Cyber Insurance and Risk Assessment: Some Insights on the Insurer Perspective

Maria Francesca Carfora and Albina Orlando(✉)

Istituto per le Applicazioni Del Calcolo "Mauro Picone". Consiglio Nazionale delle Ricerche, Via P. Castellino, 80131 Napoli, Italy
{f.carfora,a.orlando}@na.iac.cnr.it

Abstract. Cyber insurance is a crucial tool for managing risks associated with cyber threats. A challenging task for insurance companies lies in pricing cyber risk. Our study is motivated by the reasonable assumption that firms entering into cyber insurance contracts face diverse cyber threats in terms of types and magnitude. Considering these differences ensures that premiums align with the actual risk exposure of the insured. The study discusses this approach proposing a case study based on the Chronology of Data Breaches provided by the Privacy Rights Clearinghouse.

Keywords: cyber risk · cyber insurance · premium · data breaches

1 Introduction

In contemporary society, our dependence on information systems offers significant opportunities but also brings new risks. Cyber insurance emerges as a key tool for managing these risks. Insurers not only provide the opportunity to relieve insureds from the need to accumulate capital for handling catastrophic events, but they also have the potential to incentivize appropriate cybersecurity measures through premiums and proactive security screening. Over the past five years, the worldwide cyber insurance market has tripled in size and projections indicate a further increase [1]. This industry faces unique challenges and obstacles that are not commonly encountered in traditional insurance markets, such as addressing the correlation of risks, managing the geographical dispersion of risk and dealing with the lack of historical and actuarial data [2]. Several papers extensively examine the relevant literature on cybersecurity insurance, research and practice, in order to draft the current landscape and present the trends, among the others [2,3]. Very insightful contributions concern the possibility of transferring cyber risk through insurance-linked securities (see [5]). A challenging task for insurance companies concerns the pricing of cyber risk even due to the lack of comprehensive data on security breaches and losses. Information on the current industry practices for pricing risks is available in [4]. Remarkable contributions on the topic of insurance policies pricing are given in [6–8].

M. Corazza et al. (Eds.): MAF 2024, *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pp. 80–86, 2024.
https://doi.org/10.1007/978-3-031-64273-9_14

Our study is motivated by the reasonable assumption that a firm signing a cyber insurance contract faces different cyber threats in terms of types and magnitude. Taking into account these differences, enables more precise pricing of cyber risk and guarantees that premiums align with the actual risk exposure of the insured. This approach allows to mitigate the risks associated with underinsurance or overinsurance, thereby minimizing the potential for premium leakage. The rest of the paper is structured as follows. Section 2 focuses on the pricing methodology. Section 3 discusses a case study and Sect. 4 concludes.

2 Pricing Cyber Risk

The standard assumptions of classical actuarial techniques are not as applicable to price cyber risk. In actuarial mathematics a standard model is the frequency-severity approach, also called collective risk model [9]. Despite the limitations of this approach in the context of the quantification of cyber risk, in any case it can be customized to account for cyber risk-at least as a first approximation. Let us consider a policy covering a given risk. During the *policy year*, a random number N (frequency) of claims will be recorded. Each claim will cause a random loss, L_k k=1, 2, 3,...(severity), to the insured. The insurer will assess the claim amount Y_k for claim k. In case of partial cover we have $Y_k < L_k$, while $Y_k = L_k$ in case of full compensation.

Referring to a single policy, the total payout of the insurer X (aggregate claim amount) during the policy year, is defined as follows: $X = 0$, if $N = 0$ and

$$X = \sum_{k=1}^{N} Y_k, N > 0 \tag{1}$$

The *equivalence premium* (fair premium) Π is given by the expected value of the insurer's payout $E[X]$:

$$\Pi = E[X]. \tag{2}$$

Usually the following assumptions hold: the random variables L_k, k = 1, 2,...,N, are independent of the random number N and the random variables L_k, k = 1, 2,...,N are mutually independent and identically distributed. Typically, insurer adds a safety loading to the fair premium thus obtaining the so-called *net premium*, that is, before loading expenses. Resulting principles are [10]:

the *expected value principle*, $P_{ev} = (1 + \alpha)\Pi$;
the *variance principle*, $P_{var} = \Pi + \alpha Var(X)$;
the *standard deviation principle*, $P_{sd} = \Pi + \alpha\sqrt{Var(X)}$;
the *semistandard deviation principle*, $P_{ssd} = \Pi + \alpha\sqrt{E\{[max(0, X - E(X)]^2\}}$
where $\alpha > 0$ is a constant.

Other premium principles are defined via *utility theory* and incorporate the attitude towards risk of the insurer. One example is the *exponential premium principle*. Given an appropriate constant $\rho > 0$, we have $P_{ut} = \frac{1}{\rho}log[E(e^{\rho \cdot X})]$.

Another principle of the premium assessment is based on the quantiles of the distribution of X and is given by $P_Q(\varepsilon) = F_X^{-1}(1-\varepsilon)$ where F is the distribution function of X and $\varepsilon \in (0,1)$ is the confidence level. It is the quantile of order $(1-\varepsilon)$ of the loss distribution and this means that the insurer wants to get the premium that covers $(1-\varepsilon) \cdot 100\%$ of the possible loss. A reasonable range of the parameter ε is usually from 1% to 5%.

Whatever is the principle to be used for the premium assessment, we need to make realistic assumption on the distribution of both the number of claims N and the claim amounts $Y_1, Y_2, ..., Y_N$. In the following, we consider two possible solutions. In the first case, the insurer takes into account that different claims can be caused by different type of incidents. This approach is described in Sect. 3. The second solution consists in estimating the payout of the insurer considering the distribution of the total aggregate claim amount X.

In order to give an example, in the following we assume that the insurance company prices the risk that a firm may incur financial losses as a result of a data breach which is the main cause of cyber incidents [12].

3 An Illustrative Example

Privacy Rights Clearinghouse (PRC) is a nonprofit organization focusing on data privacy rights and issues. Their Chronology of Data Breaches in the US [11] includes description and type of both the breach and the breached entity, along with the breach severity, measured in terms of the number of breached records, when available. For a detailed description of this dataset see [13]. We restrict our analysis to the more recent data (breaches reported after the 1st of January, 2010) because they could better represent the current cyber threat situation; we also select only breaches with complete information on breach sizes and cause. In some previous studies [13,14] we investigated the causes of data breaches and found significant differences in the distribution of the severity of breaches caused by accidental exposure or inadequate vigilance ("negligent" breaches) with respect to breaches originating from activities that actively targeted private information ("malicious" breaches). Then we decided to model these two distributions separately. In both cases, however, we found that the best fit for the severity is given by a skew-normal distribution. Regarding "negligent" breaches, it is worth specify that many employees are often the weakest link that causes a successful cyber incident [15].

In this illustrative example, we consider a generic organization belonging to the Business typology, that includes financial and banking services, manufacturing, retail. The total amount of registered breaches with full information available for this category in the PRC dataset for the period 2010–2019 includes 732 data. In order to estimate the appropriate premium for this organization, we follow some of the suggestions provided in [16] and build a simulation pipeline to generate a large number of scenarios suitable to represent the losses distribution. Evidences from the literature and from the available data that guided us in developing this pipeline are the following:

- the probability for an organization of suffering $k \geq 0$ breaches in a year can be modeled by a geometric distribution whose parameter p has been estimated in [16] separately for the three business subcategories. A weighted average of these values results in the value $p = 0.91$;
- historical data on the type of data breaches for a generic Business type organization allows us to estimate the relative frequency of malicious (fm) and negligent (fn) events, so that a simulated event will be of malicious type with probability $fm/(fm + fn)$;
- once the severity distribution (in terms of the volume Y of breached data) has been fitted to the available data, the financial loss L for each breach event can be roughly estimated by a regression model on Y. This formula was originally derived by Jacobs [17] on Ponemon data and then improved by Farkas [16] to better represent extreme events:

$$\log(L) = 9.59 + 0.57 \log(Y). \tag{3}$$

To obtain reliable estimates for the annual losses of a Business type organization, we adopt a Monte-Carlo based simulation approach: once fixed a huge number N of scenarios

- for any $i \leq N$ we simulate a corresponding number of claims n_i by generating a random value from a geometric distribution with parameter p as mentioned before;
- then, in case $n_i > 0$, we generate a random value n_m from a Bernoulli distribution of parameter $fm/(fm + fn)$ to assign n_m events to the malicious category and the remaining $n_i - n_m$ to the negligent category;
- for each event, its severity is generated as a random value from the fitted skew-normal distribution (malicious or negligent) and the related financial loss estimated by the empirical formula (3).

Basing on the different methods described in Sect. 2, premiums can finally be evaluated from the simulated distribution of losses. Cyber insurance commonly distinguishes between “third-party” and “first-party” losses depending on whether they concern external parties or the insured itself. Jacobs transformation estimates both first and third-party losses and we make the same hypothesis. Moreover, we consider the case of full compensation.

4 Results and Conclusive Remarks

We present here some results obtained in the numerical simulations we performed according to the pipeline described in the previous section. As stated in Sect. 2, we also consider the alternative approach and proceed without splitting the simulated claims into malicious and negligent ones, but simply generate the severity of each claim as a random value from the skew-normal distribution fitted on all data. In both cases, we generated $N = 1$ million scenarios and repeated the simulation pipeline 10 times, to obtain averaged estimates of the different premiums along with their standard errors.

In the columns of Table 1 we report the values we obtained for the fair premium Π, the expected value premium P_{ev}, the standard deviation premium P_{sd}, the semi-standard deviation P_{ssd}, the exponential premium P_{ut} and the quantile premium P_Q along with their standard errors. In all simulations we chose $\alpha = 0.1$, $\rho = 0.01$, $\varepsilon = 0.05$. The first row shows results of the described pipeline. In the second row, given for comparison, the premiums has been evaluated instead by estimating frequency and severity of the breaches without splitting the malicious and negligent events.

Table 1. Estimates of the annual premiums according to the principles described in Sect. 2 with the proposed methodologies.

	Π	P_{ev}	P_{sd}	P_{ssd}	P_{ut}	P_Q
Premiums	23647 ± 38	26011 ± 42	23648 ± 38	23647 ± 38	24034 ± 41	135382 ± 2245
Premiums(alt)	23429 ± 37	25772 ± 41	23430 ± 37	23430 ± 37	23801 ± 40	124746 ± 2787

First of all we observe that in both cases (first and second row of Table 1) all the estimated premiums based on the equivalence principle (P_{ev}, P_{sd}, and P_{ssd}) exceed the fair premium. This surplus serves as a buffer to offset adverse experiences. Regarding the principles involving standard deviation, the loading is associated with the variability of the loss. Indeed, P_{ut} and P_Q are higher than Π, too. In particular, P_{ut} is defined via *utility theory* and incorporates the attitude towards risk of the insurer. As regards P_Q, that is the quantile of order $(1-\varepsilon)$ of the loss distribution, it is significantly higher than the other premiums; the reason is that the insurer wants a premium that covers $(1-\varepsilon) \cdot 100\%$ of the estimated losses.

Regarding the comparison of the two methods, it can be noticed that premiums obtained by estimating frequency and severity of the breaches without splitting the malicious and negligent events (*Premiums(alt)*, second row of Table 1), are lower than the ones obtained with the methodology described in Sect. 3. Although in this specific example the difference is minimal, this is indicative of the fact that if the insurer estimates the premium without considering the different types of cyber incidents, it could incur an underestimation of the premium itself. This difference becomes more evident in reference to the quantile premium P_Q. This means that extreme events could be significantly underestimated by the insurer. In future research, we will test the methodology on other datasets allowing us to consider a richer range of types of incidents and make more extensive comparisons.

Acknowledgments. This work was partially supported by project SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU and performed under the auspicies of the GNCS of INdAM.

References

1. SwissRe Institute: what you need to know about the cyber insurance https://market.swissre.com/risk-knowledge/advancing-societal-benefits-digitalisation/about-cyber-insurance-market.html
2. Tsohou, A., Diamantopoulou, V., Gritzalis, S., Lambrinoudakis, C.: Cyber insurance: state of the art, trends and future directions. Int. J. Inf. Secur. **22**(3), 737–748 (2023). https://doi.org/10.1007/s10207-023-00660-8
3. Marotta, A., Martinelli, F., Nanni, S., Orlando, A., Yautsiukhin, A.: Cyber-insurance survey. Comput. Sci. Rev. **24**, 35–61 (2017). https://doi.org/10.1016/j.cosrev.2017.01.001
4. Romanosky, S., Ablon, L., Kuehn, A., Jones, T.: Content analysis of cyber insurance policies: how do carriers price cyber risk? J. Cybersecur. **5**(1), 1–19 (2018). https://doi.org/10.1093/cybsec/tyz002
5. Braun, A., Eling, M., Jaenicke, C.: Cyber insurance-linked securities. ASTIN Bull. **53**(3), 684–705 (2023). https://doi.org/10.1017/asb.2023.22
6. Awiszus, K., Knispel, T., Penner, I., Svindland, G., Vob, A., Weber, S.: Modeling and pricing cyber insurance. Eur. Actuar. J. **13**, 1–53 (2023). https://doi.org/10.1007/s13385-023-00341-9
7. Zeller, G., Scherer, M.: Risk mitigation services in cyber insurance: optimal contract design and price structure. Geneva Pap. Risk Insur. Issues Pract **48**, 502–547 (2023). https://doi.org/10.1057/s41288-023-00289-7
8. Maochao, X., Lei, H.: Cybersecurity insurance: modeling and Pricing. N. Am. Actuar. J. **23**, 220–249 (2019). https://doi.org/10.1080/10920277.2019.1566076
9. Olivieri, A., Pitacco, E.: Introduction to Insurance Mathematics: Technical and Financial Features of Risk Transfers. Springer International Publishing, EAA Series (2015)
10. Embrechts, P.: Actuarial versus financial pricing of insurance. J. Risk Finance **1**(4), 17–26 (2000). https://doi.org/10.1108/eb043451
11. Privacy Rights Clearinghouse. Chronology of Data Breaches (2018). https://www.privacyrights.org/data-breaches
12. Allianz Global Corporate & Specialty(AGCS). Allianz Risk Barometer 2019. Top Business Risks for 2019. https://www.agcs.allianz.com/content/dam/onemarketing/agcs/agcs/reports/Allianz-Risk-Barometer-2019.pdf
13. Carfora, M.F., Orlando, A.: Some remarks on malicious and negligent data breach distribution estimates. Computation 12 art. number **208** (2022). https://doi.org/10.3390/computation10120208
14. Carfora, M.F., Orlando, A.: Cyber risk: estimates for malicious and negligent breaches distributions. In: Mathematical and Statistical Methods for Actuarial Sciences and Finance, Springer International Publishing (2022). https://doi.org/10.1007/978-3-030-99638-3_23
15. OECD.:Enhancing the role of insurance in cyber risk management (2017). https://doi.org/10.1787/9789264282148-en

16. Farkas, S., Lopez, O., Maud, T.: Cyber claim analysis using generalized pareto regression trees with applications to insurance. Insur. Math. Econ. **98**, 92–105 (2021). https://doi.org/10.1016/j.insmatheco.2021.02.009
17. Jacobs, J.: Analyzing Ponemon cost of data breach (2014). datadrivensecurity.info/blog/posts/2014/Dec/ponemon/