2024 Winter Seminar

Building resilience in cybersecurity : An artificial lab approach

포항공과대학교 산업경영공학과
계리모형, 리스크 관리 및 보험 연구실

배기웅

Information

- **Title:** Building resilience in cybersecurity: An artificial lab approach
- **Authors:** Kerstin Awiszus, Yannick Bell, Jan Lüttringhaus, Gregor Svindland, Alexander Voß, Stefan Weber
- **Journal:** Journal of Risk and Insurance
- **Year:** 2023

목차

1. Introduction
2. The Real world: The current state of Cybersecurity Regulation
3. The Artificial Cyber Lab – The Digital Twin of a Complex Cyber System
4. Case Study1: Security-related Interventions under Strategic Interaction
5. Case Study2: Topology-based Interventions and Cyber Pandemic Risk
6. Conclusion and Outlook

Introduction

This paper focuses on systemic cyber risk which are characterized by contagion effects in interconnected systems.

- Regulatory and macroprudential leaders are increasingly aware of the potentially catastrophic consequences of cyber risks.
- The systemic relevance of certain types of cyber threats, systemic cyber risks, is highlighted (ex. WannaCry, NotPetya)
- Focus solely on preventing attacks may be insufficient to manage and mitigate the class of systemic cyber risks.
- Therefore, building cyber resilience requires taking a more expansive approach.

Four key contributions of this paper are

1. Designing the artificial cyber lab.
2. Two exemplary case studies
   - Security-related interventions: security investment game based on the underlying dynamic contagion with Monte-Carlo simulations
   - Topology-based interventions: the effect of network heterogeneity on risk amplification
3. Discussion on selected regulatory measures.
4. A brief overview of the current regulatory framework for cybersecurity, role of private actors in shaping security standards

Current government regulations for cybersecurity

Lawmakers have enacted several regulations, including a variety legal norms due to the increasing importance of cybersecurity.

- Indeterminate legal terms when formulating security requirements. (adequate security measures, adequate technical and organizational measures.. etc.)
- Indeterminacy of the legal terms introduces a significant degree of uncertainty as to the correct cybersecurity measures to be taken.

Both the nonlegally binding nature and the complexity may prevent companies from implementing these standards in practice.

- Standards are not legally binding for private companies.
- Standards are usually characterized by a high degree of complexity.

The Artificial Cyber Lab

The data of cyber risk is scarce and non-stationary due to the rapidly evolving IT-infrastructure.

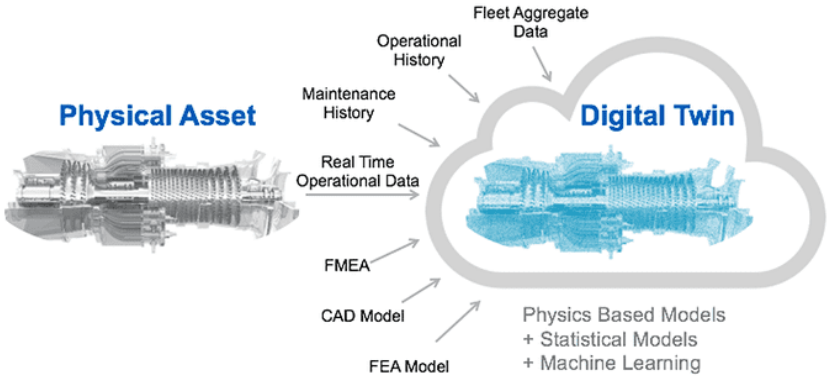
- However, classical statistical and actuarial models are insufficient to evaluate the impact of cyber resilience interventions.
- Since a frequency-severity approach relies on a sufficient amount of meaningful data.

So, this paper follows the digital twin paradigm, the artificial cyber lab.

- A novel approach based on models from network science and contagion theory.
- An experimental setup where cyber resilience measures can be implemented and tested through analysis and simulation.

Cyber lab?

- The digital twin of a complex cyber system in which possible cyber resilience measures may be implemented and tested.
- Digital twins consist of a “physical entity, a virtual counterpart, and the data links between them,”



Network, modeling, loss model

To build the virtual counterparts of real-world cyber systems, a certain degree of abstraction is necessary to provide a sufficiently complex but still tractable modeling framework.

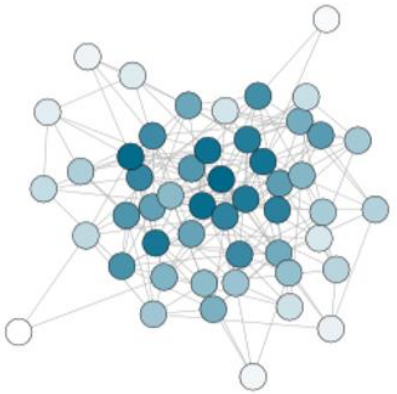
Network models for cyber risk contagion consist of three key components:

- Network : representing interaction channels between agents or entities.
- Modeling : for the spread of a certain cyber threat through interaction channels.
- Loss model : determining the monetary losses occurring at the different agents due to the spread of the considered cyber threat.

Network, modeling, loss model

Network

- This paper considered two standard classes of undirected random networks
- Agents are represented as nodes, the interaction channels between them as edges.



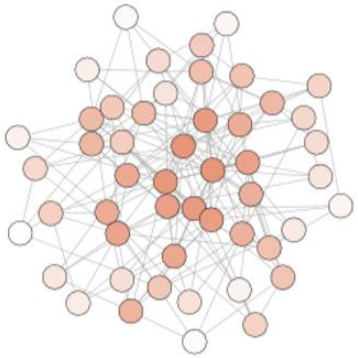
Erdős–Rényi networks

Erdős–Rényi networks (Homogeneous network, $G_p(N)$)

- Simplest random network
- N : the number of nodes
- p: each of the possible $N(N-1)/2$ edges is independently present with the same probability p.

Barabási–Albert networks (Heterogeneous network, $BA(N;m)$)

- Widely observed in the empirical analysis of networks (World Wide Web, IT networks, SNS)
- A hierarchy of nodes is observable (few nodes of high degree (hubs), a vast majority of less connected nodes.)
- N : the number of nodes
- m: starting from an initial core with n_0 nodes, $m \leq n_0$, and ϵ_0 edges, a new node i is added to the graph in each simulation step and m edges for i are randomly generated following a preferential attachment rule.



Barabási–Albert networks

Network, modeling, loss model

Centrality

- The structural importance of single nodes or edges within the network can be characterized using centrality measures C.
- Centrality is not a rigorously defined term, and a large variety of different concepts has been proposed.

Edge(betweenness) Centrality

- Centrality for network edge e.
- How many trajectories go through edge e ?
- $C^{edge}(e) = \sum_{i,j} \frac{\sigma_{ij}(e)}{\sigma_{ij}}$
- $\sigma_{i,j}$: the number of shortest paths between nodes i and j.
- $\sigma_{i,j}(e)$: the total number of these paths that go through edge e.

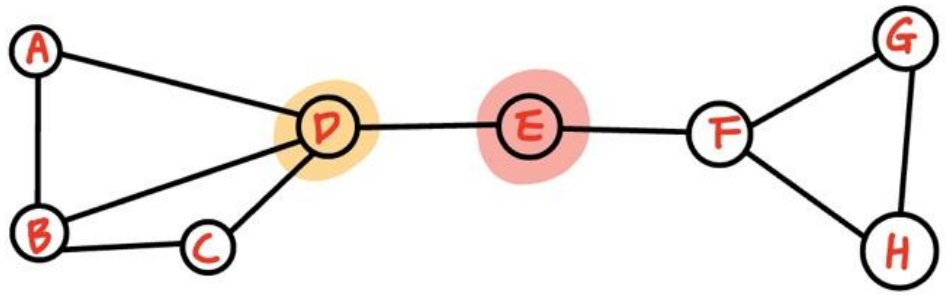
Degree Centrality

- Centrality for network node i.
- How many edges are connected to node i ?
- $C^{deg}(i) = \sum_{j=1}^{N} a_{ij} = \sum_{j=1}^{N} a_{ji}, i = 1, \dots, N$
- $a_{ij} = 1$: nodes I and j are directly connected.
- $a_{ij} = 0$: indicates no direct connection.

Betweenness Centrality

- Centrality for network node i.
- How many trajectories go through node i ?
- $C^{bet}(i) = \sum_{j,h} \frac{\sigma_{jh}(i)}{\sigma_{jh}}, i = 1, \dots, N.$
- σ_{jh}: denotes the total number of shortest paths between nodes j and h.
- $\sigma_{jh}(i)$: the particular number of these paths that go through node i.

Network, modeling, loss model



* Degree

$C^{deg}(D) = 4$

$C^{deg}(E) = 2$

* Betweenness

$$C^{bet}(E) = \frac{12}{7 \times 6} = \frac{2}{7}$$

⇒ A-F, A-G, A-H, B-F, B-G, B-H, C-F, C-G, C-H, D-F, D-G, D-H

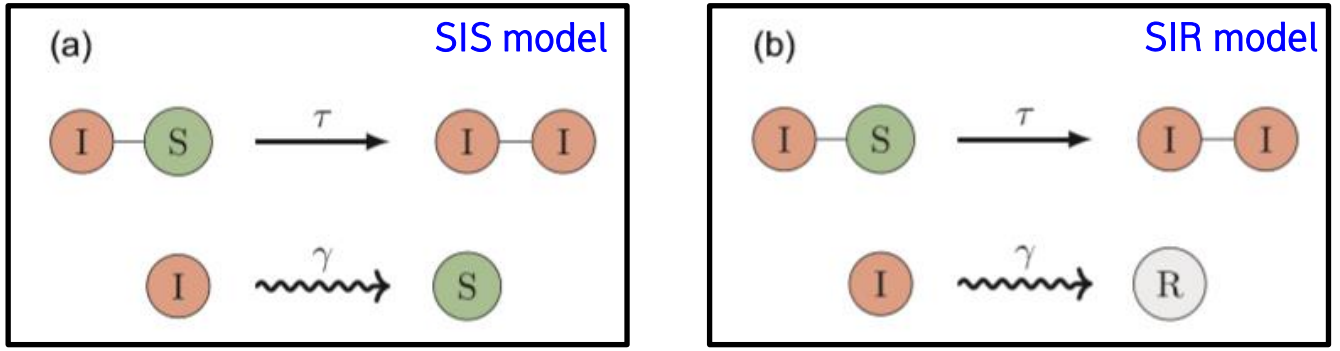
$$C^{bet}(D) = \frac{5}{7 \times 6} = \frac{5}{42}$$

⇒ A-D, A-E, A-F, A-G, A-H

Network, modeling, loss model

Modeling

- Through the interaction channels described by network, a contagious cyber risk may spread.
- Mathematical models divide the set of agents(nodes) into distinct categories
  - susceptible(S), infected(I), recovered(R)
- The SIS (a) and SIR (b) Markov models constitute frequently used epidemic-spreading models on networks
  - SIS(Susceptible, Infected, Susceptible): reinfection events are possible
  - SIR(Susceptible, Infected, Recovered): recovered individuals gain permanent immunity
- Two key parameters
  - infection rate (τ)
  - recovery rate (γ)



I : Infected node

S : Susceptible node

R : Recovered node

FIGURE 2 Infection and recovery for the (a) SIS and (b) SIR model in a network: A susceptible node is infected by its contagious neighbor with rate τ. Independent from the state of its neighbors, an infected node recovers at rate γ. SIS and SIR differ in terms of immunity: In the SIS model, a recovered node is susceptible again such that multiple infections for the same node are possible. In contrast, recovery in the SIR model means that the node is immune and cannot be infected again. SIR, susceptible-infected-recovered; SIS, susceptible-infected-susceptible. [Color figure can be viewed at wileyonlinelibrary.com]

Network, modeling, loss model

Loss Model

- Depending on the modeling purpose, the cyber loss model may emphasize different aspects of an ongoing cyber incident
  - The total number of affected network components
  - Aggregate losses of network nodes
  - The monetary losses of single entities.

An adequate model should

- reflect on the stochastic nature of risk scenarios
- capture key statistical aspects of cyber loss distributions.

Artificial cyber lab setup

What is the setup of cyber lab in this work?

- SIR model will be used (No reinfections)
- Since they consider attacks similar to the WannaCry and NotPetya attacks.
- They are both based on the EternalBlue exploit.
- Once virus is detected, the underlying security issues are easily solvable through the installation of the latest patches.

The key *in*- and *output* parameters can be summarized as follows:

- **Input**:
  - **Network**:
    - network size N (number of agents)
    - topological structure $A = (a_{ij})_{i,j=1,\dots,N}$, that is, the connectivity pattern between nodes (see Figure 1, for examples)
    - number and position of initially infected nodes
  - **Epidemic dynamics**:
    - infection rate $\tau = 0.1$ (determines the speed of the infection), assumed to be equal for all connections[12]
    - individual recovery rates γ_i for nodes $i = 1, \dots, N$ (influence the time needed for recovery —interpreted as IT security level, see Section 4)
  - **Loss distribution**:
    - stochastic modeling framework for loss formation

- **Output**:
  - **Epidemic dynamics**:
    - spread of cyber infection over time, total number of affected nodes, probability of infection for each node
  - **Loss distribution**:
    - aggregate losses for single nodes or the entire network

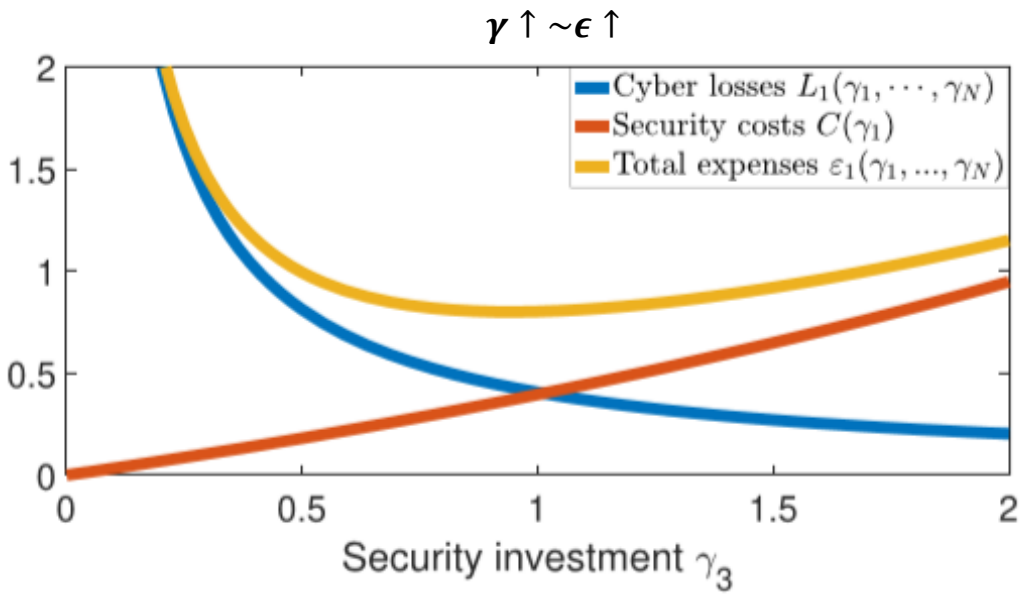
Security investments and strategic interaction

In SIR model,

- fixed homogeneous infection rate ($\tau = 0.1$)
- various individual recovery rate (interpreted as security level)
  - lower security level -> takes a longer time to detect a cyber infection or an existing security gap.

Total expenses: $\varepsilon_i(\gamma_1, \dots, \gamma_N) = L_i(\gamma_1, \dots, \gamma_N) + C_i(\gamma_i), \ \tau = 0.1.$

- A network agent i will try to minimize the total expenses, result from the trade-off between following two functions.
- Cyber loss function, $L_i := L_i(\gamma_1, \dots, \gamma_N) := E[\int_0^\infty I_i(t)dt].$
  - losses of node i as a function of all nodes' security levels due to the interconnectedness of network agents
  - Amounts of cyber losses is related to the duration of a cyberattack.
  - Representing the expected amount of time node i will spend in the infectious state I, given the security levels $\gamma_1, \dots, \gamma_N$.
- Cost function, $C(\gamma_i) = e^{k\gamma_i} - 1, k > 0.$
  - cost of the implementation of security level γ_i for node i.
  - strictly convex exponential function, satisfying $C(0) = 0$



Strategic interaction of interdependent actors

Security Investment Game

- Nodes in a network are interconnected.
- The security level choices of network agents do not only affect their individual expenses ε_t, but also the cyber losses of other network nodes L_i.
- Therefore, these nodes will in turn react to the new threat situation, initializing a cascade of strategic interactions (security investment game)

Steady state

- State with individually optimal security levels, $\gamma_i^{ind}(\gamma_{-i}) = \gamma_i$, $\forall i = 1, \ldots, N$.

Algorithm 4.2 (The security investment game).
Input: *Initial configuration* $\gamma(0) \in (0, \infty)^N$, *number of rounds* $M \in \mathbb{N}_{>0}$.

1. (Initialization) *Set* $r \to 0$.
2. *For every node* i, $i = 1, ..., N$, *calculate*

$$\gamma_i(r+1) = \underset{\gamma_i \in [0,\infty)}{\arg\min}\ \mathcal{E}_i(\gamma_1(r), ..., \gamma_{i-1}(r), \gamma_i, \gamma_{i+1}(r), ..., \gamma_N(r)).$$

More details are given in Appendix E. *Set*

$$\gamma(r+1) = (\gamma_1(r+1), \gamma_2(r+1), ..., \gamma_N(r+1)).$$

3. *If* $r < M$, *set* $r \to r+1$, *and return to Step 2; otherwise end.*

Output: *Security configuration* $\gamma(M)$ *after* M *rounds*

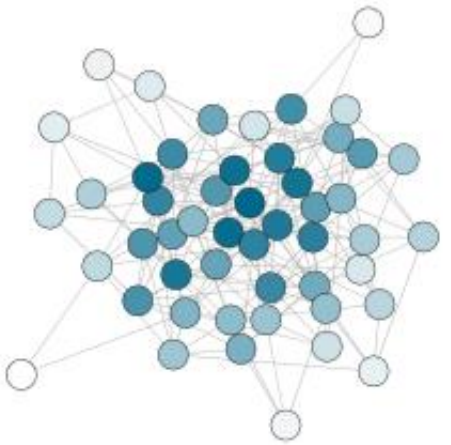
Complex network interactions

Input and output of Security Investment Game

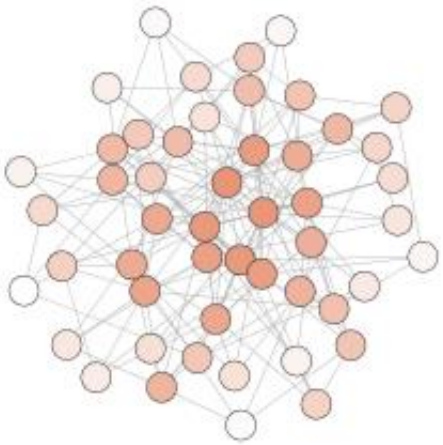
- Input:
  - Erdős–Rényi networks: $G_{0.16}(50)$
  - Barabási–Albert networks: $BA(50; 4)$
  - Number of rounds: $M = 50$
  - Initial security level: $\gamma_i(0) = 0.1$ for all nodes i.
- Output:
  - Values for the cyber losses L_i, in each round of the game, $T = 10{,}000{,}000$ trajectories of the SIR epidemic process are simulated.

Results of security investment game in the steady state

- Erdős–Rényi networks: $[0.3780 \sim 0.6526]$, $\varepsilon(\gamma^{steady}) \approx 21.66$
- Barabási–Albert networks: $[0.4719 \sim 0.7598]$, $\varepsilon(\gamma^{steady}) \approx 21.92$
- More central nodes choose higher security levels (higher costs) than nodes in periphery.



Erdős–Rényi networks



Barabási–Albert networks

- the darker the color, the higher the chosen security level.

Allocation strategies

Q. Can additional security investments further improve the situation?

- Extra amount of security is allocated among the nodes according to one of the following strategies:

1. Untargeted allocation: uniform distribution
   - $\gamma_i^{all} = \beta/N.$
2. Targeted allocation: allocation based on weights ($w_t := \frac{C(i)}{\Sigma_{j=1}^{N} C(j)}$)
   - Upper allocation strategy (higher degree, higher budget)
     - $\gamma_t^{all} := \beta\, w_t$
   - Lower allocation strategy (inverse allocation weights)
     - $\gamma_t^{all} := \beta\, (w_i^{-1}/\Sigma_{j=1}^{N} w_j^{-1})$

Allocation procedures yield a new vector of security levels

- $\tilde{\gamma}_i = \gamma_i^{stead} + \gamma_i^{all}$
- Calculate the accumulated total network expenses $\varepsilon(\tilde{\gamma}_i)$ under the new security configuration.

1. *We start from a steady state γ^{stead} of individually optimal security levels. Moreover, we fix an additional security budget $\beta > 0$.*
2. *This extra amount of security is allocated among the nodes according to one of the following strategies:*

   (a) ***Untargeted** allocation: β is uniformly distributed among all network nodes, providing an additional security investment $\gamma_t^{\text{all}} = \beta/N$ for each node I.*

   (b) ***Targeted** allocation: we choose a centrality measure C and determine the allocation weights*

$$w_t := \frac{C(i)}{\sum_{j=1}^{N} C(j)}, \quad i = 1, \ldots, N.$$

Based on these allocation weights we consider two opposing procedures:

   (i) *The **upper** allocation strategy allocates β proportionally $\gamma_t^{\text{all}} := \beta \cdot w_t$. Here a higher amount of β is assigned to nodes with a higher degree of centrality.*

   (ii) *The **lower** allocation strategy does the opposite. To this end, we calculate the inverse allocation weights*

$$\hat{w}_t := \begin{cases} w_t^{-1} & \text{if } w_t \neq 0 \\ 0 & \text{else.} \end{cases}$$

In this case, the additional security investment $\gamma_t^{\text{all}} = \beta \cdot (\hat{w}_t / \sum_{j=1}^{N} \hat{w}_j)$ for node i assigns a higher amount of β to nodes with a lower, yet positive, degree of centrality.

The proposed allocation procedures yield a new vector of security levels $\tilde{\gamma}$ with entries

$$\tilde{\gamma}_t = \gamma_t^{\text{stead}} + \gamma_t^{\text{all}}, \quad i = 1, \ldots, N.$$

3. *Finally, we calculate the accumulated total network expenses $\mathcal{E}(\tilde{\gamma})$ under the new security configuration.*

Allocation for complex networks

- Compare the different allocation strategies and centrality measure for Erdős–Rényi networks and Barabási–Albert networks by allocating an additional budget of $\beta = 5$.

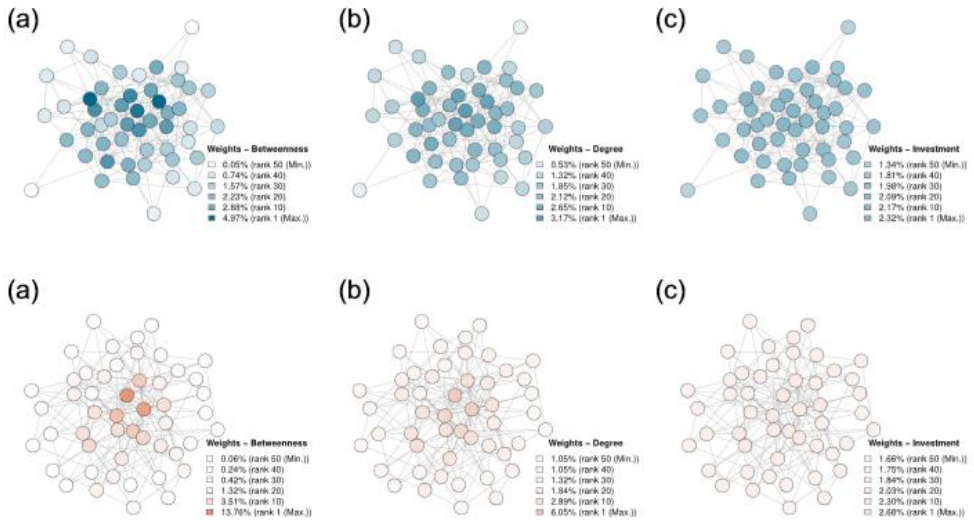


TABLE 1 Percental reduction of accumulated total expenses $\mathcal{E}$ after the allocation of the additional budget $\beta = 5$ among all network nodes.

	$c^{\deg}$	c^{bet}	c^{inv}
Upper	10.6% 11.3%	10.8% 12.3%	10.2% 9.6%
Lower	8.2% 6.7%	0.5% 3.4%	9.5% 8.3%
Untargeted	9.9% 9.0%		

Note: The three proposed allocation strategies are evaluated for each of the suggested centrality measures. Entries for the Erdős-Rényi network are colored in blue (left entries), and for the Barabási-Albert network in salmon (right entries), respectively. For each entry, cyber losses were generated from T = 10,000,000 simulations of the SIR epidemic process. Full data is given in Appendix H.

- The injection of additional network security clearly reduces the accumulated total expenses.
- The upper allocation strategy combined with topology-based centrality measures outperforms the investment-based approach.
  - The proportion of budget which is allocated to periphery is too large in both the untargeted and investment-based case.
  - The betweenness centrality of the most isolated nodes is close to zero, almost no additional security investment is allocated to these nodes.

Further centralization of upper allocations

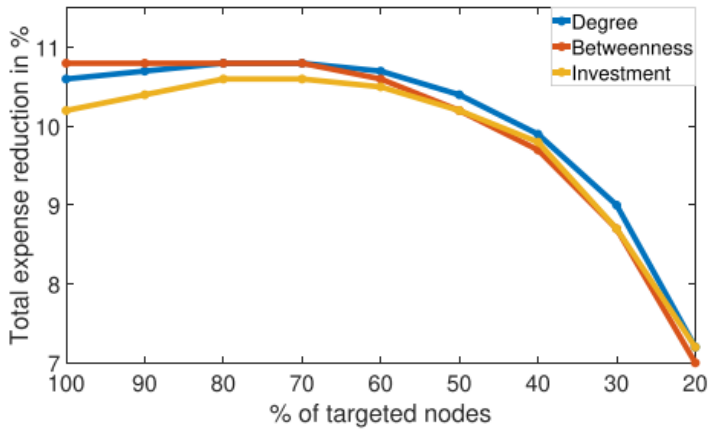
Considering additional security investments based on every node is impossible.

- Budget allocation to a certain fraction p of nodes with the highest centrality(modifying the upper allocation procedure)

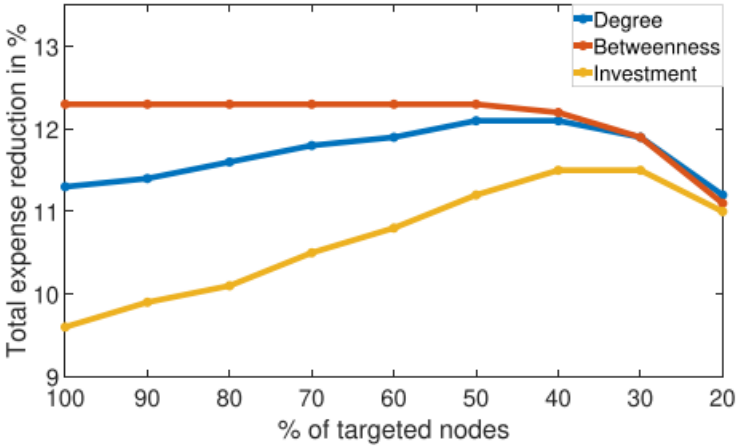
- $\gamma_i = \begin{cases} \beta\left(C(i)/\Sigma_{j\in I}C(j)\right) & if\ i \in I \\ 0, & else \end{cases}$, I is the set of corresponding node indices.

When excluding the most decentralized nodes from the allocation of the additional security budget..

- Erdős–Rényi networks (no substantial change)
- Barabási–Albert networks (different results based on measuring centrality).
- With allocations based on investment and degree centrality, only targeting nodes with a medium to high degree of centrality is even beneficial.
- In the case of betweenness-centrality-based allocation, no substantial change is observed.
  - The allocation weights of periphery nodes are close to zero.



Erdős-Rényi



Barabási-Albert

Further centralization of upper allocations

However, solely allocating the budget to a small fraction of nodes with the highest centrality does not prove to be optimal.

- Trade-off between costs and efficiency.
- Additional security investments for highly central nodes come with substantially increasing costs, since these nodes already invest a high amount in the individually optimal steady state and the cost function is strictly convex.

In view of the information gathering issue and given the comparable performance, degree-based allocations targeting the upper 50% of most central nodes may constitute a reasonable compromise.

- Determining the betweenness centrality of nodes requires information on the full network topology.
- Node degrees, the number of IT contacts of an agent in the cyber network, are local quantities
- Thus they can more easily be determined using questionnaires.

Mandatory security investments as a regulatory obligation can actually increase the overall cybersecurity in a system of interconnected agents.

- With allocations based on investment and degree centrality, only targeting nodes with a medium to high degree of centrality is even beneficial.

Demand for regulation: Network topology and cyber pandemic risk

In large-scale network, the frequency distribution of epidemic outbreak sizes in the SIR model can typically be characterized by the presence of two peaks:

- small outbreaks, affecting only a very small fraction of networks nodes,
- Proper epidemic outbreaks or pandemics, where a large number of nodes becomes infected.

For the simulation studies to access the risk of cyber pandemics

- a global infection rate of $\boldsymbol{\tau = 0.1}$, and recovery rate are assumed to be fixed and homogeneous for all nodes, $\boldsymbol{\gamma_i = \gamma = 1,\ \forall\, i = 1, \ldots, N.}$
- This parameter choice implies that detection of cyber incidents is expected to be 10 times faster than infectious transmission.
- This paper assume an overall high standard of IT security for the full network.

Cyber pandemic risk in homogeneous networks

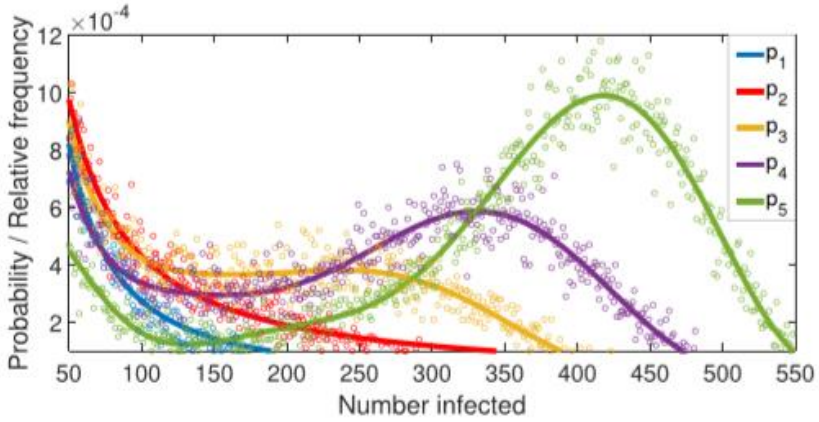
Simulation on Homogeneous networks.

- First analyzed the cyber epidemic risk exposure of homogeneous large networks with a fixed size of $N = 1{,}000$.
- Their topology is determined by the parameters p (control param of network connectivity) and N $(= 1{,}000)$.
- Observe the **outbreak size frequencies** given an initial infection of single network node, over 100,000 simulations for increasing values of p.

The resulting frequency distribution of outbreak sizes

- For low connectivity probabilities p, only small outbreaks occur.
- Outbreak size frequency is exponentially decaying.
- If a certain critical edge probability p_c is exceeded, the frequency distribution is characterized by a second peak around a characteristic large outbreak size.
- $p_1 = 0.01 < p_2 = 0.011 < \boldsymbol{p_c} < p_3 = 0.012 < p_4 = 0.013 < p_5 = 0.014$

The regulator should aim to keeping the network connectivity below the critical threshold p_c.



The heterogeneous case: Cyber pandemic risk in scale-free networks

Simulation on Heterogeneous networks.

- On a larger scale, many real-world networks are characterized by a preferential attachment principle, Barabási–Albert networks.
- A more heterogeneous topology is often observed.
- New random variable K, representing the degree k_i of a randomly chosen network node i.

Degree distribution under preferential attachment follows a power-law.

- $P(K = k) \sim k^{-\alpha}$, with degree exponent $\alpha \in R_+$.
- Node arrangement with $\boldsymbol{\alpha = 3}$ can be modeled using the Barabási–Albert class, so-called scale-free networks.
- Scale-free networks has heavily connected high-degree hubs in their center and less connected nodes in their periphery.

The heterogeneous case: Cyber pandemic risk in scale-free networks

Representative networks from both Erdős–Rényi and Barabási–Albert class, highlighting the different degree distributions

FIGURE 9 Erdős–Rényi $G_{0.01}(1000)$ (left) and Barabási–Albert $BA(1000; 5)$ (right). In both cases, the node size of node i is given by $100 \cdot \sqrt{k_i / \sum_{j=1}^{1000} k_j}$, an increasing function of the node's relative degree $k_i / \sum_{j=1}^{1000} k_j$. [Color figure can be viewed at wileyonlinelibrary.com]

The difference in the network topology has a strong impact on the epidemic vulnerability -> different result of outbreaks

- Over 100,000 simulations on $G_{0.01}(1000)$ and $BA(1000; 5)$. (similar numbers of edge)
- A clear second peak in the frequency distribution of outbreak sizes is observed for the Barabási–Albert class.
- The heterogeneity in the topology of Barabási–Albert networks remarkably lowers the critical connectivity threshold(p_c) for cyber pandemics, amplifying the epidemic spread and triggers the emergence of large-scale outbreaks.

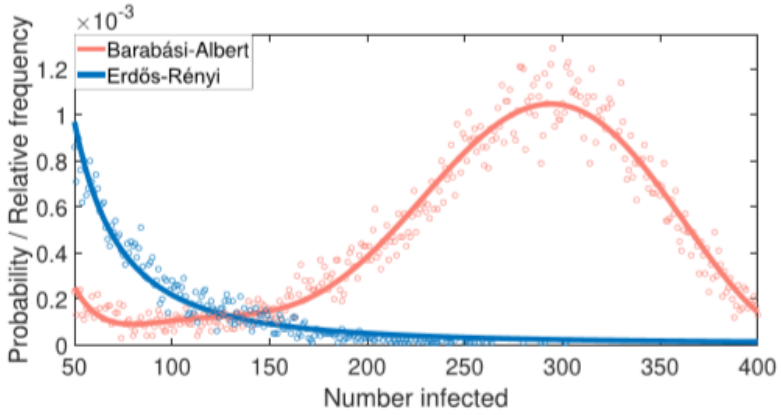


FIGURE 10 Final outbreak size frequencies given an infection of a single network node for the Barabási–Albert $BA(1000; 5)$ and Erdős–Rényi networks $G_{0.01}(1000)$ from Figure 9 over 100,000 simulations. Exact data points from the simulation and a regression curve (power law for Erdős-Rényi, polynomial of degree 8 for Barabási-Albert) are plotted. [Color figure can be viewed at wileyonlinelibrary.com]

The heterogeneous case: Cyber pandemic risk in scale-free networks

Vulnerability of heterogeneity in the topology of Barabási–Albert class is related to the distribution of node degrees.

- For Erdős–Rényi random graphs, in the limit the degree distribution is Poisson with param λ.
- cyber pandemics can be prevented in the infinite limit if the network security/recovery rate γ satisfies $\gamma \geq \tau(\lambda - 1)$.

For scale-free networks, it may be difficult to prevent cyber pandemics by solely improving the network security or reducing the overall network connectivity.

- Large-scale pandemic outbreaks are possible if and only if $N \to \infty$, $\frac{\tau}{\tau+\gamma}\frac{E[K^2-K]}{E[K]} > 1$.
- In the infinite size limit, the second moment $E[K^2]$ of the degree distribution diverges to ∞ while the first moment $E[K]$ stays finite.
- Hence, with growing N, the security parameter γ must be substantially increased to prevent the occurrence of cyber pandemics. This come with massively increasing cost.

In scale-free networks, cyber pandemics are thus an inherent risk of the underlying network topology

- The risk of cyber pandemic outbreaks cannot be controlled by security-related interventions by increasing the recovery rate γ, but requires a manipulation of the degree distribution, that is, the topological network arrangement.

Implementing suitable interventions

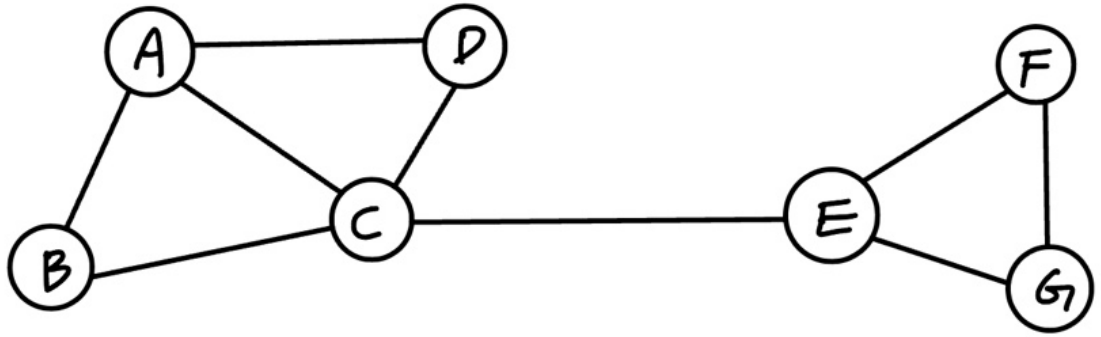
Manipulation of the topological network arrangement with limiting or controlling critical network connections and nodes.

- Edge removal: physical deletion of connections and edge hardening which corresponds to strong protection of network connections via firewalls, the closing of open ports, or the monitoring of data flows using specific detection systems.
- Node Splitting: separate contagion channels, let them pass through two different nodes with the same operational risk.

Classical measure for network functionality is the average shortest path length $< l >$.

- To identify critical network connections and nodes in a way which reduces negative effects on the network functionality to a minimum.
- The minimum number of edges connecting i and j.
- $< l > = \Sigma_{i,j} \frac{1}{N(N-1)} l_{i,j}$
- A small value of $< l >$ is a measure for fast and efficient data flow, corresponding to a high network functionality.

Implementing suitable interventions



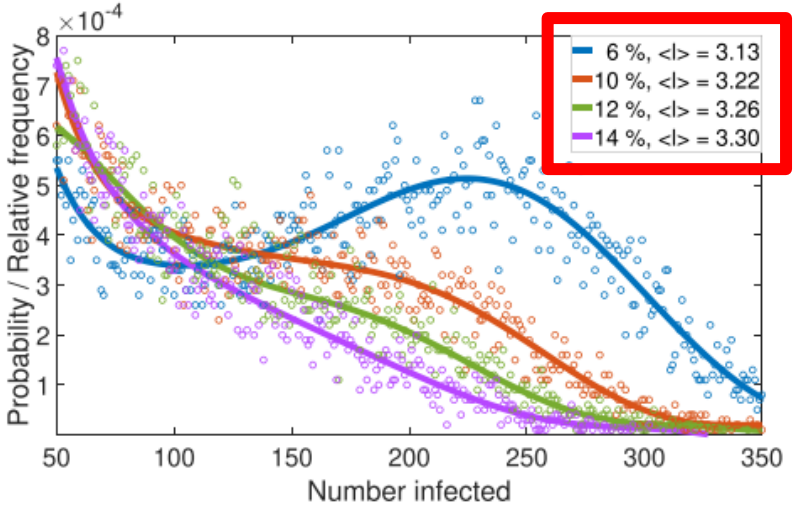
$$\langle \ell \rangle = \sum_{i,j} \frac{\ell_{ij}}{N(N-1)}$$

$$= \frac{2 \times (1+1+1+2+3+3+1+2+2+3+3+1+1+2+2+2+3+3+1+1+1)}{7 \times 6} \approx 1.86$$

Edge removal and node splitting

Edge removal

- Utilized the edge centrality to identify epidemically critical edges.
- Consequently, delete the most central network edges. Stop the deletion process, if the resulting network does not exhibit a cyber pandemic outbreak anymore.
- The remaining network possesses a higher functionality represented by $< l_c >$ than in the case of random edge removals.
- Initial functionality of $< l > \approx 2.96$.



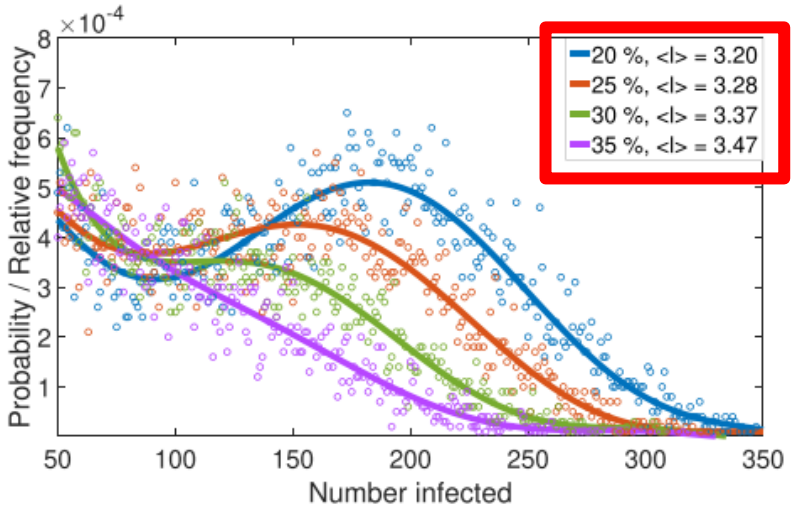
Edge centrality

Edge centrality-based removals

- The percentage of critical links is found to be about 14% ($< l > = 3.30$).

Random edge removals

- Approximately 30%-35% of edges need to removed here to eliminate the risk of cyber pandemics ($< l > = 3.47$).

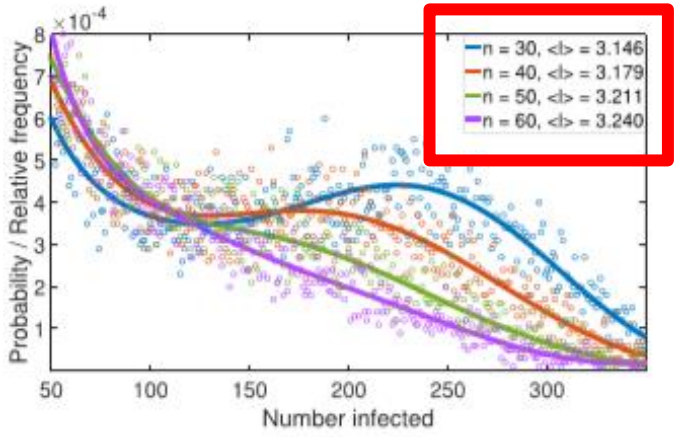


Random

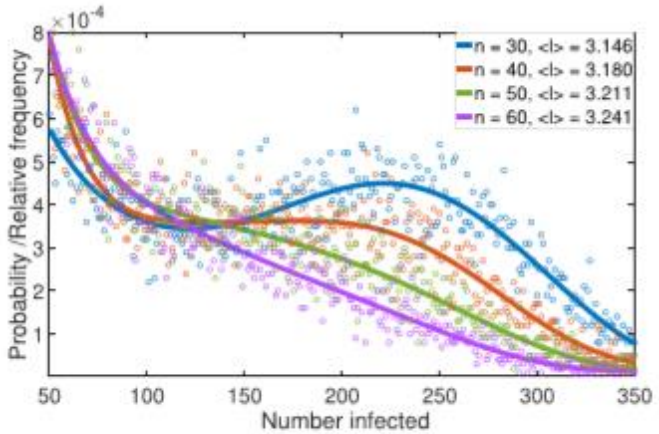
Edge removal and node splitting

Node Splitting

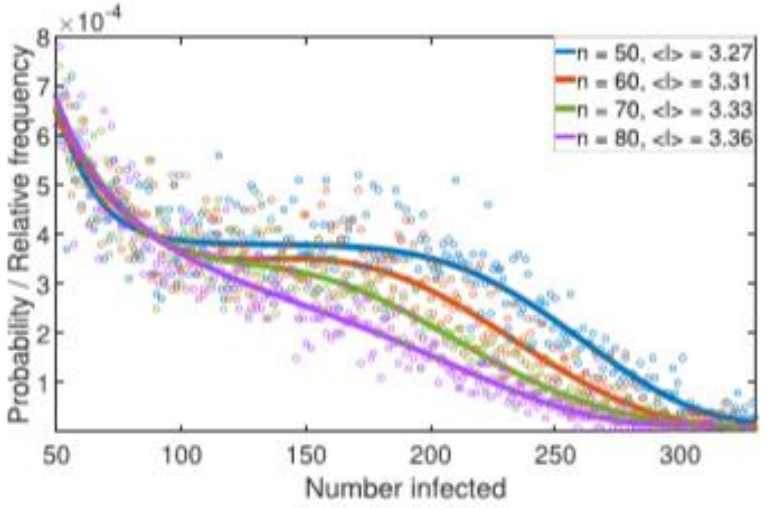
- A splitting procedure which is based on the suitable choice of a node centrality measure C.
- Create an order of node i' s network neighbors when nodes are sorted according to their centrality.
- Node splitting is even more effective than edge removals: Only about 6% of the most central nodes need to be splitted to control the risk of cyber pandemics.
- Nodes with highest centrality are splitted in an iterative manner, that is, centralities are re-evaluated after each split.
- Hence, nodes resulting from a split can be splitted again if they still exceed the rest of the network in terms of centrality.
- The number of critical splits is found to be about $n = 60$ which corresponds to 6% of the nodes.
- Further, the functionality of $< l >\approx 3.24$ of the resulting network is better than in the case of edge removals $< l >\approx 3.30$.



Degree-based



Betweenness-based



Risk allocation and design of contractual obligations.

Risk allocation

- G : original network, G_c : new network
- The network connections of G_c should note warrant further regulatory action.
- Suitable allocation schemes and possible obligations should be derived from the set of deleted or rewired connections.

Contact coefficient on Edge removals

- ϵ_i : the number of critical connections of node i.
- $c_i = \frac{\epsilon_i}{2|\varepsilon_c|}$: contact coefficient. measuring the cyber pandemic risk contribution of the single node i.

Contact coefficient on Node splitting

- $I \subset \{1, \dots, N\}$: the set of nodes from the initial network G which are splitted during the procedure.
- $c_i = \begin{cases} \left(C(i)/\Sigma_{j \in I} C(j)\right) & if\ i \in I \\ 0, & else \end{cases}$

Risk allocation and design of contractual obligations.

Insurance-related obligations.

- A major problem of regulators (insurance companies): not be able to directly control connections within cyber networks.
- In that case, contractual obligations, like surcharges and insurance risk premiums, may incentivize the deletion or protection of critical contagion channels.

Fixed surcharges

- π_i : a cyber premium, not yet accounting for systemic cyber risks
- $\widetilde{\boldsymbol{\pi_i}} = \boldsymbol{\pi_i} + \boldsymbol{c_i} \cdot \boldsymbol{f} \geq \boldsymbol{\pi_i}$
- contact coefficient could serve to determine the fraction of a fixed systemic risk surcharge $f > 0$ which has to be borne by node i.
- If node i possesses no critical network connections, $c_i = 0$.

Risk premia

- L : the random total loss in the original network G
- L_c : the total loss in the new network G_c
- $L_e := L - L_c$: cyber pandemic loss
- ρ : risk measure such as Var or Expected Shortfall
- $\rho(L_e)$: corresponding risk capital, considering topology-based premium
- $\boldsymbol{\pi}(\boldsymbol{c_i}) = \boldsymbol{c_i} \cdot \boldsymbol{\rho}(\boldsymbol{L_e})$

Evaluation of topology-based interventions

Homogeneous networks

- Connectivity(p_c) plays a major role in the emergence of cyber pandemic risk.

Heterogeneous networks

- Preferential attachment principle
- Highly connected network hubs may amplify risk propagation compared to homogeneous networks.
- $N \to \infty$, cyber pandemics cannot be prevented by strengthening the security of network participants but requires manipulating the degree distribution of underlying network topology.

Centrality and contact coefficients

- An effective way to measure an agent's relative topological importance and allocate the cyber pandemic risk of the system to its individual nodes.
- Determining these coefficients requires information on the full network topology

Topology-based interventions

- Only need to focus on a small group of highly central nodes.

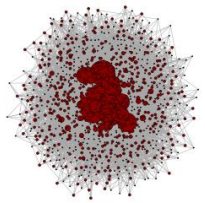


FIGURE 14 Visualization of contact coefficients based on edge removals in the Barabási–Albert network introduced in Figure 9: Here, node size of node i equals $100 \cdot \sqrt{c_i / \sum_{j=1}^{1000} c_j}$, an increasing function of the node's importance with respect to its contact coefficient c_i. [Color figure can be viewed at wileyonlinelibrary.com]