Designing stochastic accumulation scenarii for cyber-insurance.

Caroline Hillairet[1], Olivier Lopez[2].

April 22, 2020

Abstract

In this paper, we develop stochastic models to determine the impact of a massive cyber attack on an insurance portfolio. The model is based on the classical SIR (Susceptible - Infected - Recovered) epidemiological model. For a given type of attack, we provide a general framework to quantify the impact on the portfolio of such an event, and calibrate response policies for the insurance company (such as prevention and reaction time to the attack). We also consider the possibility of a "collapse" of the response system, that could happen if too many policyholders are affected simultaneously. In which case, the insurance company is unable to bring assistance to the whole affected policyholders. We provide sharp bounds for the probability that such an event occur, and a way to simulate "cyber-hurricanes". As an illustration, we replicate the impact of a Wannacry-type event on an insurance portfolio.

Key words: cyber insurance; emerging risks; epidemiological models; risk theory.

Short title: Accumulation scenarii for cyber-insurance.

[1] Ensae Paris, Crest, 5 Avenue Le Chatelier, F-91120 Palaiseau, France.

[2] Sorbonne Université, CNRS, Laboratoire de Probabilités, Statistique et Modélisation, LPSM, 4 place Jussieu, F-75005 Paris, France,

E-mails: caroline.hillairet@ensae.fr, olivier.lopez@sorbonne-universite.fr

1 Introduction

Cyber-insurance is a rising field responding to the growing weight of cyber-risk. Apart from evaluating the risk in a classical frequency/severity paradigm, many massive cyber events (like Wannacry or NotPetya for example) raised a major concern for insurers and risk managers: how is it possible to absorb an accumulation scenario? By accumulation, one means that, due to the systemic characteristic of cyber-risk, loss of mutualization can occur if a large number of policyholders is attacked in a short amount of time. In this paper, we propose a general framework to design and simulate accumulation scenarii, in order to understand their impact on an insurance portfolio. In particular we provide a framework that allows to quantify the impact of prevention and fast reaction to diminish the cost of such an episode.

The rest of the paper is organized as follows. In Section 2, we describe the general model. Section 3 provides theoretical results allowing to approximate the evolution of the portfolio. Section 4 presents the results of the simulation of a Wannacry-type event.

2 Portfolio

2.1 Modeling the effects of a cyber attack on policyholders

Our model to describe the effect of a cyber-incident on a portfolio is decomposed in three parts:

- an "infection" model, describing how likely the policyholders may be stroke by the attack;
- a "recovery" model, to consider the time during which the policyholder requires assistance;
- a model for the reaction and prevention, that is how long does it take to identify the threat and how fast the policyholders react in implementing the patches.

Considering n policyholders, the way the $j-$th policyholder is affected by the attack is therefore described by the following three time random variables, each corresponding to one of the three parts of the model as listed above:

- T_j = time at which he is infected (may be infinite);

- U_j = duration of the recovery period (for him, the crisis ends at time $T_j + U_j$);
- C_j = time at which he implements security changes that make him immune to the attack.

Let us define $Y_j = \inf(T_j, C_j)$ and $\delta_j = \mathbf{1}_{T_j \leq C_j}$. Y_j is the time at which the j-th policyholder is no longer susceptible to be affected, either because immunity has been acquired, or because contamination has occurred. Moreover, if $\delta_j = 1$, then the policyholder has been affected by the attack. Otherwise, the reaction has been fast enough, preventing him from damages. We assume that $(T_j, U_j)_{1 \leq j \leq n}$ are i.i.d. This independence assumption may not be entirely true, since some of the policyholders may be in contact, and thus can be able to transmit a malware more easily to one another. However, if we assume the portfolio to be large enough and if the subscription policy has avoided to constitute significant clusters of connected policyholders, such phenomena should be marginal and thus can be neglected. Typically this independence assumption reflects the fact that, if n is much smaller that the size of the national population (or even global population), the infection is more likely to come from outside the portfolio that from inside.

The three variables (T, U, C) can be modeled using their hazard rate. For a continuous random variable Z, the corresponding hazard rate function, denoted λ_Z in the following, is defined as

$$\lambda_Z(t) = \lim_{dt \to 0^+} \mathbb{P}(Z \in [t, t + dt] | Z \geq t).$$

Modeling λ_T corresponds to modeling the dynamic evolution of the cyber episode. It reflects the severity of the crisis at a global level. On the other hand, λ_U and λ_C reflects the reaction of the insurance to the events: how fast the company is able to assist its customers (λ_U) and how fast and efficiently it can perform prevention to diminish the intensity of the crisis.

2.2 Measuring the impact on the portfolio

To measure the impact of a cyber event at a portfolio level, let us introduce the following notations,

$$\begin{aligned}\mathfrak{N}_t &= \sum_{j=1}^{n} \delta_j \mathbf{1}_{Y_j \leq t},\\ \mathfrak{R}_t &= \sum_{j=1}^{n} \delta_j \mathbf{1}_{Y_j + U_j \leq t},\\ \mathfrak{I}_t &= \mathfrak{N}_t - \mathfrak{R}_t.\end{aligned}$$

$\mathfrak{N}_t$ denotes the cumulative number of infected policyholders at time t, while $\mathfrak{R}_t$ is the number of infected who have recovered before time t, and $\mathfrak{I}_t$ is the number for which the crisis is still ongoing at time t. By recovery, we do not mean "full recovery" (that is retrieving the same level of activity and having compensated the losses): the timescale for full recovery may be much larger than the duration of the crisis (weeks or months, compared to days). This recovery time only refers to the immediate help that is required by the policyholder.

Indeed, our purpose is to evaluate the risk of being unable to deliver the assistance because too many policyholders are affected. Since assistance to victims is a key element of many cyber-insurance contracts, if $\mathfrak{I}_t$ exceeds some threshold (which is the total capacity of response of the insurance company), important penalties can strike the company which becomes unable to correctly execute the contract. Moreover, the more policyholders have to be assisted at a given time t, the higher the costs of assistance may be. This is why, in Section 3, we focus on approximations of the distribution of $(\mathfrak{I}_t)_{t\geq 0}$.

3 Approximation of the evolution of the portfolio through Gaussian processes

Let us introduce some notations before stating our theoretical results. Let $S_C(t) = \mathbb{P}(C \geq t)$ denote the survival function of the variable C, $F_U(t) = \mathbb{P}(U \leq t)$, and

$$\begin{aligned}
\phi_{Y,U}(t,h) &= \int_0^t S_C(u)\left\{F_U(t+h) - F_U(u)\right\} f_T(u)du, \\
\nu(t) &= \int_0^t S_C(u) f_T(u)du, \\
\rho(t) &= \int_0^t S_C(u) f_V(u)du,
\end{aligned}$$

where $V = T + U$ and f_T (resp. f_V) is the density of T (resp. V).

The following result shows that, if the size of the portfolio n is large enough, the distribution of the process $(\mathfrak{I}_t)_{t\geq 0}$ can be approximated by a Gaussian process.

Theorem 3.1 *Let $\iota_1(t) = \nu(t) - \phi_{Y,U}(t,h)$, $\iota_2(t) = \nu(t) - \rho(t)$.*
Then, define

$$\mathfrak{Z}_{n,t} = \frac{(\mathfrak{I}_t - n\iota_2(t))}{n^{1/2}}.$$

The process $(\mathfrak{Z}_{n,t})_{t\geq 0}$ converges in distribution towards a centered Gaussian process with covariance function $\sigma_{\mathfrak{Z}}(t, t+h) = \iota_1(t) - \iota_2(t)\iota_2(t+h)$, for all $t \geq 0$, $h \geq 0$.

In other words, the central scenario for the evolution of $(\mathfrak{I}_t)_{t\geq 0}$ is $t \to n\iota_2(t)$, with Gaussian errors around this trend. That is for a large portfolio, the proportion of infected policyholders is closed to $\iota_2(t) = \int_0^t S_C(u)(f_T(u) - f_V(u))du$. In this expression, the difference of the density of the infection times and the recovery times, is weighted by the survival function of the security implementation : the faster the vaccination, the smaller the proportion of infected policyholders. Besides, a rough approximation for $\sup_{t\geq 0} \mathfrak{I}_t$ is $\sup_{t\geq 0} n\iota_2(t)$, and distribution-free deviation bounds as in Proposition 3.2 below can help to quantify the potential error in such an approximation.

Proposition 3.2 *For all $x \geq 0$,*

$$\mathbb{P}\left(n^{-1/2}\sup_{t\geq 0}|\mathfrak{I}_t - n\iota_2(t)| \geq x\right) \leq 2.5\exp\left(-2x^2 + Cx\right),$$

for some absolute constant C.

4 Simulation of a Wannacry-type event

The Wannacry attack hit the world in May 2017, see Mohurle and Patil (2017). It is particularly emblematic due to the important number of computers infected around the world (more than 300,000 according to Chen and Bridges (2017)). The attack consisted in a ransomware introduced in the systems through a well documented vulnerability of Microsoft Windows (EternalBlue exploit, see Kao and Hsiao (2018)). Wannacry blocked the system, preventing the users to access data. A ransom was asked to unblock the system, which had to be paid in bitcoins. Wannacry lead to significant business interruptions and losses - the global amount of paid ransoms, which is approximatively 80,000 euros according to Willman (2017), is negligible compared to the total estimated damages (the loss is estimated as 92 billions pounds for the sole UK's National Health Service (NHS) according to Field (2018)).

4.1 Model for λ_T

To mimic the Wannacry incident and its propagation, we rely on a classical epidemiological SIR model. This SIR model is used to determine the hazard rate λ_T, describing the evolution of the strength of the contagion in the global population. These types of compartmental models are commonly used to describe (biological) epidemics since M'Kendrick (1925) and Kermack and M'Kendrick (1927), see also Lefèvre and Picard (1996). These types of models have already been applied to several actuarial applications, see e.g. Chen and Cox (2009), Feng and Garrido (2011) or Lefèvre et al. (2017).

The SIR model (for Susceptible - Infected - Recovered) describes the evolution of $(s_t, i_t, r_t)_{t\geq 0}$, where s_t is the number of exposed victims in the global population at time t, i_t is the number of infected entities that are still contagious at time t, and r_t is the number of entities who are no longer contagious. The global size of the population N does not evolve through time (which is reasonable since the crisis only lasts a few days), that is, for all t, $N = s_t + i_t + r_t$. The evolution of (s_t, i_t, r_t) is guided by the following system of differential equations,

$$\frac{ds_t}{dt} = -\beta s_t i_t \tag{4.1}$$

$$\frac{di_t}{dt} = \beta s_t i_t - \gamma i_t \tag{4.2}$$

$$\frac{dr_t}{dt} = \gamma i_t, \tag{4.3}$$

where β is the contagion rate and γ the "recovery" rate. The function $t \to i_t$ reflects the evolution of the strength of the epidemics, and we take $\lambda_T(t) = \beta i_t$, since $\beta i_t dt = \beta s_t i_t dt/s_t$ represents the proportion of newly infected people between t and $t + dt$ in the global population. Determining reasonable values for the parameters (β, γ, N) is difficult due to the lack of public data on the real-time evolution of the crisis and on the strength of the contagion. We now describe the heuristic we develop to determine such set of parameters.

We can impose $\gamma = 1$, since it seems reasonable to assume that, in approximatively 1 day, containment measures are developed to prevent the infection to spread. On the other hand, the force of the contagion, described by β, and the total size of the exposed global population N, require more delicate assumptions. To calibrate them, we used the relationships between the parameters and $r_\infty = \lim_{t\to\infty} r_t = 300,000$ (the total number of victims), and $i_{\max} = \max_{t\geq 0} i_t$. Again, the quantity $i_{\max}$ is unknown. Nevertheless, we can estimate it from the peak of paid ransoms, which occurs on 15th May 2017 (93 - 29% - ransoms out of 320 between 12th and 21th May, the payments are linked to three bitcoin adresses for which the time of transactions is available, see Willman (2017)). Using this information, we consider the (rough) approximation $i_{\max} = 29\% \times r_\infty$. This leads to the following set of parameters described in Table 1.

	Value
β	2.556×10^{-7}
γ	1
N	4064279
R_0	1.04
$i_{\max}$	87188
r_∞	300,000

Table 1: Parameters and main characteristic for a SIR model calibrated from the Wannacry ransomware attack. R_0, the basic reproduction number (classical indicator in epidemiology), is defined as $R_0 = N\beta/\gamma$.

4.2 Response to the attack

In our simulations, we consider an exponential distribution of the time of intervention U, that is $\lambda_U(t) = \lambda_0$. Again, U does not represent the time before full recovery. The mean

value of U, $1/\lambda_0$, should be taken as a few days, to be of the same magnitude as the length of the episode (10 days in our simulations). In our simulation, we take $\lambda_0 = 1/3$.

For the variable C, we distinguish three types of reaction to the crisis:

- a translated exponential distribution, $\lambda_C^{(1)}(t) = c_1 \mathbf{1}_{t \geq \tau_1}$. This means that, once the response has begun, the proportion of policyholders per time who update their security system is constant through time;
- a Pareto-type distribution, $\lambda_C^{(2)}(t) = c_2(t - \tau_2 + 1/2)^{-\alpha_2} \mathbf{1}_{t \geq \tau_2}$, for $\alpha_2 > 0$. This corresponds to a situation where the vigilance of the policyholders decreases through time: the more careful perform update short after the date of response τ_2, while the ones who did not instantly perform this update are more likely to ignore the threat;
- a Weibull-type situation where there is a progressive attention devoted to this threat among policyholders, that is $\lambda_C^{(3)}(t) = c_3(t - \tau_3)^{\alpha_3} \mathbf{1}_{t \geq \tau_3}$, for $\alpha_3 > 0$.

In each case, a parameter $(\tau_j)_{1 \leq j \leq 3}$ represents the reactivity of the response.

4.3 Simulation results

We consider two portfolios of respectively $n = 5000$ and $n = 10,000$ exposed policies. For each portfolio, we perform $10,000$ simulations of the impact of a cyber epidemics with the same attack intensity as Wannacry. For each type of response, we consider three delays of reaction: a fast response (3 days after the start of the event), a medium response (5 days), and a slow response (7 days). For each replication, we focus on the maximum number of policyholders requiring immediate assistance, that is $\sup_t \mathfrak{I}_t$. The values of the parameters of the three types of responses described above are taken so that $E[C_j - \tau_j | C_j \geq \tau_j] = 1$.

Two typical simulated trajectories of $(\mathfrak{I}_t)_{t \geq 0}$ are shown in Figure 1. Some empirical statistics on $\sup_{t \geq 0} \mathfrak{I}_t$ are shown in Tables 2 and 3 below.

Additionally, Figure 2 and Figure 3 represent the histograms for the variable $\sup_t \mathfrak{I}_t$ in the different settings, for $n = 10,000$.

Type of reaction	Mean	Standard deviation	95% confidence interval
No reaction	114.7122	8.91404	(98;133)
Slow Exponential	108.8625	9.241057	(92,128)
Slow Pareto	109.8391	9.269941	(93,129)
Slow Weibull	110.0568	9.213971	(93,129)
Medium Exponential	98.1915	9.220256	(81,117)
Medium Pareto	99.9944	9.378708	(82,119)
Medium Weibull	100.5861	9.292267	(83,119)
Fast Exponential	78.3415	8.359442	(63,95)
Fast Pareto	80.9131	8.712046	(64,98)
Fast Weibull	82.4775	8.667168	(66,100)

Table 2: Summary statistics on $\sup_{t\geq 0} \mathfrak{I}_t$, $n = 5000$.

Type of reaction	Mean	Standard deviation	95% confidence interval
No reaction	222.469	12.78206	(198,248)
Slow Exponential	212.6108	13.34208	(187,240)
Slow Pareto	214.622	13.38092	(189,242)
Slow Weibull	214.7712	13.2952	(190,242)
Medium Exponential	192.366	13.09089	(167,219)
Medium Pareto	196.3496	13.29188	(171,223)
Medium Weibull	197.0575	13.19889	(172,224)
Fast Exponential	153.595	11.87992	(131,177)
Fast Pareto	159.5131	12.33206	(136,184)
Fast Weibull	162.0158	12.23092	(138,187)

Table 3: Summary statistics on $\sup_{t\geq 0} \mathfrak{I}_t$, $n = 10,000$.

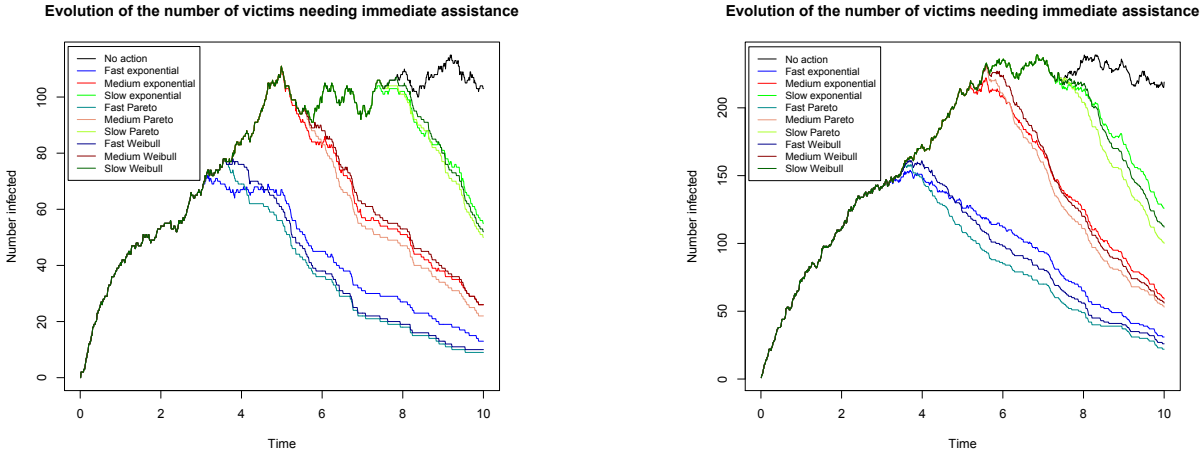

Figure 1: Two simulated trajectories of $t \to \mathfrak{I}_t$ (left-hand side: size of the portfolio $n = 5000$, right-hand side $n = 10,000$.)

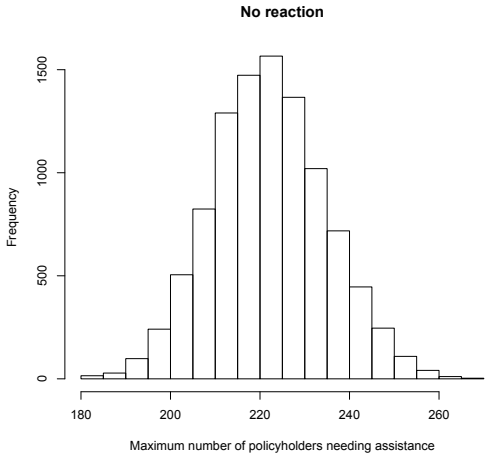

Figure 2: Histogram for $\sup_t \mathfrak{I}_t$ (10,000 simulations, size of the portfolio $n = 10,000$) in case of absence of response to the attack.

We see that all three types of responses lead to a similar impact on $\sup_t \mathfrak{I}_t$ (which is not entirely surprising since the expectation of these three distributions has been taken identical). Some differences in terms of variance still exist. The main parameter seems to be the time of response. A slow response will hardly diminish the burden of the assistance teams, while a fast response in 3 days significantly reduces the magnitude of $\sup_t \mathfrak{I}_t$.

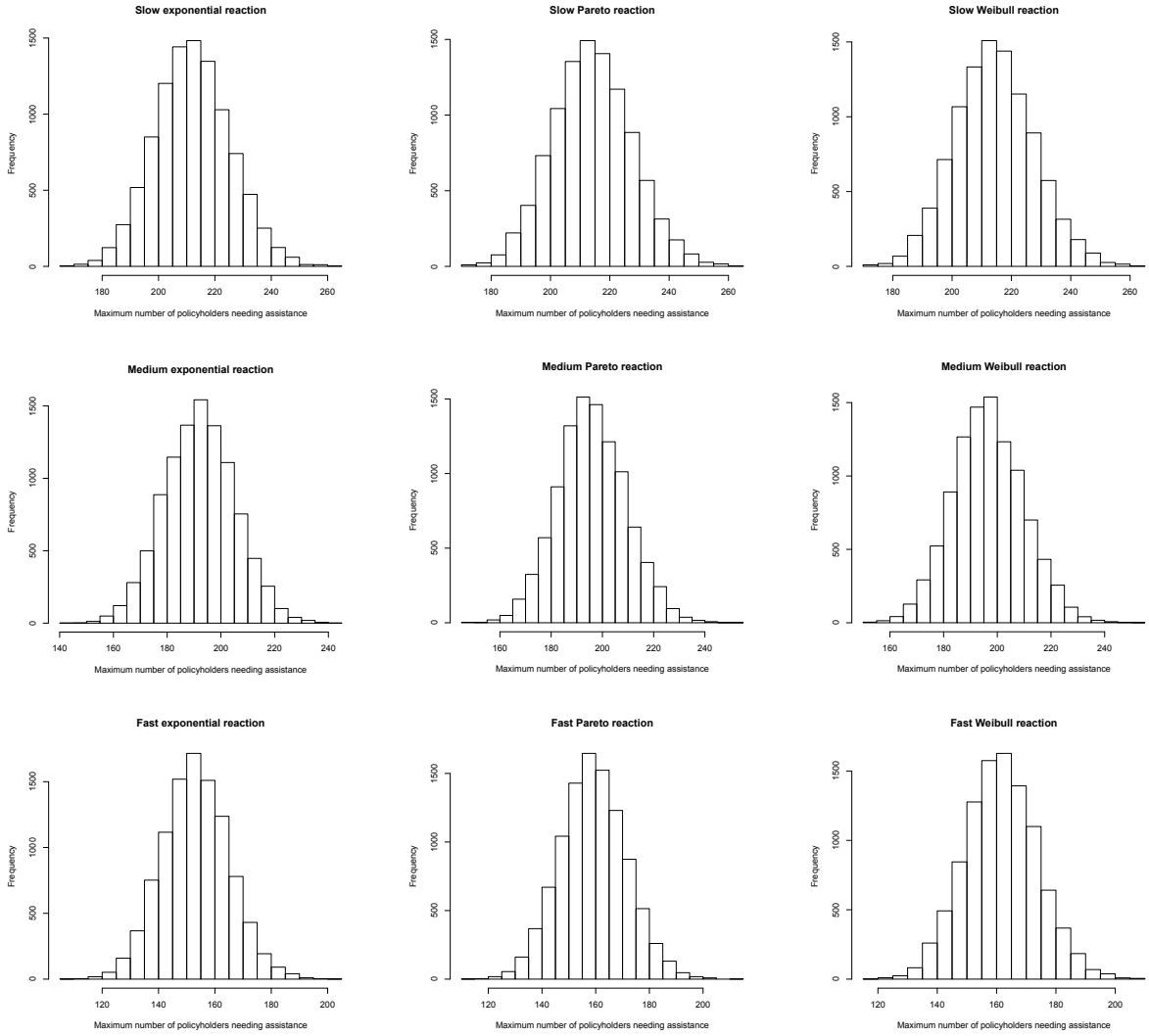

Figure 3: Histogram for $\sup_t \mathfrak{I}_t$ ($10,000$ simulations, size of the portfolio $n = 10,000$) for different type of responses and delays.

Acknowledgement: *The authors acknowledge funding from the project Cyber Risk Insurance: actuarial modeling, Joint Research Initiative under the aegis of Risk Fundation, with partnership of AXA, AXA GRM, ENSAE and Sorbonne Université.*

5 Conclusion

In this paper, we propose a general model for evaluating the damages caused by a cyber-hurricane on an insurance portfolio. We consider a particular setting in the simulation study, trying to mimic an event similar to the famous Wannacry episode, but every part of the model can be adapted to take various scenarii into account. This model can be used to quantify the benefits of a reaction to such a crisis. In the numerical study, we considered three patterns of response, where the most determinant parameter seems to be the fastness of the reaction. Behavioral studies on how policyholders perform prevention may be determinant to calibrate this response and evaluate the risk of collapse of the system.

References

Chen, H. and Cox, S. H. (2009). An option-based operational risk management model for pandemics. *North American Actuarial Journal*, 13(1):54–76.

Chen, Q. and Bridges, R. A. (2017). Automated behavioral analysis of malware: A case study of wannacry ransomware. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 454–460. IEEE.

Feng, R. and Garrido, J. (2011). Actuarial applications of epidemiological models. *North American Actuarial Journal*, 15(1):112–136.

Field, M. (2018). Wannacry cyber attack cost the nhs£ 92m as 19,000 appointments cancelled. *The Telegraph*, page 2018.

Kao, D.-Y. and Hsiao, S.-C. (2018). The dynamic analysis of wannacry ransomware. In *2018 20th International Conference on Advanced Communication Technology (ICACT)*, pages 159–166. IEEE.

Kermack, W. and M'Kendrick, A. (1927). A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A.*, 115:700–721.

Lefèvre, C. and Picard, P. (1996). Collective epidemic models. *Mathematical Biosciences*, 134(1):51 – 70.

Lefèvre, C., Picard, P., and Simon, M. (2017). Epidemic risk and insurance coverage. *Journal of Applied Probability*, 54(1):286–303.

M'Kendrick, A. G. (1925). Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematical Society*, 44:98–130.

Mohurle, S. and Patil, M. (2017). A brief study of wannacry threat: Ransomware attack 2017. *International Journal of Advanced Research in Computer Science*, 8(5).

Willman, R. (2017). Wannacry outbreak data. `https://github.com/rwillmann/WannaCry-Outbreak-Data-12-May-2017---19-May-2017-`.