Cyber insurance: actuarial modeling

Caroline Hillairet[1], Olivier Lopez[2]

[1] Ensae Paris, [2] Sorbonne Université

Joint Research Initiative "Cyber insurance: actuarial modeling"
supported by Risk Fundation, AXA Research Fund

Webinar Cyber Risk, Cyber Insurance and Financial Stability
7th December 2021

Cyber-risk

- Various types of attacks (ransomware, phishing, classic frauds...)
- Strike states, companies, people.
- Huge costs : estimated to 1% of the global GDP.
- In France, numbers of ransomwares reported to ANSSI multiplied by 3 between 2019 and 2020.
- Report LUCY of AMRAE ("Association pour le Management des Risques et des Assurances de l'Entreprise") published in 2021 :
  - Study of all French companies that take out a cyber insurance contract through a broker (2019 and 2020).
  - Volume of premiums increased of 49% between 2019 and 2020.
  - The ratio Claims / Premiums is 167% for 2020, 84% for 2019.
  - 2020 significant increase in Losses is due to the occurrence of only 4 XXL claims.

Colonial Pipeline

The Colonial Pipeline Cyberattack | What to Know | Pipeline Resumes | $5 Million Ransom | DarkSide Says Shutting Down | Weaknesses in U.S. Cybersecurity

Cyberattack Forces a Shutdown of a Top U.S. Pipeline

The operator, Colonial Pipeline, said it had halted systems for its 5,500 miles of pipeline after being hit by a ransomware attack.

- "Double extorsion": ransomware attack combined with blackmail.
- 4.2% increase of WTI and Brent.
- Authors: the hacker group "Darkside" (Ransomware as a service).

Research Initiative
Cyber Risk: Actuarial Modeling

Cyber-risk specificities and mutualization issues

- Similarities with operational risk...
- ... but specificities in the structure of cyber events.

1. The risk is **new** and **evolves fast**.
2. "**Silent cyber**": non-cyber policies may content guarantees that can be triggered by cyber events if not excluded.
3. Extreme events (huge losses can occur)
4. Accumulation risk : potential concentration of incidents which leads to loss of mutualization.

These features may endanger mutualization of cyber risk.

Outline

Pricing and reserving

- Frequency / Severity modeling:
  - N is the number of claims ("frequency"),
  - Y is the cost of a claim ("severity").
- Pricing:

$$\pi = E[N]E[Y],$$

  where:
  - π is the premium of the insurance contract,
  - one assumes that Y and N are independent.
- Reserving:
  one needs to understand the whole distribution of N and Y.
- Insurance is interesting only if π is small enough.

Extreme Value Theory

- In reserving, the "tail of the distribution" matters.
- The tail index, often denoted γ, determines the heaviness of the tail.
- If $\gamma > 0.5$, variance is infinite. If $\gamma > 1$, the average loss is not properly defined (one sometimes says it is "non insurable").

Impact of γ on the policies

- If a limit to the policy is introduced, γ becomes negative, which seems to avoid any problem with the tail of the distribution.
- This solution is artificial and does not completely solve the problem: if the "original" γ is large, the limit to put to obtain a reasonable price/reserve is small.

- General idea : the value of this γ parameter is linked to the quality of the coverage. If γ is high, more restrictions have to be introduced, and the coverage becomes poorer.

γ and heterogeneous population

- Example : two types of incidents :
  - type 1 : can induce severe losses, but still insurable ($\gamma_1 < 1$).
  - type 2 : can induce very severe losses with $\gamma_2 > 1$.
- A statistician tries to determine γ based on data on these losses, without knowing if there are related to a type 1 or type 2 risk.
- Consequence : the estimated γ will be close to γ_2.
- Alternative : if one can identify the two populations and the two values of γ, one can propose a better contract to type 1 without penalizing too much type 2.
- See Farkas et al. (2021): *Cyber claims analysis through Generalized Pareto Regression Trees with applications to insurance pricing and reserving.*
- A reliable and accurate database is an essential prerequisite for a better analysis of the tail distribution.

Frequency: Hawkes modeling

- One observes autocorrelation between events and clustering effect: Poisson process modeling is not adapted, since inter-arrival times $(\tau_n - \tau_{n-1})$ are not iid.
- Alternative: **Hawkes modeling**.
  - Self-exciting model with stochastic intensity $\lambda(t)$, fully specified by the point process itself

$$\mathbb{P}_t(\tau_n - \tau_{n-1} \in [t, t+dt] \mid \tau_n - \tau_{n-1} \geq t) = \lambda(t)dt$$

$$\lambda(t) := \lambda_0(t) + \sum_{\tau_n < t} \Phi(t - \tau_n),$$

λ_0 base intensity and Φ excitation kernel.

  - Capture auto-excitation fetaures: every event increases the probability for a new event to occur.
- *Reference : Bessy-Roland, Boumezoued, Hillairet "Multivariate Hawkes process for cyber insurance," Annals of Actuarial Science, 2020.*

Outline

Research Initiative
Cyber Risk: Actuarial Modeling

Loss of mutualization : when there is no independence

- Example in insurance : natural catastrophes and portfolios with spatial correlations:

- Cyber risk: how to define proximity?

Contagion models with networks effects

- Multi-group SIR (Susceptible-Infected-Recovered) models with different subpopulations.

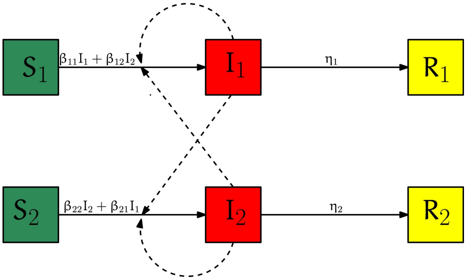
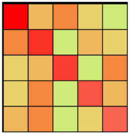


Figure from Magal et al. (2018)

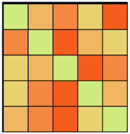
- $\mathcal{B} = (\beta_{i,j})_{1\leq i,j\leq k}$ matrix of infection rates : $\beta_{i,j}$ materializes how j contaminates i.
- We also introduce a flexible framework to model the initial attacks that trigger the contagion.

Some examples of comparisons that show the impact of the topology of the network

- Two classes of matrices $\mathcal{B}$:
  - "Clustered" : the transmission is essentially intern to a class.



  - "Non-clustered" : the transmission is stronger from one class to another than within a given class.



$\rightarrow$ the "Non-clustered" situation is worse, since the outbreak rapidly spreads from one class to any others.

Calibration of connections on OECD data

- Calibration of the model based on macroeconomic data: OECD data to identify the dependence/connectivity between some sectors of activity (more details in d'Oultremont, Lopez, Spoorenberg (2021) and Hillairet et al. (2021)).
- Contagion matrix

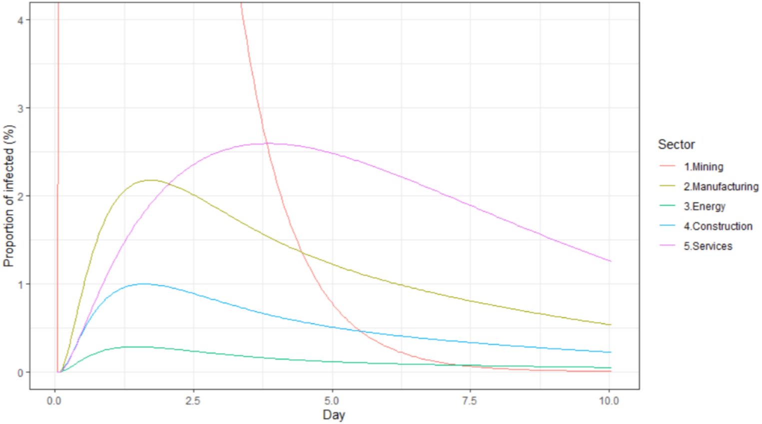
	Mining	Manufacturing	Energy	Construction	Services	Total
Mining	0,0634	0,2927	0,0449	0,1427	0,1255	0,6692
Manufacturing	0,0063	0,0527	0,0027	0,0108	0,0351	0,1076
Energy	0,0135	0,0370	0,0571	0,0150	0,0452	0,1679
Construction	0,0019	0,0068	0,0007	0,0141	0,0091	0,0326
Services	0,0003	0,0042	0,0004	0,0017	0,0161	0,0227
Total	0,0855	0,3934	0,1057	0,1844	0,2309	1

Table 3: Normalized Interaction matrix $\mathbf{B}_0$.

- Calibration of a Wannacry-type scenario $\mathcal{B} = \beta\mathcal{B}_0$
- Disclaimer: this contagion matrix does not completely reflect the true connectivity between sectors.

Example of epidemic dynamics of Wannacry type

Evolution of the proportion of infected - Attack on Mining.



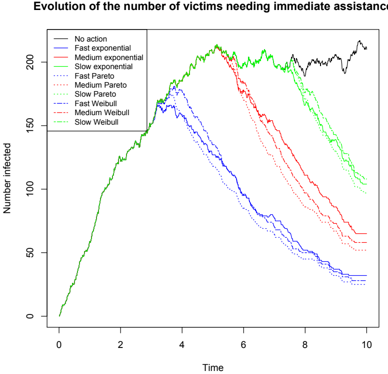
The value of the peak for the Mining sector is at 70% (after 10 hours)

How can we use these models?

- "Ranking" of sectors of activity: one can identify which sector is more "systemic" than others in the sense that, if attacked, it will lead to a higher number of victims.
- Quantifying the "peak": helps to identify how many "tech" assistance will be required by the policyholders during such a crisis.
- Diversification: design a portfolio that may resist to such contagious episode.
- Identify the benefits of protection:
  - of a given sector: protecting some key sectors may help to prevent the infection from spreading, even if this sector is not directly attacked.
  - from the reaction of the targets.

Illustration: simulation of a Wannacry-type scenario

- Simulation on a portfolio of 10 000 policyholders.
- We model the reaction of the policyholders (i.e. their capacity to protect themselves once they are informed on the ongoing threat).
- Different forms and types of responses (in blue fast reaction, in red medium reaction, in green slow reaction)



References

- Joint Research Initiative "Cyber-insurance: actuarial modeling" : `https://sites.google.com/view/cyber-actuarial/home?authuser=0`
- On contagion scenarii:

Hillairet C., Lopez O. (2021) **Propagation of cyber incidents in an insurance portfolio: counting processes combined with compartmental epidemiological models**, in *Scandinavian Actuarial Journal*

D'Oultremont L., Lopez O., Spoorenberg B. (2021) **Modeling accumulation scenarios in cyber risk**, Detra Note (Detralytics, `https://detralytics.com/detra-notes/`)

Hillairet C., Lopez O., D'Oultremont L., Spoorenberg B. (2022) **Cyber contagion: impact of the network structure on the losses of an insurance portfolio** Preprint

References

- On extreme events:

Farkas, S., Lopez, O., Thomas, M., (2021) **Cyber claims analysis through Generalized Pareto Regression Trees with applications to insurance pricing and reserving.**, in *Insurance: Mathematics and Economics, Elsevier, vol. 98(C), pages 92-105*

- Frequency modeling and clustering effect

Bessy-Roland Y., Boumezoued A., Hillairet C., (2020) **Multivariate Hawkes process for Cyber Risk Insurance** in *Annals of Actuarial Science>Volume 15 Issue 1*

Hillairet C., Réveillac, A., Rosenbaum M. (2021) **An expansion formula for Hawkes processes and application to cyber-insurance derivatives** *preprint*