Predicting Cyber-Attacks using Hawkes processes

Caroline Hillairet, Ensae Paris, Crest

Joint work with Alexandre Boumezoued, Milliman R&D

and Yannick Bessy-Roland, Axa.

May 11th – May 15th 2020

IAA AAI INSTITUT DES ACTUAIRES

SECTIONS VIRTUAL COLLOQUIUM 2020

With the support of Axa Risk Fundation

RESEARCH INITIATIVE
Cyber Risk: Actuarial Modeling

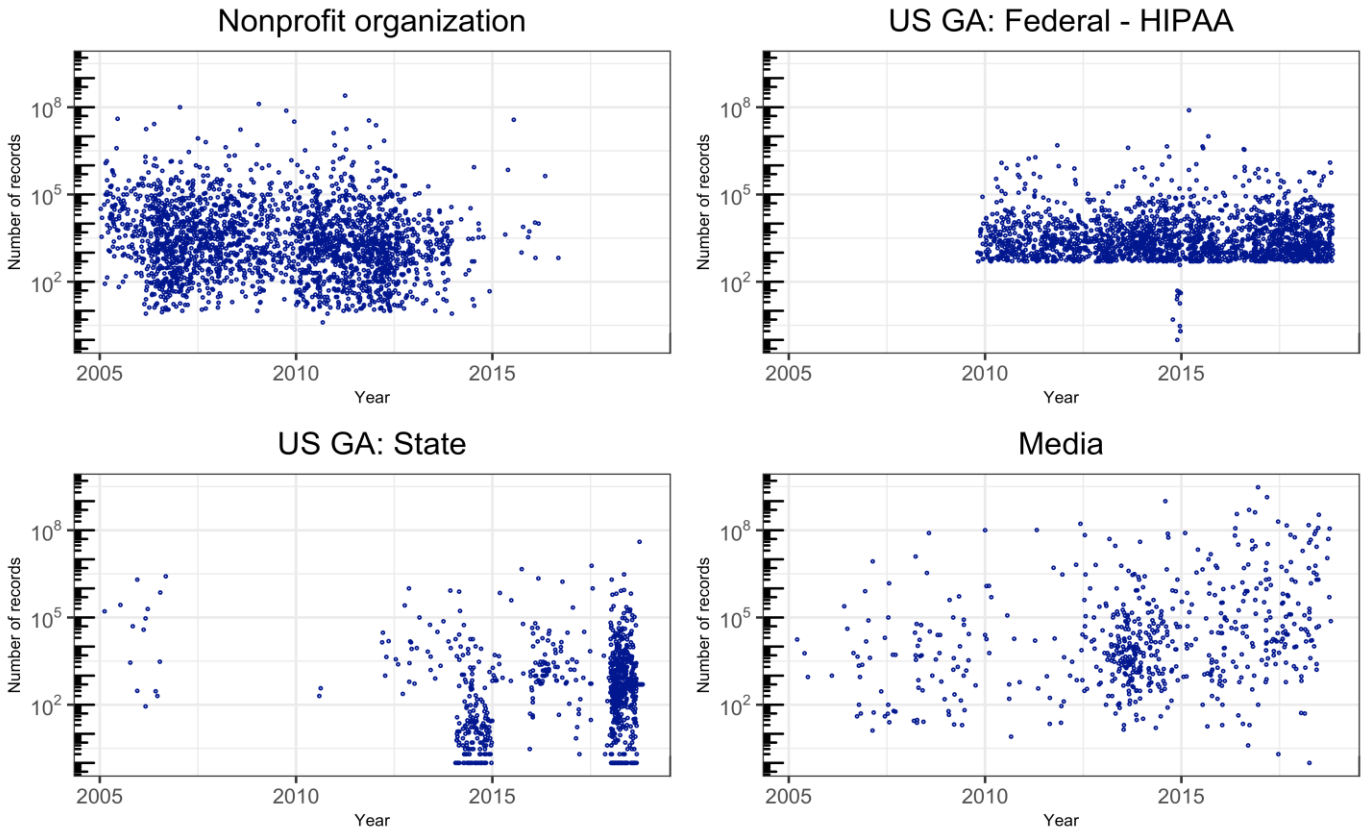
Outline of the Talk

- Data breaches dataset
- Hawkes model
- Fitting and prediction

Data breaches dataset

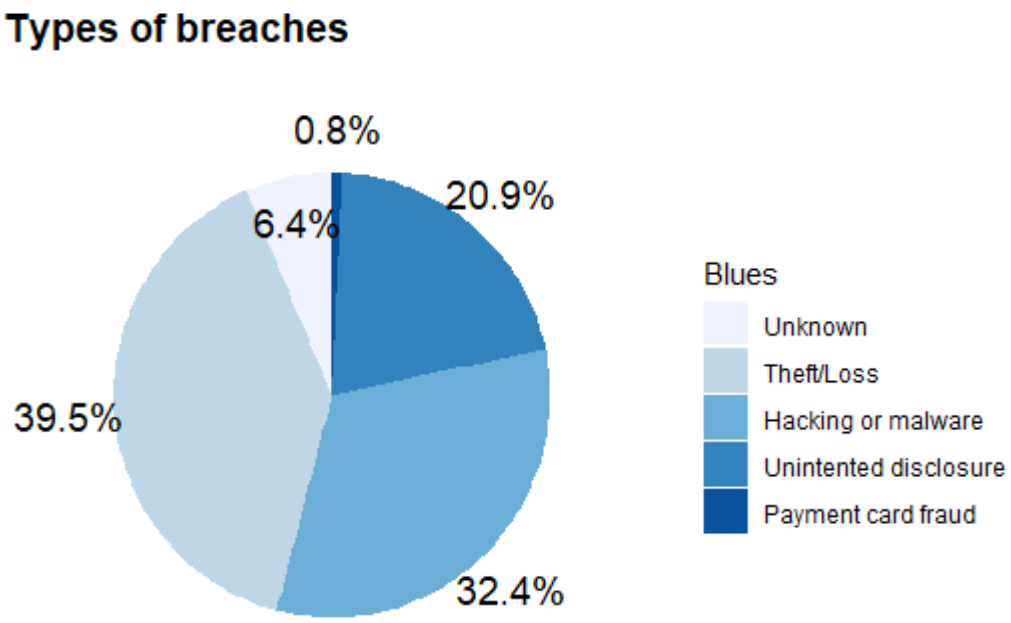
Privacy Rights Clearinghouse

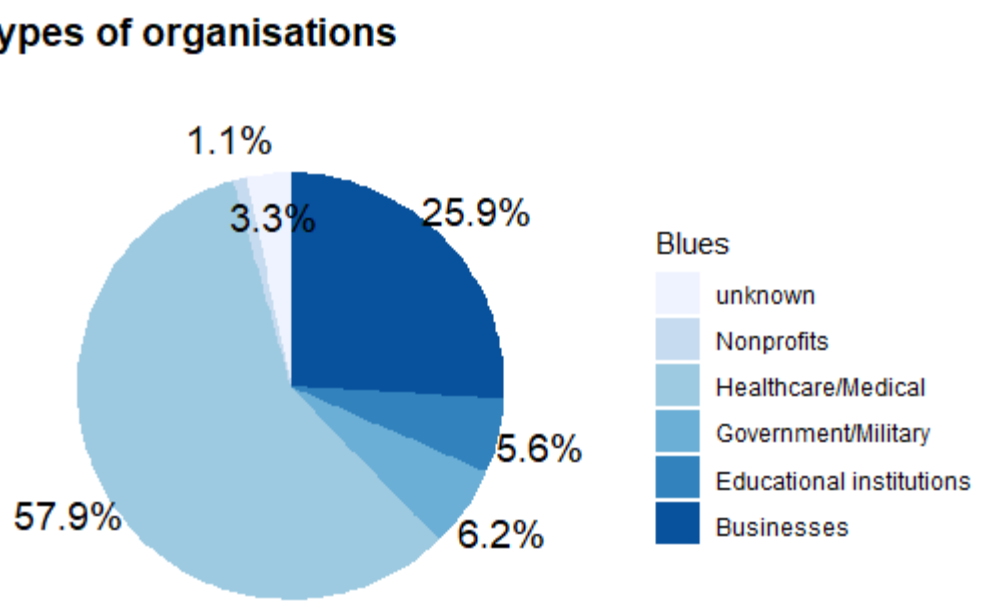
- A public database that contains 8800 data breaches in the US over the period 2005-2019
- Different types of sources reporting the cyber breaches to the database



Data breaches dataset

Descriptive statistics over 2010-2018



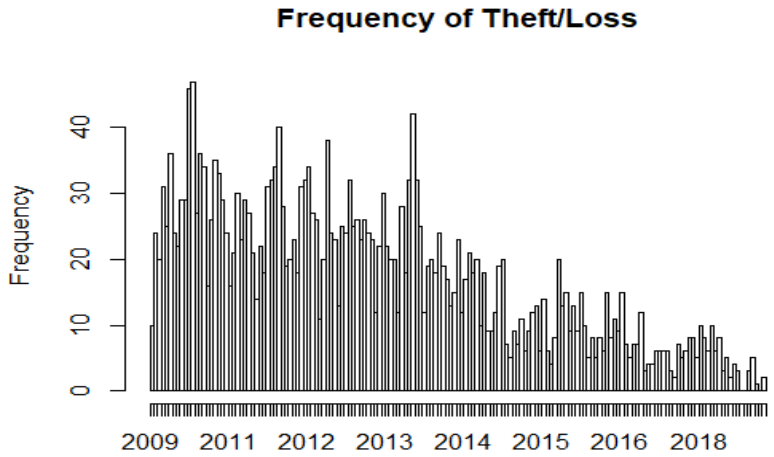


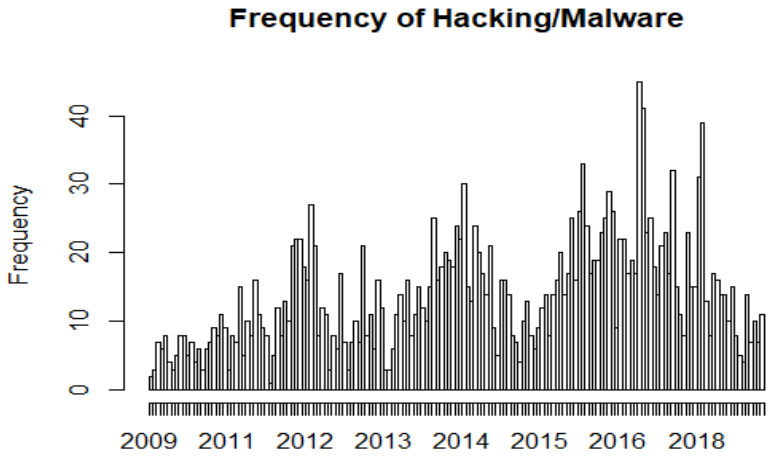
- A majority of **Theft/Loss** and **Hacking/Malware**
- 21% of Unintented disclosure

- A majority in **Healthcare/Medical**
- Businesses are well represented too

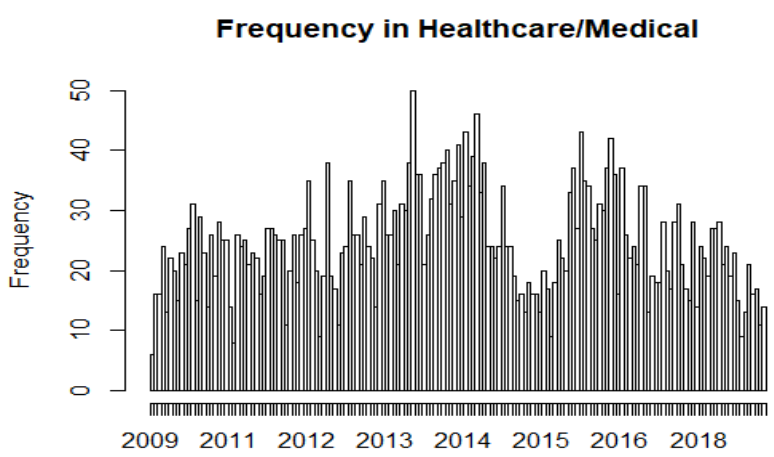
Data breaches dataset

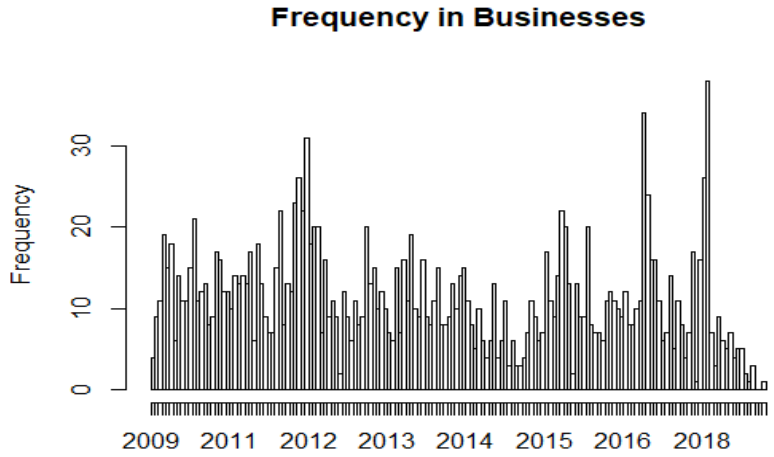
Cyber attacks frequencies by type and organization





- Apparent **clustering** by type of attacks
- Deterministic **trends** or stochastic regimes?



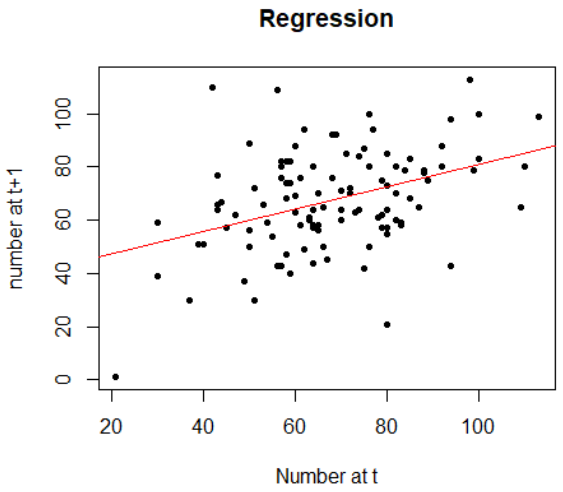


- Apparent **clustering** by type of organization attacked
- No clear **trends**

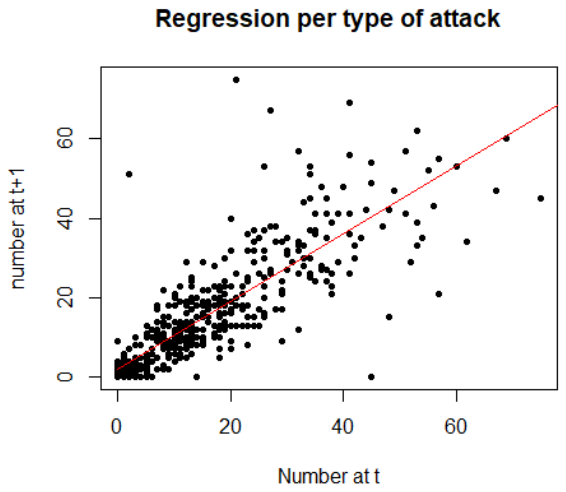
Data breaches dataset

Autocorrelation of the number of incidents

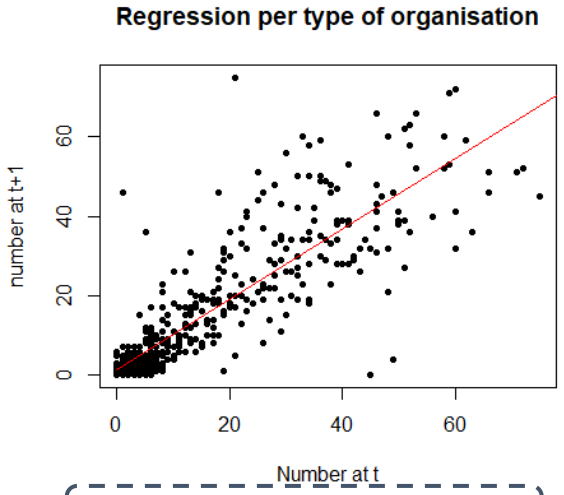
- Regression of the number of event during the **following month $t+1$** as a function of the number of event during the current month t → should be independent for a Poisson process model to be valid
- Autocorrelation dramatically increases when focusing on attacks and/or organizations of **the same type**



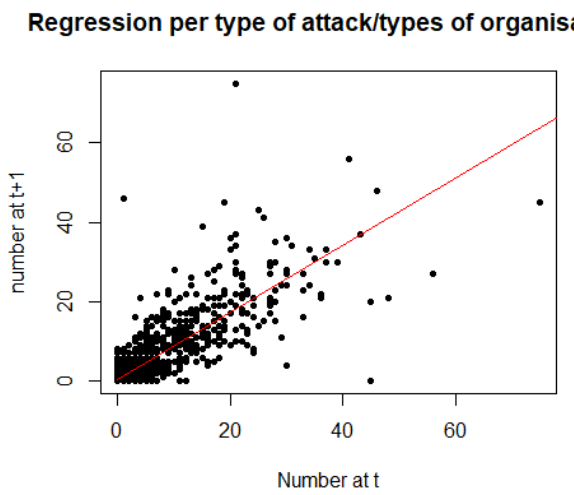
- R-squared : 0.154
- Confidence interval (95%) $[0.030, 0.278]$



- R-squared : 0.726
- Confidence interval (95%) $[0.687, 0.766]$



- R-squared : 0.780
- Confidence interval (95%) $[0.750, 0.810]$



- R-squared : 0.718
- Confidence interval (95%) $[0.702, 0.735]$

Hawkes model

Choice of the Hawkes model

- **Taking into account autocorrelation**
  - Cox model : Poisson model with stochastic intensity → difficulty to specify the stochastic intensity dynamics
  - Hawkes model : Self-exciting model with stochastic intensity, fully specified by the point process itself

- **Choice of the Hawkes model:**
  - Self-excitation: every event increases the probability for a new event to occur within a given group (same organization or attack type) → Clustering
  - Inter-excitation: in the case of multi-dimensional Hawkes process, every attack in one group increases the occurrence probability of new events in the other groups

- **Related references:**
  - Giesecke et al. (2010), Bacry et al. (2015) (finance)
  - Peng et al. (2017), Baldwin et al. (2017) (cyber risk)

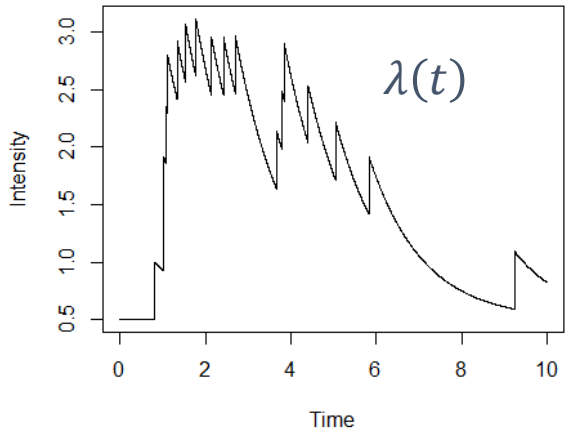
Univariate Hawkes process

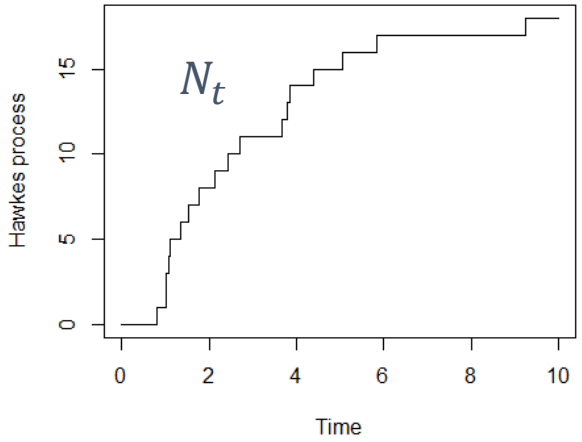
- An univariate Hawkes process with **exponential kernel** is a counting process $N_t = \sum_{n\geq 1} 1_{T_n \leq t}$ with intensity:

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

$\mu: \mathbb{R}_+ \to \mathbb{R}_+$ is a deterministic baseline intensity

- The sum represents the **impact of past events;** it captures the **self-excitation property**





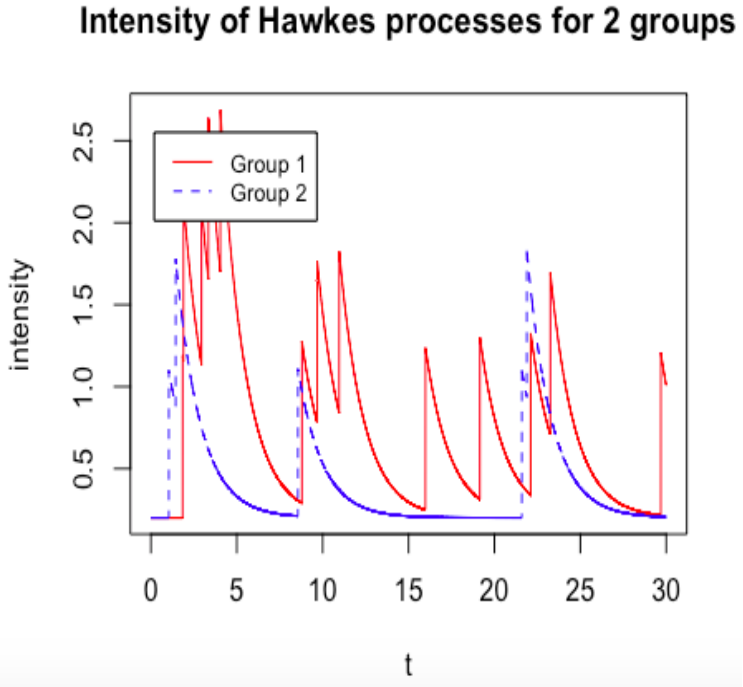
- **Each jump** represents an attack
- **Clustering phenomena**
- Intensity decreases **exponentially** between jumps

Multivariate Hawkes process

- Multivariate Hawkes process allows to model interactions **between types of entities/attacks/states**:
  $\left(N_t^{(1)}\right)_{t\geq 0}, \ldots, \left(N_t^{(K)}\right)_{t\geq 0}$, K counting processes with **jump times** $\left(T_n^{(1)}\right)_{n\geq 1}, \ldots, \left(T_n^{(K)}\right)_{n\geq 1}$
  The **intensity process** with exponential kernel of the counting process (i) is defined as:

$$\lambda_i(t) = \mu_i(t) + \sum_{j=1}^{K} \sum_{T_n^{(j)} < t} \alpha_{i,j} \exp\left\{-\beta_{i,j}(t - T_n^{(j)})\right\}$$

$\alpha_{i,j}$, $\beta_{i,j}$: **Impact of group j on group i**



Group 1 self-excitation

Impact of Group 2 on Group 1

Matrix of excitation:

$$\alpha = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{bmatrix} = \begin{bmatrix} 0.0 & 0.99 \\ 0.0 & 0.90 \end{bmatrix}$$

- Group 2 is **purely self-excited**
- Group 1 is **fully influenced by Group 2**

Kernels of multivariate Hawkes process

- « Classical » exponential kernel:

$$\phi_{i,j}(s) = \alpha_{i,j} \exp(-\beta_{i,j} s)$$

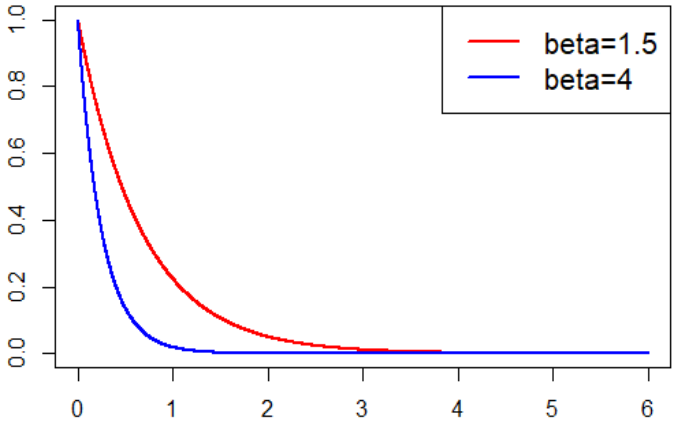
  - Instantaneous excitation
  - **Complexity**: we assume all $\beta_{i,j}$ (excitation memory of i from j) depend on the groups and are to be calibrated
  - The intensity process is not Markov (for dimension ≥ 2)

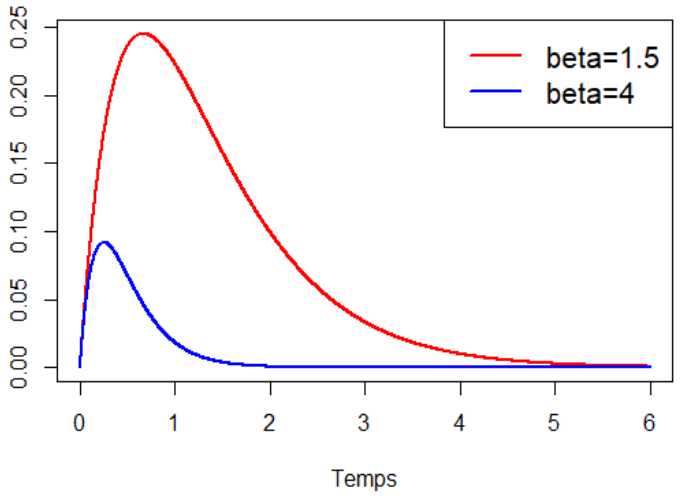
- Kernel with **delay**:

  - $\phi_{i,j}(s) = \alpha_{i,j} s \exp(-\beta_i s)$
  - The intensity process is not Markov (even in dimension 1)
  - **Complexity**: we assume all β_i (excitation memory of i from any other group) depend on the groups and are to be calibrated

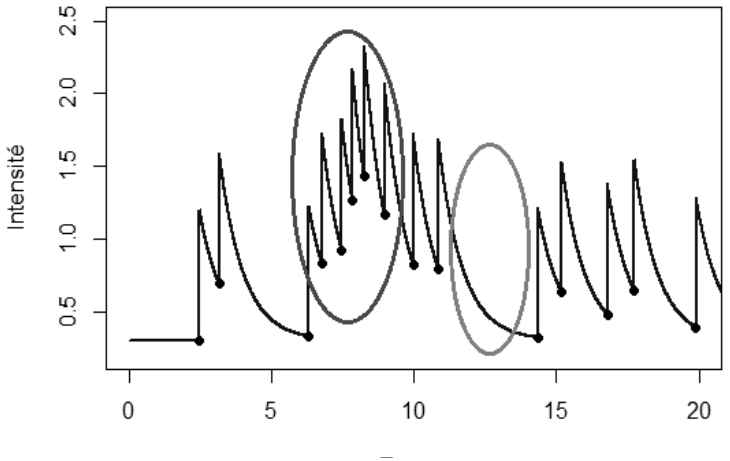
Development of closed-form formulas for the expected number of claims for such multivariate Hawkes processes

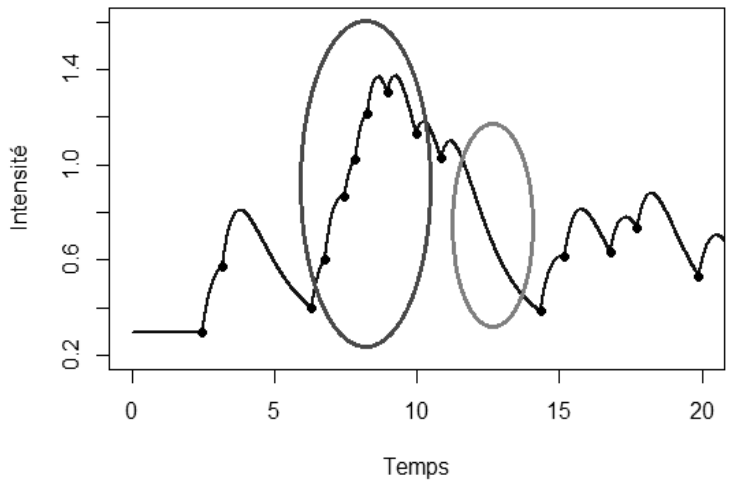
Kernels





Intensity process





Fitting and prediction

Data grouping

- **Crossing variables: attack type, sector, state**
  - Retaining groups with more than 200 attacks and remaining in OTHER
  - Total: **six groups**

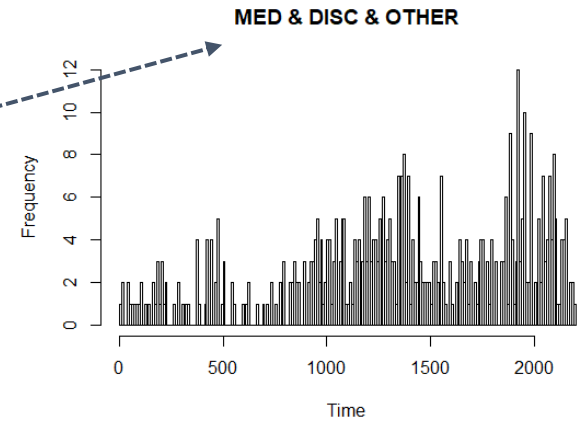
Group	Number of breaches
OTHER (1)	2046
MED & DISC & OTHER (2)	497
BUSINESSES & HACK & OTHER (3)	386
MED & HACK & OTHER (4)	472
MED & THEFT/LOSS & CALIFORNIA (5)	214
MED & THEFT/LOSS & OTHER (6)	943

Model specification

- **The three Hawkes kernels considered**

	Kernel 1 (exp)	Kernel 2 (exp)	Kernel 3 (delay)
$\phi_{i,j}(s)$	$\alpha_{i,j}\exp(-\boldsymbol{\beta_i} s)$	$\alpha_{i,j}\exp(-\boldsymbol{\beta_{i,j}} s)$	$\alpha_{i,j}\boldsymbol{s}\exp(-\boldsymbol{\beta_i} s)$
Nb parameters	54	84	54

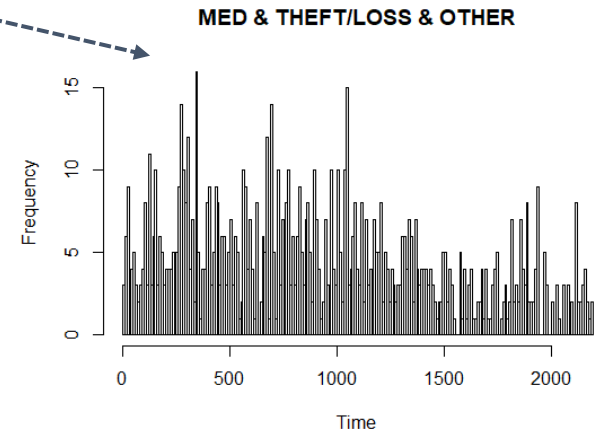
- **Baseline intensity**: $\boldsymbol{\mu_i(t) = \mu_{0,i} + \gamma_i t}$ to account for **trends** in the dataset



- **Possibility to add a Lasso penalty** (not discussed here)

$$L(\mu,\phi)_{penalized} = L(\mu,\phi) - \boldsymbol{\nu} \sum_{\boldsymbol{1\le i,j\le d}} |\boldsymbol{\alpha_{i,j}}|$$

  - Reduce complexity
  - Improve prediction capacity



Calibration results

- **Likelihood:** kernel 3 with delay best fits the data

	Kernel 1 (exp)	Kernel 2 (exp)	Kernel 3 (delay)
$\phi_{i,j}(s)$	$\alpha_{i,j}\exp(-\boldsymbol{\beta_i} s)$	$\alpha_{i,j}\exp(-\boldsymbol{\beta_{i,j}} s)$	$\alpha_{i,j}\boldsymbol{s}\exp(-\boldsymbol{\beta_i} s)$
Nb parameters	54	84	54
-Likelihood (2011-2015)	6513	6172	**6153**
-Likelihood (2011-2016)	7639	7516	**7485**

- **Adequacy tests (Kolmogorov-Smirnov):** adequacy is satisfactory, except for group (4)

OTHER (1)	**0.0503**	**0.0865**	**0.9060**
MED & DISC & OTHER (2)	**0.5546**	**0.1300**	**0.5173**
BUSINESSES & HACK & OTHER (3)	**0.5558**	**0.5966**	**0.3363**
MED & HACK & OTHER (4)	0.0024	0.0361	0.0370
MED & THEFT/LOSS & California (5)	**0.1146**	**0.5669**	**0.4246**
MED & THEFT/LOSS & OTHER (6)	**0.0733**	**0.6341**	**0.5379**

Calibration analysis

Captures the baseline non-excited intensity

Same orders of magnitude

captures the historical trend

	$\mu_0^{(i)}$	β_i	γ_i
OTHER (1)	0.87	5.39	-2.53e-04
MED & DISC & OTHER (2)	0.02	6.88	9.52e-05
Businesses & HACK & OTHER (3)	0.12	7.31	-3.56e-06
MED & HACK & OTHER (4)	0.02	5.75	9.65e-05
MED & Theft/Loss & CALIFORNIA (5)	0.05	5.96	-7.26e-06
MED & Theft/Loss & OTHER (6)	0.36	5.84	-1.07e-04

Table 7: Parameters $(\mu_0^{(i)})_{1\leq i\leq 6}$, $(\beta_i)_{1\leq i\leq 6}$ and $(\gamma_i)_{1\leq i\leq 6}$

	1	2	3	4	5	6
1	6.04	6.06	4.36	3.51	2.54	2.95
2	1.48	6.28	1.82	4.70	3.31	0.83
3	1.45	1.34	3.17	1.84	0.14	1.15
4	0.31	2.83	1.74	8.37	0.32	0.12
5	0.38	0.62	0.12	1.19	7.80	0.99
6	2.03	2.57	3.15	1.63	0.83	6.70

Captures the major self/external interactions

Table 8: Parameters $(\alpha_{i,j})_{1\leq i,j\leq 6}$

Strong reciprocal self-excitation of groups (2) and (4)

«Causal» excitation of (2) by (5)

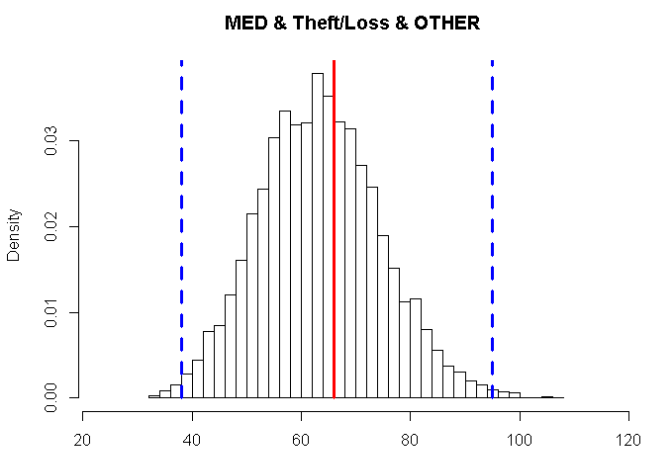
Strong self-excitation in these MED groups

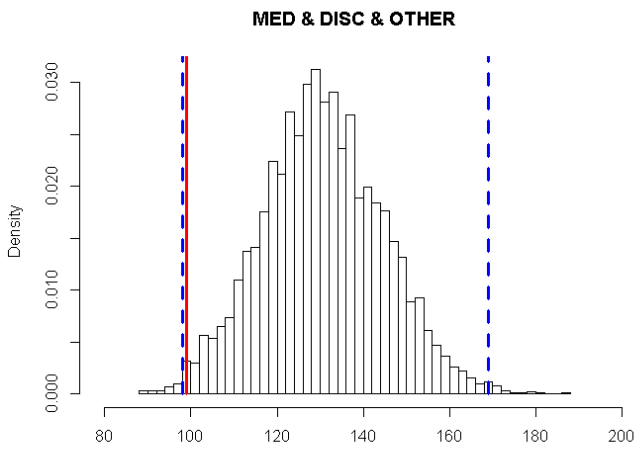
Matrix of ratios between maximal excitation and baseline intensity:

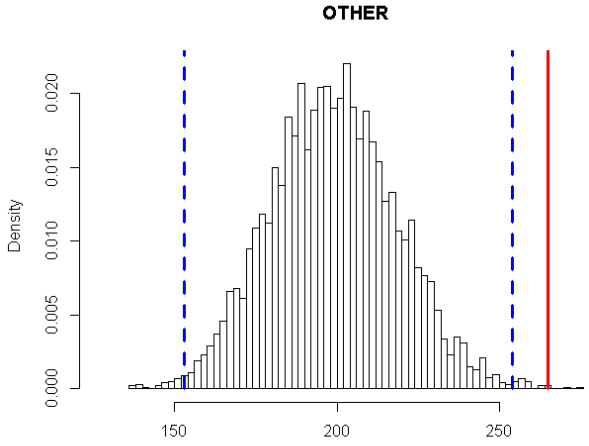
	1	2	3	4	5	6
1	0.47	0.47	0.3	0.28	0.19	0.23
2	4	17	5	12.5	9	2
3	0.58	0.58	1.33	0.75	0.08	0.5
4	1	9	5.5	26.5	1	0.5
5	0.4	0.8	0.2	1.4	9.6	1.2
6	0.36	0.44	0.56	0.28	0.14	1.17

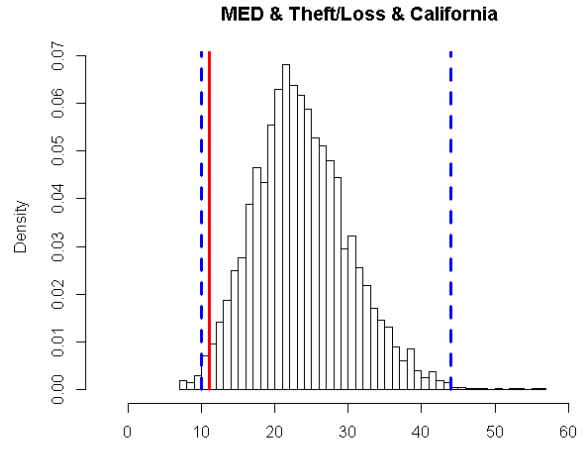
Table 10: Ratios Maximum excitation/basic intensity $(\frac{\Gamma_{i,j}}{\mu_0^{(i)}})_{1\leq i,j\leq 6}$

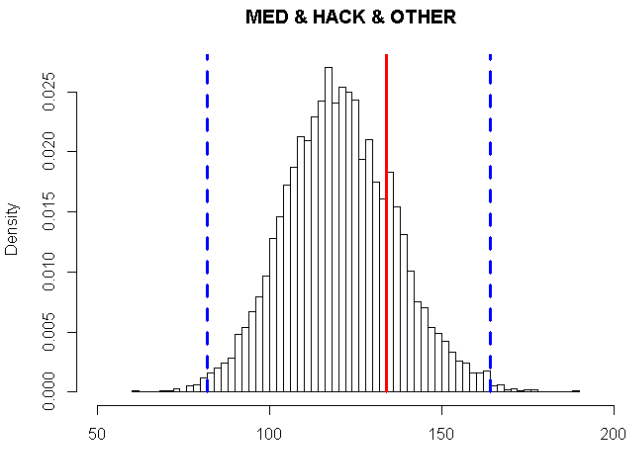
Out-of-sample prediction results for 2017 (kernel 3)

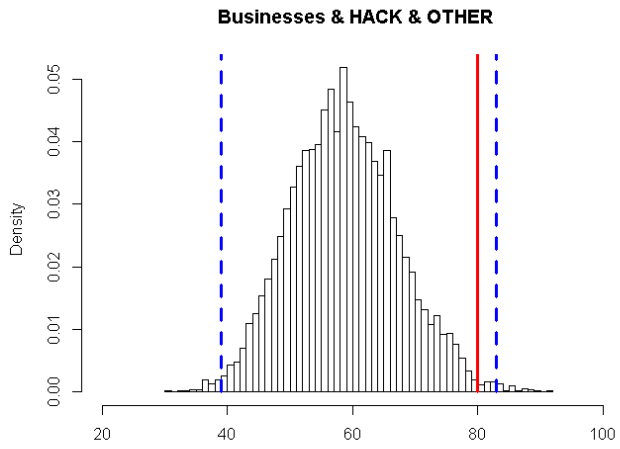












Simulation based on the thinning algorithm for point processes

Predictions with mean and (0.5%, 99.5%) percentiles

Joint prediction of all groups capturing the causal and asymmetric interactions

Parameter uncertainty can be added

Conclusion

- **Take-away message**
  - Heterogeneity of the database: the choice of the groups is determinant for the prediction accuracy.
  - Auto/inter-excitation, calibrated using multivariate Hawkes processes (and kernel with delay)
  - Projection: whole joint distribution of the events' arrivals (and not only the marginal distributions)

- **For further study**
  - Risk Exposure
  - Severity of cyber risk: see Sébastien Farkas' talk

References

Paper available at https://hal.archives-ouvertes.fr/hal-02546343

- Bacry, E., Mastromatteo, I., Muzy J.F. (2015). Hawkes processes in finance. Market Microstructure and Liquidity, 1(01):1550005,
- Baldwin, A., Gheyas, I., Ioannidis, C., Pym, D., & Williams, J. (2017). Contagion in cyber security attacks. Journal of the Operational Research Society, 68(7), 780-791.
- Böhme, R., & Kataria, G. (2006, June). Models and Measures for Correlation in Cyber-Insurance. In WEIS.
- Boumezoued, A. (2016). Population viewpoint on Hawkes processes. Advances in Applied Probability, 48(2), 463-480.
- Daley, D. J., & Vere-Jones, D. (2007). An introduction to the theory of point processes: volume II: general theory and structure. Springer Science & Business Media.
- Edwards, B., Hofmeyr, S., & Forrest, S. (2016). Hype and heavy tails: A closer look at data breaches. Journal of Cybersecurity, 2(1), 3-14.
- Errais, E., Giesecke, K., & Goldberg, L. (2010) Affine point processes and portfolio credit risk. SIAM Journal on Financial Mathematics, 1(1):642– 665.
- Farkas, S., Lopez, O., & Thomas M. (2020) Cyber claim analysis through generalized pareto regression trees with applications to insurance. https://hal.archives-ouvertes.fr/hal-02118080v2/document.
- Hawkes, Alan G. (1971). Spectra of some self-exciting and mutually exciting point processes. Biometrika, 58(1), 83-90.
- Hawkes, Alan G, David Oakes. (1974). A cluster process representation of a self-exciting process. J. of Applied Probability 493–503.
- Oakes, David. (1975). The Markovian self-exciting process. Journal of Applied Probability 69–77.
- Peng, C., Xu, M., Xu, S., & Hu, T. (2017). Modeling and predicting extreme cyber attack rates via marked point processes. Journal of Applied Statistics, 44(14), 2534-2563.
- Xu, M., & Hua, L. (2019). Cybersecurity Insurance: Modeling and Pricing. North American Actuarial Journal, 1-30.

Thank you for your attention

Contact details :

Caroline Hillairet

Ensae Paris, Crest
5 Avenue Le Chatelier,
91120 Palaiseau
France

Caroline.Hillairet@ensae.fr

https://www.actuarialcolloquium2020.com/

Disclaimer:

The views or opinions expressed in this presentation are those of the authors and do not necessarily reflect official policies or positions of the Institut des Actuaires (IA), the International Actuarial Association (IAA) and its Sections.

While every effort has been made to ensure the accuracy and completeness of the material, the IA, IAA and authors give no warranty in that regard and reject any responsibility or liability for any loss or damage incurred through the use of, or reliance upon, the information contained therein. Reproduction and translations are permitted with mention of the source.

Permission is granted to make brief excerpts of the presentation for a published review. Permission is also granted to make limited numbers of copies of items in this presentation for personal, internal, classroom or other instructional use, on condition that the foregoing copyright notice is used so as to give reasonable notice of the author, the IA and the IAA's copyrights. This consent for free limited copying without prior consent of the author, IA or the IAA does not extend to making copies for general distribution, for advertising or promotional purposes, for inclusion in new collective works or for resale.