Diffusion and Networks: *From simple models to applications*

Ingve Simonsen Department of Physics NTNU, Trondheim, Norway



Part I:

41th Winter School in Theoretical Physics Ladek, 2005

Acknowledgments

Collaborators:

- Halvor Lund, NTNU
- Lubdvig Lizana, New York
- Stefan Bornholdt, Bremen
- Lubos Buzna, Zilina
- Karsten Peters, Dresden
- Dirk Helbing, Zurich
- Kim Sneppen, Copenhagen
- Rafal Weron, Wroclaw
- Sergei Maslov, Brookhaven
- Kasper Eriksen, Copenhagen

Financial Support

- Norwegian Research Council
- EU project IRRIIS
 - IRRIIS : "Integrated Risk Reduction of Informationbased Infrastructure Systems"

ERC

COST

What are we going to discuss?

• Topic of the school:

Simple Models + Complex Systems

• What we will do is:

Random Walk + Complex Networks => Some Application



- A simple model:
 - The Random Walk Model

Applications

- Structures of Complex Networks
- Cascading Failures on Networks
- Epidemology on Networks

Conclusions

Random Walks on Networks

The Model

The Classic Random (RW) Walk Model

- The term RM was coined by Karl Pearson (1905)
- RW is useful in many branches of sciences
- The walker goes up or down with equal probability



 $\sigma(t) \propto \sqrt{t}$



Recommend book:



A Primer on Complex Networks?

• A network consists of a collection of

- Nodes (vertices)
- Links (edges)
- Adjacency matrix : $W = [W_{ij}]$
 - W_{ij} is the weight on link from node j to i (our convention)
 - $W_{ij} = 0$ means <u>no</u> link



Random Walks on Complex Networks

- Random walkers (i.e. particles)"live" on the nodes
- They are moving (flowing) around on the network!
- In each time step, a walker moves towards one of the neighboring nodes chosen by random
- This process is repeated over and over again.....
- Note: The number of walkers is constant in time



The Master Equation for No. of Particle

- <u>Convention</u> : W_{ij} refers to the link from node j to *i*;
- Define the outgoing link weight from node j :

$$w_j = \sum_i W_{ij} \qquad \left(k_j = \sum_i A_{ij}\right)$$

• The change in no. of particle at node *i* from *t* to t+1

$$n_{i}(t+1) - n_{i}(t) = \sum_{j} W_{ij} \frac{n_{j}(t)}{w_{j}} - \sum_{j} W_{ji} \frac{n_{i}(t)}{w_{i}} + n_{i}^{\pm}(t),$$

Or....

$$n_{i}(t+1) = \sum_{j} \frac{W_{ij}}{w_{j}} n_{j}(t) + n_{i}^{\pm}(t)$$
Indee Simonsen

The Master Equation for Node Density

Node Density of walkers at node *i*

$$\rho_i(t) = \frac{n_i(t)}{N}$$

The (master) equation

$$\rho_i(t+1) = \sum_j T_{ij} \rho_j(t) + \rho_i^{\pm}(t), \qquad T_{ij} = \frac{W_{ij}}{W_j}$$

• ... or using matrix-multiplication

$$\vec{\rho}(t+1) = T\vec{\rho}(t) + \vec{\rho}^{\pm}(t)$$

The Master Equation for Link Current

The outgoing link current (per link/weight unit)

$$c_i(t) = \frac{\rho_i(t)}{w_i} = \frac{n_i(t)}{w_i N}$$

Hence, it follows (if the network is undirected)

$$\vec{c}(t+1) = T^{T}\vec{c}(t) + \vec{j}^{\pm}(t);$$
$$T_{ij} = \frac{W_{ij}}{W_{j}} \qquad j_{i}^{\pm}(t) = \frac{\rho_{i}(t)}{W_{i}}$$

Ingve Simonsen

Solution of the Master Equation (assuming $\vec{\rho}^{\pm}(t) = 0$)

- We start from some arbitrary walker dist. : $\rho_i(t=0)$
- Question : What will happen when $t \rightarrow \infty$?
- If T was symmetric, we know the answer:

$$\vec{\rho}(t) = T^t \vec{\rho}(0) = V \Lambda^t V^{-1} \vec{\rho}(0)$$

- but T is NOT symmetric...so what do we do...
- However, T is similar to a symmetric matrix

$$S = KTK^{-1}$$
 with $K_{ij} = \frac{\delta_{ij}}{\sqrt{k_i}} \implies S_{ij} = \frac{\delta_{ij}}{\sqrt{k_i k_j}}$

Solution of the Master Equation (cont...)

- Since S is symmetric, the transfer matrix T has
 - real eigenvalues, $\lambda^{(\alpha)}$ (sorted in decreasing order)
 - due to the conservation of walkers $|\lambda^{(\alpha)}| \le l$ ($\lambda^{(1)}=1$)
 - These eigenvalues will control the walker dynamics via the factor : [λ^(α)]^t
 - Principal eigenvalue : $\lambda^{(1)}=1$ (non-decaying)
 - Slowest decaying mode : $\lambda^{(2)}$ *etc*.

• What does the stationary state look like?

Solution of the Master Equation (cont...)

• What does the stationary state look like?

• Solve : $\rho_{st} = T\rho_{st}$

$$\left[\rho_{st}\right]_{i} = \frac{k_{i}}{\sum_{n} k_{n}} \propto k_{i}$$

$$\rho_{st} = \rho(\infty) = \rho^{(\alpha=1)}$$

- Vertices of high degree will have more walkers in the st. state
- Introduce the outgoing walker current from vertex *i* :

$$c_i^{(\alpha)}(t) = \frac{\rho_i^{(\alpha)}(t)}{k_i}$$

$$c_i^{(1)} = \frac{1}{\sqrt{N}}$$

Ingve Simonsen

Structure of Complex Networks

Application I

K. Eriksen et al., Phys. Rev. Lett. 90, 148701 (2003).



 Idea: The slowest decaying mode will probe the large scale structure of the network



Current Projection Technique

- In the stationary state the currents are all the same
- Idea : For the slowly decaying modes, the currents are similar for nodes belonging to the same module!
- Plot all nodes in current space
- Eg. : d=2 (dimension of the projection)

$$\left|P_{i}^{(d=2)} = (c_{i}^{(2)}, c_{i}^{(3)})\right|$$

Results for Real-World Networks

Two examples:

- Zachary's Karate club Network (friendship network)
 - The university karate club breaks up due to an internal conflict
 - A small network (N=34; L=72)
 - The modular structure is known!
 - Reference :
 - W.W. Zachary, J. Antropol. Res. **33**, 452473 (1977).

Results for Real-World Networks (cont..)

- Course-grained Internet Network (Autonomous Systems)
 - Medium sized network (N=6,474; L=12,572)
 - The modular structure is not well-known in advance
 - Reference :
 - National Laboratory of Applied Network Research
 - http://moat.nlanr.net/AS/



The Zachary Friendship Network

- It is a workbench for community-finding algorithms
- W.W. Zachary (1977) studied a university karate club, and weighted the friendship among its members
 - Nodes : 34
 - Links : 78
- Known community structure:
 - Trainer (node 1)
 - Administrator (node 34)



Source : PNAS 99, 7821 (2002)

The Zachary Friendship Network (cont..)



Ingve Simonsen

Zachary's friendship Network, cont.



[From PNAS 99, 7821 (2002)]

The known modular structure is well reproduced by the diffusion model !

Projection in higher dimensions



The known modular structure is well reproduced by the diffusion model !

Ingve Simonsen

Network Scientists



Ingve Simonsen

Fig. 4. A "network" scientists collaboration network (see Refs. [9,19] for details). (a) The largest component, consists of N = 145 vertices, of this network. Figure after Ref. [9], where colors correspond to communities found there. (b) Same as Fig. 3, but for the network of Fig. 4(a). The optimal partitioning is found for 14 communities characterized by the modularities Q = 0.78 and $Q_A = 0.70$.

Autonomous System Network



We now consider a larger network!

N = 6474L = 12572

Definition: An Autonomous System (AS) is a connected segment of a network consisting of a collection of subnetworks interconnected by a set of routers.

Nodes : AS numbers

Links : Information sharing

Ingve Simonsen

Autonomous System Network (cont.)



The Internet is *indeed* structured The structure follows roughly the national/political structure

The *extreme edges* of the Internet are represented by *Russian and US military* sites !

These country modules *could not* have been detected using spectral analysis of A

Ref : PRE 64, 026704 (2001)

Ingve Simonsen

Summary : Network Topology

 The current projection technique probes large scale structures of the network

References

- K. Eriksen et al., Phys. Rev. Lett. 90, 148701 (2003)
- I. Simonsen, Physica A **357**, 317 (2005)

Cascading Failure of Networks

Application II

I. Simonsen, et al., Phys. Rev. Lett. 100, 218701 (2008).

Motivation

New York, August 14, 2003







Rome, September 28, 2003







Diffusion and Networks

Ingve Simonsen

Motivation



Blackout in parts of the USA and Canada (2003), an impressive example of the long-reaching accompaniments of supply network failures.

Ingve Simonsen

Risk of Power Blackouts



Power Blackouts: Real-Life examples *Europe Nov. 2006: What happened...?*

State of the power grid shortly

before the incident



Sequence of events on November 4, 2006



1,3,4,5 - lines switched off for construction work

2 - line switched off for the transfer of a ship by *Meyer-Werft*

Source : Report on the system incident of November 4, 2006, E.ON Netz GmbH

Ingve Simonsen

Power Blackouts: Real-Life examples

Failure in the continental European electricity grid on November 4, 2006



EU project IRRIIS: E. Liuf (2007) Critical Infrastructure protection, R&D view

Power Blackouts: The Domino Effect (Cascading failure)



"Under certain conditions, a network component shutting down can cause current fluctuations in neighboring segments of the network, though this is unlikely, leading to a cascading failure of a larger section of the network. <u>This may</u> <u>range from a building, to a</u> <u>block, to an entire city, to the</u> <u>entire electrical grid."</u>

Source : Wikipedia

Ingve Simonsen
Cascading Failures Exist in Real Systems

- Examples
 - The power grid
 - Telecommunication networks
 - Transportation systems
 - Computer networks/ the Internet
 - Pipe line systems (water/gas/oil)
- They can be very *costly*
- They typically affect many people

Question : How can one protect (supply) network systems against cascading failures?

A few words on System Design

- The systems are designed with a *given load* in mind
- To ensure stability, the engineering approach, is to introduce some *overcapacity* into the system (security margins)
- ... but overcapacity is *costly*!
- System robustness is often ONLY evaluated locally
- <u>Cascading failure</u>: When an initial perturbation occurs, loads have to redistributes. If the resulting loads exceed the capacities of link/nodes, new failures can result.... "the Domino effect"

Why do we have blackouts....?

System load (throughput)

- optimized to get the maximum out of the system
- high load means small operating margins
- has impact on interactions and component failures
- Tradeoff between load and risk of failure
 - at system level
 - for system components
- What is the role of the deregulation?



Some Terminology

Node Capacity:

$$(1+\alpha)L_i^{(0)}$$

- Load in the stationary state : $L_i^{(0)}$
- Overcapacity (tolerance) : α
- Overload when :

$$L_i(t) > (1 + \alpha) L_i^{(0)}$$

Fraction of nodes remaining (in the Giant Component)

$$G(\alpha) = \frac{N_{GC}^{(\alpha)}(t \to \infty)}{N_{Diffusion and Networks}}$$

Ingve Simonsen

Previous physics works : Cascading Failures Motter and Lai: PRE 66, 065102R (2002)

Load calculated as betweeness centrality

 Overload checked only for the stationary state

No sinks/sources



Ingve Simonsen

Previous works : Summary/Open Questions

- Previous works of cascading failures exclusively considered the <u>stationary state</u>
- But...why should the system *not* experience additional failures due to overloading during the transient period?
- Question to address:
 - What is the role played by the dynamics in cascading failures on complex networks
- A dynamical model is needed for such a study
 - But which one to choose?

Expected difference between a static and a dynamic model for flow redistribution Dynamic model Stationary model Initial failure t = 1 t = 2

Model : Requirements

• It should be:

- Generic : no particular physical process is addressed
- As simple as possible, but not simpler...
- Important ingredient (in our opinion)
 - The flowing quantity should be <u>CONSERVED</u>



Our solution : A Random Walk (or Diffusion type) model !

The Master Equation

- Our simple dynamical model incorporates:
 - Flow conservation
 - Network topology
 - Load redistribution

$$\vec{c}(t+1) = T^T \vec{c}(t) + \vec{j}^{\pm}(t); \qquad T_{ij} = \frac{W_{ij}}{W_j}$$

$$c_i(t) : \text{The outgoing current from node } i \text{ per link weight unit}$$

Model Dynamics: Is it realistic?



EUROSTAG power simulator : wwwaurostag.epfl.ch



• Link current on the link from node *j* to *i*

$$C_{ij}(t) = W_{ij}c_j(t)$$

Loads (on the same link)

$$L_{ij}(t) = C_{ij}(t) + C_{ji}(t) = W_{ij}c_{j}(t) + W_{ij}c_{i}(t)$$

Stationary Solution

Equation

$$(1-T^T)\vec{c}(\infty)=\vec{j}^{\pm};$$

Solution

$$\vec{c}(\infty) = \vec{c}^{(0)}(\infty) + (1 - T^T)^+ \vec{j}^{\pm}$$

where

Homogeneous solution $c_i^{(0)}(\infty) = 1/N$ $(1-T^T)^+$ is the generalized inverse of $1-T^T$

• Link capacities
$$(1+\alpha)L_{ij}^{(0)}$$

Model Dynamics: UK high voltage power grid (300-400kV)



At t=0, link 0 is broken!



When does a link/node fail?

 Link/node capacities relative to the <u>undisturbed</u> state (L_{ij}) via a *tolerance parameter* α

 $(1+\alpha)L_{ij}^{(0)}$

 A link/node fails whenever its *current* load, L_{ij}(t) exceeds the capacity of *that* link/node





Main steps of the simulations

The simulations consist of the following steps:

- 1. A *triggering event* (t=0) [remove a random link]
- 2. Calculate the link loads $L_{ij}(t)$
- 3. Check if any links are *overloaded* via $L_{ij}(t) > (1+\alpha)L_{ij}^{(0)}$
 - 1. If so remove such overloaded links
- 4. Repeat step 2 and 3 till no more links are overloaded
- 5. Average the results over the triggering event of pnt. 1 (and source and slinks locations)

Stationary Model vs. Dynamic Model : The northwestern US power transmission grid

 $|\mathcal{N}|$

number of nodes (5000)



 $|\mathcal{L}_R|$

number of links



number of remaining links

The stationary model can <u>overestimates robustness</u> by more than 80% (in this case)

Overload exposure times may be relevant and will increase the robustness.....



Ingve Simonsen

Stationary Model vs. Dynamic Model : The role of the two time-scales

• There are two characteristic time-scales in the problem:

- Overload exposure time (protection system response time): τ
- Typical transient time for the dynamics: τ_0
- Control parameter : $\chi = \frac{\tau}{\tau_0}$
 - Static cascading failure model: $\chi >>1$
 - Dynamical ($\tau=0$) cascading failure model : $\chi=0$

• The real situation is probably somewhere in between....

Summary: Cascading Failure



- *The dynamical* process on the network <u>may be</u> important to consider when evaluating network robustness (cascading)
 - Using a stationary model may dramatically overestimate (by 80-95%) the robustness of the underlying network
 - The actual overestimation does depend on the actual overload exposure time
- In a dynamical model:
 - links may fail that otherwise would not have done so (overshooting)
 - The proximity to a disturbance is more important in a dynamical model





- Dynamical model :
 - I. Simonsen, L. Buzna, K. Peters, S. Bornholdt, D. Helbing, Phys. Rev. Lett. 100, 218701 (2008).
- See also : Phys. Rev. Lett. 90, 14870 (2003).
 Physica A 357, 317 (2005).
- Stationary models:
 - Motter and Lai, Phys. Rev. E 66, 065102R (2002).
 - Bakke *et al.* Europhys. Lett. **76**, 717 (2006).

Ingve Simonsen

Epidemology on Networks

Application III

Colizza et al., Nature Physics 3, 276 (2007).

Lund et al. Work in progress (2010)

Epidemiology: The Questions

• What is the ratio of infected people in a population?

- Does the ratio remain small or is it close to unity
- What is the steady state behavior

• Other important questions:

- How fast does an infection spread?
- Can we stop it and how?

Reaction-Diffusion on Networks



(Diffusion constants)

- More then one type of particle A, B, ...
- Two processes
 - Reactions <u>on</u> nodes
 - Diffusion <u>between</u> nodes

The SIS-Model

The SIS Model

- Individuals can become infected with the decease and recover from it with no immunity
- Two types of individuals: <u>Susceptible</u> (A) and <u>Infected</u> (B)

Reactions



В

SIS-Reactions Only
No diffusion
$$A + B \rightarrow 2B$$

 $B \rightarrow A$

• No. of A and B-particles at node *i* at time *t*

$$\partial_t N_A(i,t) = \mu N_B(i,t) - \beta \Gamma_V(i,t) = -\partial_t N_B(i,t)$$

Reaction Kernels

$$\Gamma_{v}(i,t) = \frac{N_{A}(i,t)N_{B}(i,t);}{N(i)}; \quad v = 1 \qquad \text{Type I} \quad \text{``the village''}$$
$$\Gamma_{v}(i,t) = \frac{N_{A}(i,t)N_{B}(i,t)}{N(i)}; \quad v = 2 \qquad \text{Type II} \quad \text{``the city''}$$
$$\frac{N_{A}(i,t)N_{B}(i,t)}{N_{\times} + N(i,t)}; \quad v = 3 \qquad \text{Type M (for Mixed)}$$

Ingve Simonsen

Reactions-Diffusion on Networks

- Each time step consists of two processes
 - SIS-reactions on a node
 - Diffusion to neighboring nodes

$$\vec{N}_{A}(t+1) = \overbrace{(1-D_{A})\vec{N}_{A}(t) + D_{A}T\vec{N}_{A}(t)}^{\text{Diffusion const. } D_{A}}$$
$$+ \underbrace{\mu \vec{N}_{B}(t) - \beta \vec{\Gamma}_{v}(t)}_{\text{Reaction}}$$
$$\vec{N}_{B}(t+1) = (1-D_{B})\vec{N}_{B}(t) + D_{B}T\vec{N}_{B}(t)$$
$$- \mu \vec{N}_{B}(t) + \beta \vec{\Gamma}_{v}(t)$$

The Stationary State for Infected "particles"

Quantities:

$$\rho = \frac{N}{V}; \qquad \rho_B(t) = \frac{N_B(t)}{V}; \qquad \frac{\rho_B(t)}{\rho} = \frac{N_B(t)}{N}$$

Population ratio of infected in the stationary state

$$\frac{\overline{\rho}_B}{\rho} = \lim_{t \to \infty} \frac{N_B(t)}{N}$$



- Cases
 - DA=0; DB=0; (not so interesting)
 - DA=1; DB=0; (not so interesting)
 - DA=0; DB=1
 - **D**A=1; **D**B=1
- Assumptions
 - Mean field approximation
 - Uncorrelated network

Phase TransitionsAnalytic Mean Field Results(Colizza et al. 2007)

• Type I : "the village"

$$\overline{\rho}_{B} = \begin{cases} \rho - \rho_{c} & \rho \ge \rho_{c} \\ 0 & \rho < \rho_{c} \end{cases}$$

$$\rho_{c} = \begin{cases} \frac{\mu}{\beta} & D_{A} = 0\\ \frac{\mu}{\beta} \frac{\langle k \rangle^{2}}{\langle k^{2} \rangle} & D_{A} = 1 \end{cases}$$

 $D_{A} \in \{0,1\}$

• Type II: "the city"

$$\overline{\rho}_{B} = \begin{cases} \rho \left(1 - \frac{\mu}{\beta} \right) & \mu < \beta \\ 0 & \mu \ge \beta \end{cases}$$

Analytic Mean Field Results (Lund et al. 2010)

Phase Transition

$$\overline{\rho}_{B} = \begin{cases} \rho - \rho_{c} & \rho \geq \rho_{c} \\ 0 & \rho < \rho_{c} \end{cases}$$

where the critical nodal density is $(D_B=1)$



Inave Simonsen

Analytic Results



Figure 4.1: Analytical expressions for stationary states of type I, II and M ($\rho_x = 1$) with diffusing B particles ($D_B = 1$), from Eqs. (4.14), (4.16), (4.18), (4.20), (4.21) and (4.22). Solid lines are type I, dashed lines are type II and dash-dot lines are type M. Circles indicate $D_A = 0$, cross (+) indicates $D_A = 1$. The network is scale-free with V = 1000 nodes and $\gamma = 2.5$.

Ingve Simonsen

Diffusion and Networks

 $\frac{\beta}{2} = 2$

μ

Numerical Results : Type I "the village"



lation results read directly from figure in Ref. [8], analytical results superimposed by us.



Figure 4.2: Comparison of (a) Colizza's [8] and (b) Lund's [2] simulation results for type I reactions, with $\mu/\beta = 2$ and $D_B = 1$. Lines indicate analytical results (Eqs. (4.14) and (4.20)), symbols are simulation results. Each symbol with properties (γ, V, D_A) : Circles \diamond (2.5, 10⁵, 1), squares \Box (2.5, 10⁴, 1), diamonds \diamond (2.5, 10³, 1), downward triangle \bigtriangledown (2.5, 10⁴, 0), upward triangle \triangle (2.5, 10⁵, 0), leftward triangle \triangleleft (3.0, 10⁴, 0).

Ingve Simonsen

Numerical Results : Type II "the city"



(a) Colizza (stochastic, μ and β unkown). Simulation results read directly from figure in Ref. [8], analytical results superimposed by us.



Figure 4.3: Comparison of (a) Colizza's and (b) Lund's simulation results for type II reactions, with $\rho = 20$ and $D_B = 1$. Lines indicate analytical result (Eq. (4.21)), symbols are simulation results. Each symbol with properties (γ, V, D_A) : Circles \bigcirc (2.5, 10⁴, 1), squares \square (3.0, 10⁴, 1), diamonds \diamondsuit (2.5, 10⁴, 0) and triangles \bigtriangledown (3.0, 10⁴, 0).

Ingve Simonsen

Type M : The Mixed Reaction



Figure 4.6: Simulation results for type M reactions, with $D_B = 1$ and $\beta/\mu = 2$. Lines indicate analytical result (Eqs. (4.18) and (4.22)), symbols are simulation results. Each symbol with properties (V, D_A) : Circles \circ (10⁴, 1), squares \Box (10², 1), diamonds \diamond (10⁴, 0) and triangles ∇ (10², 0).

Ingve Simonsen

Diffusion and Networks

71

Summary: Epidemology on Networks

Reaction-Diffusion can be used to study epidemic spreading on complex networks

Phase transitions will occur depending on

- Diffusion constants
- Reaction rates
- Network properties
- Nodal densities





- Random walk (RW) models are simple and powerful
- Combined with complex networks (CN) they shine
- The combination RW+CN shown useful for
 - Networks topology studies
 - Cascading failures
 - Epidemiology studies
Take home message

A model is almost never too simple to be useful!

Thank you for your attention!





- Networks topology
 - K. Eriksen, I Simonsen., S Maslov, and K. Sneppen Phys. Rev. Lett. 90, 148701 (2003).
- Cascading failures
 - I. Simonsen, L. Buzna, K. Peters, S. Bornholdt, D. Helbing, Phys. Rev. Lett. 100, 218701 (2008).
- Epidemiology on networks
 - Colizza et al., Nature Physics **3**, 276 (2007)
 - Lund et al. In preparation (2011)

The End!