The Epidemics of Corruption

Philippe Blanchard
Andreas Krueger
Tyll Krueger
Peter Martin

arxiv:physics/0505031
www.AndreasKrueger.de/networks

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 ~ total prevalence
3. Mean field desinfection ~ number of uninfected

e.g.:

- opinions, fashions, …
- waves of scientific hypes, discussed topics…
- transition to a democratic society, sustainable society, …
- innovation processes
- ***Corruption…***

Main features and findings

- generalized epidemic process
- on the graph of social relationships
- strong nonlinear dependence of transmission probability on local density of corruption
- additional mean field influence of the overall prevalence of corruption in society
- Global-Local interaction, fatal resonance
- Existence of an infection threshold
- important role:
  - network clustering (local redundancy of contact paths)
  - degree-degree correlation (hierarchical vs. more democratic)
- we study:
  - phase transitions
  - interaction of the processes
  - influence of graph structure

Corruption:
few attempts to model mathematically

1. Microeconomics, game theory,
   maximizing profit functional,
   mean corruption, stability analysis
2. cellular automata, simple state variables,
   local interaction dynamics,
   often only on 1-dim lattice (nevertheless complex dynamics)

Our way:
   Similar to 2., but on *complex networks*

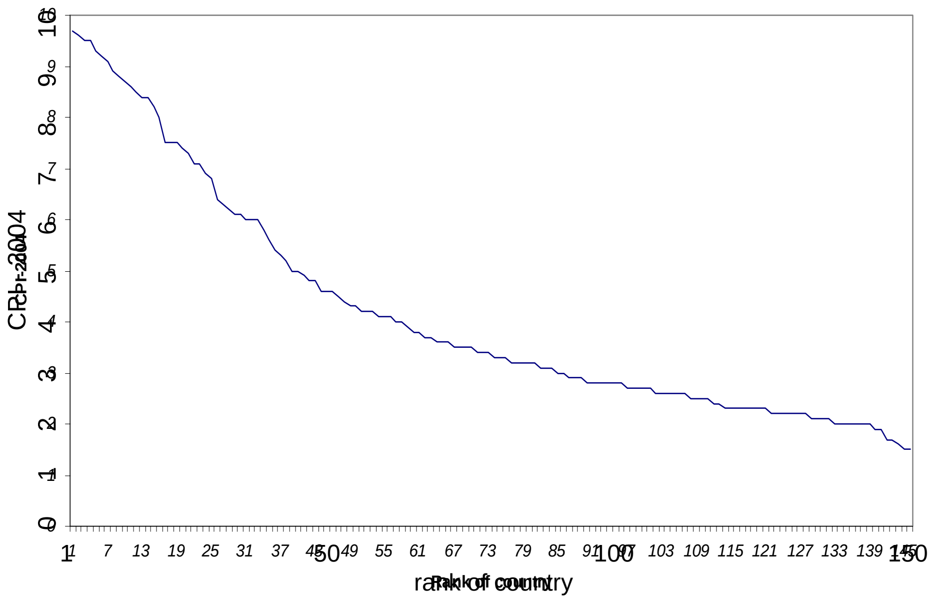
Corruption

- Human social interaction
  Deviation from fair play (cultural context)
- Misuse of (public) power
  Gain profit in a more or less illegal way
- Criminal act, but also **state of mind**
- Typology of corrupt actors:
  highly educated, well positioned,
  not thinking to have done s.th. wrong
- Notorious problem to get empirical data

The Corruption Perception Index CPI-2004 for the first 40 countries (from Transparency International)

Rang	Länder	CPI2004	Rang	Länder	CPI2004
1	Finnland	9.7	21	Barbados	7.3
2	Neuseeland	9.6	22	Frankreich	7.1
3	Dänemark	9.5	23	Spanien	7.1
4	Island	9.5	24	Japan	6.9
5	Singapur	9.3	25	Malta	6.8
6	Schweden	9.2	26	Israel	6.4
7	Schweiz	9.1	27	Portugal	6.3
8	Norwegen	8.9	28	Uruguay	6.2
9	Australien	8.8	29	Oman	6.1
10	Niederlande	8.7	30	Vereinigte Arabische	6.1
11	Großbritannien	8.6	31	Botswana	6
12	Kanada	8.5	32	Estland	6
13	Österreich	8.4	33	Slowenien	6
14	Luxemburg	8.4	34	Bahrain	5.8
15	Deutschland	8.2	35	Taiwan	5.6
16	Hongkong	8	36	Zypern	5.4
17	Belgien	7.5	37	Jordanien	5.3
18	Irland	7.5	38	Katar	5.2
19	USA	7.5	39	Malaysia	5
20	Chile	7.4	40	Tunesien	5

The Corruption Perception Index CPI-2004
for the first 40 countries (from Transparency International)



Corruption state variables

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

$\Omega(x, t)$
= number of infected neighbours at time t

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

b_t = total prevalence of corruption at time t

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

$|V|$ = number of nodes

Implemented LOCAL Processes:

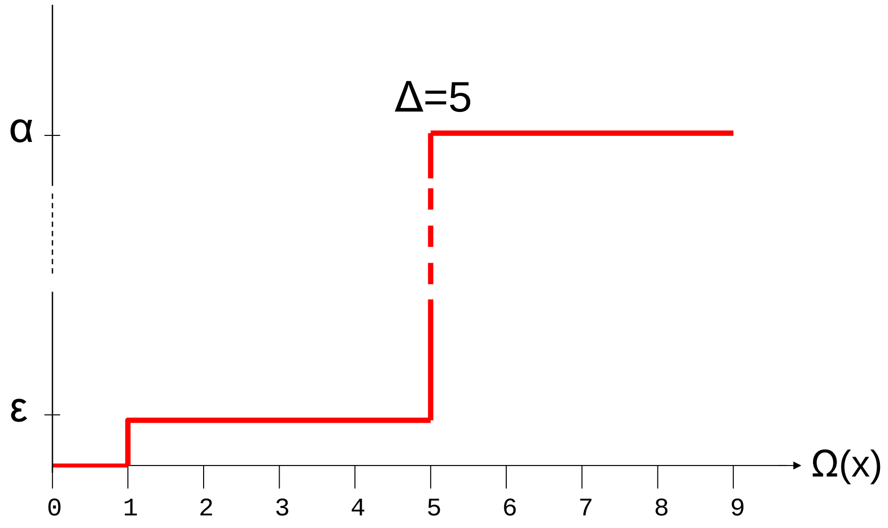
α-process: "if enough neighbours…"

- 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 …"

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

Local infection probabilities, (absolute Δ=5 threshold)



The threshold Δ or δ:

Let d(x) be the degree of node x
and Ω(x) the number of infected neighbours

Absolute **threshold Δ**

The α-process can happen if $\Omega(x) \geq \Delta$
regardless of the degree d(x) of a node
→ nodes with d(x) < Δ are irrelevant for the α-process
→ hubs with d(x) >> Δ are easily infected by the the α-process

Relative **threshold Δ**

The α process can happen if $\Omega(x) / d(x) \geq \delta$
→ Any node can be taken by the α-process
→ hubs with d(x) >> Δ are more difficult to infect

We have concentrated on the *absolute threshold* Δ
because easier to treat analytically – and due to lack of time ☺

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?**

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

GLOBAL processes, cntd.

β-process: mean-field infection

- proportional to total prevalence b_t
- Individual has to *overcome* “fear”
  fear is proportional to *uninfected part* (1- b_t)

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

γ-process: mean-field des-infection

- Only the healthy can cure
  →proportional to 1-total prevalence (1-b_t)

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

The networks

- Erdösz-Renyi Random Graphs (RGs)
- Triangle-Modified RGs:
  Throw triangles first, then fill up with edges
- MolloyReed Algorithm to get arbitrary degree distribution, e.g. scale-free
  ! always multiplicative degree-degree correlation !
- Modified MolloyReed Algorithm: choose ~equal outdegree for all nodes
  → additive degree-degree correlation
- “Real world” empirical networks of the EU-funded R&D projects (projects and organizations from CORDIS database of the European commission)
- Set graphs, intersection graphs, “affiliation networks”, unipartite projections of such bipartite graphs …

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

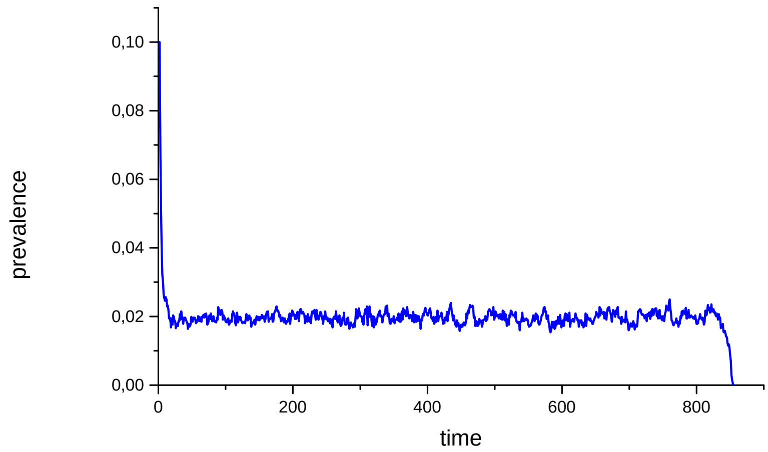
→Transition finder: vary b_0 to locate $b_{undercrit}$ and $b_{overcrit}$ and get (mean value) b_{crit}

→ sweep (network property)

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

Single run until stagnation 1

Low semistable prevalence in a real collaboration network

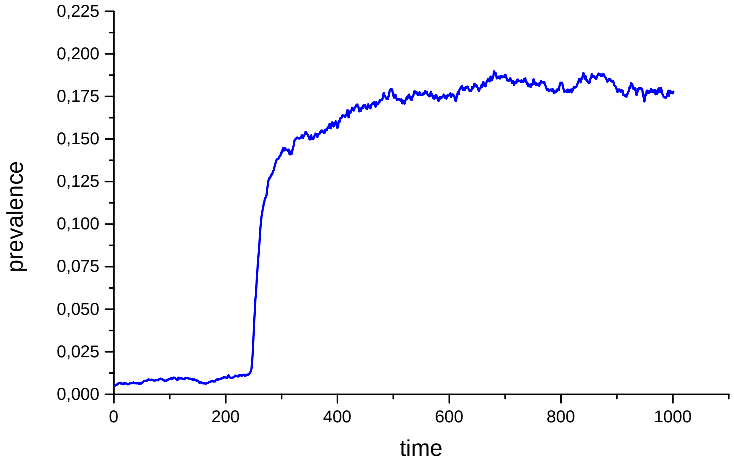


Δ=30 ε=0 α=0.99 β=0.09 γ=0.545 b_0=0.10

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

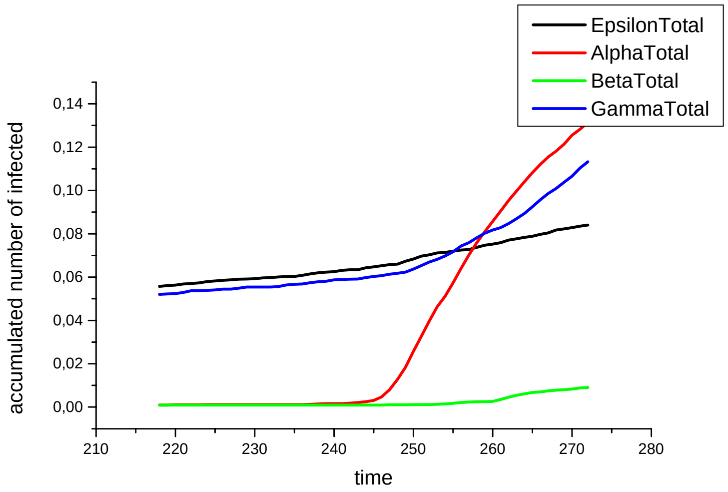


Δ=25 ε=0.001 α=0.20 β=0.04 γ=0.03 b_0=0.005

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

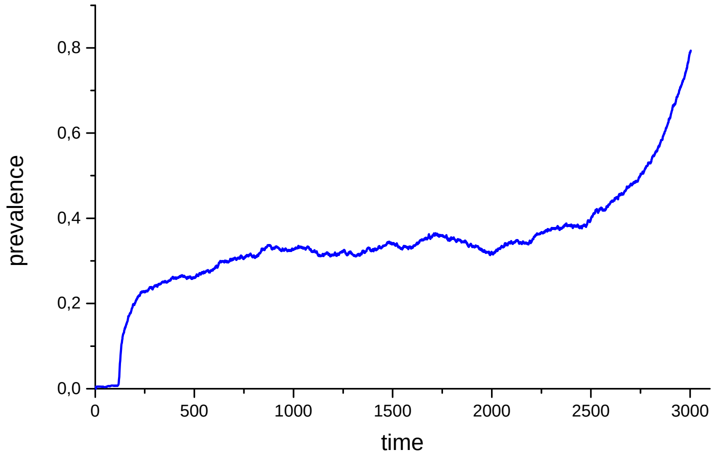
Single run until stagnation 2

Contribution of the 4 (des)infection paths for the “jump”



Single run until stagnation 3

Slow increase of prevalence until collapse

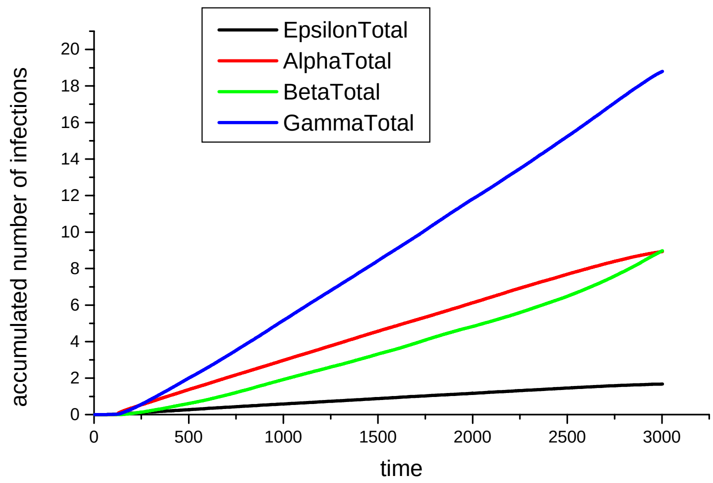


Δ=20 ε=0.001 α=0.20 β=0.04 γ=0.03 b_0=0.005

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

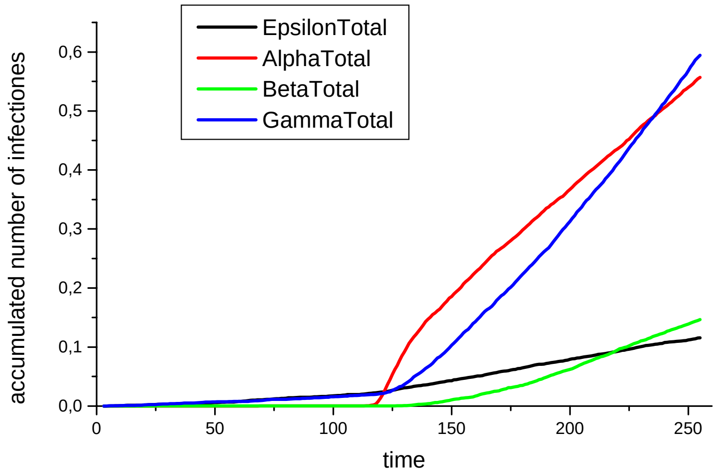
Single run until stagnation 3

Contribution of the 4 (des)infection paths for the run before



Single run until stagnation 3

Zoom for the first jump to the intermediate level



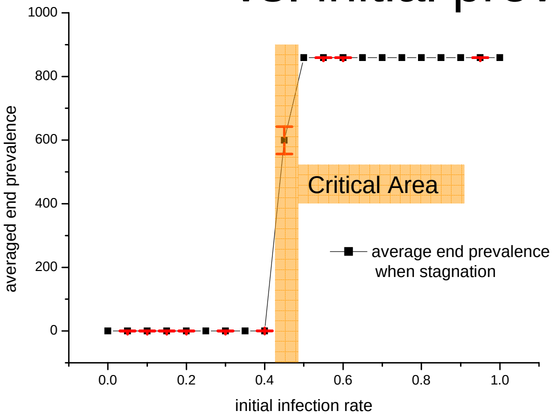
…(end of) single runs

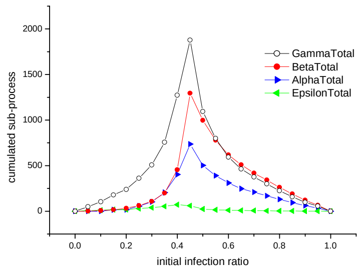
- Possibility for a fatal resonance between local (ε, α) and mean-field process (β)
- Single-Process plots very useful to understand the rich behaviour

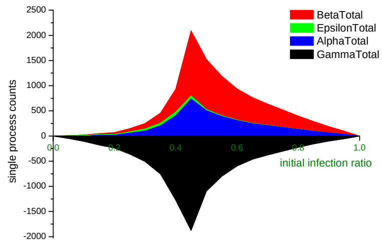
critical density …

- …will now be THE interesting parameter
- How does it change when we vary parameters or graph properties ?

end prevalence vs. initial prevalence



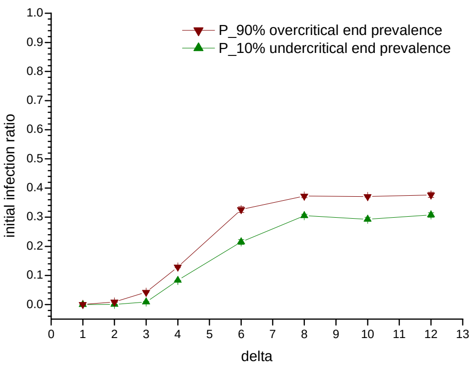




Δ=8 ε=0.001 α=0.20 β=0.08 γ=0.06

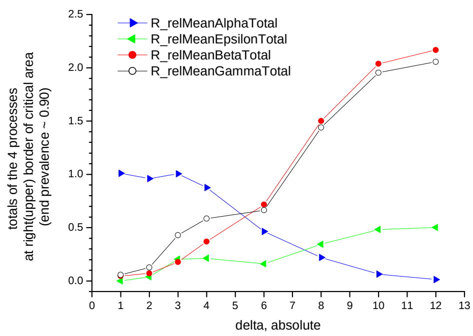
N=859 M=3368 by Random Set Model (O=1000 P=500)
degree: min=2 max=33 mean=7.8
44% nodes have degree ≥ 8

Critical Density over Δ = neighbour infection threshold



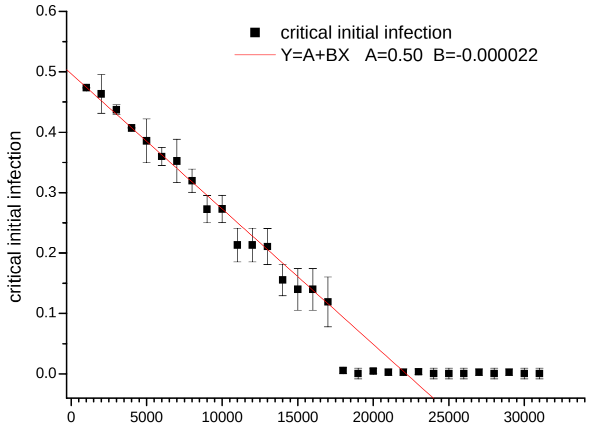
Lower and upper bounds for b_{crit}
as a function of neighbour infection threshold Δ

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

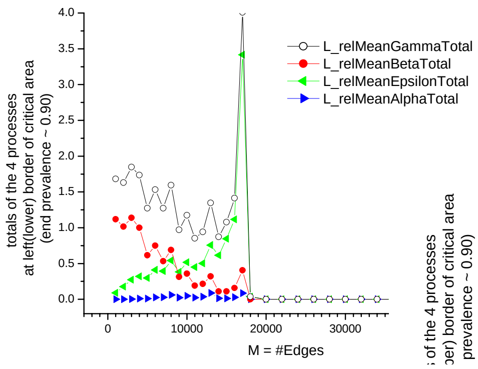
Critical Density over M = #edges

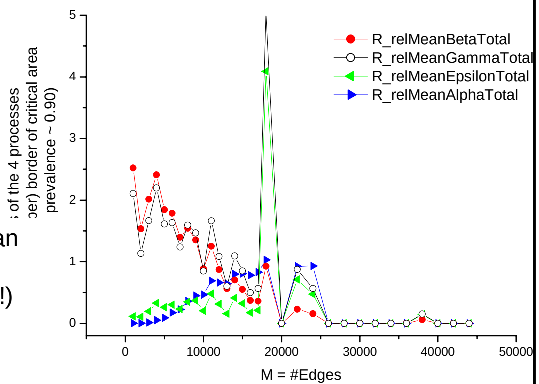


Δ=5 ε=0.005 α=0.35 β=0.08 γ=0.04 (20 runs averages)

Network GNM: N=4000

b_{crit} over M single processes

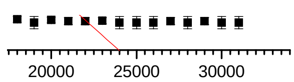




Up to edge density 2 (M=8000, mean degree=4), β-process > α-process, then α –process > β-process (Δ = 5!)

At edge density 4.5 a sharp peak (where ε-process ~= γ-process), in classical terms, this corresponds to reproduction number R_0=1

At the same time collapse of critical initial corruption density



Phase transition with respect to Initial Number of Infected Nodes

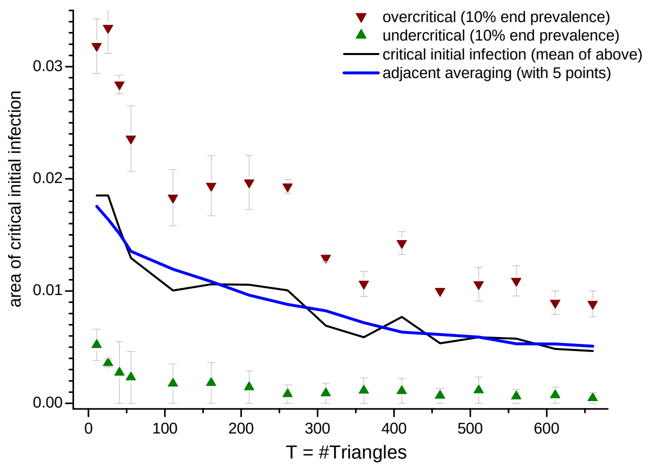
Classical epidemic process is:

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

In corruption infection:

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

Critical Density over T = #triangles



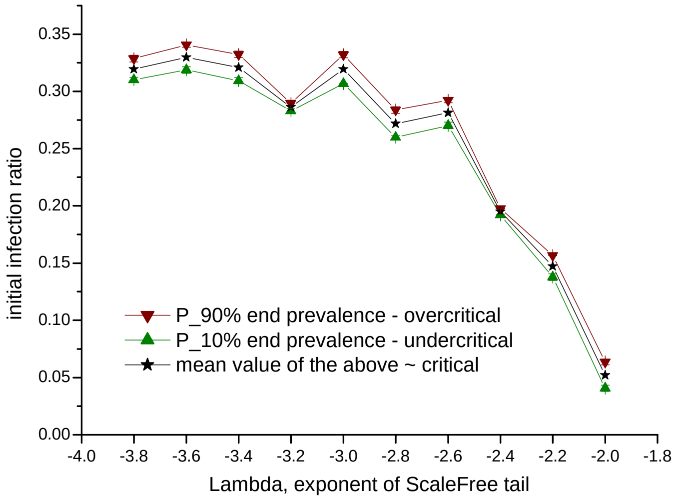
Δ=2 ε=0.005 α=0.30 β=0.08 γ=0.04 (20 runs averages)

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

Critical Density over λ = SF-exponent



$\Delta=5$ $\varepsilon=0$ $\alpha=0.35$ $\beta=0.08$ $\gamma=0.04$

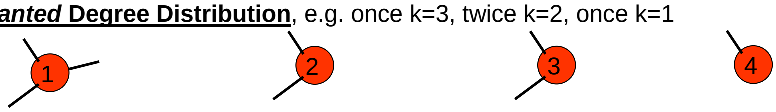
N=20000 M=50000 Network generated by MolloyReed with given Degree-Histogram

Critical Density over λ = SF-exponent

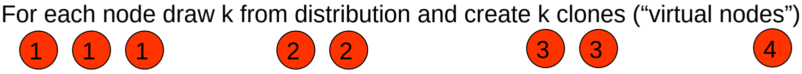
- There seems to be a phase transition around λ=2.4 for Δ=5 (around λ=2.9 for Δ=2)
- not at λ=3 (where structural phase transition, λ<3: expected pathlength finite), probably due to finite size effects (N=20000 only)

Molloy Reed Graph Generator

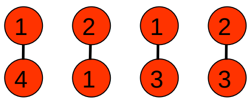
1) ***Wanted* Degree Distribution**, e.g. once k=3, twice k=2, once k=1



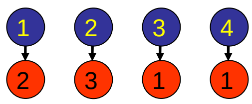
2) For each node draw k from distribution and create k clones (“virtual nodes”)



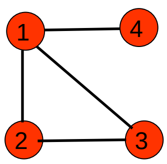
3) Random Pairing

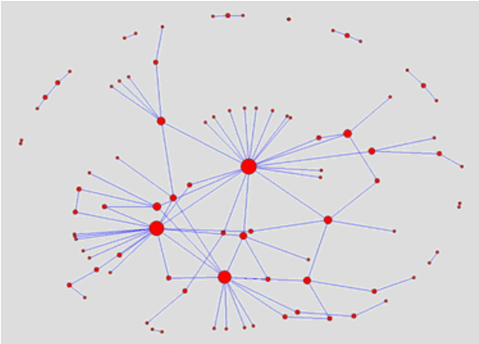


Our modification:
For **adddditive** (instead of **multiplicative**)
DegreeDegree Correlation:
3b) Pairing with ~equal “Out-Degree” for each node



4) Identify again all virtual nodes from same originator
5) Remove double- and self-links
6) The result: A network with the ~wanted degree distribution:





Molloy Reed algorithm

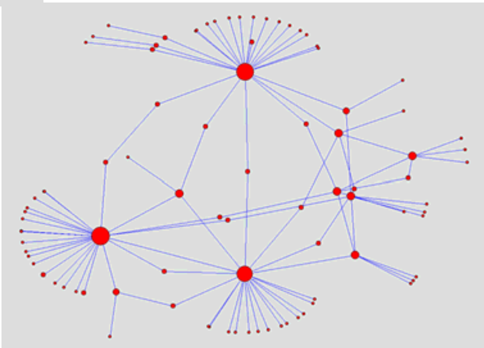
Out-Degree ~ Degree
In-Degree ~ Degree

→ **Multiplicative**
DegreeDegree Correlation

Modified Molloy Reed algorithm

Out-Degree ~ 2M/N
In-Degree ~ (Degree – (2M/N)) ~ Degree

→ **Additive**
DegreeDegree Correlation



Multiplicative DegreeCorrelations: Chains of almost sure linkages from high degree to low degree vertex sets

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

multiplicative DegreeDegreeCorrelation

Why in graphs with such a correlation the threshold b_{crit} → 0 for N→∞ ?

For fixed $b_0 > N^{\frac{1}{\lambda}-\nu}$ and ν>0 the vertices x with $d(x) \geq k_0 >> \Delta/b_0$ get almost surely infected via the α-process (as long as γ< α). Let A_{k0} be the set of such vertices.

A vertex y with $d(y)=k<k_0$ is linked to the set A_{k0} with probability q_k

$$q_k \sim 1 - \prod_{k' \geq k_0}^{k_{\max} \sim N^{\frac{1}{\lambda}}} \left(1 - \frac{const \cdot k \cdot k'}{N}\right)^{const \cdot \frac{N}{(k')^{\lambda}}}$$

$$\sim 1 - e^{-const \cdot \frac{k}{N} \sum_{k' \geq k_0} N \cdot \frac{k'}{(k')^{\lambda}}} \sim 1 - e^{-const \cdot k \frac{1}{k_0^{\lambda-2}}}$$

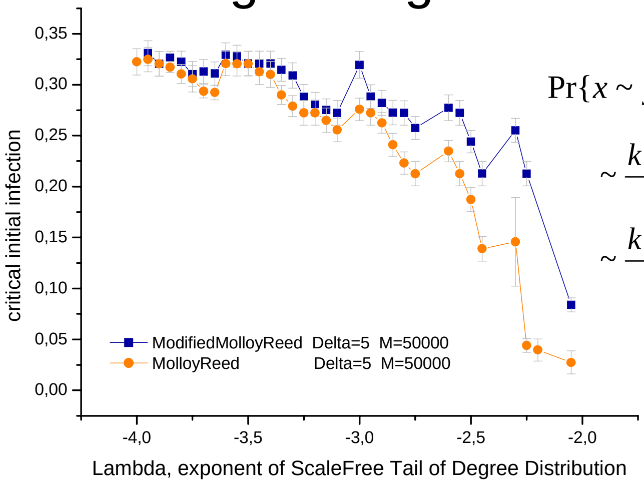
For vertices y with $d(y) > k_0^{\lambda-2}$ one has almost sure linkage to the set A_{k0} (because q_k close to 1). Now infection via α –process…

→ Positive N-independent infection density $b_t >> b_0$ such that the β–process is overcritical, and finally the whole vertex set corrupt

! N has to be large, possible reason for our λ~2.4 instead of λ~3 transition

For SF-graphs with additive degree correlation this argument about chains of almost sure linkages from high degree to low degree cannot be adopted. One therefore expects a higher value of the critical density b_{crit} for additive DegreeCorrelation, so the system is not as susceptible for corruption.

Additive vs. Multiplicative DegreeDegree Correlation



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

$\sim \frac{k \bullet k'}{N}$ multiplicative (normal MolloyReed)

$\sim \frac{k + k'}{N}$ additive (modified MolloyReed)

DegreeDegreeCorrelation

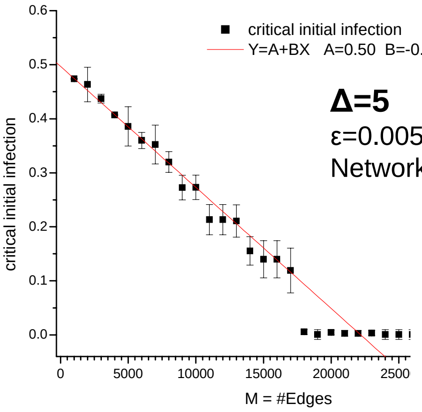
Δ=5 ε=0 α=0.35 β=0.08 γ=0.04

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

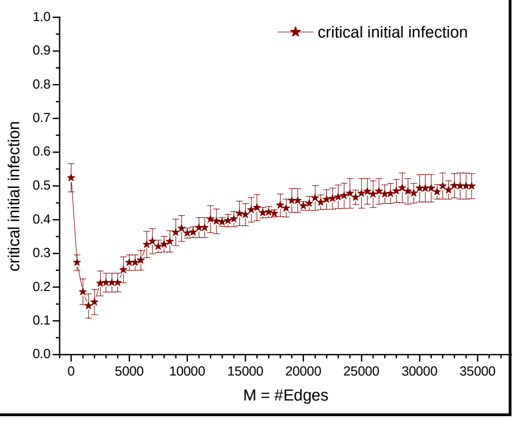
Additive vs. Multiplicative DegreeDegree Correlation

- 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 can exist).

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

Absolute and relative threshold **Δ**

Degree dependent threshold:
Δ(x) = 0.8 * degree(x)

- There is still a critical density, but the value ***increases*** with the edge density
- because the mean threshold increases proportionally to the mean degree

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!

Perspectives

- Faster, faster, faster (bigger systems, esp. SF!)
- Deeper understanding of α-process
  (already non-trivial on trees!)
- Quenched disorder in all parameters
- More topology-dependent processes (like relative Δ=d(x)*0.8)
  - e.g. hubs react differently from leaves
  - cliquish-people react stronger to neighbours ($\alpha\uparrow$ $\beta\downarrow$)
    than lowly connected people who react stronger to mass media ($\alpha\downarrow$ $\beta\uparrow$)
- Corruption state variable not only 0 or 1
- Evolving networks, interaction of process and structure,
  e.g. selection of (non)corrupt neighbours
- Weighted networks
- Degree weighted total corruption prevalence
- Application to other fields (e.g. innovation dynamics)

- Your suggestions?

Thank you for your attention!

The Epidemics of Corruption

Philippe Blanchard
Andreas Krueger
Tyll Krueger
Peter Martin

arxiv:physics/0505031
www.AndreasKrueger.de/networks