Actuarial Modeling of Cyber Risk

Caroline Hillairet, ENSAE Paris, CREST

Based on joint works with Olivier Lopez and Anthony Réveillac.
Joint Research Initiative "Cyber risk: actuarial modeling"
supported by Risk Fundation, AXA Research Fund

Annual Assembly of the Swiss Association of Actuaries,
September 2023

Outline

Cyber-risk

- An important growing risk
  - According to J. Powell (President of the U.S. Federal Reserve) cyber-attacks constitute the main threat to the global financial system.
  - Huge costs : estimated to 1 % of the global GDP.
- A multifaceted risk
  - Various types of attacks (ransomware, phishing, classic frauds...)
  - A cyber incident can be voluntary (cyber attack) or not.
  - Multiple consequences of a cyber-incidents: Business interruption (sometimes months before retrieving the same level of activity), Loss of data, Indirect damages (in some cases, destruction or death).
  - Strike states, companies, public administrations, individuals.
- Role of Cyber-insurance
  - a fundamental tool to improve the resilience of the economy.
  - Cyber insurance includes various guarantees: financial reparation, immediate assistance to restart the activity, prevention and risk analysis. protection against regulation issues caused by leaks of data, crisis communication...

Cyber-risk specificities

1. The risk is new and constantly evolving with a fast adaptation of the attackers (in case of malicious cyber). Very few data available
2. Changes through time in the reporting behavior, due to regulation and evolution in the perception of the risk.
3. Extreme events (huge losses can occur): cyber-risk has a catastrophic component. But unlike natural disaster, it is not stable since relying on human behavior. This behavior changes rapidly through time.
4. Accumulation risk : cyber-risk has a systemic component. Potential concentration of incidents which leads to loss of mutualization.

These features may endanger risk pooling.

Difficult quantification of the economic losses due to cyber risk

Endangering risk pooling

- Risk pooling relies on the Law of Large Number.
- TCL : control the gap between the total losses and the premium.
- Risk pooling is endangered as soon as:
  - the risk is so volatile that variance is infinite.
  - the risk is "heavy tailed" and the average cost may not be defined.
  - variance is finite, but very large, because of the heterogeneity of the population.
  - policyholders are not independent.
  - the number of policyholders is too low.
- The event insured should be sufficiently rare.
- Imprecision related to statistical estimation
  (few available data + data quality issue)

Outline

Extremes events and heavy tailed distribution

- Cyber losses: typically the case of heavy tail distributions.
  - corresponds to distribution with high dispersion, i.e. with slow rate of decrease of the density function.
  - Can take high values with significantly high probability
- In statistics : corresponds to Extreme Value Theory.
- This theory allows to identify common behaviors in the tail of distributions.
- The tail index, often denoted γ, allows to determine the heaviness of the tail.

How to adapt classification and regression techniques to this context?

ENSAE IP PARIS

RESEARCH INITIATIVE Cyber Risk: Actuarial Modeling

Peaks over threshold

- Asymptotically ($u \to \infty$), exceedances $X = Y - u$ over the threshold u occur according to (univariate) generalized Pareto distribution
- **Theorem of Pickands** (1975): If there exists $(a_u) > 0$, (b_u) and a cumulative distribution function H such that

$$\lim_{u\to\infty} \mathbb{P}[Y - u \geq a_u x + b_u \mid Y > u] \to_{u\to\infty} 1 - H(x),$$

then H is a Generalized Pareto distribution (GPD)

$$H_{\sigma,\gamma}(x) = \begin{cases} 1 - \left(1 + \frac{\gamma}{\sigma}x\right)^{-1/\gamma} & \text{if } \gamma \neq 0 \\ 1 - \exp\left(-\frac{x}{\sigma}\right) & \text{if } \gamma = 0. \end{cases}$$

- σ = scale parameter, γ = shape parameter (called the tail index)

Tail index

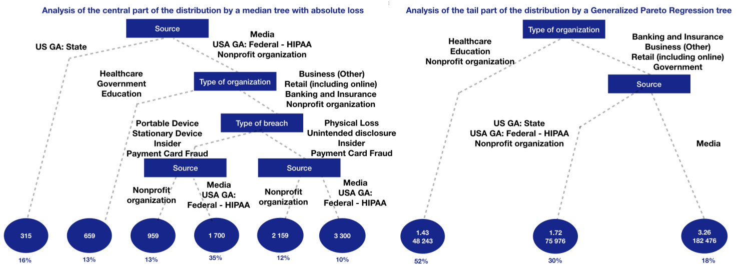
- Typically three types of behaviors depending on γ
  - $\gamma < 0$: « Weibull domain », light tail distributions
  - $\gamma = 0$: « Gumbel domain », also light tail, like normal distribution or log-normal.
  - $\gamma > 0$: « Fréchet domain »: heavy tail, Pareto-like distributions.
- A way to classify a distribution with respect to its tail behavior.
- If $\gamma > 0.5$, variance is infinite: mutualisation is less efficient (a much larger value of policyholders is required).
- If $\gamma > 1$, the average loss is not properly defined (one sometimes says it is "non insurable").

Generalized Pareto CART

- **Building risk classes**, focusing on the tail characteristic of the distribution (Farkas, Lopez, Thomas (2021))
- Using modified CART (Clustering And Regression Tree) introduced by Breiman et al. (1984): Generalized Pareto regression trees
- **Applications**:
  - classification of vulnerabilities/risk factors,
  - help to separate types of incidents or circumstances according to whether they can be covered without endangering risk pooling.
  - design of parametric insurance products.

Illustration on the PRC database

- Privacy Rights Clearing House data-base (PRC). 8800 events over the period 2005-2019.
- Left: **standard CART**, with splitting rule using an absolute loss (median regression) → risk factors for classification for the central part of the distribution,
- Right: **Generalized Pareto regression tree**, with splitting rule using a GDP-log-likelihood loss → risk factors for classification for the tail



Outline

Loss of risk pooling : when there is no independence

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

- But for cyber risk: how to define proximity?
- Tool to test diversification of a cyber portfolio: accumulation scenarios based on epidemiological models with network effects.

Contagion models with networks effects

- Multi-group SIR (Susceptible-Infected-Removed) models with different sub-populations.

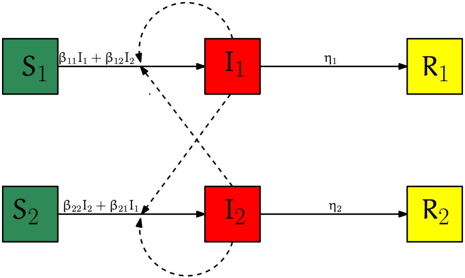


Figure from Magal et al. (2018)

- $\mathcal{B} = (\beta_{i,j})_{1 \leq i,j \leq K}$ 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.

Multigroup compartmental epidemiological model

- Multi-group SIR: consider K **subpopulations**: $1 \leq i \leq K$

$$\frac{dS_i(t)}{dt} = -\left(\alpha_i(t) + \sum_{j=1}^{K} \beta_{i,j} I_j(t)\right) S_i(t)$$

$$\frac{dI_i(t)}{dt} = \left(\alpha_i(t) + \sum_{j=1}^{K} \beta_{i,j} I_j(t)\right) S_i(t) - \gamma_i I_i(t)$$

$$\frac{dR_i(t)}{dt} = \sum_{i=1}^{K} \gamma_i I_i(t).$$

- $\mathcal{B}$ matrix of infection rate :$\beta_{i,j}$ materializes how j contaminates i.
  $\rightarrow$ **network effects**.
- $\alpha_i(t)$ represents an intensity of attacks in class i.
  Example: single initial burst $\alpha_i(t) = \alpha 1_{0 \leq t < 1}$ for some i.

Impact of protection measures

- Multi-group SIR: consider K subpopulations:

$$\frac{dS_i(t)}{dt} = -\eta_i(t)\left(\alpha_i(t) + \sum_{j=1}^{K} \beta_{i,j} I_j(t)\right) S_i(t)$$

$$\frac{dI_i(t)}{dt} = \eta_i(t)\left(\alpha_i(t) + \sum_{j=1}^{K} \beta_{i,j} I_j(t)\right) S_i(t) - \gamma_i I_i(t)$$

$$\frac{dR_i(t)}{dt} = \sum_{i=1}^{K} \gamma_i I_i(t).$$

- $\eta_i(t)$ represents how group i is protected against the threat.
- Example: $\eta_i(t) = 1 - \lambda 1_{I_i(t) \geq s}$, or $\eta_i(t) = 1 - \lambda 1_{\sum_k I_k(t) \geq s}$

ENSAE IP PARIS

RESEARCH INITIATIVE Cyber Risk: Actuarial Modeling

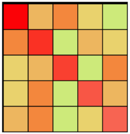
How can we use these models?

- "Ranking" of sectors : one can identify which group is more "systemic" than others in the sense that, if attacked, it will lead to a higher number of infected.
- Quantifying the "peak": helps to identify how many "tech" assistance will be required at the peak of the crisis. Saturation risk which can cause an increase of the costs.
- Diversification
- Identify the benefits of protection:
  - of a given group: protecting some key groups may help to prevent the infection from spreading.
  - from different reaction.

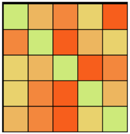
Topology of the network

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.



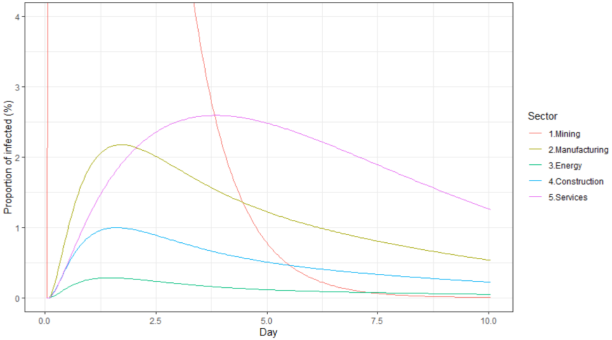
  - "Non-clustered" : the transmission is stronger from one class to another than within a given class.



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

Example of epidemic dynamics of Wannacry type

- Calibration of a Wannacry-type scenario $\mathcal{B} = \beta \mathcal{B}_0$
- Contagion matrix $\mathcal{B}_0$ based on macroeconomic data: OECD data to identify the connectivity between some sectors of activity.

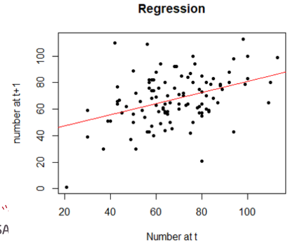


Evolution of the proportion of infected - Attack on Mining.

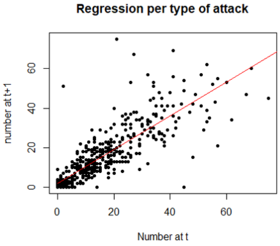
Outline

Auto-excitation and clustering of cyber-events

- Privacy Rights Clearing House data-base (PRC).
- Regression of the number of event during the following month $t+1$ as a function of the number of event during the current month t (should be independent for a Poisson process model to be valid)
- Auto-correlation dramatically increases when focusing on attacks of the same type



R^2=0.154 [0.030, 0.278]



R^2=0.726 [0.687, 0.766]

Hawkes model

- Hawkes processes to model contagion of cyber events, cascading phenomenon in the supply chain : Self-exciting model with stochastic intensity, fully specified by the point process itself (equivalently its jump times $(\tau_n)_n$)
- H Hawkes process with (deterministic) excitation kernel Φ and base intensity λ_0 is the counting process ($H_0 = 0$) with intensity process

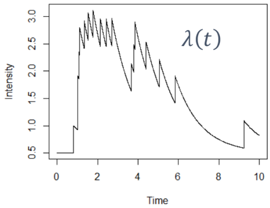
$$\lambda(t) := \lambda_0(t) + \int_{(0,t)} \Phi(t-s)dH_s = \lambda_0(t) + \sum_{\tau_n < t} \Phi(t-\tau_n) \quad t \in [0, T],$$

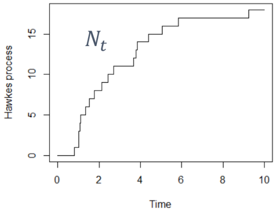
- Hawkes models used in finance, such as credit risk (Errais, Giesecke and Goldberg (2010)...), high-frequency finance (Bacry et al. (2015)...), in cyber-security (Baldwin et al. (2017)) and many others fields.

Toy example of Hawkes process with exponential kernel

- $(H_t)_{t\geq 0}$ counting process with jump times $(\tau_n)_{n\geq 1}$
- Intensity process of the counting process with exponential kernel

$$\lambda(t) = \mu + \sum_{\tau_n < t} \alpha \exp\left(-\beta(t - \tau_n)\right)$$





- Each jump represents an attack/claim
- Intensity decreases exponentially between jumps

Pricing Expansion formula

- Evaluation of quantities/insurance contracts (such as Stop Loss contracts) with underlying cumulative loss processes indexed by a Hawkes process (compound Hawkes process)

$$L_t = \sum_{i=1}^{H_t} X_i.$$

- 2 key ingredients : Thinning algorithm (Poisson Imbedding) + Malliavin calculus (Mecke Formula)
- **Expansion formula** : compromise between simplicity/tractability of approximate formulas and accuracy.
- Control of the error if standard valuation formulas are used (Poisson model with independence): correcting term due to the self-exciting property.

Evaluation of insurance contracts

of the form (such as Stop Loss contracts)

$$\mathbb{E}[K_T\ h(L_T)] = \mathbb{E}\left[\int_{(0,T]} Z_t dH_t\ F\right]$$

- K_T: effective loss covered by the (re)insurance company,

$$K_T := \sum_{i=1}^{H_T} g(\eta_i, \vartheta_i) e^{-\kappa(s-\tau_i)} = \int_{(0,T]} Z_t dH_t$$

Z a $\mathbb{F}$-predictable process, $(\eta_i, \vartheta_i)_{i\geq 1}$ iid r.v. independent of H,

$$Z_s := \sum_{i=1}^{+\infty} g(\eta_i, \vartheta_i) e^{-\kappa(T-s)} \mathbf{1}_{(\tau_{i-1}, \tau_i]}(s), \quad s \in [0, T]$$

- $L_T := \sum_{i=1}^{H_T} f(\eta_i) e^{-\kappa(s-\tau_i)}$: loss that activates the contract.

  $F := h(L_T) (= \mathbf{1}_{\{m \leq L_T \leq M\}})$ is a functional of the Hawkes process.

Malliavin IPP formula (Mecke formula)

Aim: transformation of $\int ...dH_t$ into $\int ...dt$.

- If $H = N$ is an homogeneous Poisson process with intensity $\mu > 0$

$$\mathbb{E}\left[\int_{(0,T]} Z_t dN_t F\right] = \mu \int_0^T \mathbb{E}\left[Z_v F \circ \varepsilon_v^+\right] dv$$

  - $F \circ \varepsilon_v^+ =: F^v$ denotes the functional on the Poisson space where a deterministic jump is added to the paths of N at time v
  - **adding a jump at some time v = adding "artificially" a cyber event at time v** (stress test).
- In case of a Poisson process N: the additional jump at some time v only impacts the payoff of the contract by adding a new event in the cumulative loss
- **In case of a Hawkes process H: it also impacts the dynamic (after time v) of the counting process H.**

Thinning Algorithm

Representation of a Hawkes process in terms of a Poisson measure N on $[0, T] \times_{+}$ (known as "**Poisson imbedding**" or "Thinning Algorithm")

$$\begin{cases} H_t = \int_{(0,t]} \int_{+} \mathbf{1}_{\{\theta \leq \lambda_s\}} N(ds, d\theta), \\ \lambda_t = \mu + \int_{(0,t)} \Phi(t-u) dH_u. \end{cases} \quad (1)$$

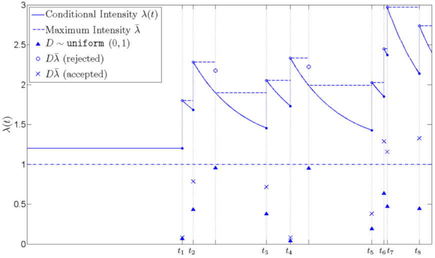


Illustration from Ogata (1981)

Expansion formula for the Hawkes process

Assume Z bounded $\mathbb{F}^H$-predictable process, F bounded $\mathcal{F}_T^N$-measurable r.v. and $\|\Phi\|_1 < 1$.

$$\mathbb{E}\left[F \int_{[0,T]} Z_t dH_t\right] = \mu \int_0^T \mathbb{E}\left[Z_v F^v\right] dv$$
$$+\mu \sum_{n=2}^{+\infty} \int_0^T \int_0^{v_1} \cdots \int_0^{v_{n-1}} \prod_{i=2}^{n} \Phi(v_{i-1} - v_i) \mathbb{E}\left[Z_{v_1}^{v_n,\dots,v_2} F^{v_n,\dots,v_1}\right] dv_n \cdots dv_1 .$$

- the first term corresponds to the formula for a Poisson process (setting Φ at zero)
- the sum in the second term can be interpreted as a **correcting term due to the self-exciting property** of the counting process.
- Extensions to intensity process depending of the claims' sizes.

Concluding remarks and Extensions

- We proposed models and developed methodologies for a better assessment of cyber-risk and to contribute to the viability of the cyber-insurance economic model.
  - Taking into account the specificities of cyber-risk (high volatility in claims, accumulation risk...)
  - with a concern to their practical implementation/calibration
  - But the relevance of such modeling is currently constrained by the limited data available: a need to nourish them with consistent and reliable data (on policyholders, on claims), for a better risk analysis.
- Future works
  - Study of the behavioral aspects of the different actors (insurers, but also policyholders and hackers).
  - Transfer of risk, e.g. parametric insurance.
  - Similarities and connections with other risks : disruption of the supply chain; conjunctions between different risks (shortage of raw materials, geopolitics, health...)

Bibliography

- 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., Reveillac A., Rosenbaum M., (2023) **An expansion formula for Hawkes processes and application to cyber-insurance derivatives** in *Stochastic Processes and Their Applications*
- Hillairet C., Lopez O. (2021) **Propagation of cyber incidents in an insurance portfolio: counting processes combined with compartmental epidemiological models**, in *Scandinavian Actuarial Journal*
- Hillairet C., Lopez O., D'Oultremont L., Spoorenberg B. (2022) **Cyber contagion: impact of the network structure on the losses of an insurance portfolio** in *Insurance: Mathematics and Economics*
- 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*

Joint Research Initiative **Cyber-risk: actuarial modeling** :

`https://sites.google.com/view/cyber-actuarial/home`

Thank you for your attention !

ENSAE IP PARIS

RESEARCH INITIATIVE Cyber Risk: Actuarial Modeling