# Pricing of Cyber Insurance Contracts in a Network Model

**Stefan Weber**

Leibniz Universität Hannover

www.stochastik.uni-hannover.de

(joint work with Matthias Fahrenwaldt & Kerstin Weske)

WU Wien – December 13, 2017

# Motivation

- Cyber risks pose a large threat to businesses and governments
- Estimated global loss per year $\approx$ 400 billion USD[1]
- **Dimensions of cyber risk**
  - **Causes**: Human errors; technical failures; insider/hacker attacks
  - **Damage**: Lost, stolen or corrupted data; damage to firms' or governments' operations, property and reputation; severe disruption of critical infrastructure; physical damage, injury to people and fatalities
  - **Risk assessment**: Analysis of critical scenarios; stochastic cyber model and statistical evaluation
  - **Mitigation**: Modify system technology; develop emergency plan; insurance solutions

[1] Center for Strategic & International Studies (2014)/ Llloyds of London CEO Inga Beale (2015)

# Motivation (2)

- **Actuarial challenges of cyber risk**
  - **Data**:
    Data is not available in the required amount or in the desired granularity
  - **Non-stationarity**:
    Technology and cyber threats are evolving fast
  - **Accumulation risks**:
    The typical insurance independence assumption does not hold, but there is no simple geographical distinction between dependent groups as, for example, in the case of NatCat

# Motivation (3)

- We consider the special case of infectious cyber threats, e.g., viruses and worms
- **Example**:
  *WannaCry* infected more than 230.000 computers in 150 countries in May 2017
- **Our main contribution**
  A mathematical model for infectious cyber threats and cyber insurance
  - Stochastic model based on IPS and marked point processes
  - We suggest higher-order mean-field approximations
  - Insurance application: premiums can be calculated
  - Systemic risk: we analyze the influence of the network structure

# Model Idea

- **Infection spread process:**
  - Agents are connected in a network
  - Infections spread from neighbor to neighbor and are cured independently

  $\rightarrow$ Continuous time Markov process, i.e., SIS/contact process

- **Insurance claims processes:**
  - Infected nodes are vulnerable to cyber attacks that occur at random times and generate losses of random size

  $\rightarrow$ Marked point process

- A (re-)insurance company covers a function of the nodes' losses

# Outline

# Outline

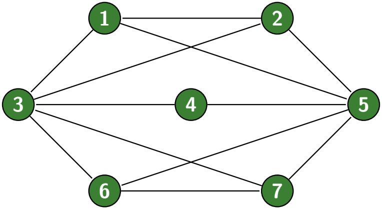
# Network of Agents

- $N$ interconnected agents, labeled $1, 2, \dots, N$
  (e.g., corporations, systems of computers, or single devices)

- **Connections:** Network without self-loops, represented by a (symmetric) adjacency matrix $A \in \{0,1\}^{N \times N}$ $(a_{ii} = 0)$
  - $a_{ij} = 1$: connection between node $i$ and $j$,
  - $a_{ij} = 0$: $i$ and $j$ are not directly connected

- **Example:**

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}$$



# Spread Process (1)

- **SIS-model** (Susceptible-Infected-Susceptible)

  At each point in time, node $i$ can be in one of **two states** $X_i(t) \in \{0, 1\}$:

  - $X_i(t) = 1$: node $i$ is infected = **vulnerable** to cyber attacks,
  - $X_i(t) = 0$: node $i$ is susceptible at time $t$

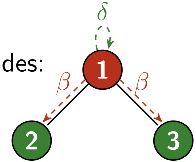
- Each node changes its state at a random time with a rate that may depend on the states of other nodes

  Key parameters:

  - $\beta > 0$ (infection rate),
  - $\delta > 0$ (curing rate)

  Nodes are infected by their infected neighbors, and infected nodes are cured independently from other nodes:

  - $X_i : 0 \to 1;\ \beta \sum_{j=1}^{N} a_{ij} X_j(t)$ (Infection),
  - $X_i : 1 \to 0;\ \delta$ (Curing)



# Spread Process (2)

## Definition

The **spread process** $X$ is a Feller process on the configuration space $E = \{0,1\}^N$ defined by the generator $G : C(E) \to \mathbb{R}$ with

$$Gf(x) = \sum_{i=1}^{N} \left( \beta(1-x_i) \sum_{j=1}^{N} a_{ij} x_j + x_i \delta \right) (f(x^i) - f(x)), \quad x \in E, \ f \in C(E),$$

where $x_j^i = x_j$ for $i \neq j$ and $x_i^i = 1 - x_i$

# Outline

# Claims Process

- **Mechanism**
  - The spread process $X$ does not directly cause any damage
  - The system as a whole is subject to randomly occurring cyber attacks
  - A node is affected by a cyber attack at time $t$ if and only if it is infected = **vulnerable** at time $t$

- **Mathematical Model**
  - Number of attacks: counting process $M = (M(t))_{t \geq 0}$
    - ... with stochastic intensity $(\lambda(t))_{t \geq 0}$
    - ... independent of $X$
  - Loss sizes: nonnegative process $L = (L(t))_{t \geq 0}$
    - ... independent of $X$
    - ... with $L(t) = (L_1(t), \ldots, L_N(t))^\top$
    - Losses of an attack at time $t$ are captured by:

$$L(t) \circ X(t) = (L_1(t)X_1(t), \ldots, L_N(t)X_N(t))^\top$$

# Expected Aggregate Losses

- For any time $t$, the insurance contract is characterized by a function $f(\cdot;\cdot) : \mathbb{R}_+ \times \mathbb{R}_+^N \to \mathbb{R}_+$:
- The insurance company covers $f(t; L(t) \circ X(t))$, if a loss event occurs at time $t$

→ The expected aggregate losses of the insurance company over time window $[0, T]$ are given by:

$$\mathbb{E}\left[\int_0^T f(t; L(t) \circ X(t)) dM(t)\right] = \mathbb{E}\left[\int_0^T f(t; L(t) \circ X(t)) \lambda(t) dt\right] \quad (1)$$

- **Question:** Explicit calculation?

# Example: Proportional Insurance

Let $f$ describe a proportional insurance contract, i.e.,

$$f(t; L(t) \circ X(t)) = \sum_{i=1}^{N} \alpha_i L_i(t) X_i(t)$$

In this case, eq. (1) becomes

$$\begin{aligned} \mathbb{E}\left[\int_0^T f(t; L(t) \circ X(t)) dM(t)\right] &= \mathbb{E}\left[\int_0^T f(t; L(t) \circ X(t)) \lambda(t) dt\right] \\ &= \int_0^T \sum_{i=1}^{N} \alpha_i \cdot \mathbb{E}[X_i(t)] \cdot \mathbb{E}[L_i(t)\lambda(t)]\, dt \end{aligned}$$

$\rightarrow$ For **linear claim functions**, only the first moments $\mathbb{E}[X_i(t)]$ of the spread process are needed in order to calculate the expected aggregate losses

# General Claims

- **Non-linear claim functions** $f$ can be uniformly approximated by polynomials of a chosen degree $n_p$ in probability
- **Basic idea:**
  - By the theorem of **Stone-Weierstraß**, any continuous $f$ can be uniformly approximated by polynomials on any compact set
  - The compact set is chosen such that the probability of the argument being outside the compact is sufficiently small

This leads to expressions of the following form:

$$\int_0^T \mathbb{E}\Bigg( \mathbb{1}_{[0,u]}(\Lambda(L)) \cdot \lambda(t) \cdot \sum_{i=1}^N \Bigg[ a_0 + a_1 \sum_{i_1=1}^N b_{i_1} L_{i_1} \mathbb{E}[X_{i_1}] + a_2 \sum_{i_1=1}^N \sum_{i_2=1}^N b_{i_1} b_{i_2} L_{i_1} L_{i_2} \mathbb{E}[X_{i_1} X_{i_2}]$$

$$+ \ldots + a_{n_p} \sum_{i_1=1}^N \sum_{i_2=1}^N \cdots \sum_{i_{n_p}=1}^N b_{i_1} b_{i_2} \cdots b_{i_{n_p}} \cdot L_{i_1} L_{i_2} \cdots L_{i_{n_p}} \cdot \mathbb{E}[X_{i_1} X_{i_2} \cdots X_{i_{n_p}}] \Bigg] \Bigg) \, dt$$

→ Only moments up to order $n_p$ of the spread process (i.e., $\mathbb{E}[X_{i_1}(t) \cdots X_{i_k}(t)]$ for $i_j \in \{1, \ldots, N\}$ and $k \leq n_p$) are required for the computation of the expected aggregate losses

# General Claims (2)

For both linear and non-linear claim functions:

- **Key issue** when computing the expected aggregate losses:
  - Calculate moments of $X$
  - Due to Kolmogorov's equations, these are characterized by ODE systems
- **Challenge:**
  - Direct calculation of moments is hardly tractable for realistic network sizes due to very large ODE systems
- **Suggestion**

  Mean-field approximation of the moments of the spread process

# Outline

# First Order Mean-Field Approximation (1)

- ODEs of time-derivatives of first moments $\mathbb{E}[X_i(t)]$:

$$\frac{d\mathbb{E}[X_i(t)]}{dt} = -\delta\mathbb{E}[X_i(t)] + \beta\sum_{j=1}^{N} a_{ij}\mathbb{E}[X_j(t)] - \beta\sum_{j=1}^{N} a_{ij}\mathbb{E}[X_i(t)X_j(t)], \quad i = 1, 2, \ldots, N$$

- **Problem:** Joint second moments keep the system from being closed
- **Ansatz:**
  Incorrectly factorize the second moments

$$\mathbb{E}[X_i(t)X_j(t)] \approx F(\mathbb{E}[X_i(t)]) \cdot F(\mathbb{E}[X_j(t)])$$

  with a suitably chosen function $F : [0, 1] \to [0, 1]$, e.g., $F(x) = x$

# First Order Mean-Field Approximation (2)

## Definition

The first order mean-field approximation $z_i^{(1)}$ corresponding to the mean-field function $F$ is defined as the solution to the following system of ODEs:

$$\frac{dz_i^{(1)}(t)}{dt} = -\delta z_i^{(1)}(t) + \beta \sum_{j=1}^{N} a_{ij} z_i^{(1)}(t) - \beta \sum_{j=1}^{N} a_{ij} F(z_i^{(1)}(t)) \cdot F(z_j^{(1)}(t)),$$

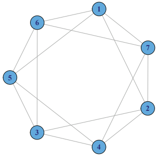
for $i = 1, \ldots, N$

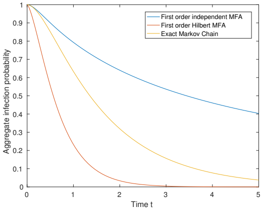
- The choice of $F(x) = x$ leads to an upper bound, the choice of $F(x) = \sqrt{x}$ to a lower bound approximation of the exact moment
- For certain parameter choices, the approximation error decreases exponentially in time

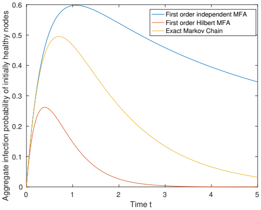
# First Order Mean-Field Approximation (3)

- The accuracy of first order mean-field approximations is typically low, if interaction is sufficiently strong
- **Example:**
  We consider a regular network with $N = 7$ nodes and degree $d = 4$

$$A := \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$



# $n$-th Order Mean-Field Approximation (1)

- In order to achieve higher accuracy, we extend this idea and construct mean-field approximations of order $n$: $(z_I^{(n)})_{I \subseteq \{1,2,\ldots,N\},\ |I| \leq n}$
- This increases the complexity of the approximation
- **Methodology**
  - Define the product $X_I := \prod_{i \in I} X_i$ for $I \subseteq \{1, 2, \ldots, N\}$.
    Since the components of $X$ are commutative and idempotent, we may neglect the order of the indices or powers of its components
  - As a consequence of Kolmogorov's forward equations, the dynamics of the moments $(E[X_I])_{I \subseteq \{1,2,\ldots,N\}}$ are described by a coupled system of $2^N - 1$ ODEs
  - **Approximation**
    Focus only on $(E[X_I])_{I \subseteq \{1,2,\ldots,N\},\ |I| \leq n}$
    1. $|I| < n$:
       ODE for $\frac{d}{dt} z_I^{(n)}$ is exact ODE for $\frac{d}{dt} E[X_I]$
    2. $|I| = n$:
       ODE for $\frac{d}{dt} z_I^{(n)}$ is approximation obtained by (incorrectly) factorizing moments of order $n + 1$

# $n$-th Order Mean-Field Approximation (2)

## $|I| = n$

- Choose the following two objects:
  1. a *mean-field function* $F : [0,1] \to [0,1]$ and
  2. a *partition scheme* $(I_1, I_2)$ such that for $j \notin I$ we have $I \cup \{j\} = I_1(I,j) \cup I_2(I,j)$ with non-empty $I_1(j) = I_1(I,j), I_2(j) = I_2(I,j)$
- This leads to the following approximation:

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}[X_I] &= -n\delta\mathbb{E}[X_I] + \beta\sum_{i\in I}\sum_{j=1}^{N} a_{ij}\mathbb{E}[X_{I\setminus\{i\}\cup\{j\}}] - \beta\sum_{i\in I}\sum_{j=1}^{N} a_{ij}\mathbb{E}[X_{I\cup\{j\}}] \\
&\approx -n\delta\mathbb{E}[X_I] + \beta\sum_{i\in I}\sum_{j=1}^{N} a_{ij}\mathbb{E}\left[X_{I\setminus\{i\}\cup\{j\}}\right] - \beta\sum_{i\in I}\sum_{j=1, j\in I}^{N} a_{ij}\mathbb{E}[X_I] \\
&\quad -\beta\sum_{i\in I}\sum_{j=1, j\notin I}^{N} a_{ij}\cdot F\left(\mathbb{E}[X_{I_1(j)}]\right)\cdot F\left(\mathbb{E}[X_{I_2(j)}]\right).
\end{aligned}$$

# $n$-th Order Mean-Field Approximation (3)

## $|I| < n$

- In the approximate ODE system, the ODE for $\frac{d}{dt} z_I^{(n)}$ is the exact ODE for $\frac{d}{dt} E[X_I]$:

$$\frac{d}{dt}\mathbb{E}[X_I] \quad = \quad -n\delta\mathbb{E}[X_I] + \beta \sum_{i \in I} \sum_{j=1}^{N} a_{ij} \mathbb{E}\left[X_{I \setminus \{i\} \cup \{j\}}\right] - \beta \sum_{i \in I} \sum_{j=1}^{N} a_{ij} \mathbb{E}\left[X_{I \cup \{j\}}\right]$$

$\longrightarrow$ **n-th order approximation** with

$$|I| = n: \quad \dot{z}_I^{(n)} = -\left(n\delta + \beta \sum_{i \in I} \sum_{j=1, j \in I}^{N} a_{ij}\right) z_I^{(n)} + \beta \sum_{i \in I} \sum_{j=1}^{N} a_{ij} z_{I \setminus \{i\} \cup \{j\}}^{(n)}$$

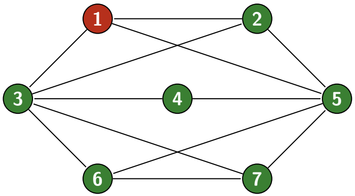
$$- \beta \sum_{i \in I} \sum_{j=1, j \notin I}^{N} a_{ij} F\left(z_{I_1(j)}^{(n)}\right) \cdot F\left(z_{I_2(j)}^{(n)}\right)$$

$$|I| < n: \quad \dot{z}_I^{(n)} = -\, n\delta z_I^{(n)} + \beta \sum_{i \in I} \sum_{j=1}^{N} a_{ij} z_{I \setminus \{i\} \cup \{j\}}^{(n)} - \beta \sum_{i \in I} \sum_{j=1}^{N} a_{ij} z_{I \cup \{j\}}^{(n)}$$

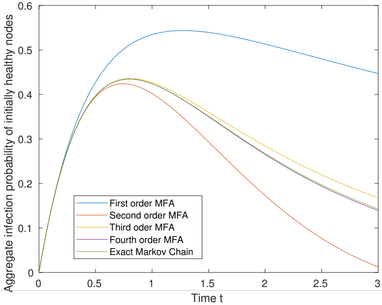
# $n$-th Order Mean-Field Approximation (4)

- The $n$-th order mean-field approximation yields approximations of all moments of $X$ up to order $n$:
  - $n$-th moments enable us to compute expected aggregate losses for non-linear claim functions
  - The $n$-th order approximation also yields improved approximations of the first order moments, i.e., infection probabilities of each node

**Example:** Aggregate infection probability of initially healthy nodes in the $n$-th order mean-field approximation for $n = 1, 2, 3, 4$, $F(x) = x$, $\beta = 0.5$ and $\delta = 1.817$




Initial state of the infection

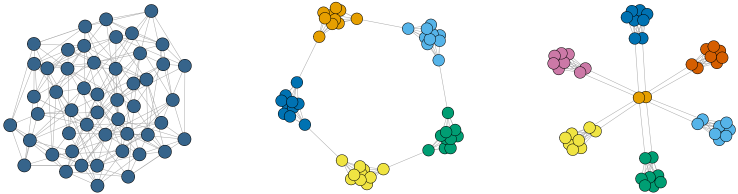


# Outline

# Network Scenarios

- We consider three different stylized network scenarios

**Homogeneous** | **Clustered** | **Star-shaped**

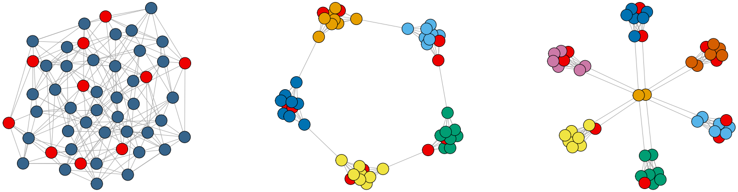


- The number of nodes and the degree of each node are equal in all scenarios ($N = 50, d = 7$)

→ We are comparing the impact of the network topology

# Simulation Setup

- We initially infect 20% of the nodes in the networks:



- For the spread process, we choose: $\beta = 0.5$, $\delta = 3.51$
- Cyber attacks occur at the jumps of a homogeneous Poisson process with rate $\lambda = 3$
- Losses at each vulnerable node are exponentially distributed with mean $\mu = 2$
- Approximation of expected aggregate losses of the insurance company in $[0, 3]$ on the basis of
  - mean-field approximations for the moments of the spread process,
  - Monte-Carlo simulations of the claims processes

# Example: Aggregate Losses

Total loss coverage, i.e., the treaty function $f(t,\cdot)$ is given by

$$f(t, L(t) \circ X(t)) := \sum_{i=1}^{N} L_i(t) X_i(t)$$

→ Estimated expected aggregate losses:

| *Losses: Total coverage* | Homogeneous | Clustered | Star |
|---|---|---|---|
| First order MFA | 96.4671 | 97.6170 | 96.5425 |
| Second order MFA | 51.4911 | 39.7776 | 39.4127 |
| Third order MFA | 77.8349 | 70.6588 | 68.0767 |
| **Fourth order MFA** | 68.0676 | 61.3693 | 59.9005 |

# Example: Excess of Loss per Risk – XL

XL, i.e., the treaty function $f(t,\cdot)$ is given by

$$f(t, L(t) \circ X(t)) := \sum_{i=1}^{N} \min\{L_i(t), 2\} \cdot X_i(t)$$

→ Estimated expected insurance losses:

| *Losses: XL* | Homogeneous | Clustered | Star |
|---|---|---|---|
| First order MFA | 60.9795 | 61.7036 | 61.0247 |
| Second order MFA | 32.5475 | 25.1401 | 24.9105 |
| Third order MFA | 49.2010 | 44.6618 | 43.0300 |
| **Fourth order MFA** | 43.0265 | 38.7894 | 37.8615 |

# Outline

# Conclusion

- Model for pricing cyber insurance
- Cyber losses that are triggered by two underlying risk processes:
  - a cyber infection $\leftrightarrow$ interacting Markov chain
  - cyber attacks on vulnerable sites $\leftrightarrow$ marked point process
- Due to the large dimension of the system, the computation of expected aggregate insurance losses and pricing of cyber contracts is challenging:
  - polynomial approximation of non-linear claim functions
  - $n$-th order mean-field approximation of moments of the spread process
- Numerical case studies demonstrate:
  - Significant impact of network topology
  - Higher order mean-field approximations improve accuracy

**Thank you for your attention!**