Insurance: Mathematics and Economics 107 (2022) 88–101

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Cyber-contagion model with network structure applied to insurance

Caroline Hillairet [a], Olivier Lopez [b,*], Louise d'Oultremont [c], Brieuc Spoorenberg [c]

[a] Ensae Paris, Center for Research in Economics and Statistics, 5 avenue Henry Le Chatelier, 91120 Palaiseau, France
[b] Sorbonne Université, CNRS, Laboratoire de Probabilités, Statistique et Modélisation, LPSM, 4 place Jussieu, F-75005 Paris, France
[c] Detralytics, Rue Belliard 2A, 1000 Bruxelles, Belgium

ARTICLE INFO

Article history:
Available online 17 August 2022

JEL classification:
C51
C63

Keywords:
Cyber insurance
Cyber risk
Compartmental models
Multi-SIR
Network structures

ABSTRACT

In this paper, we provide a model that aims to describe the impact of a massive cyber attack on an insurance portfolio, taking into account the structure of the network. Due to the contagion, such an event can rapidly generate consequent damages, and mutualization of the losses may not hold anymore. The composition of the portfolio should therefore be diversified enough to prevent or reduce the impact of such events, with the difficulty that the relationships between actor are difficult to assess. Our approach consists of introducing a multi-group epidemiological model which, apart from its ability to describe the intensity of connections between actors, can be calibrated from a relatively small amount of data, and through fast numerical procedures. We show how this model can be used to generate reasonable scenarios of cyber events, and investigate the response to different types of attacks or behavior of the actors, allowing to quantify the benefit of an efficient prevention policy.

1. Introduction

1.1. Accumulation in cyber risk

Cyber risk is a major challenge in a digital era where industries and public services strongly rely on numerical tools. The number of cyber incidents and attacks in the recent years (see Kshetri (2020), Agence Nationale de la Sécurité des Systèmes d'Information (2021)) has even increased with the Covid-19 pandemic and the expansion of teleworking, see Lallie et al. (2021). In this context, cyber insurance is an essential tool for the industry to develop a proper protection against this threat, see Xie et al. (2020). These products typically mix financial compensation, prevention, and assistance in case of incident (see Romanosky et al. (2019)). But a major concern is the potentially "systemic" nature of the risk. A single massive attack, striking simultaneously (or in a short lapse) a large number of victims, could lead to losses of mutualization, endangering the viability of the insurance mechanism, see Welburn and Strong (2019). The Wannacry (see for example Mohurle and Patil (2017)) or NotPetya (see Fayi (2018)) episodes[1] are warning signs, whose estimated costs represent billions of dollars of losses (see Chen and Bridges (2017)). In this paper, we aim to provide a flexible framework to describe, model, and project the impact of the accumulation of cyber events on a portfolio. We particularly focus on how network effects can contribute to the spread of a cyber epidemic.

The network structure is known to have a significant influence on the spread of cyber attacks. This is true when one focuses on a single company, which wants to monitor the spread of the infection through the nodes of its information systems (see Adams and Heard (2014)), but also at an higher level when one looks at a set of different actors that can be affected through the connections between each other (see Welburn and Strong (2019), Böhme et al. (2010)). Recently, Fahrenwaldt et al. (2018) illustrated the influence of different types of networks in cyber insurance and showed how their shape leads to a faster propagation of an attack or not. In all these approaches, the contagion is assumed to spread among the policyholders. Hillairet and Lopez (2021) develop an alternative approach to model the dynamic of the cyber contagion, and the reaction of the actors. It assumes that contamination is more likely to come from outside the portfolio than from inside. This seems reasonable, based on the fact that a portfolio is in fact of small size, compared to the global population among which the cyber epidemic spreads. Moreover, this choice leads to a simple model,

* Corresponding author.
E-mail addresses: caroline.hillairet@ensae.fr (C. Hillairet), olivier.lopez@sorbonne-universite.fr (O. Lopez).

[1] Wannacry and NotPetya are worldwide ransomware attacks which took place just a few weeks apart (May 2017 and June 2017 respectively), and that both exploited the vulnerability Eternal Blue. They are two outstanding examples of massive attacks: in a few days, Wannacry hit around 300,000 computers in 150 countries, and NotPetya several million in 65 countries, and billion of dollars of losses for each one.

https://doi.org/10.1016/j.insmatheco.2022.08.002
0167-6687/

which allows to model separately the dynamic of the contagion and the time for recovery after being hit.

1.2. Contributions

The approach we develop in the present paper is in the continuity of Hillairet and Lopez (2021), but focusing on the way to incorporate the network topology and evaluate its impact. We aim to complete previous analysis through a multi-group SIR model, whose connectivity components are calibrated on OECD data, and only require macro-level data. Our contributions are the following:

- To model the contagion, we rely on compartmental models widely used in epidemiology, namely multi-group SIR models, see for example Beretta and Capasso (1988), Guo et al. (2006), or Magal et al. (2018), and also Diekmann and Heesterbeek (2000), Brauer et al. (2019), Daley and Gani (2001), Andersson and Britton (2012) for more properties on classical epidemiological models. We introduce specific terms in the system of differential equations describing the dynamic, in order to capture some specificities of cyber risk. In particular, we model the reaction to the crisis, namely the introduction of countermeasures as soon as the threat has been identified by the community. More precisely, we take into account that, even after the detection of a specific attack, some potential victims do not manage to implement a perfect protection against the threat, but only reduce the risk of being hit. In the application, we focus on the situation where the contamination of policyholders comes from outside the portfolio, while an independence assumption holds inside the portfolio. The possibility of combining this approach with interactions between individuals is mentioned but not explored here in detail. The choice of these deterministic models is driven by the need to rely on very few data to calibrate them for practical use, due to the present weakness of information on cyber risk.
- We derive some theoretical properties of this model. One of the key results is to provide a methodology to easily evaluate the outcome of a cyber pandemic by solving a fixed point equation to compute the total number of victims in each group. We explain how this model can be used to quantify the impact of a contagious cyber event on an insurance portfolio, as well as the impact of protection measures.
- We address the issue of the model's calibration, which is of major concern, although it is hardly treated in the literature cyber risk, since data on digital connections are very hard to get. We show how this model can be calibrated using a proxy computed from a relatively small amount of macro-economic data.[2] An example of such calibration is conducted from OECD data, in order to mimic a Wannacry episode. It allows us to quantify the contagiousness of each industrial network, and how their reaction may help to reduce the spread of the attack. Adding a model for the cost of each infection, we quantify the impact on an insurance portfolio and its premium, using data from studies conducted on the French market of insurance brokers.
- We investigate how the connectivity of the network can have influence on the spread of the attack and the final number of victims.

The rest of the paper is organized as follows. Section 2 introduces the multi-group SIR model and derives some theoretical properties. Section 3 provides an example of calibration from OECD data in order to mimic a Wannacry episode. A detailed quantitative analysis is then conveyed concerning the contagiousness of each industrial sector, the impact of their reactions, as well as the impact on an insurance portfolio. Finally, the impact of the network is investigated through simulations in section 4. Technical materials and proofs, as well as a sensitivity analysis, are postponed in the Appendix.

[2] Although we consider economic flows, we do not claim to capture any financial contagion: these data have to be understood as a proxy to identify digital links.

2. Multi-group SIR model

In this section, we introduce the compartmental models that will allow us to model how connections between the actors impact the contagion of a cyber pandemic. The multi-group SIR we use is introduced in section 2.1. We also explain how our model can be adapted to take into account the fact that failures in the supply chain may be generated through a cyber event and increase the impact of an incident. Section 2.2 studies the consequences of the spread of the cyber attack on an insurance portfolio.

2.1. Definition of the model

Compartmental models, since their introduction by McKendrick (1925) and Kermack and McKendrick (1927), have become common tools in epidemiology. See for example Brauer et al. (2012) for a review on this topic, or Di Domenico et al. (2020) for recent work in the context of the Covid-19 pandemic. These types of models have been introduced in actuarial sciences by Feng and Garrido (2011) to study epidemiological risk, see also more recent developments in Feng et al. (2022). The core of these models is to describe the different states of an individual in a population stroke by a disease. The most simple version, the SIR model (for Susceptible - Infected - Recovered) splits the population into three groups:

- the "Susceptibles" are exposed to the risk of developing a pathology. In our case, the susceptibles will be the entities that can be stroke by the ongoing cyber attack; for example, in the Wannacry or NotPetya episodes, susceptibles are computers vulnerable to the Eternal Blue exploit, see Kao and Hsiao (2018). As we can see from this example, the total number of susceptibles is hard to track - in fact, even the exact number of computers equipped with a given operating system is impossible to obtain;
- a susceptible may then become "Infected" by the pathology (here, by the cyber virus), and is contagious. An infected will contribute to the contagion;
- after some time, an infected becomes "Recovered", or "Removed": it means that this individual or entity can not transmit the virus anymore. In the context of cyber risk, the term removed must be understood as the fact that the victim stops participating to the contamination, because countermeasures have been adopted. The time before full recovery of a cyber attack can be very long, of a different scale to the length of the attack itself (the attack lasted around a week for Wannacry, to be compared with months or years of recovery according to Low (2017)).

2.1.1. The dynamics of the epidemic

The dynamic of the epidemic is then described by a system of differential equations, governing the rate at which individuals in each of the compartments move from one state to another. Here, we want to take into account the fact that the population on which the attack spreads is heterogeneous. Typically, we do not expect the contagion to spread identically on industries from the health sector or from the financial sectors: the nature of the assets that can be targeted by hackers, how easy to get a ransom from a given

type of victim, and the difference in terms of level of security in different sectors will indeed have an impact on contagion (see Alrimy et al. (2018)). For this reason, we will consider a multi-group SIR model (see for example Magal et al. (2018)). In this case, the population of victims is decomposed into d groups (for example representing different sectors of activities for industrial actors, but these groups may also be constituted from the type of behaviors of some categories of victims). For $j = 1, ..., d$, and at each instant $t \geq 0$, $s_j(t)$ (resp. $i_j(t)$, resp. $r_j(t)$) is the number of susceptibles (resp. infected, resp. removed) in category j at time t. Next, the evolution of each of these compartments is governed by

$$\frac{ds_j(t)}{dt} = -\eta_j(t)\left\{\alpha_j(t) + \sum_{k=1}^{d}\beta_{k,j}i_k(t)\right\}s_j(t), \tag{2.1}$$

$$\frac{di_j(t)}{dt} = \eta_j(t)\left\{\alpha_j(t) + \sum_{k=1}^{d}\beta_{k,j}i_k(t)\right\}s_j(t) - \gamma_j i_j(t), \tag{2.2}$$

$$\frac{dr_j(t)}{dt} = \gamma_j i_j(t). \tag{2.3}$$

The matrix $\mathbf{B} = (\beta_{k,j})_{1\leq k,j\leq d}$ conveys the information on how class k contaminates class j. This matrix is the key element of the model to capture the network topology. It materializes the connections between the different groups. An oriented graph can be associated to $\mathbf{B}$ by constituting the adjacency matrix $\mathfrak{G}$ with coefficients $g_{j,k} = \mathbf{1}_{\beta_{j,k}\neq 0}$. Let us emphasize that the matrix $\mathbf{B}$ (hence $\mathfrak{G}$) is not necessarily symmetric - that is why we use the term "oriented" graph - since the contamination may not flow identically from both sides: group j may strongly contaminate group k, while group k may be less contagious for group j since group j developed more security measures to reduce contamination from group k.

Compared to the most classical version of the multi-group SIR as described in Magal et al. (2018), we introduce some additional terms in order to take into account specificities of cyber attacks. First, the vector $\mathbf{A}(t) = (\alpha_j(t))_{1\leq j\leq d}$ represents a latent form of attacks, i.e. not contagious. Through the introduction of this term, we want to consider a mechanism which is not only contagious, but which can be caused by successive attacks on different categories of victims. At the same time, we introduce a protection component against the threat, materialized by the vector $\mathbf{H}(t) = (\eta_j(t))_{1\leq j\leq d}$. This vector diminishes the rate of new infections through time, meaning that on the contrary to Hillairet and Lopez (2021), it models here an imperfect protection, that is not 100 % efficient. Indeed, perfect protection is not always possible. For example, large organizations may have difficulties to implement a correction throughout their systems in a short amount of time. Moreover, some attacks may rely on human factors, like in the case of phishing or attacks based on fraudulent mails. In those cases, one may increase the awareness on the threat, and thus reduce the transmission and the risk of infection, but without achieving perfect protection.

Remark 2.1. The framework that we develop is adapted to a contagious mechanism which is not necessarily purely cyber. Indeed, a possible consequence of cyber attacks is business interruption (see Hobbs (2021), Romanosky et al. (2019), Cashell et al. (2004)). Hence, the contagion may not only be caused by the transmission of a virus, but through breaking the supply chain (see Ghadge et al. (2019), Boyes (2015)). The matrix $\mathbf{B}$ can for example materialize a chain of dependence between sectors of activity, one being hit because it relies on the production of another that has been stroke by the cyber attack.

2.1.2. Basic reproduction number

For given values of the parameters of this model, a natural question is to determine if the contagion is likely to spread or if the cyber-epidemic is about to collapse. The basic reproduction number is a way to measure the evolution of the dynamic. Its definition in the context of a general system of differential equations like (2.1)-(2.3) has been introduced by Diekmann et al. (1990). See also Heffernan et al. (2005). It relies on the so-called "next-generation matrix".

Let $\mathbf{s} = (s_1, ..., s_d)$ a vector representing the susceptibles in the different categories, and let $\mathcal{F}_j(\mathbf{s}, i_1, ..., i_d)dt$ (resp. $\mathcal{V}_j(\mathbf{x}, i_1, ..., i_d)dt$) denote the number of new infections (resp. of infected entities entering the "removed" status) between t and $t + dt$. We consider a given composition of the population $\mathbf{s}^{(0)}$ (namely the initial composition of the population if one wishes to understand, at the beginning of the episode, if the attack is able to spread). Next, define

$$F_{j,k} = \frac{\partial \mathcal{F}_j(\mathbf{s}^{(0)}, i_1, ..., i_d)}{\partial i_k} = \beta_{k,j}s_j^{(0)}, \quad \text{and}$$

$$V_{j,k} = \frac{\partial \mathcal{V}_j(\mathbf{s}^{(0)}, i_1, ..., i_d)}{\partial i_k} = \gamma_j \mathbf{1}_{k=j},$$

considering the special case where $\eta_j(t) \equiv 1$. Then, the next-generation matrix is FV^{-1}. The (j,k) entry of this matrix materializes the rate at which infected entities in the k-th category generate new infected in category j, multiplied by the average length of stay in the j-th infected compartment. The basic reproduction number R_0 is then the spectral radius of FV^{-1}. If $R_0 > 1$, then the epidemic expands, while the epidemic fades if $R_0 < 1$ (see for example van den Driessche and Watmough (2002), Andreasen (2011), Theorem 2.2 in Perasso (2018) or section 5.2 in Brauer et al. (2019)).

2.2. From the multi-group SIR to the impact on an insurance portfolio

The multi-group SIR defined in section 2.1 describes a dynamic on a large population. On the other hand, an insurance portfolio is of smaller size, introducing some randomness in the result. We introduce this randomness by considering that the portfolio is a random sample from a much larger population for which the deterministic epidemiological model holds. Let us note that an alternative way would be to consider stochastic dynamics even for the larger population. Different systemic interaction models could be used, at a global level (using Hawkes processes, see e.g. Bessy-Roland et al. (2021)) or at a local level, using stochastic SIR models. Such alternative would be particularly interesting in a situation where the larger population in which the attack propagates is of reasonable size. In this case, the deterministic approximation of the law of large numbers would be seen as a too rough approximation, requiring to turn towards, for example, stochastic SIR models. Examples of use of stochastic SIR models in insurance can be found for example in Lefèvre et al. (2017), or Lefèvre and Simon (2021). Let us point that, although the technique we propose could be adapted to stochastic SIR up to some modifications, the calibration would be harder. Due to the weakness of databases on cyber events, we chose to focus on the simpler model described below, which requires a minimum amount of data. Let us describe the model at a portfolio level more precisely. Each policyholder m is described by a random variable T_m and a deterministic characteristic $\mathbf{x}_m$ where

- T_m is the random time at which the m-th policyholder is infected by the cyber virus;
- $\mathbf{x}_m$ is deterministic, and represents the category of the multi-group SIR to which the m-th policyholder belongs (that is $\mathbf{x}_m \in \{1, ..., d\}$).

Let us note that T_m can be infinite with non-zero probability. The distribution of T_m is linked to the dynamic of the cyber attack. Let us introduce the hazard rate function

$$\lambda_{T_m}(t) = \lim_{dt\to 0^+} \frac{\mathbb{P}(T_m \in [t, t+dt] | T_m \geq t)}{dt}. \tag{2.4}$$

The value $\lambda_{T_m}(t)$ quantifies the risk of being infected at time t, which depends on the current circulation of the cyber virus, and on the protection level.

In the following, we consider the simplest case where the policyholders are considered as independent from each other with respect to the spread of the attack. The idea is that, due to the small size of the portfolio, infection is more likely to come from outside: it does not propagate within the community of policyholders directly, but because of the connections of the policyholders to a wider attacked network. Introducing a dependence structure and/or introducing stochastic epidemiological models is a possible extension that is left to future research. Under this simplifying assumption, and since the portfolio can be understood as a random sample of individuals from the global population, the infection rate $\lambda_{T_m}(t)dt$ of the individual m belonging to group $\mathbf{x}_m$ should be equal to the probability of selecting a newly infected individual among the individuals of population $\mathbf{x}_m$ that were not infected before t (that is $s_j(t)$ if $\mathbf{x}_m = j$). Between t and $t+dt$, there are $\eta_j(t)\{\alpha_j(t) + \sum_{k=1}^d \beta_{k,j} i_k(t)\} s_j(t)dt$ new contaminations in population j. This leads to

$$\lambda_{T_m}(t) = \lambda(t, j) = \eta_j(t) \left\{ \alpha_j(t) + \sum_{k=1}^d \beta_{k,j} i_k(t) \right\}, \text{ if } \mathbf{x}_m = j.$$

From this hazard rate function, one can deduce the average number of infected policyholders in the portfolio, which directly derives from the quantities in (2.1)-(2.3). If we have n_j policyholders from category j, the expectation of the number of victims in class j is

$$n_j \left(1 - \exp\left\{ -\int_0^\infty \lambda(t, j)dt \right\} \right) = n_j \times \nu.$$

Moreover, the variance is then $n_j \nu(1-\nu)$. More details on approximations of this number can be found in Hillairet and Lopez (2021).

Through (2.4), it is assumed that the portfolio behaves like the global population. This may not be true in practice, for example due to adverse selection. Statistical analysis of portfolio data can potentially help to quantify the distortion between the two populations, which can be incorporated into the model. Another possible modification is the introduction of dependence between policyholders, for example through a copula function. Again, additional statistical analysis may help to identify such phenomena.

Remark 2.2. In Hillairet and Lopez (2021), an additional random variable U_j was introduced to describe the length of immediate assistance required after a victim is hit. This variable is important if one wishes to understand how many policyholders have to be assisted at a given time. This question has important consequences, because if this number becomes too high, a saturation of the response capacity can lead to additional damages. In the present paper, we do not focus on this problem, since we are more motivated by understanding the impact of the network topology on the spread of the infection. Nevertheless, an approach similar to the one of Hillairet and Lopez (2021) can easily be added to complement this model.

2.3. Total number of victims

In this section, we provide theoretical results that rely on the total number of victims from a cyber incident to the parameters of the multi-group SIR model. The total number of victims in the global population (hence the average number of infected policyholders) can be easily determined by (numerically) solving the system of differential equations (2.1)-(2.3). We refer the interested reader to Amann (2011) for extensive theoretical results on systems of ODE, and to Brauer et al. (2008) for additional results on SIR models for a single population. By measuring the total number of infected individuals in each group of the population depending on the starting point of the infection, we will have the ability to better understand the impact of connectivity between classes. In the case of no reaction from the attacked community (that is protection coefficients $\eta_j = 1$ for all j), the total number of victims in the global population can be determined in a simple and fast way, by solving a fixed point problem. Let $r_j(\infty) = \lim_{t\to\infty} r_j(t)$. Since every infected ultimately becomes removed after a finite amount of time, $r_j(\infty)$ represents the total number of infected in class j. In Theorem 2.3 below, we show that $\mathbf{r}(\infty) = (r_j(\infty))_{1\leq j\leq d}$ is the solution of an equation of the type $\mathbf{r}(\infty) = \Phi(\mathbf{r}(\infty))$ (the definition of the function Φ is given in the statement of the theorem). The arguments of the proof are similar to the one in Magal et al. (2018), but the function Φ is not the same. This difference comes from the fact that the model we consider is more general, but also from the fact that the path of the proof is slightly different, and leads to a simpler function.

Theorem 2.3. *For $j = 1, ..., d$, assume that $\eta_j = 1$ and let $\mathcal{A}_j = \int_0^\infty \alpha_j(t)dt$. Assume that if for all j, $i_j(0) = 0$ then there exists j_0 such that $\mathcal{A}_{j_0} \neq 0$. Then, for $\mathbf{x} = (x_1, ..., x_d)^{tr}$, where ^{tr} denotes the transpose, let*

$$\Phi_j(\mathbf{x}) = i_j(0) + s_j(0) \left\{ 1 - \exp\left(-\left(\mathcal{A}_j + \sum_{k=1}^d \frac{\beta_{k,j}}{\gamma_j} x_k \right) \right) \right\},$$

and $\Phi(\mathbf{x}) = (\Phi_j(\mathbf{x}))_{1\leq j\leq d}$. The vector $\mathbf{r}(\infty)$ is the unique solution of the equation

$$\mathbf{r} = \Phi(\mathbf{r}), \quad \text{on } \mathcal{R} = \{\mathbf{r} : 0 \leq r_j \leq s_j(0) + i_j(0)\}.$$

The proof of Theorem 2.3 is postponed to the Appendix (subsection A.2). Note that the case where for all j, $i_j(0) = 0$ and $\mathcal{A}_{j_0} = 0$ corresponds to the trivial situation with no infected at time zero and no initial burst of attacks, leading then to the static situation where the multi-group SIR system is stuck at $\mathbf{r} = 0$ (which is clearly a fixed point in this situation). Theorem 2.3 characterizes the total number of victims as the solution of a fixed point equation, in case there is no reaction from the attacked community (that is $\eta_j = 1$ for all j). This result allows to quickly calibrate or assess the impact of such an episode. In section A.3, we show that, in some situations, the solution of this fixed point problem can be obtained from a fast converging iterative algorithm.

Theorem 2.3 can be easily generalized to constant protection parameter η_j, by multiplying α_j and the $\beta_{k,j}$ by η_j, for all j. Similarly, one can also take into account constant cross-categories protection effects $(\eta_{jk})_{1\leq j,k\leq d}$. In section 3.4, numerical experiments are used to investigate the case of time-varying protection coefficients, that may for example depend on the current proportion of infected in each group.

3. Illustration on a particular example

In this section, we give an example of calibration of the model based on macroeconomic data. Our aim is to show that plausible

Table 1
Exchange of added value between sectors - OECD data, 2015. A line represents the flow of added value sent from the corresponding sector to the sectors in columns.

	Mining	Manufacturing	Energy	Construction	Services
Mining	225.52	1026.27	154.72	506.18	412.55
Manufacturing	14.86	8654.41	94.61	1709.06	1362.29
Energy	4.92	342.46	674.89	165.10	284.47
Construction	1.41	58.85	12.55	3685.20	197.56
Services	33.62	4396.65	249.46	2164.84	22206.97

parameters may be obtained through the use of a relatively small amount of data.

We consider a population composed of five categories of potential policyholders, namely

- Mining and quarrying;
- Manufacturing;
- Electricity, gas, water supply, sewerage, waste and remediation services;
- Construction;
- Total business sector services.

These classes correspond to categories used by OECD to identify the dependence between some sectors of activity. We consider a particular form of contagion matrix $\mathbf{B} = \beta \mathbf{B}_0$, where $\mathbf{B}_0$ reflects the connectivity between actors. This section is organized as follows. In section 3.1, we calibrate this matrix $\mathbf{B}_0$ which somehow contains information on the topology of the network formed by the potential victims. Section 3.2 then shows how a Wannacry type episode can be calibrated for this particular structure of population. In section 3.3, we investigate the vulnerabilities of the different categories of the population by focusing on the impact on a cyber attack targeting one single sector. Section 3.4 shows how to quantify the benefit of certain type of interventions during the crisis to reduce its impact. Finally, in section 3.5, we illustrate the use of this model to measure the impact of a cyber episode on a insurance portfolio.

3.1. Connectivity between sectors

The groups considered in this study are split into different sectors of activity. To assess the connectivity between those sectors, we use an OECD study on the origin of value added in final demand, see OECD (2018). The statistics of this study are shown in Table 1. They represent a way to model how a category depends on another, via the flow of traded added value. Of course, this does not reflect the digital dependence between these sectors, which would be a much more accurate information if available. Therefore, we do not aim here to produce a very accurate vision of the connectivity between these sectors, but only to determine a reasonable benchmark. Following this objective, the (strong) assumption that we make is that the digital flow between these categories is, somehow, proportional to the economical flow reflected by Table 1.

More precisely, we want our matrix $\mathbf{B}$ to be of the following form, $\mathbf{B} = \beta \mathbf{B}_0$, where the parameter β is here to describe the strength of the contagion (its calibration is discussed in Section 3.2), and $\mathbf{B}_0$ is a normalized matrix (the sum of all of its coefficients is equal to one) containing only the information regarding the connectivity between actors. The matrix $\mathbf{B}_0$ is calibrated using the OECD data of Table 1 and 2. Indeed, the volume of exchanges between sectors (given in Table 1) has to be normalized by the number of companies in each sector, which is the information given in Table 2, from OECD (2015).

The contagion matrix $\mathbf{B}_0$ of Table 3 is obtained by dividing the value of a given line of Table 1 by the number of companies of the corresponding sector (from Table 2), before normalizing the values in order to ensure that the sum of all coefficients is equal to 1.

Table 2
Distribution of companies between sectors - OECD data, 2015.

Sector	Number of companies	Percentage
Mining	66,492	0.20%
Manufacturing	3,068,178	9.02%
Energy	220,892	0.65%
Construction	4,874,747	14.34%
Services	25,768,765	75.79%

According to this matrix, we see that the Mining & Quarrying sector would be the most contagious one, followed by the Energy sector. This high contagiousness is however to be tempered by the small population size of these sectors. Services and Manufacturing are the sectors that receive more cross-infections than the others. As expected, the manufacturing sector, strongly dependent from the supplies from other sectors, also achieves a high level of dependence.

Let us also note that this high contagiousness of the Mining and Energy sectors can also make sense from a supply-chain modeling perspective: the approach that we develop is focused on a case of contagious cyber event. But a cyber event targeting a sector can also have consequences on another one that is highly dependent, triggering business interruptions for companies that are not directly stroke by the virus. The framework we develop can also be used to take this type of phenomenon into account, through a proper design of the matrix $\mathbf{B}_0$, see also Remark 2.1.

Remark 3.1. We would like to emphasize that this contagion matrix $\mathbf{B}_0$, computed from economic flows quantities of OECD, is only a proxy of cyber connectivity. Those macroscopic data have the advantage to be public and provide a kind of link between sectors, while digital connections among sectors are much harder to get. Section A.1 evaluates the sensitivity of the results to the parameters of the contagion matrix $\mathbf{B}_0$. We made the choice to rely on public data, although imperfect, in order to show the simplicity of the model calibration. In practice, several cyber security firms provide data on connectivity between some categories of actors. In the meantime, the multi-SIR model we develop can work with a relatively high scale of aggregated data, while competing approaches like in Antonio et al. (2021) need a very detailed graph at a microscopic level.

3.2. Calibration of a Wannacry type episode

In the dynamics described by (2.1)-(2.3), we consider the contagion matrix $\mathbf{B} = \beta \mathbf{B}_0$, where the parameter β helps to design the intensity of the contagion. We here develop how to calibrate this parameter so that we may obtain a cyber event relatively similar to Wannacry. For this calibration, we first investigate the case without reaction, that is $\eta_j = 1$ for all j. We follow the path of Hillairet and Lopez (2021), in which a calibration of a Wannacry type event was proposed, based on indirect information about its dynamic (namely the timeline of the payments of ransoms, which

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

	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 4
Parameters used to simulate a Wannacry-type episode based on a single-type population. The parameters γ and N (total size of the victim population) have been taken as in Hillairet and Lopez (2021).

Parameter	Value
α_0	7×10^{-3}
β	1.845×10^{-5}
γ	1
N	4,064,279

Table 5
Comparison of the sectors through different attack scenarios. The sectors are ordered from the one leading to the highest epidemic, to the lowest. We consider a total population of $N = 4'064'279$ of potential victims, with the same distribution between sectors as in Table 2.

Targeted sector	β	α	Total infected	Peak
Mining	1.845×10^{-5}	3.5	714,347	89,984
Manufacturing	1.845×10^{-5}	0.078	587,338	70,815
Energy	1.845×10^{-5}	1.077	450,824	50,759
Services	1.845×10^{-5}	0.0049	256,833	27,483
Construction	1.845×10^{-5}	0.009	223,744	26,233

is publicly available due to the use of the Bitcoin protocol). Here, the calibration is different in two ways:

- we consider a non homogeneous population with contagion matrix $\mathbf{B} = \beta\mathbf{B}_0$, the total size of the population of potential victims being the same as in Hillairet and Lopez (2021) ($N = 4,064,279$) with distribution given by the proportions of Table 2;
- the initialization of the epidemic is done in a different way (see below).

Indeed, in Hillairet and Lopez (2021), a small number of initially infected i_0 spreads the cyber attack. Here, we do not need to use this number (which has to be chosen arbitrarily), because we prefer to use the functions α_j to ignite the epidemic, which seems more consistent with the patterns of cyber attacks. To calibrate the value of β, we consider that the attack strikes all classes at the same rate, $\alpha_j(t) = \alpha_0 \mathbf{1}_{t\leq 1}$ for all j : during one day, there is a burst of infections caused by the hackers that strike the victims at uniform rate α_0.

We follow the approach of Hillairet and Lopez (2021), where a model is chosen from its ability to replicate the peak of the epidemic - maximum number of victims affected at a given moment - and the total number of victims over 10 days, which is the approximate length of the episode. This leads to the parameters of Table 4.

The evolution through time of the infections in each category is reported in Fig. 1. We can observe that the peak of infections is not located at the same time (we see that this peak is achieved later for services, with a slower decay). The size of this peak can be of some concern: as pointed in Hillairet and Lopez (2021), this represents the number of victims needing assistance at a given time. Since many cyber insurance contracts are supposed to provide immediate assistance to their policyholders when hit, a too high peak could lead to an impossibility to deliver the service that was contractually guaranteed (also if assistance comes too late due to saturation, this could increase the amount of damages).

Let us recall that N does not represent the number of policyholders in an insurance portfolio, which tends to be, in present cyber insurance portfolios, much smaller than N from Table 4. The number of infected given here represents how the infection spreads on a population that is much larger (at a national level or even at a global level). To obtain the (average) impact of such an episode on a portfolio, one can retrieve the proportion of victims in each sector and apply it to the number of policyholders of this category in the portfolio.

3.3. Measuring the vulnerability of the different sectors

We now use our matrix $\mathbf{B} = \beta\mathbf{B}_0$ (where β is given by Table 4 and $\mathbf{B}_0$ is given by Table 3) to investigate which sector seems the most vulnerable and can potentially trigger a systemic event. Under this configuration, the basic reproduction number can be computed as explained in Section 2.1.2, leading to $R_0 = 1.02$. This value is slightly higher than 1, which means that the cyber epidemic will spread. To measure the vulnerability of sector j, we concentrate on the initial attack on it (that is $\alpha_j(t) = \alpha^{(j)}\mathbf{1}_{t\leq 1}$, and $\alpha_k(t) = 0$ for $k \neq j$). Taking $\alpha^{(j)} = \alpha_0$ would not make things comparable: if the size of the population of sector j is small compared to N, this would result in a small number of initial infections through these direct attacks (approximately α_0 times the size of this sub-population). Therefore, we take $\alpha^{(j)} = \alpha_0/p_j$, where p_j is the proportion of sector j in the total population (see Table 2), which seems more appropriate. The values of the coefficients $\alpha^{(j)}$ are given in Table 5, where we also gather results on the total size of the epidemic in each attack scenario, and the peak of infections (that is the highest number of currently affected victims at a given time).

From the coefficients of Table 3, it is logical to find that an attack targeting the Mining and Quarrying sector leads to the most important impact: we already mentioned the high contagiousness of this sector according to $\mathbf{B}_0$. Moreover, as the population of this sector is small, almost all companies in this category are infected by the initial attack. On the other hand, an attack on the Manufacturing sector, although this category is larger (9.02% of the total population compared to the 0.20% of Mining), also leads to a number of victims which is higher than the Wannacry episode (same property for the Energy sector, but with a smaller number of companies and a smaller impact than for Manufacturing). Again, this is caused by the high contagion spread by this sector.

Table 6 provides more precision on how a category contaminates another in the different scenarios. One can observe that the Mining sector suffers fewer from contagion when not attacked directly. We can also observe that the largest proportion of victims is not always achieved in the category where the attack was initiated: in the case of an attack on Manufacturing, Services are hit at a high rate (16.58%), which corresponds to 510'716 victims in this category, compared to 58'692 in the Manufacturing sector. On the

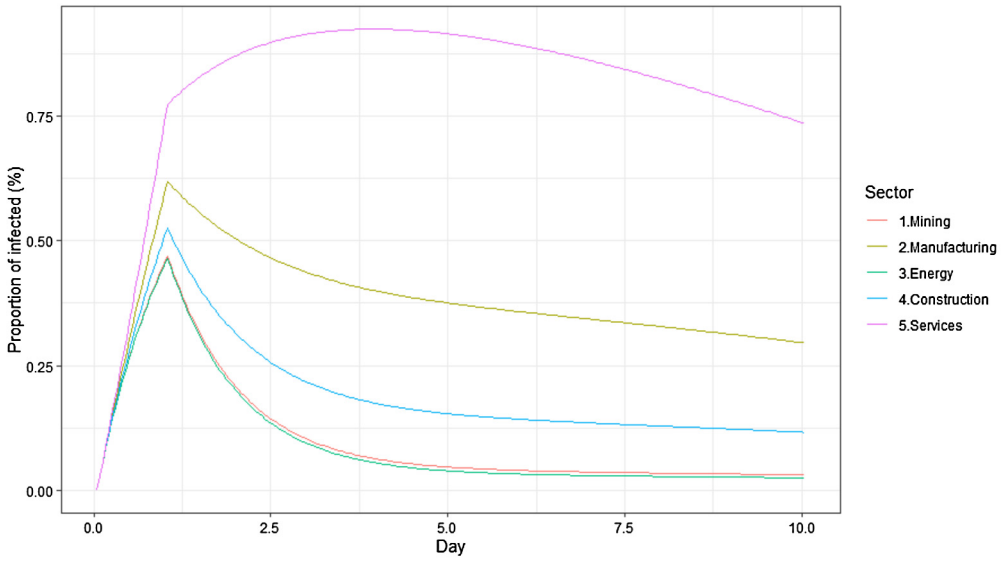

Fig. 1. Evolution of the proportion of infected - Uniform attack. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Table 6
Proportion inside each sector of companies affected by the epidemic, depending on the targeted sector.

Targeted sector	Mining	Manufacturing	Energy	Construction	Services
Uniform attack	1.06%	4.11%	0.99%	2.07%	8.86%
Attack on Mining	99.70%	12.69%	1.36%	5.49%	20.37%
Attack on Manufacturing	1.02%	16.01%	0.66%	3.05%	16.58%
Attack on Energy	0.93%	5.96%	64.08%	2.35%	12.93%
Attack on Construction	0.33%	2.49%	0.21%	6.60%	5.72%
Attack on Services	0.25%	2.59%	0.21%	1.01%	7.84%

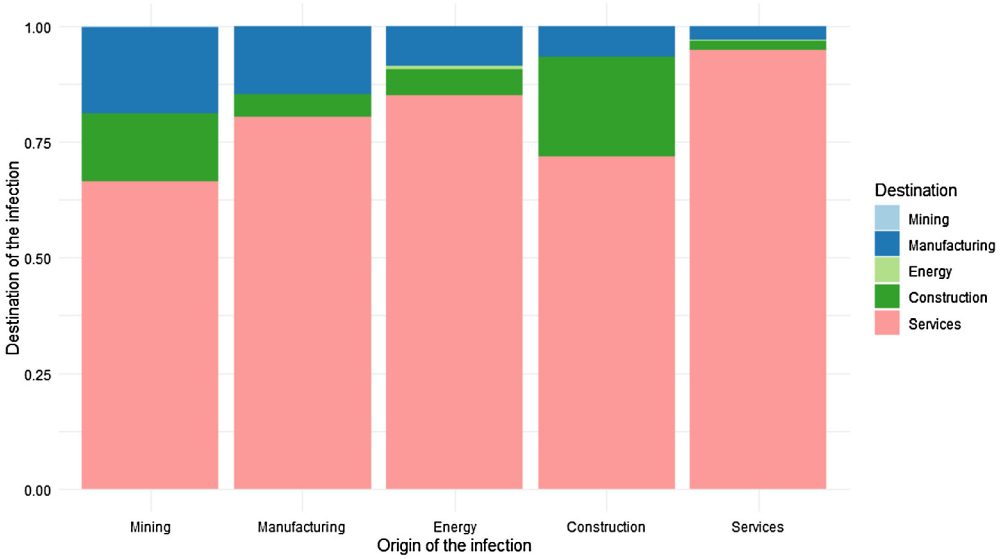

Fig. 2. Cross-infection between sectors after a uniform burst (proportion by destination).

other hand, we see that the Service sector is less affected when directly hit (except if we compare with the case of an attack on Construction) than when the initial burst of infection strikes another sector.

Fig. 2 gives another illustration of this phenomenon. If we look at a scenario of an uniform attack, we can see that each sector - since each of them is highly connected to Services - generates a large number of infection in this Services sector. To a lesser extent, we see that Manufacturing and Construction are also affected by this contagion effect.

Additionally, we provide an example of the dynamic evolution of the number of victims in Fig. 3, corresponding to the worst case scenario of an attack on Mining. Compared to Fig. 1, the percentages of victims at a current time are much higher in each sector. We also can observe that the peak is achieved a little bit later than in the first situation (except for the Mining sector, which is fast completely contaminated).

This postponed peak is bad news for the total number of victims (which is typically related to the area under each curve). On the other hand, it also creates an opportunity: this gives more time

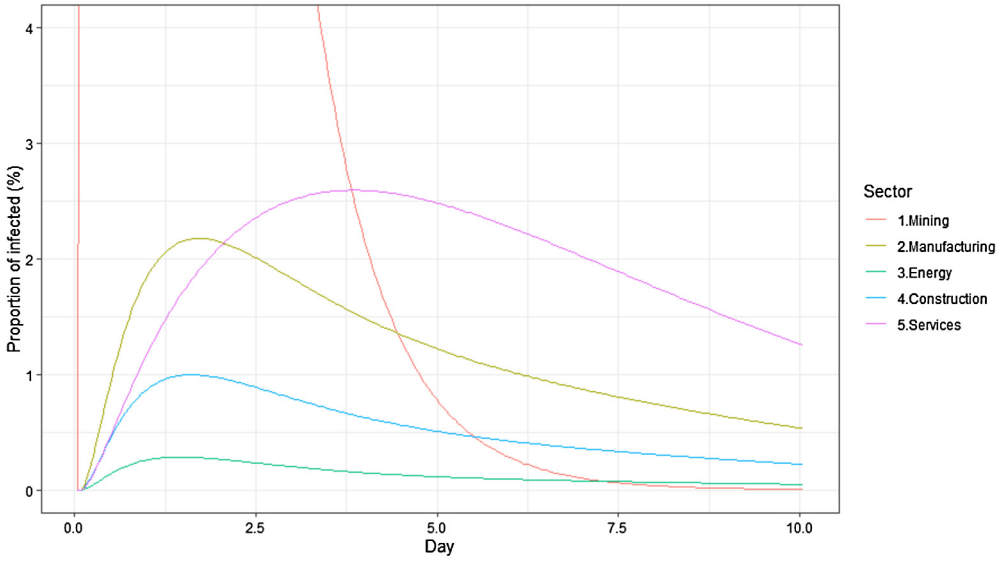

Fig. 3. Evolution of the proportion of infected - Attack on Mining. Note that the y-axis has been bounded to 4%, so that the effects of the infection remain readable for all sectors. Indeed, the value of the peak for the Mining sector is very high at 70% of infected companies after 10 hours of epidemic.

to react to the attack by providing countermeasures. This question of measuring the benefits of such a reaction is the purpose of the following Section.

3.4. Impact of a reaction during the crisis

We now consider the case where, during the crisis, a collective reaction of the victims occurs and reduces the impact of the episode. The impact of the reaction of the population is at the core of several modeling and scenario projections developed in the context of the Covid-19 pandemic, see for example Di Domenico et al. (2020). In the case of cyber events, the length of the attack (a few days) tends to limit the impact of this reaction, but efficient measures to prevent infection by the digital virus seems easier to implement. In the Wannacry case, for example, a « kill switch » was identified (see Mohurle and Patil (2017)) that allowed to diminish the severity.

In model (2.1)-(2.3), we introduce a function η_j that corresponds to the reaction of category j. We here investigate the impact of two particular shapes of reaction functions,

$$\eta_j^{(1)}(t) = 1 - \lambda \mathbf{1}_{i_j(t) \geq s}, \tag{3.1}$$

$$\eta_j^{(2)}(t) = 1 - \lambda \mathbf{1}_{\sum_{k=1}^d i_k(t) \geq s}. \tag{3.2}$$

In each case, a threshold s triggers the reaction: the threat draws the attention and is considered worth taking measures only if a sufficient number of victims have been hit. The difference is that, in the first case, the category j only bothers when its members are hit: a threat making lots of victims in the other sectors does not lead to a reaction as long as category j is preserved. The opposite case is the situation where category j pays attention to what happens to others and reacts accordingly.

We consider three levels of protection, $\lambda = 0.1$, $\lambda = 0.3$ and $\lambda = 0.5$, and three different thresholds of reaction $s = 10,000$, $s = 50,000$ and $s_3 = 100,000$. Tables 7 and 8 show the impact of reactions (3.1) and (3.2) respectively, in case of an uniform initial attack, and when only one single sector j_0 reacts. Denoting $r_j^{(\lambda)}(\infty)$ (resp. $r_j(\infty)$) the total number of victims in category j if sector j_0 reacts (resp. if there is no reaction), the two following ratios are reported in Tables 7 and 8:

- The ratio on the "total" sector $\mathfrak{R}^T := \frac{\sum_{j=1}^d r_j^{(\lambda)}(\infty)}{\sum_{j=1}^d r_j(\infty)}$
- The ratio on the "collateral" sectors that do not react $\mathfrak{R}^c := \frac{\sum_{j \neq j_0} r_j^{(\lambda)}(\infty)}{\sum_{j \neq j_0} r_j(\infty)}$.

In Table 7, some cells have been darkened to reflect the fact that the reaction thresholds are sometimes too large to trigger a reaction of the corresponding sector (this is the case when the threshold exceeds the number of companies in the sector). This situation does not occur in Table 8 since the reaction (3.2) is not only based on what happens in the sector itself, but also on the observation of what happens to the other categories. Clearly, this second type of reactions is more efficient, since it allows to detect quicker that something happens. For some sectors, warning comes sometimes even too late for reaction (3.1) even if the threshold is less than the number of companies in this sector. This is no surprise, but Table 8 helps to quantify the gain obtained through (3.2). We can also observe that the reaction having the most important impact is the one on the Services sector. Let us recall that this sector contains the largest number of companies. This reduction of the size of the cyber epidemic is first of all caused by the fact that less companies are infected, in this sector, due to the reaction. But it is also interesting to notice that this induces effects in the other sectors also, since the collateral gains are quite important too.

To conclude this section, let us mention that these results are only an illustration, and, again, the effects of the attack on these different sectors are related to the particular way the matrix $\mathbf{B}_0$ has been calibrated (focusing of exchanges of added value). We want to emphasize that the contribution we aim to provide is to illustrate, as precisely as possible, a calibration strategy of the model, and what can be learned from it. In the present case, we can observe that maximizing the effect of prevention is not only a matter of finding the most contagious sector regarding the parameters of matrix $\mathbf{B}$. For example, the mining sector has been identified as highly contagious, but its size is small, while prevention on the larger services sector has more effect, even on other segments.

3.5. Cost of claims

In this section we consider an insurance portfolio composed of n policyholders. Following the notations of section 2.1, the m-

Table 7
Impact of the reaction (3.1) on the number of victims, depending on the sector which reacts (only one sector at a time) and on the threshold s activating the reaction, in case of an uniform initial attack. The column "Total" shows the ratio $\mathfrak{R}^T$ and the column "Collateral" shows $\mathfrak{R}^c$.

$\lambda = 10\%$	$s = 10,000$		$s = 50,000$		$s = 100,000$	
Reaction from sector:	Total	Collateral	Total	Collateral	Total	Collateral
Mining						
Manufacturing	96.12%	98.01%	99.54%	99.89%	100%	100%
Energy	100%	100%				
Construction	99.14%	99.72%	100%	100%	100%	100%
Services	74.45%	78.88%	83.41%	87.07%	92.04%	94.50%
$\lambda = 30\%$	$s = 10,000$		$s = 50,000$		$s = 100,000$	
Reaction from sector:	Total	Collateral	Total	Collateral	Total	Collateral
Mining						
Manufacturing	89.15%	94.34%	98.66%	99.67%	100%	100%
Energy	100%	100%				
Construction	97.49%	99.17%	100%	100%	100%	100%
Services	46.16%	53.84%	63.20%	70.30%	80.99%	86.46%
$\lambda = 50\%$	$s = 10,000$		$s = 50,000$		$s = 100,000$	
Reaction from sector:	Total	Collateral	Total	Collateral	Total	Collateral
Mining						
Manufacturing	83.17%	91.05%	97.84%	99.47%	100%	100%
Energy	100%	100%				
Construction	95.93%	98.65%	100%	100%	100%	100%
Services	32.86%	41.16%	52.48%	60.76%	74.08%	81.10%

Table 8
Impact of the reaction (3.2) on the number of victims, depending on the sector which reacts (only one sector at a time) and on the threshold s activating the reaction, in case of an uniform initial attack. The column "Total" shows $\mathfrak{R}^T$ and the column "Collateral" shows the ratio $\mathfrak{R}^c$.

$\lambda = 10\%$	$s = 10,000$		$s = 50,000$		$s = 100,000$	
Reaction from sector:	Total	Collateral	Total	Collateral	Total	Collateral
Mining	99.80%	99.99%	99,83%	99.99%	99.87%	99.99%
Manufacturing	94.60%	96.99%	95.82%	97.82%	97.09%	98.63%
Energy	99.81%	99.98%	99.84%	99.98%	99.88%	99.99%
Construction	98.51%	99.40%	98.87%	99.59%	99.18%	99.74%
Services	73.10%	77.62%	80.40%	84.36%	86.79%	90.05%
$\lambda = 30\%$	$s = 10,000$		$s = 50,000$		$s = 100,000$	
Reaction from sector:	Total	Collateral	Total	Collateral	Total	Collateral
Mining	99.39%	99.95%	99.49%	99.97%	99.61%	99.98%
Manufacturing	84.99%	91.44%	88.33%	93.78%	91.83%	96.07%
Energy	99.42%	99.93%	99.52%	99.95%	99.64%	99.97%
Construction	95.66%	98.24%	96.70%	98.79%	97.62%	99.23%
Services	43.66%	51.37%	57.35%	64.73%	69.97%	76.61%
$\lambda = 50\%$	$s = 10,000$		$s = 50,000$		$s = 100,000$	
Reaction from sector:	Total	Collateral	Total	Collateral	Total	Collateral
Mining	98.97%	99.90%	99.14%	99.93%	99.35%	99.96%
Manufacturing	76.87%	86.55%	81.92%	90.19%	87.25%	93.77%
Energy	99.03%	99.88%	99.21%	99.92%	99.40%	99.95%
Construction	92.99%	97.14%	94.66%	98.03%	96.14%	98.75%
Services	30.04%	38.29%	45.65%	54.04%	60.53%	68.54%

th policyholder experiences an infection at time T_m (with, again, $\mathbb{P}(T_m = \infty) \neq 0$, since some policyholders may not be stroke by the attack). The distribution of T_m depends on the category $\mathbf{x}_m$ to which this policyholder belongs. We assume T_m independent from $T_{m'}$ for $m \neq m'$. It corresponds to the simplest situation where the contagion comes from outside the portfolio and when there is no interaction between policyholders. We consider the five categories of section 3, with the contagion matrix of sections 3.1-3.2 and we assume the same distribution of the categories than the one of the OECD data (see Table 2). In full generality, the composition of the portfolio does not necessarily reflect the proportion of categories in the whole population, since the portfolio results from a selection process. There is no difficulty in handling this situation: the values of $\mathbf{x}_m$ are inputs of the model, and will allow to distort the final results.

The number of policyholders stroke by the event is

$$N_{tot} = \sum_{m=1}^{n} \mathbf{1}_{T_m < \infty}.$$

We then consider a vector of potential losses $(Z_m)_{m=1,\dots,n}$. The total cost of the event is

$$C_{tot} = \sum_{m=1}^{n} Z_m \mathbf{1}_{T_m < \infty}.$$

To simulate the distribution of the variable C_{tot}, we assume that the random variables $(Z_m)_{m=1,\dots,n}$ are independent and identically distributed, and independent from $(T_m)_{m=1,\dots,n}$. We consider an exponential distribution for the random variable Z_1, with mean μ. In order to obtain reasonable values for the loss, we consider two different values for μ, corresponding to studies conducted on the French market of insurance brokers, and in particular the study

Table 9
Mean value of the cost per policyholder, and associated standard deviation. The amounts are given in k€.

Targeted sector	μ_1 $E[C_{tot}]/n$	$\sqrt{Var(C_{tot})/n}$	μ_2 $E[C_{tot}]/n$	$\sqrt{Var(C_{tot})/n}$
Uniform attack	0.35	1.77	0.87	4.47
Attack on Mining	0.81	2.63	2.05	6.66
Attack on Manufacturing	0.67	2.41	1.69	6.10
Attack on Energy	0.55	2.20	1.40	5.58
Attack on Services	0.17	1.25	0.43	3.18
Attack on Construction	0.05	0.68	0.12	1.72

"LUCY"[3] conducted in 2021 by the AMRAE.[4] According to this analysis, the average cost of a cyber claim in France was $\mu_1 = 4.68$ k€ in 2019 and $\mu_2 = 11.84$ k€ in 2020. Let us nevertheless stress that these statistics have been gathered on a population of policyholders which is mainly composed of large companies of the private sector.

The expectation and variance of C_{tot} are easy to obtain, since

$$E[C_{tot}] = \lambda E[N_{tot}] = n\mu\mathbb{P}(T_1 < \infty), \text{ and}$$
$$Var(C_{tot}) = n\mu^2\mathbb{P}(T_1 < \infty)\{2 - \mathbb{P}(T_1 < \infty)\}.$$

Let us note that $E[N_{tot}]$ is supposed to be close to the total number of infected from Table 5 times n/N (since n/N is the proportion of the total affected population represented by the portfolio). The values of $E[C_{tot}]/n$ and $Var(C_{tot})/n$ are given in Table 9. Moreover, the whole distribution of C_{tot} can be approximated by a Gaussian distribution (this is a consequence of assuming that Z_m are i.i.d. with a second order moment), allowing to approximate the quantiles of C_{tot}.

Let us note that $E[C_{tot}]$ does not represent a pure premium in our case. Indeed, C_{tot} represents the total cost of the contagion episode, which means conditionally that this particular contagion scenario occurs. Although we chose not to emphasize it in order to simplify the notations, $E[C_{tot}]$ depends on the contagion matrix $\mathbf{B} = \beta\mathbf{B}_0$, and the intensity of initial attacks on each class, say $\mathbf{A}(\cdot) = (\alpha_1(\cdot), ..., \alpha_d(\cdot)) \in \mathfrak{A}$, where $\mathfrak{A}$ is a functional space representing the possible scheme of initial attacks. In other words, $E[C_{tot}] = f_{\mathbf{B}_0}(\beta, \mathbf{A})$. If one assumes that, from one attack to another, the matrix $\mathbf{B}_0$ (representing the connectivity) stays the same, the pure premium becomes

$$\pi = \int f_{\mathbf{B}_0}(\beta, \mathbf{A}) dp(\beta, \mathbf{A}),$$

where p is a probability distribution on $[0, \infty] \times \mathfrak{A}$. Computing π is then feasible only if one adds some prior expertise (contained in the distribution p) on the possible types of attacks. Let us recall that the aim of our model is less to compute a premium than to allow to simulate stress scenarios that helps to understand how the portfolio reacts to some generic types of cyber episodes.

Remark 3.2. In the present analysis, we assume that the cost of a claim has the same distribution independently from the category to which the policyholder belongs. This is of course a strong assumption. In practice, this cost would be dependent of the sector of activity. It would also dependent on several other factors like the size of the targeted company (as mentioned above, the policyholders in the perimeter of the AMRAE study are mostly large companies), and, of course, the perimeter of the contract. Since the present paper essentially focuses on a model for counting the number of victims, we do not explore more deeply this path. Moreover, introducing an heterogeneity in the cost would blur the conclusions on the analysis of the impact of contagion, which is our main purpose.

[3] https://www.amrae.fr/bibliotheque-de-amrae/lucy-lumiere-sur-la-cyberassurance-amrae-mai-2021.

[4] Association pour le Management des Risques et les Assurances de l'Entreprise, French association of risk managers, see https://www.amrae.fr/.

4. Simulations

In this section, we try to evaluate via simulations how the connectivity of the network (namely, the structure of the matrix $\mathbf{B}$) can have influence on the final number of victims. As in Section 3, we consider matrices of the form $\mathbf{B} = \beta\mathbf{B}_0$. The matrix $\mathbf{B}_0$ is chosen in different classes of matrices, each corresponding to some properties that we want the network to satisfy. To simplify the discussion, we consider the effect of the network in absence of reaction.

The different classes of matrices (which all are normalized, in the sense that the sum of the values of all their coefficients is equal to 1) are the following:

- $\mathcal{B}_1$: an "homogeneous" case, where there is no particular structure, coefficients are generated randomly (independent uniform distributions are used to simulate each coefficient, then the matrix is normalized);
- $\mathcal{B}_2$: a "clustered" case, where for all $j = 1, ..., d$, $\sum_{k\neq j} \beta_{k,j} \leq \beta_{j,j}$. This situation corresponds to the case where the contagion occurs mostly within a given sector, and can extend to others with less intensity;
- $\mathcal{B}_3$: a "non-clustered" case, where for all $j = 1, ..., d$, $\beta_{j,j} \leq \min_{k\neq j} \beta_{k,j}$. This corresponds to a situation where contagion mostly occurs from outside a given sector, and where there are few contagions among susceptibles of the same category;
- $\mathcal{B}_4$: a "cascade" case where all the coefficients of the matrix are 0 except for $\beta_{j,j}$ for $j = 1, ..., d$, and $\beta_{j,j+1}$ for $j = 1, ..., d-1$. This corresponds to a potential cascade effect, since an infection coming from the first category must first contaminate the second, before infecting the third, and so on. The last category does not contaminate any other class.

The simulation procedure is exposed in section 4.1. We next introduce in section 4.2 a way to measure the impact of each scenario that we consider. Results are gathered in section 4.3.

4.1. Simulation procedure

For each class of matrices $\mathcal{B}_j$, we generate randomly (uniformly over $\mathcal{B}_j$) $B = 100,000$ matrices, $(\mathbf{B}_0^{(b)})_{b=1,...,B}$. For each b, we study the impact of different types of attacks on different structures of population. The value β is always taken as $\beta = 2 \times 10^{-5}$ (close to the value of Section 3). Regarding the structure of the population of potential victims (decomposed into 5 categories), we consider three different compositions, whose total size is taken as $N = 5,000,000$ (same scale as N calibrated in Section 3):

- Configuration 1: homogeneous, that is each category is made of 1,000,000 potential victims;
- Configuration 2: a class is larger than the others, with size 1,800,000 (size 800,000 for the others);
- Configuration 3: a class is smaller than the others, with size 600,000 (while the 4 others have size 1,100,000).

In Configurations 2 and 3, we call the "special class" the category which has not the same population as the others (which are denominated "standard classes" in the following). Next, we consider different type of attacks, targeted on a single class j_0 with intensity attack of the type $\alpha_{j_0}(t) = \alpha \mathbf{1}_{t\leq 1}$. In Configuration 1, since all classes are similar in terms of composition, it does not matter which class is initially stroke. In Configuration 2 and 3, we distinguish two cases: attack on the special class or on a standard class.

4.2. Metric used to measure the impact of each scenario

For each attack, we evaluate the value of α which allows to achieve the same number of victims as Wannacry (estimated at 300,000). For two populations of the same size, a higher α shows that the hackers need to make stronger efforts to achieve the same effect. In other words, the structure of the network is more favorable, in the sense that it slows down the epidemic. To better identify this effect, we compare this value of α (say α^W) to the value α^* that would be required if the whole population behaved as the initially targeted population. More precisely, if we are considering a cyber episode obtained from a matrix $\beta \mathbf{B}_0^{(b)}$ and an attack on the class j, we consider a benchmark case where we consider a single homogeneous population of size N (that is a one-dimensional case where $d = 1$ with contagion parameter $\beta \times B^{(b)}_{0,(j,j)}$, where $B^{(b)}_{0,(j,j)}$ is the j-th diagonal coefficient of $\mathbf{B}_0^{(b)}$. In this benchmark case, no contagion effect occurs, only contamination inside a single isolated category similar to the one where the attack was launched. The value α^* is the one required so that the initial attack allows to achieve 300,000 victims. We next compute the ratio, $\rho = [N_a \alpha^W - N\alpha^*][N\alpha^*]^{-1}$, with N_a the size of the population that is attacked in the corresponding scenario. A high value for ρ indicates that the type of structure considered tends to slow down the transmission and to mitigate the impact of the episode.

4.3. Simulation results

We report in Table 10 the mean value and the median value of the indicator ρ (over these 100,000 replications), for different configurations and different targeted classes, depending on the contagion matrix class. In each case, we see that the network structure seems to slow down the infection, compared to the situation where all the infected belong to the same group. $\mathcal{B}_1$ provides some kind of benchmark case (since there is no particular structure in the network). From these results, we can also observe that a matrix of type $\mathcal{B}_3$ seems the less favorable situation, since the intensity of attacks required to trigger a Wannacry-type event is lower. This is no big surprise: with a structure such as $\mathcal{B}_3$, the cyber attack has a low propagation rate inside a given sector. But the contagion quickly spreads to all other sectors, having the ability to rapidly expose the whole population from cross-infections. Indeed, the size of the initially infected category here seems determinant: let us recall that, in Configuration 2 (resp. 3), the "special" (resp. "standard") category is the largest sector, and we see that the losses are more important if it is the initial target, no matter the class of matrices considered (and especially in the case of $\mathcal{B}_2$). With this fast propagation of the attack to other sectors for networks of type $\mathcal{B}_3$, the attack spreads quickly to the whole population.

5. Conclusion

Multi-group SIR models are simple tools to describe an epidemic, and are widely used in epidemiology. Their adaptation to the study of contagious cyber events seems relevant due to the ability of such models to take into account differences of connectivity between groups of actors. They can also be easily modified to capture particular shapes of attacks. In this paper, we showed how this model can be used to investigate the impact of a particular shape of network, allowing to identify potential weaknesses in a portfolio. Moreover, we want to emphasize that, although this model may not provide the most accurate physical description of the phenomenon, its calibration is relatively easy. Indeed, obtaining precise information about how policyholders are connected with each other (and with potential other sources of infection) is really hard to get. In the example that we provided, we showed that the model we develop can be calibrated from a small amount of data at a macroscopic level. We point out that this illustration is only a rough example, based on publicly available data related to some kind of economic connectivity between actors. The aim of this example is only to show a methodology of calibration, and what can be obtained from it.

To conclude, let us also mention that the model can be adapted to take into account not only cyber risk, but also its consequences when it comes to breaking the supply chain. Indeed, a cyber infection can then contaminate other companies not only digitally: a business interruption in a given sector, from which another sector is very dependent in terms of supply, could generate some losses for victims that are not directly targeted by the initial attack. The model we develop can also capture these types of situation, after adaptation of the parameters (taking for example into account that some reserves of the product whose supply has been disturbed can delay the propagation of the infection).

Declaration of competing interest

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Caroline Hillairet and Olivier Lopez acknowledge funding from the project *Cyber Risk Insurance: actuarial modeling*, Joint Research Initiative under the aegis of Risk Foundation, with partnership of AXA, AXA GRM, ENSAE and Sorbonne Université. The authors thank the anonymous referees of this paper for their insightful remarks and comments.

Appendix A

A.1. Sensitivity analysis

To evaluate the sensitivity of the result to the parameters of the contagion matrix, we applied a shock on every coefficient of $\mathbf{B}_0$ separately. The result that we show below correspond to a shock of magnitude 10%, before renormalizing the coefficients of matrix $\mathbf{B}_0$. The effects of these augmentations of the contagiousness are shown in Table 11 through the number of infected policyholders during the crisis.

One can observe that an increase of the contagion coming from the mining sector (line "Mining" of Table 11) does not lead to a strong change in the result, although an attack on mining was identified as particularly concerning according to Table 9. On the other hand, changing the coefficients corresponding to services has a stronger impact: this is caused by the fact that contagion is not

Table 10
Mean and Median values of ρ computed from 100,000 simulations of matrices from the classes $\mathcal{B}_j$ for $j = 1, \cdots, 4$.

Configuration	Targeted class	Special		Standard	
		Mean of ρ	Median of ρ	Mean of ρ	Median of ρ
$\mathcal{B}_1$	Configuration 1			5.01	6.02
	Configuration 2	4.55	5.48	5.41	6.49
	Configuration 3	6.05	7.23	5.05	6.06
$\mathcal{B}_2$	Configuration 1			2.86	1.67
	Configuration 2	2.42	1.44	6.51	6.52
	Configuration 3	7.39	7.40	2.79	1.62
$\mathcal{B}_3$	Configuration 1			0.50	0.37
	Configuration 2	0.20	0.10	0.37	0.25
	Configuration 3	0.50	0.38	0.21	0.10
$\mathcal{B}_4$	Configuration 1			2.89	1.90
	Configuration 2	2.57	1.69	3.12	2.10
	Configuration 3	3.53	2.39	2.82	1.88

Table 11
Marginal effect of an increase of 10% of a single coefficient of matrix $\mathbf{B}_0$. The value in cell (k, j) corresponds to a shock on the coefficient (k, j) of Table 3, and is shown as a relative variation in the number of total victims (in percentage of the total number of victims of the scenario).

	Mining	Manufacturing	Energy	Construction	Services
Mining	0.36%	0.02%	0.12%	0.10%	0.005%
Manufacturing	0.10%	2.31%	0.11%	0.12%	4.68%
Energy	0.12%	0.08%	0.12%	0.11%	0.02%
Construction	0.11%	0.15%	0.12%	0.16%	0.99%
Services	0.10%	2.92%	0.10%	0.50%	3.55%

only a matter of a large value for $\beta_{k,j}$, but also results from the number of companies directly affected by this change, that is the number of companies in categories k and j.

A.2. Proof of Theorem 2.3

We provide the proof of Theorem 2.3 that characterizes the total number of victims as the solution of the fixed point equation $\mathbf{r} = \Phi(\mathbf{r})$, on $\mathcal{R} = \{\mathbf{r} : 0 \le r_j \le s_j(0) + i_j(0)\}$, in the case there is no reaction (that is $\eta_j = 1$ for all j). We recall the notations:

$$\Phi_j(\mathbf{x}) = i_j(0) + s_j(0)\left\{1 - \exp\left(-\left(\mathcal{A}_j + \sum_{k=1}^{d}\frac{\beta_{k,j}}{\gamma_j}x_k\right)\right)\right\},$$

$$\text{with} \quad \mathcal{A}_j = \int_0^\infty \alpha_j(t)dt.$$

From (2.3) we have $r_j(\infty) = \gamma_j I_j$, with $I_j = \int_0^\infty i_j(t)dt$. On the other hand, $r_j(\infty) = s_j(0) - s_j(\infty) + i_j(0)$, since all infected become recovered in a finite time and, from (2.1),

$$s_j(\infty) = s_j(0)\exp\left(-\left(\mathcal{A}_j + \sum_{k=1}^{d}\beta_{k,j}I_k\right)\right).$$

We see that $\mathbf{r}(\infty) = (r_j(\infty))_{1\le j\le d}$ is a fixed point of Φ, that is

$$\mathbf{r}(\infty) = \Phi(\mathbf{r}(\infty)).$$

The question is now to prove the unicity of this fixed point. For two vectors $\mathbf{x}$ and $\mathbf{y}$, we write $\mathbf{x} \le \mathbf{y}$ if all the components of $x_j \le y_j$ for all j. If, in addition $x_j < y_j$ for at least one j, we say that $\mathbf{x} < \mathbf{y}$. Note that $\Phi_j(\mathbf{x}) < s_j(0) + i_j(0)$ for all $\mathbf{x} \in \mathcal{R}$. Besides, since we have excluded the trivial case where for all j, $i_j(0) = 0$ and $\mathcal{A}_j = 0$ we get $\mathbf{0} < \Phi(\mathbf{0})$.

Clearly, if $\mathbf{x} \le \mathbf{y}$, $\Phi(\mathbf{x}) \le \Phi(\mathbf{y})$. By induction, we therefore get that

$$\mathbf{0} < \Phi(\mathbf{0}) \le \Phi^{(2)}(\mathbf{0}) \le \cdots \le \Phi^{(k)}(\mathbf{0}) \le \Phi^{(k)}(\mathbf{n}(0)) \le \cdots \le \Phi(\mathbf{n}(0)) < \mathbf{n}(0),$$

where $\Phi^{(k)}(\mathbf{x}) = \Phi(\Phi^{(k-1)}(\mathbf{x}))$ (with $\Phi^{(0)}(\mathbf{x}) = \mathbf{x}$), and $\mathbf{n}(0) = (s_j(0) + i_j(0))_{1\le j\le d}$. This shows that both sequences $\Phi^{(k)}(\mathbf{0})$ and $\Phi^{(k)}(\mathbf{n}(0))$ converge to a finite limit, respectively denoted by $\mathfrak{l}_0$ and $\mathfrak{l}_n$. Necessarily, since Φ is continuous, $\Phi(\mathfrak{l}_0) = \mathfrak{l}_0$, $\Phi(\mathfrak{l}_n) = \mathfrak{l}_n$. Moreover, $\mathfrak{l}_0 \le \mathfrak{l}_n$.

The next step consists in showing that $\mathfrak{l}_0 = \mathfrak{l}_n$. We will proceed by contradiction, assuming that

$$\mathfrak{d} = \mathfrak{l}_n - \mathfrak{l}_0 > 0. \tag{A.1}$$

Let $J_\Phi(\mathbf{x}) = (\partial_j\Phi_k(\mathbf{x}))_{1\le j,k\le d}$ denote the Jacobian matrix of Φ, where ∂_k denotes the partial derivative with respect to the k-th component. If $\mathfrak{l}_0 < \mathfrak{l}_n$, we could write

$$\mathfrak{d} = \Phi(\mathfrak{l}_n) - \Phi(\mathfrak{l}_0) = \int_0^1 J_\Phi(\mathfrak{l}_0 + t\mathfrak{d})\mathfrak{d}dt. \tag{A.2}$$

Observe that

$$\partial_k\Phi_j(\mathbf{x}) = \frac{\beta_{k,j}}{\gamma_j}\big((s_j(0) + i_j(0)) - \Phi_j(\mathbf{x})\big) > 0.$$

Hence, the differential of Φ inherits some monotonicity properties of Φ, in the sense that, for all $\mathbf{x} \le \mathbf{y}$ and $\mathfrak{h} \ge \mathbf{0}$,

$$J_\Phi(\mathbf{x})\mathfrak{h} \ge J_\Phi(\mathbf{y})\mathfrak{h}.$$

This, combined with (A.2), leads to

$$\mathfrak{d} \le J_\Phi(\mathfrak{l}_\mathbf{0})\mathfrak{d}. \tag{A.3}$$

Let ρ denote the spectral radius of $J_\Phi(\mathfrak{l}_\mathbf{0})$. Since all the coefficients of $J_\Phi(\mathfrak{l}_\mathbf{0})$ are positive, hence the matrix is irreducible, and we can apply the Perron-Frobenius Theorem to ensure that there

exists some eigenvector $\mathbf{y}_0 > 0$ such that $\mathbf{y}_0^{tr} J_\Phi(\mathfrak{l}_\mathbf{0}) = \rho \mathbf{y}_0^{tr}$. Hence, from (A.3),

$$\mathbf{y}_0^{tr}\mathfrak{d} \leq \rho \mathbf{y}_0^{tr}\mathfrak{d},$$

which implies that $\rho \geq 1$. On the other hand, we have

$$\mathfrak{l}_0 = \Phi(\mathfrak{l}_0) = \int_0^1 J_\Phi(0 + t\mathfrak{l}_\mathbf{0})\mathfrak{l}_0 dt > J_\Phi(\mathfrak{l}_\mathbf{0})\mathfrak{l}_0.$$

This implies that

$$\mathbf{y}_0^{tr}\mathfrak{l}_0 > \mathbf{y}_0^{tr} J_\Phi(\mathfrak{l}_\mathbf{0})\mathfrak{l}_0 = \rho \mathbf{y}_0^{tr}\mathfrak{l}_0. \tag{A.4}$$

Since $\mathbf{y}_0 > 0$ and $\rho \geq 1$, (A.4) contradicts the fact that $\mathfrak{l}_0 \geq \Phi(\mathbf{0}) > \mathbf{0}$. Hence, necessarily, (A.1) is wrong and $\mathfrak{l}_n = \mathfrak{l}_0$, which shows the unicity of the fixed point.

A.3. Solving the fixed point problem

We study here the fixed point problem of Theorem 2.3 to determine the total number of victims $\mathbf{r}(\infty)$. According to Theorem 2.3, we can approximate $\mathbf{r}(\infty)$ using a recurrent sequence $\mathbf{u}_{n+1} = \Phi(\mathbf{u}_n)$ initialized for example at $\mathbf{u}_0 = \mathbf{n}(0) = (s_j(0) + i_j(0))_{1\leq j\leq d}$. In some situations, the rate of convergence can be shown to be geometric.

Consider the special case where the matrix $\mathbf{B}$ is diagonally dominant, that is for all j, $\beta_{j,j} \geq \sum_{k=1,k\neq j}^d \beta_{k,j}$. This corresponds to the special case where the contagion is stronger within each given group than with respect to other actors. In this case, if the intensity of attacks is strong enough, one can derive a rate of convergence for $\mathbf{u}_n$. Indeed, the differential of Φ is a contracting application. First of all, from (2.1),

$$s_j(t) \leq s_j(0) \exp\left(-\int_0^t \alpha_j(s)ds\right),$$

which leads to

$$s_j(\infty) = s_j(0) - r_j(\infty) \leq s_j(0) \exp\left(-\int_0^\infty \alpha_j(s)ds\right),$$

therefore $\mathcal{R}$ can be replaced by $\widetilde{\mathcal{R}} = \{\mathbf{r} : s_j(0)[1 - \exp(-\mathcal{A}_j)] \leq r_j \leq s_j(0) + i_j(0)\}$.

We have, for $\mathbf{h} \in \mathbb{R}^d$ with $\|\mathbf{h}\|_\infty = 1$, and $\mathbf{x} \in \widetilde{\mathcal{R}}$

$$\begin{aligned}
|(J_\Phi(\mathbf{x})\mathbf{h})_j| &= \left|\sum_{k=1}^d \partial_k \Phi_j(\mathbf{x}) h_k\right| \\
&= \left|\sum_{k=1}^d \frac{\beta_{k,j}}{\gamma_j}(s_j(0) + i_j(0) - \Phi_j(\mathbf{x}))h_k\right| \\
&= \left|\sum_{k=1}^d \frac{\beta_{k,j}}{\gamma_j} s_j(0) \exp\left(-\left(\mathcal{A}_j + \sum_{k=1}^d \frac{\beta_{k,j}}{\gamma_j} x_k\right)\right) h_k\right| \\
&\leq 2\frac{\beta_{j,j} s_j(0)}{\gamma_j} \exp\left(-\left(\mathcal{A}_j + \frac{\beta_{j,j}}{\gamma_j} x_j\right)\right) \\
&\leq 2\frac{s_j(0)}{x_j}\frac{\beta_{j,j} x_j}{\gamma_j} \exp\left(-\left(\mathcal{A}_j + \frac{\beta_{j,j}}{\gamma_j} x_j\right)\right) \\
&\leq \frac{s_j(0)}{x_j} \exp(-\mathcal{A}_j) \leq \frac{\exp(-\mathcal{A}_j)}{1 - \exp(-\mathcal{A}_j)},
\end{aligned}$$

where the first inequality comes from the diagonally dominance condition on $\mathbf{B}$ and the third one from the inequality $2x\exp(-x) < 1$ for all $x \geq 0$. As a consequence, if $\mathcal{A}_j > \log 2$ for all j, $\|\Phi(\mathbf{x}) - \Phi(\mathbf{y})\|_\infty \leq M\|\mathbf{x} - \mathbf{y}\|_\infty$, with $M = \sup_{j=1,\dots,d} \exp(-\mathcal{A}_j)[1 - \exp(-\mathcal{A}_j)]^{-1} < 1$. This leads to

$$\|\mathbf{u}_n - \mathbf{n}(0)\|_\infty \leq M^n \|\mathbf{u}_1 - \mathbf{n}(0)\|_\infty.$$

Even if the assumptions of this particular case do not hold, the convergence is nevertheless quite fast in the applications we consider in the paper.

References

Adams, N.M., Heard, N.A., 2014. Data Analysis for Network Cyber-Security. World Scientific.

Agence Nationale de la Sécurité des Systèmes d'Information, 2021. Etat de la menace rançongiciel. https://www.cert.ssi.gouv.fr/uploads/CERTFR-2021-CTI-001.pdf.

Al-rimy, B.A.S., Maarof, M.A., Shaid, S.Z.M., 2018. Ransomware threat success factors, taxonomy, and countermeasures: a survey and research directions. Computers & Security 74, 144–166.

Amann, H., 2011. Ordinary Differential Equations: An Introduction to Nonlinear Analysis, vol. 13. Walter de Gruyter.

Andersson, H., Britton, T., 2012. Stochastic Epidemic Models and Their Statistical Analysis, vol. 151. Springer Science & Business Media.

Andreasen, V., 2011. The final size of an epidemic and its relation to the basic reproduction number. Bulletin of Mathematical Biology 73 (10), 2305–2321.

Antonio, Y., Indratno, S.W., Saputro, S.W., 2021. Pricing of cyber insurance premiums using a Markov-based dynamic model with clustering structure. PLoS ONE 16 (10), e0258867.

Beretta, E., Capasso, V., 1988. Global stability results for a multigroup SIR epidemic model. In: Hallam, T.G., Gross, L.J., Levin, S.A. (Eds.), Mathematical Ecology, pp. 317–342.

Bessy-Roland, Y., Boumezoued, A., Hillairet, C., 2021. Multivariate Hawkes process for cyber insurance. Annals of Actuarial Science 15 (1), 14–39.

Böhme, R., Schwartz, G., et al., 2010. Modeling cyber-insurance: towards a unifying framework. In: WEIS.

Boyes, H., 2015. Cybersecurity and cyber-resilient supply chains. Technology Innovation Management Review 5 (4), 28.

Brauer, F., Castillo-Chavez, C., Castillo-Chavez, C., 2012. Mathematical Models in Population Biology and Epidemiology, vol. 2. Springer.

Brauer, F., Castillo-Chavez, C., Feng, Z., 2019. Mathematical Models in Epidemiology, vol. 32. Springer.

Brauer, F., Van den Driessche, P., Wu, J., Allen, L.J., 2008. Mathematical Epidemiology, vol. 1945. Springer.

Cashell, B., Jackson, W.D., Jickling, M., Webel, B., 2004. The economic impact of cyber-attacks. Congressional research service documents, CRS RL32331 (Washington, DC), 2.

Chen, Q., 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). IEEE, pp. 454–460.

Daley, D.J., Gani, J., 2001. Epidemic Modelling: An Introduction. Cambridge University Press. Number 15.

Di Domenico, L., Pullano, G., Sabbatini, C.E., Boëlle, P.-Y., Colizza, V., 2020. Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies. BMC Medicine 18 (1), 1–13.

Diekmann, O., Heesterbeek, J.A.P., 2000. Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation, vol. 5. John Wiley & Sons.

Diekmann, O., Heesterbeek, J.A.P., Metz, J.A., 1990. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. Journal of Mathematical Biology 28 (4), 365–382.

Fahrenwaldt, M.A., Weber, S., Weske, K., 2018. Pricing of cyber insurance contracts in a network model. ASTIN Bulletin: The Journal of the IAA 48 (3), 1175–1218.

Fayi, S.Y.A., 2018. What Petya/NotPetya ransomware is and what its remidiations are. In: Information Technology-New Generations. Springer, pp. 93–100.

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

Feng, R., Garrido, J., Jin, L., Loke, S.-H., Zhang, L., 2022. Epidemic Compartmental Models and Their Insurance Applications. Springer International Publishing, Cham, pp. 13–40.

Ghadge, A., Weiß, M., Caldwell, N.D., Wilding, R., 2019. Managing cyber risk in supply chains: a review and research agenda. Supply Chain Management.

Guo, H., Li, M.Y., Shuai, Z., 2006. Global stability of the endemic equilibrium of multigroup SIR epidemic models. The Canadian Applied Mathematics Quarterly 14 (3), 259–284.

Heffernan, J.M., Smith, R.J., Wahl, L.M., 2005. Perspectives on the basic reproductive ratio. Journal of the Royal Society Interface 2 (4), 281–293.

Hillairet, C., Lopez, O., 2021. Propagation of cyber incidents in an insurance portfolio: counting processes combined with compartmental epidemiological models. Scandinavian Actuarial Journal 2021 (8), 671–694.

Hobbs, A., 2021. The Colonial Pipeline hack: Exposing vulnerabilities in us cybersecurity.
Kao, D.-Y., Hsiao, S.-C., 2018. The dynamic analysis of Wannacry ransomware. In: 2018 20th International Conference on Advanced Communication Technology (ICACT). IEEE, pp. 159–166.
Kermack, W., McKendrick, A., 1927. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London, Series A 115, 700–721.
Kshetri, N., 2020. The evolution of cyber-insurance industry and market: an institutional analysis. Telecommunications Policy 44 (8), 102007.
Lallie, H.S., Shepherd, L.A., Nurse, J.R., Erola, A., Epiphaniou, G., Maple, C., Bellekens, X., 2021. Cyber security in the age of COVID-19: a timeline and analysis of cyber-crime and cyber-attacks during the pandemic. Computers & Security 105, 102248.
Lefèvre, C., Simon, M., 2021. Ruin problems for epidemic insurance. Advances in Applied Probability 53 (2), 484–509.
Lefèvre, C., Picard, P., Simon, M., 2017. Epidemic risk and insurance coverage. Journal of Applied Probability 54 (1), 286–303.
Low, P., 2017. Insuring against cyber-attacks. Computer Fraud & Security 2017 (4), 18–20.
Magal, P., Seydi, O., Webb, G., 2018. Final size of a multi-group SIR epidemic model: irreducible and non-irreducible modes of transmission. Mathematical Biosciences 301, 59–67.
McKendrick, A.G., 1925. Applications of mathematics to medical problems. Proceedings of the Edinburgh Mathematical Society 44, 98–130.
Mohurle, S., Patil, M., 2017. A brief study of Wannacry threat: ransomware attack 2017. International Journal of Advanced Research in Computer Science 8 (5), 1938–1940.
OECD, 2015. Structural business statistics. ISIC rev. 4. https://www.oecd-ilibrary.org/content/data/8e34f7e7-en.
OECD, 2018. Origin of value added in final demand. https://www.oecd-ilibrary.org/content/data/data-00827-en.
Perasso, A., 2018. An introduction to the basic reproduction number in mathematical epidemiology. ESAIM: Proceedings and Surveys 62, 123–138.
Romanosky, S., Ablon, L., Kuehn, A., Jones, T., 2019. Content analysis of cyber insurance policies: how do carriers price cyber risk? Journal of Cybersecurity 5 (1).
van den Driessche, P., Watmough, J., 2002. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Mathematical Biosciences 180 (1), 29–48.
Welburn, J.W., Strong, A.M., 2019. Systemic cyber risk and aggregate impacts. Risk Analysis (1).
Xie, X., Lee, C., Eling, M., 2020. Cyber insurance offering and performance: an analysis of the US cyber insurance market. The Geneva Papers on Risk and Insurance. Issues and Practice 45 (4), 690–736.