t) in [0,1]
= node x is corrupt/non-corrupt at time t



Ω(x, t)
= number of infected neighbours

$$= \sum_{y \sim x} \omega(y,t)$$



7 / 20 = 0.35

b_t = total prevalence of corruption at time t

$$b_t = \frac{1}{|V|} \sum_{y \in V} \omega(y,t)$$

Implemented LOCAL Processes:

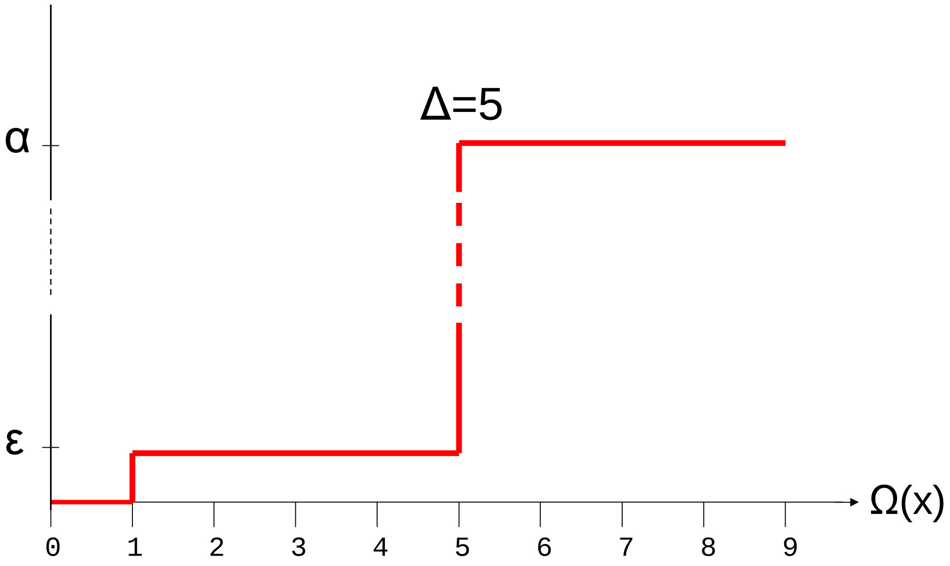
α-process: "if enough neighbours…"

- The **local** transmission probability for # of corrupt neighbours ≥ Δ
- Typical value: α>>ε,β,γ
- Possible translations: **Influenceability by others, Decisiveness**

ε-process: "if at least one neighbour …"

- (similar to classical) **local** epidemic probability for # of corrupt neighbours < Δ
- Typical value: ε<<α,β,γ (very small)
- Possible translations: **Naiveity, Openness**

Local infection probabilities, (absolute Δ=5 threshold)



Implemented GLOBAL Processes:

β-process: "**infection** through public opinion"

- The mean field transmission process due to the total (believed) prevalence of corruption
- Typical value: ε<β<γ
- Possible translations:
  "Random" infection: How informed are you? How much do you belief in mass media?

$$\Pr_\beta(\omega_{t+1}=1 \mid \omega_t=0) = \beta(b_t)(1-(1-b_t)) = \beta(b_t)^2$$

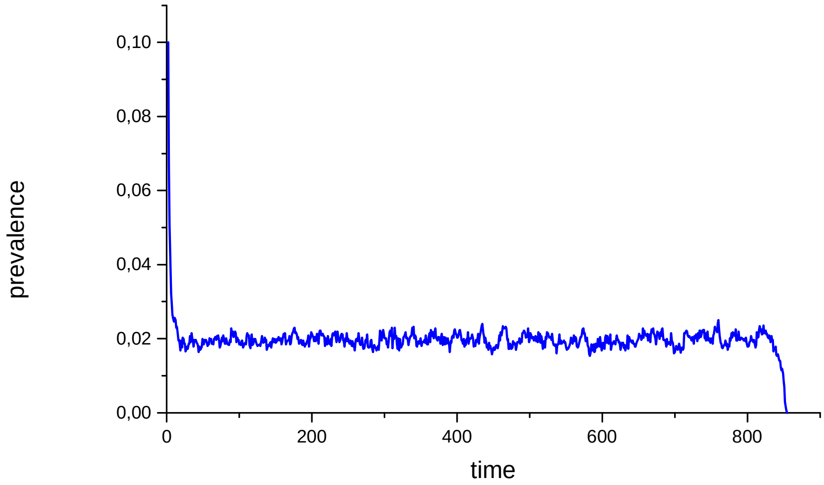
γ-process: "(only) the healthy can **cure** others"

- The (mean field!) corruption recover process due to the fight of the (healthy) society against corruption
- Typical value: β<γ<α
- Possible translations:
  random resistance / recovering / cleaning

$$\Pr_\gamma(\omega_{t+1}=0 \mid \omega_t=1) = \gamma(1-b_t)$$

Single run until stagnation 1

Low semistable prevalence in a real collaboration network

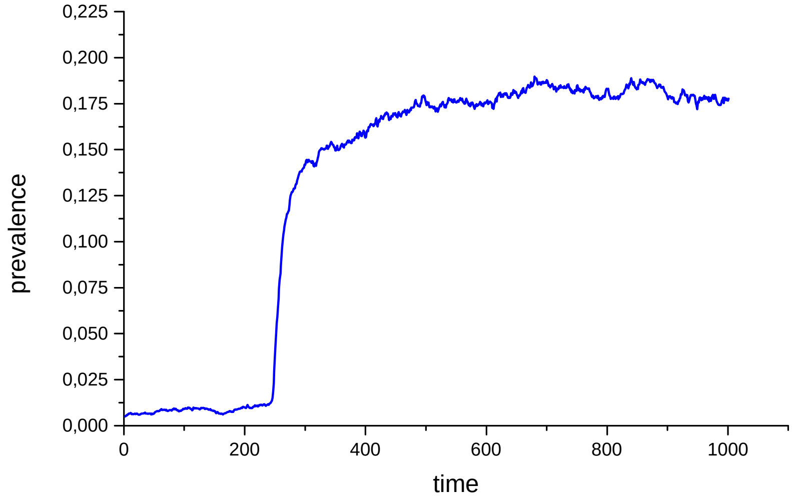


Δ=30 ε=0 α=0.99 β=0.09 γ=0.545 b_0=0.10

Network FP2: N=4879 M=57633 mean degree=23.6 mean triangles=256.9

Single run until stagnation 2

Jump from very low prevalence to low prevalence

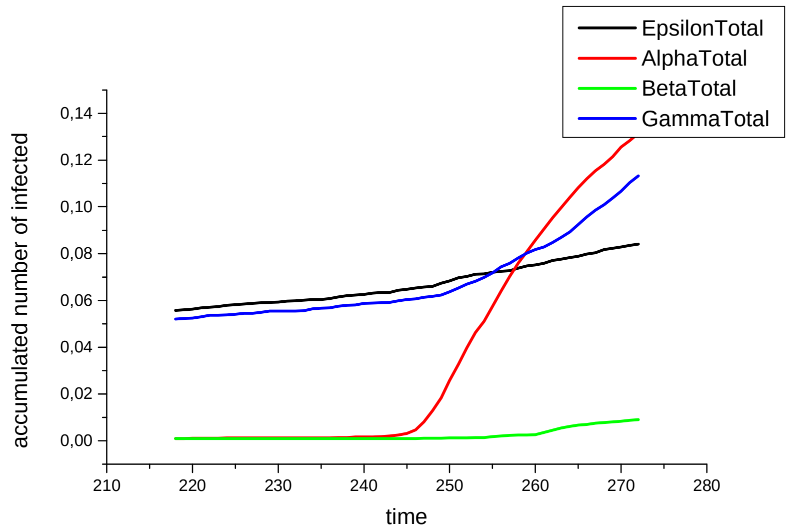


Δ=25 ε=0.001 α=0.20 β=0.04 γ=0.03 b_0=0.005

Network FP3: N=7710 M=93852 mean degree=24.3 mean triangles=418.1

Single run until stagnation 2

Contribution of the 4 (des)infection paths for the "jump"



…(end of) single runs: some insights

- The Single-Process plots are very useful to understand the rich behaviour
- Exists possibility for a fatal resonance between local (ε, α) and mean-field process (β)
- small accumulation of infection unnoticed until a critical density (a point of no return) of corruption is reached
  → then quickly complete saturation of society

Similar to other complex systems with hidden phase transitions (e.g. climate)

Structure of the Python Program

Initial infection with b_0 corrupt nodes (random/ball)

→update one vertex

→all vertices, sync update

→do until stagnation Example Legend

get: end prevalence (usually ~0 or ~1)

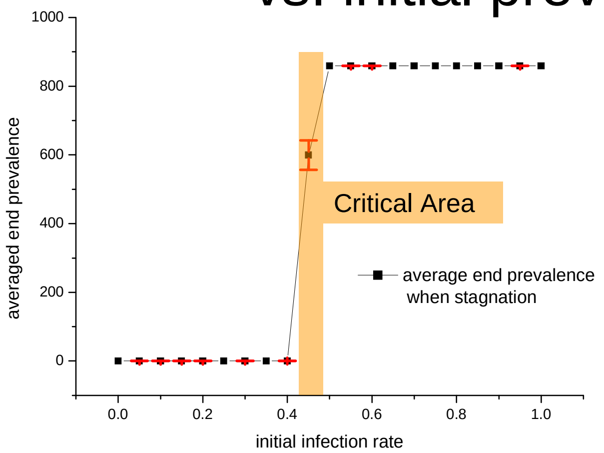
→ many runs to get *average* end prevalence

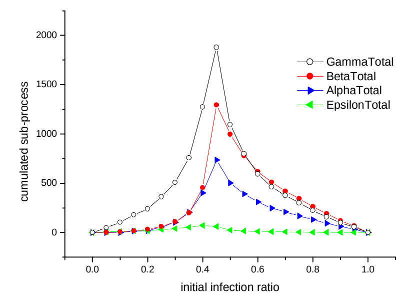
→Transition finder: vary b_0 to locate $b_{undercrit}$ and $b_{overcrit}$ and get (mean value) b_{crit}

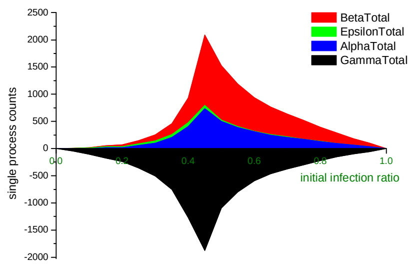
→ sweep (network property)

X=N, M, T or λ to plot b_{crit} over X

end prevalence vs. initial prevalence



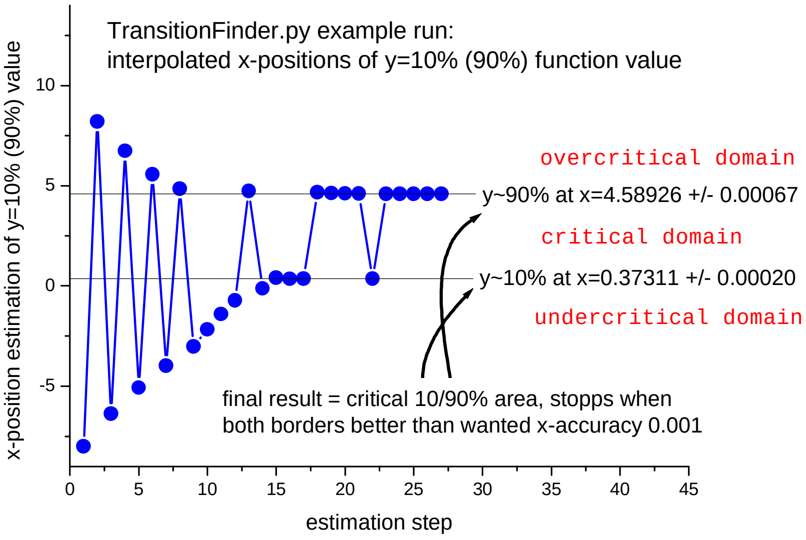




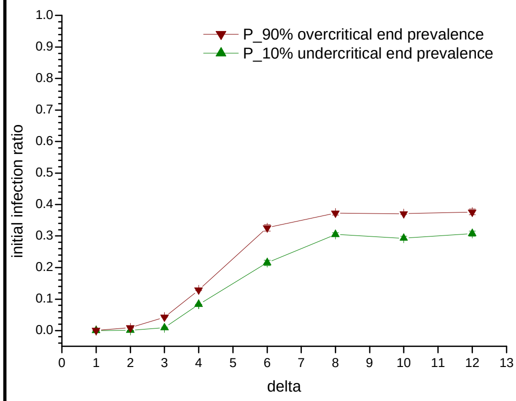
Δ=8 ε=0.001 α=0.20 β=0.08 γ=0.06

N=859 M=3368 by Random Set Model (O=1000 P=500)
degree: min=2 max=33 mean=7.8
44% nodes have degree ≥ 8

Find x_{crit} but avoid critical area by linear interpolation of 10% / 90% y-value

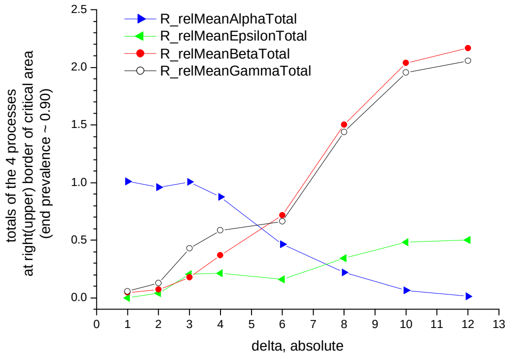


Critical Density over
Δ = neighbour infection threshold



Lower and upper bounds for b_{crit}
as a function of neighbour infection threshold Δ

ε=0.005 α=0.35 β=0.08 γ=0.04
GNM with N=1500 M=5000
→ mean degree ~ 6.7



Total number of state changes
splitted into the 4 different sub-processes

Phase transition with respect to **Initial Number of Infected Nodes**

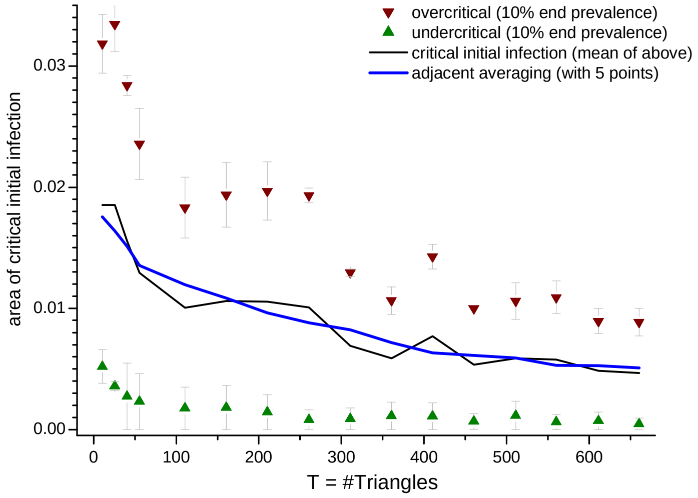
In corruption infection:

- both (mean-field and local) processes can have phase transitions **with respect to the *initial density of corrupt vertices***

<u>Classical epidemic process</u> is:

- either overcritical (reproduction number $R_0>1$) → and *a single initially infected node* infects a positive fraction with positive probability
- or undercritical ($R_0<1$) → all infected will die out, everyone's healthy

Critical Density over T = #triangles



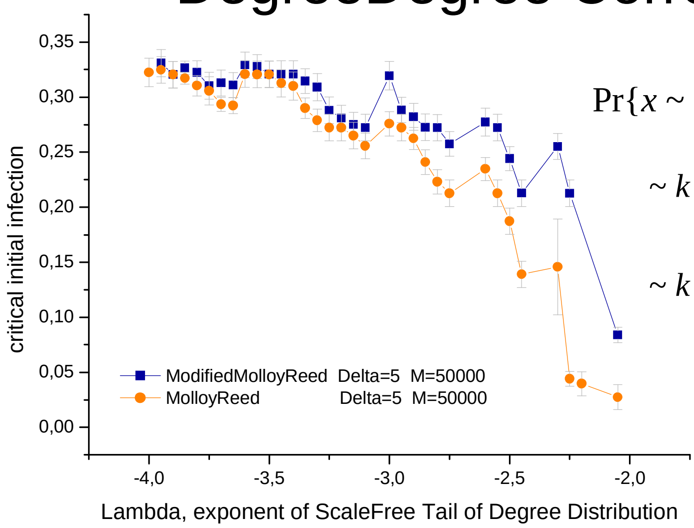
Δ=2 ε=0.005 α=0.30 β=0.08 γ=0.04 (20 runs averages)

Network GNTM: N=1000 M=2000

Clustering helps corruption

- In classical epidemics, local clique-clustering slows down the disease spreading because of **re-infection** instead of the infection of healthy
- Here, though, the highly clustered, medium-degree vertices are especially well-suited for the spread of corruption, because a threshold Δ of neighbours has to be corrupt to trigger the α-process

Additive vs. Multiplicative DegreeDegree Correlation



$$\Pr\{x \sim y \mid d(x) = k \wedge d(y) = k'\}$$

$\sim k \bullet k'$ multiplicative (normal MolloyReed)

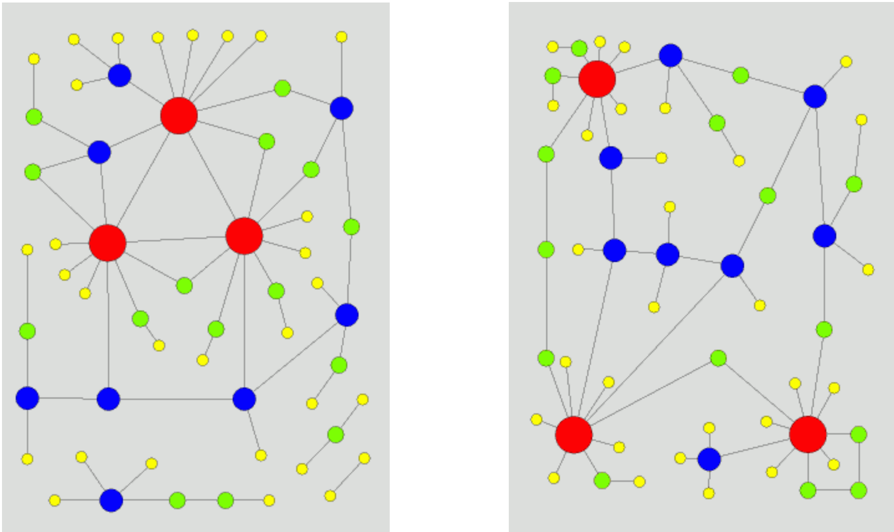
$\sim k + k'$ additive (modified MolloyReed)

DegreeDegreeCorrelation

Δ=5 ε=0 α=0.35 β=0.08 γ=0.04

N=20000 M=50000 generated by (modified) MolloyReed with given Degree-Histogram

Multiplicative vs. Additive Degree-*Correlation*



N.B.: identical degree *distribution:*
3 reds (degree 10); 8 blues (degree 4), 15 greens (degree 2), …

Additive vs. Multiplicative DegreeDegree Correlation

- SF-Networks with multiplicative DegreeCorrelation (hierarchical, …) are more easily corrupted than those with additive DegreeCorrelation (polycentric, democratic)
- Especially true for low λ < 3 (where *very big hubs* exist).

Epidemic Control

- This is an ABSTRACT model!
  Only structural & schematic tendencies!
- Positively correlated to corruption:
  α & ε = strength of influence of others
  β = strength of e.g. mass media
- Negatively correlated to corruption:
  Δ=How many neighbours have to be corrupt?
  γ=How strong does the society fight back?
- avoid high clustering
- “Transparency”: $\Delta\uparrow$ $\alpha\downarrow$ $\beta\downarrow$
- “Police”: $\beta\downarrow$(increase of fear), $\gamma\uparrow$(uncovering rate)
  but $\gamma > \alpha, \beta$ is a “total police state”
- Moral resistance: $\Delta\uparrow$ $\alpha\downarrow$
- (Hierarchical) Decision Systems should be as flat, independent, polycentric as possible!

Knowledge diffusion: first thoughts

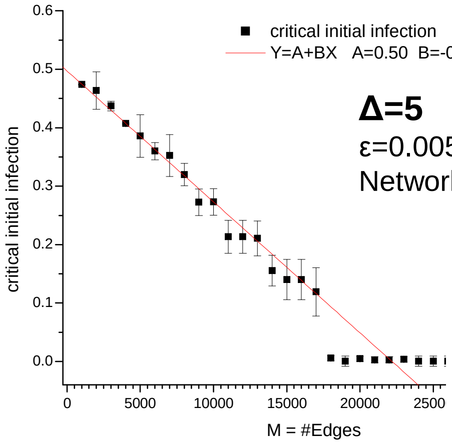
- Infection: classical (epsilon) & threshold (alpha)
- no / passive / active knowledge
- activation triggered: local / global / spontaneous
- mean field:
  - the “active” publish
  - there is a time lag
- Forgetting: spontaneous & local effect

Thank you very much

→ questions?

www.AndreasKrueger.de/networks
→ dissertation/disputation

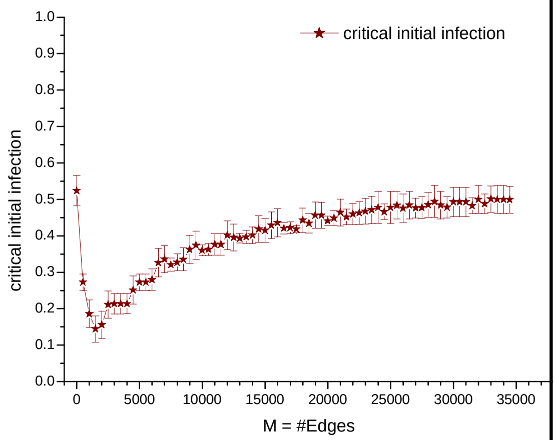
Absolute vs. Relative Threshold Δ



Δ=5

ε=0.005 α=0.35 β=0.08 γ=0.04

Network GNM: N=4000



Δ(x) = 0.8 * degree(x)

ε=0 α=0.35 β=0.08 γ=0.04

Network GNM: N=4000

global observable

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

local observables

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

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