Predicting Cyber-Attacks using the 'Hawkes' R Package

Alexandre Boumezoued, Milliman R&D

Joint work with Caroline Hillairet (ENSAE) and Yannick Bessy-Roland (Milliman R&D)

Insurance Data Science Conference, June 14th

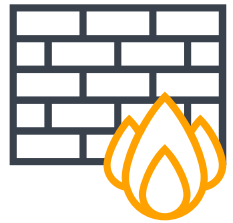
The authors thank for their help: Tahayacine Bourassine, Louis François, Henri Perillat, Rayan Sanhaj

Milliman

Agenda

Introduction

Example of types of cyber attacks



Phishing

- The attacker sends a document appearing reliable (mainly e-mail) in order to collect sensitive information

Malware

- Software that is specifically designed to disrupt, damage, or gain unauthorized access to a computer system

Man in the middle

- The attacker secretly relays and possibly alters the communications between two parties who believe they are directly communicating with each other

- **Definition: The risk of an attack on digital data as well as the consequences on the information system**

Denial of service

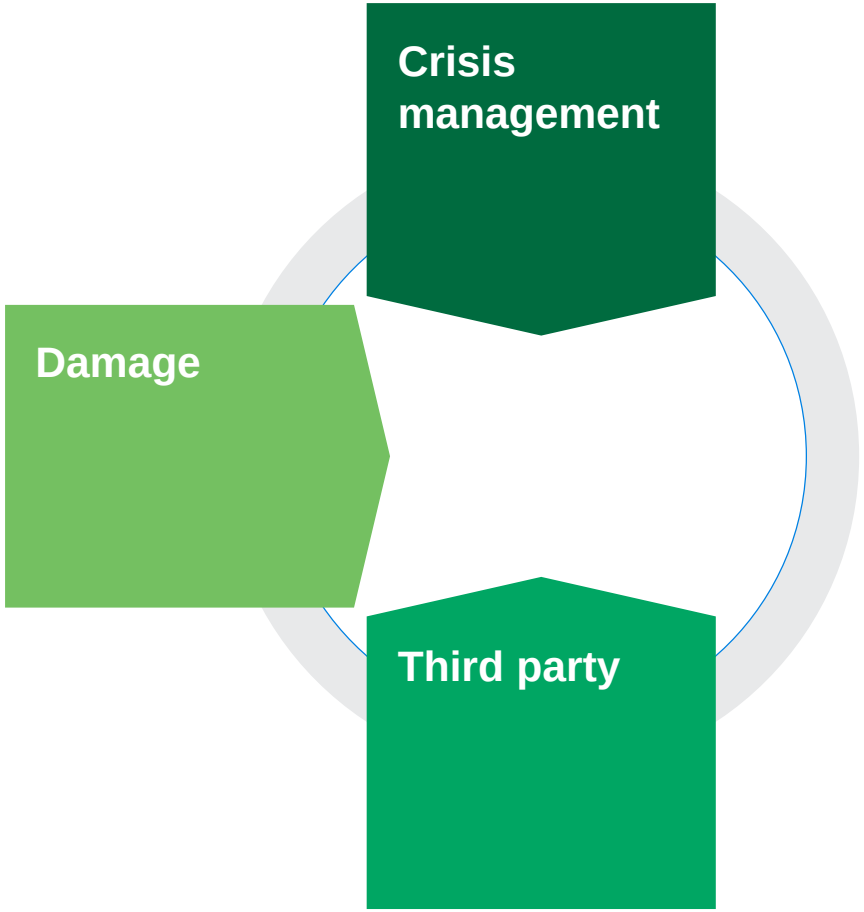
- Attack in which the perpetrator seeks to make a machine or network resource unavailable to its intended users by disrupting services of a host connected to the Internet

Zero day

- Attack that occurs on the same day a weakness is discovered in software. At that point, it's exploited before a fix becomes available from its creator

Introduction

Cyber insurance covers



Cyber insurance

- **Crisis management:**
  - Costs of investigation
  - Costs of assistance
  - Other costs of crisis management
- **Damage:**
  - Data cleaning
  - Data restauration
  - Payment of ransom
  - Operating loss
- **Third party liability:**
  - Virus transmission
  - Personal liability insurance
  - Denial of service

Agenda

1. Introduction
2. **Data breaches dataset**
3. Hawkes model
4. Fitting and prediction with R
5. References

Data breaches dataset

The Privacy Rights Clearinghouse database

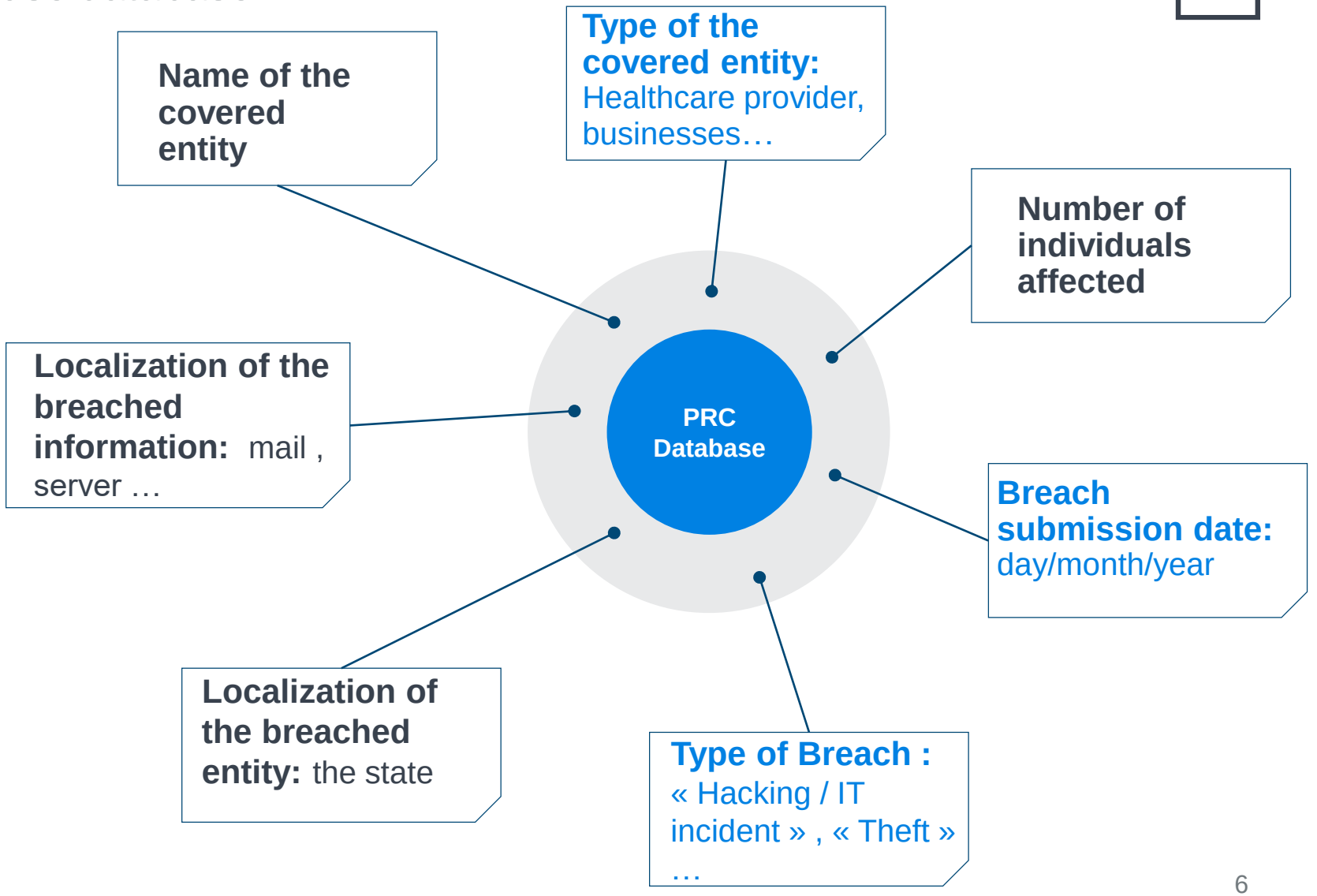
- **Description**

A public database that contains 8871 data breaches in the US over the period 2005-2018

https://www.privacyrights.org/data-breaches

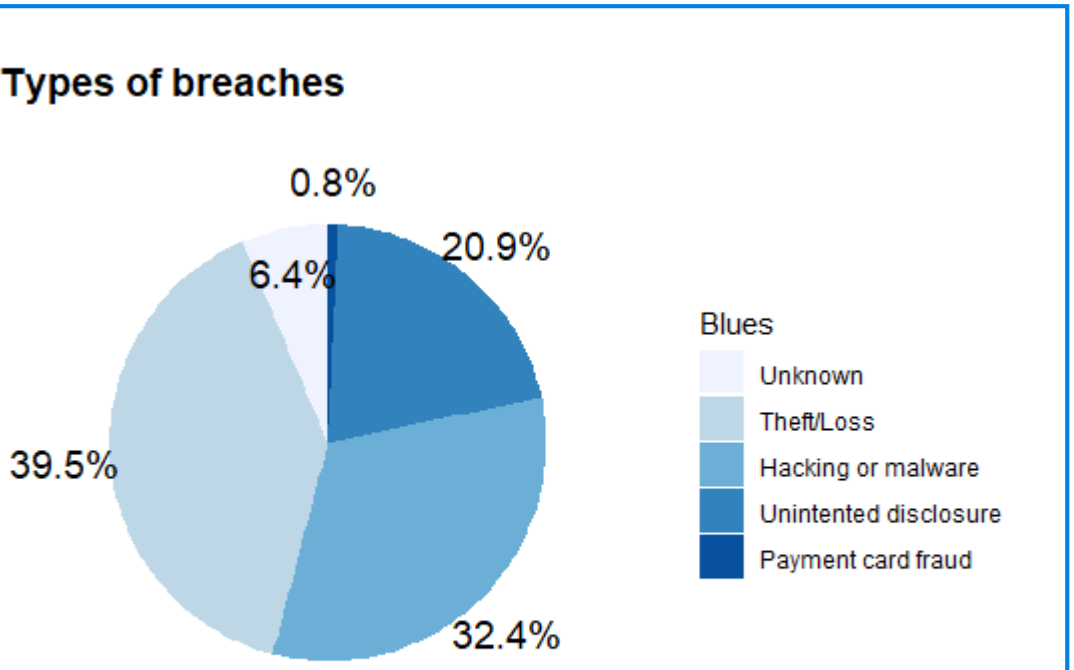
- **Covariates**

Those used in the results presented are highlighted in blue

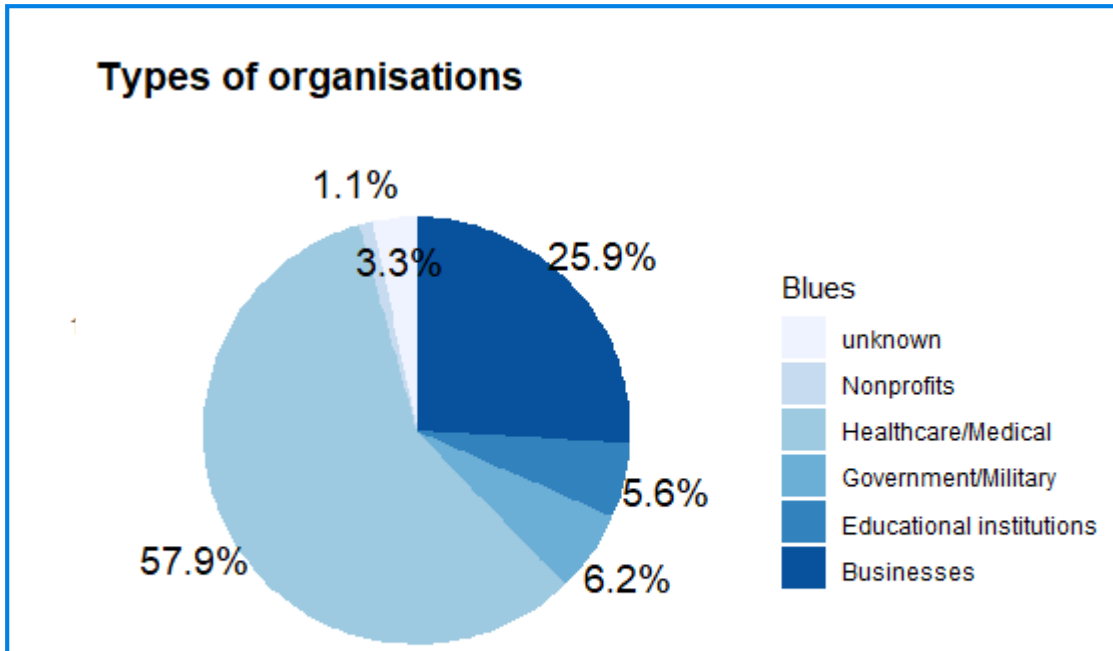


Data breaches dataset

Descriptive statistics over 2010-2018



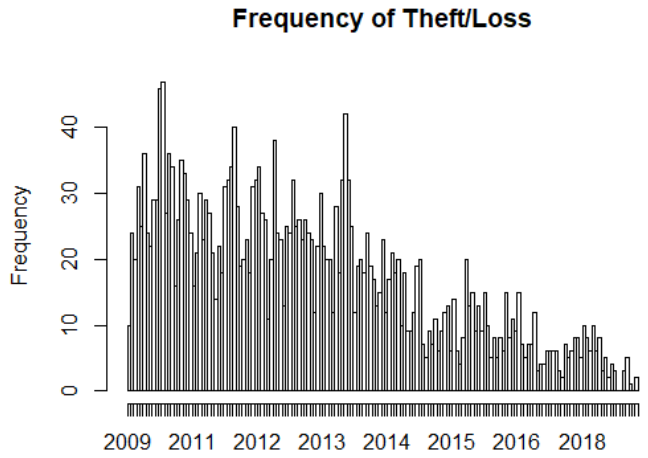
- A majority of **Theft/Loss** and **Hacking/Malware**
- 21% of Unintented disclosure

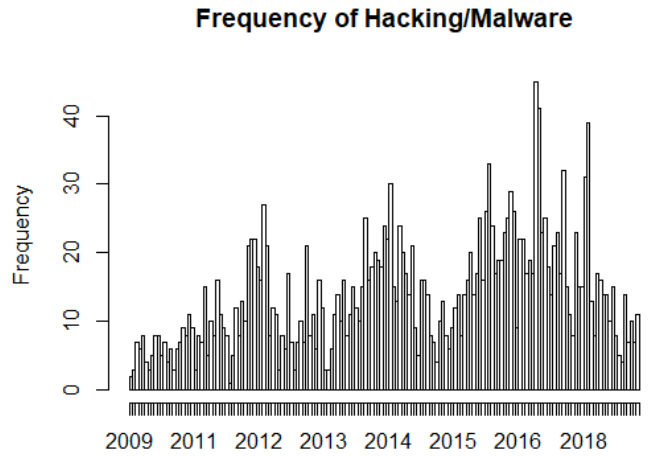


- A majority in **Healthcare/Medical**
- Businesses are well represented too

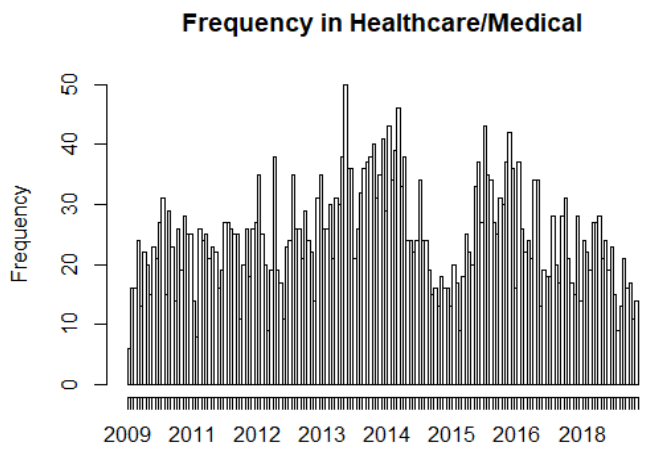
Data breaches dataset

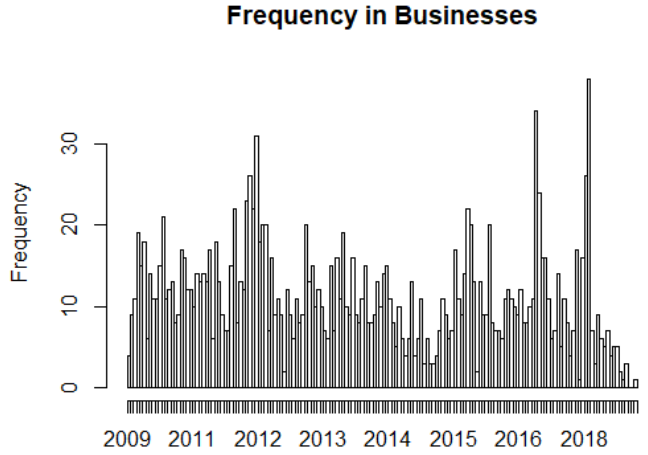
Cyber attacks frequencies by type and organization





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



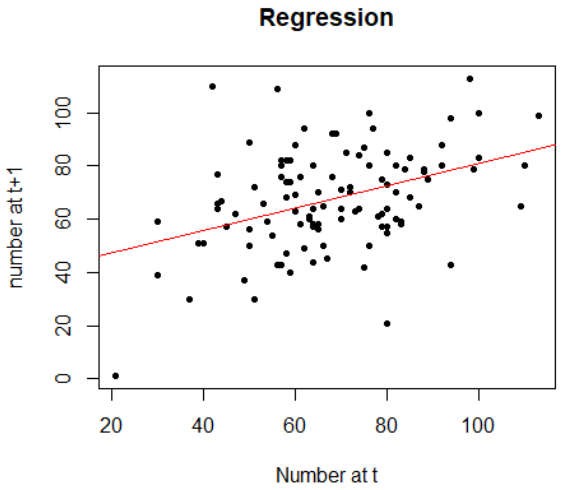


- Apparent **clustering** by type of organization attacked
- No clear **trends**

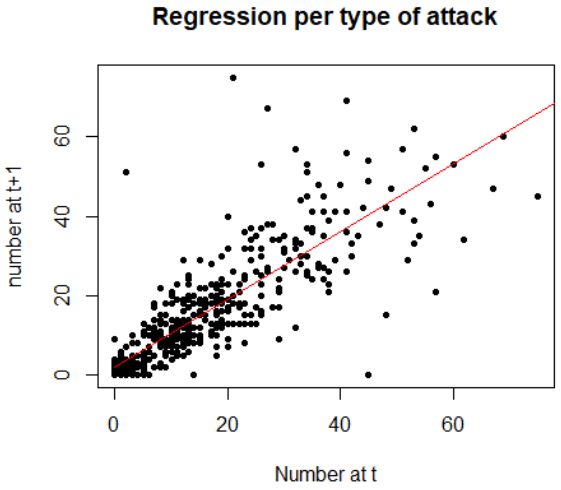
Data breaches dataset

Autocorrelation of the number of incidents

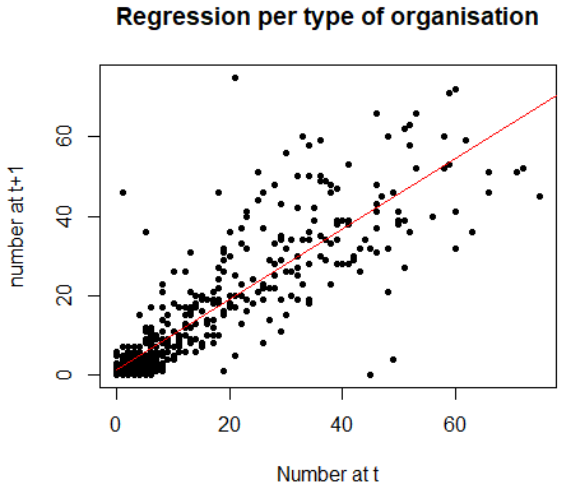
- Regression of the number of event during the **following month $t+1$** as a function of the number of event during the current month t → should be independent for a Poisson process model to be valid
- Autocorrelation dramatically increases when focusing on attacks and/or organisations of **the same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



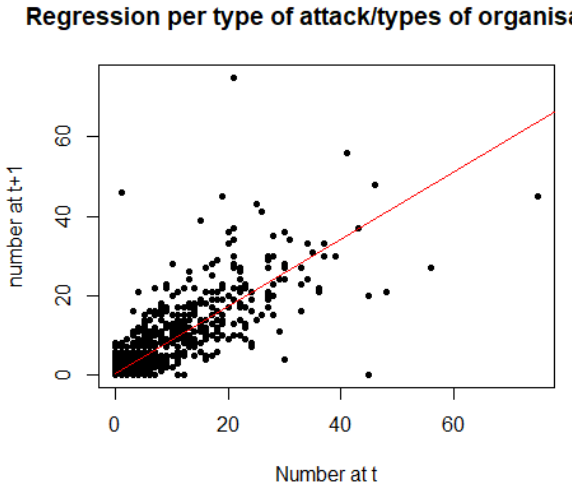
- R-squared : 0.154
- Confidence interval (95%) $[0.030, 0.278]$



- 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%) $[0.702, 0.735]$

Agenda

1. Introduction
2. Data breaches dataset
3. **Hawkes model**
4. Fitting and prediction with R
5. References

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: the self-exciting property allows to model cluster effect (groups of attacks – same origin!)
  - 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
- **Use of the 'hawkes' R package:**
  - *Riadh Zaatour (2014). hawkes: Hawkes process simulation and calibration toolkit. R package version 0.0-4.*
- **Related references:**
  - Peng et al. (2017), Baldwin et al. (2017)

Hawkes model

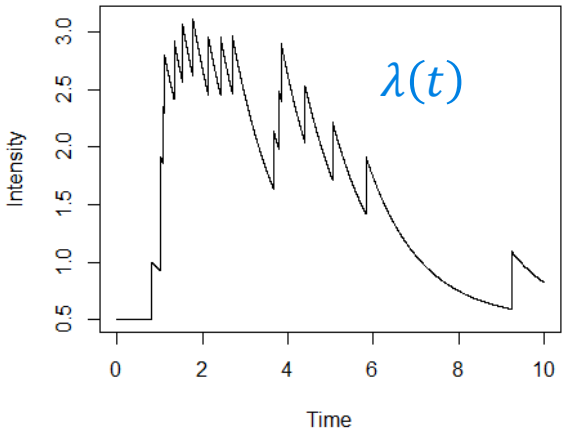
Univariate Hawkes process

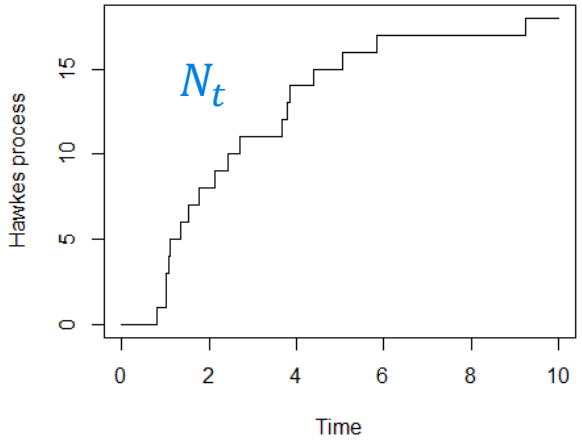
- A Hawkes process with **exponential kernel** is a counting process $N_t = \sum_{n\geq 1} 1_{T_n \leq t}$ with intensity:

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

  - $\mu: \mathbb{R}_+ \to \mathbb{R}_+$ 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

Hawkes model

Multivariate Hawkes process

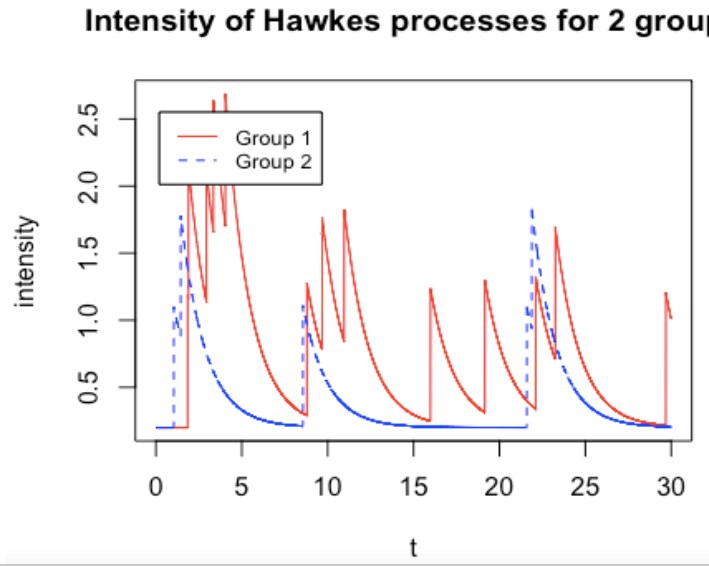
- Multivariate Hawkes process allows to model interactions **between types of entities/attacks**:
  - $\left(N_t^{(1)}\right)_{t\geq 0}, \ldots, \left(N_t^{(K)}\right)_{t\geq 0}$, K counting processes with **jump times** $\left(T_n^{(1)}\right)_{n\geq 1}, \ldots, \left(T_n^{(K)}\right)_{n\geq 1}$
  -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

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

Impact of Group j on Group i

Group 1 self-excitation

Impact of Group 2 on Group 1



Matrix of excitation:

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

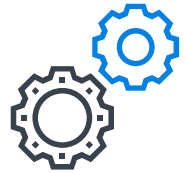
- Group 2 is **purely self-excited**
- Group 1 is **fully influenced by Group 2**

Agenda

1. Introduction
2. Data breaches dataset
3. Hawkes model
4. **Fitting and prediction with R**
5. References

Fitting and prediction with R

Using the ‘hawkes’ R package for calibration



- The function **likelihoodHawkes** returns the opposite of the likelihood given parameters and times

- Given a set of times from the dataset, one can maximize the likelihood using for instance the **constrOptim** function, which allows optimization under constraints (to ensure non-negative parameters)

```
likeliHawkes=function(parameters){
mu0=parameters[1]
alpha0=parameters[2]
beta0=parameters[3]
return(likelihoodHawkes(mu0,alpha0,beta0,observed_times))
}
constrOptim(theta=c(2,0.2,1),f=likeliHawkes,ui=cbind(c(1,0,0),c(0,1,0),c(0,0,1)),ci=c(0,0,0),grad = NULL)$par
```

Fitting and prediction with R

Calibration results on 2010-2016

By type of attack:

Theft/Loss $\left(N_t^{(1)}\right)_{t\geq 0}$; Hacking/Malware $\left(N_t^{(2)}\right)_{t\geq 0}$

μ_1	μ_2	$\alpha_{1,1}$	$\alpha_{2,2}$	β_1	β_2	$\alpha_{1,2}$	$\alpha_{2,1}$
0.58	0.12	0.76	0.05	1.94	0.06	0.1	<e-6

- Limited background intensity for Hacking/Malware in comparison to Theft/Loss : $\mu_2 \ll \mu_1$
- Self-excitation for Theft/Loss is initially significantly higher than that of Hacking/Malware: $\alpha_{1,1} \gg \alpha_{2,2}$, but vanishes more rapidly: $\beta_1 \gg \beta_2$
- Hacking/Malware is a leading attack type as it triggers Theft/Loss events, but the reverse does not hold: $\alpha_{1,2} \gg \alpha_{2,1} \approx 0$

By type of entity:

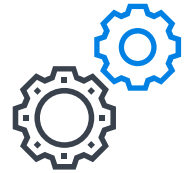
Healthcare/Med. $\left(N_t^{(1)}\right)_{t\geq 0}$; Businesses $\left(N_t^{(2)}\right)_{t\geq 0}$

μ_1	μ_2	$\alpha_{1,1}$	$\alpha_{2,2}$	β_1	β_2	$\alpha_{1,2}$	$\alpha_{2,1}$
0.73	0.38	0.98	0.38	2.71	2.03	0.41	0.14

- Significant background intensities for both Healthcare/Med. and Businesses, with $\mu_1 > \mu_2$
- In terms of self-excitation, both significant; the initial Healthcare/Med. self-excitation is higher $\alpha_{1,1} > \alpha_{2,2}$ but vanishes more rapidly: $\beta_1 > \beta_2$
- Businesses attacks initially trigger more significantly potential future Healthcare/Med. attacks in comparison ($\alpha_{1,2} > \alpha_{2,1}$), although the effect vanishes more rapidly (in comparison again): $\beta_1 > \beta_2$

Fitting and prediction with R

Using the ‘hawkes’ R package for simulation



- The function **simulateHawkes** generates a Hawkes process given the parameters

- An example of simulation of a monovariate Hawkes process is provided below:

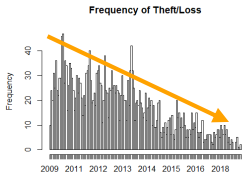
```
alpha0=0.5
beta0=0.8
mu0=1
theta0=0.03
tau=5
simulated_times=simulateHawkes(mu0, alpha0, beta0, tau)[ [ 1 ] ]
```

Fitting and prediction with R

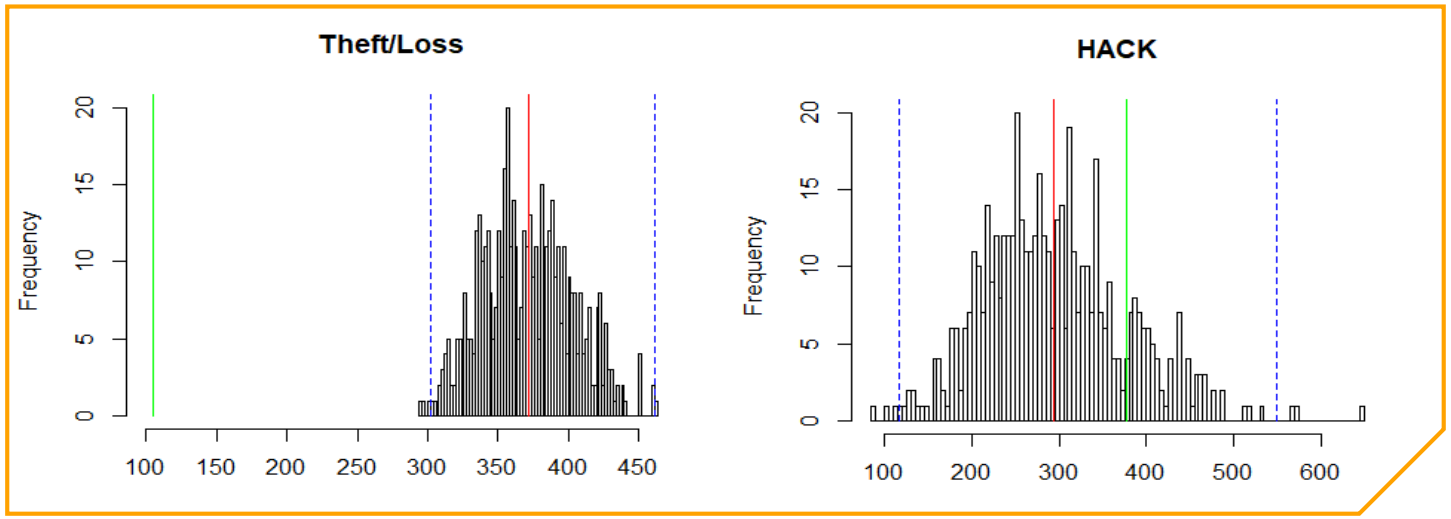
Out-of-sample prediction results for 2017

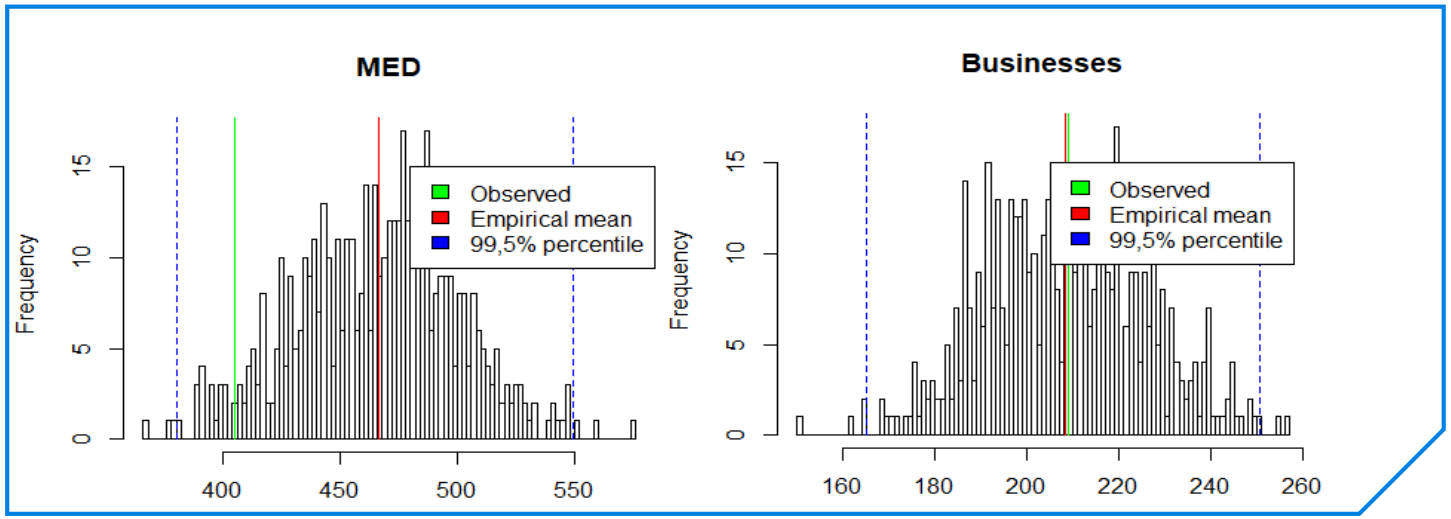
- **Joint forecasts** based on multivariate Hawkes model for either types of attacks or types of entity
- Good prediction results for Hacking/Malware, but poor results for Theft/Loss due to the non account for the clear trend in the data

→ possibility to specify downward linear trend $\mu(t)$



- Good prediction results by sector, where auto-correlation is high and no clear deterministic trend appears
- **Overall, ability for the model to capture the average magnitude (pricing/reserving) and provides a full distribution (capital requirement)**





Milliman

Agenda

1. Introduction
2. Data breaches dataset
3. Hawkes model
4. Fitting and prediction with R
5. **References**

References

- **Hawkes and point processes:**
  - 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.
  - 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.
- **Cyber risk modelling:**
  - Böhme, R., & Kataria, G. (2006, June). Models and Measures for Correlation in Cyber-Insurance. In WEIS.
  - Edwards, B., Hofmeyr, S., & Forrest, S. (2016). Hype and heavy tails: A closer look at data breaches. Journal of Cybersecurity, 2(1), 3-14.
  - 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.
  - 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.
- **Hawkes R package:**
  - Riadh Zaatour (2014). hawkes: Hawkes process simulation and calibration toolkit. R package version 0.0-4.

Disclaimer