How to measure cybersecurity risk
Misurare il cyber rischio

Silvia Facchinetti, Paolo Giudici and Silvia Angela Osmetti

Abstract In the last years there have been a scholars increasing interest in cybersecurity risk measurement. This paper proposes a methodology to measure cyber risk, using ordinal data, to prioritise appropriate interventions. The method relies on the construction of a Criticality index: a new measure of risk based on the cumulative probabilities of the ordinal variable that represents the level of severity for different risk event. Its properties are derived and compared with alternative measures employed in operational risk measurement. we apply our proposal to real data of a telecommunication service company. The proposed measure is found to be quite effective to rank cyber risk types and, therefore, allow selective preventive actions.
Abstract *Negli ultimi anni si è registrato un interesse crescente da parte degli studiosi al problema del cyber rischio e alla sua misurazione. Dato che in tale contesto i dati sono spesso di natura ordinale, nel presente lavoro proponiamo un indice per stimare il cyber rischio utilizzando dati qualitativi riguardanti il livello di severità dei cyber attacchi. Noi analizziamo le caratteristiche dell'indice, e lo confrontiamo con misure alternative utilizzate nella valutazione del rischio operativo. La nostra proposta viene applicata ai dati di una compagnia di telecomunicazioni, riguardanti la severità dei cyber attacchi subiti dai clienti della compagnia. L'indice risulta efficace per classificare le diverse linee di business in base alla loro rischiosità e quindi per consentire tempestive azioni preventive.*

Key words: criticality index, risk measure, ordinal variables, cyber attacks

---

Silvia Facchinetti
Department of Statistical science, Università Cattolica del Sacro Cuore, Milano, e-mail: silvia.facchinetti@unicatt.it

Paolo Giudici
Department of Economics and Management, University of Pavia, Pavia e-mail: paolo.giudici@unipv.it

Silvia Angela Osmetti
Department of Statistical science, Università Cattolica del Sacro Cuore, Milano, e-mail: silvia.osmetti@unicatt.it

1 Introduction

In the last years the number of cyber attacks in information technology (IT) systems is surging. Therefore, cybersecurity has become more of a concern for businesses. Among operational risks caused by IT systems, cyber risks are gaining increasing importance, due to technological advancements and to the globalisation of financial activities (see e.g. [2],[3]). Therefore the cybersecurity risk measurement is an increasing interest for scholars (see e.g. [1],[6]).

We propose a methodology to measure cyber risks, starting from ordinal random variables, that represent the levels of severity for different risk events (expressed on ordered categories, such as "high", "medium" or "low" risk), in different business lines. In particular we propose, for different business lines, a cybersecurity risk index that is based on the relative frequencies of the severity levels. As a result, we obtain an ordinal measure of risk which will be used to compare different events and business lines, producing an ordering among risks useful to prioritise intervention in process controls.

Besides the theoretical proposal, we will present empirical evidences on the performance of our index, using a real data set, that concerns cyber risk measurement in a telecommunication company.

2 Proposal

Let $X \sim \{x_k, p_k; k = 1,2,\ldots,K\}$ be a categorical random variable with ordered categories x_k and probabilities $p_k = P(x_k)$, that represents a severity variable, with decreasing levels, $k = 1,2,\ldots,K$.

We define a *Criticality Index* for the categorical random variable X with the following expression:

$$I = \frac{1}{K-1} \sum_{k=1}^{K-1} (K-k) p_k = \frac{\sum_{k=1}^{K} F_k - 1}{K-1}, \tag{1}$$

where $F_k = \sum_{l=1}^{k} p_l$ are the values of the cumulative distribution function of the ordinal variable X, for $k = 1,2,\ldots,K$.

It is a natural measure of risk for ordinal variables, with values in $[0,1]$. It thus provides a risk measure easy to interpret and suitable for the comparison of the risk level between different risk events and/or business lines.

We propose to estimate cybersecurity risk in each business line/event type combination by the sample version of the *Criticality Index*, obtained by replacing the probabilities p_k with their estimators $\hat{p}_k = r_j/n$ for $k = 1,2,\ldots,K-1$:

$$\hat{I} = \frac{1}{K-1} \sum_{k=1}^{K-1} (K-k) \frac{r_k}{n} = \frac{\sum_{k=1}^{K} \tilde{F}_k - 1}{K-1}, \tag{2}$$

where

$$\tilde{F}_k = \sum_{l=1}^{k} \frac{r_l}{n} \text{ for } k = 1,2,\ldots,K,$$

is the empirical cumulative distribution function, $r_l = \sharp(\tilde{x}_i \equiv x_l)$ is the number of observations in the sample equal to the category x_l, with $r_l \in \mathbf{N}$ and $\sum_{l=1}^{K} r_l = n$ (n is the total number of risk events observed for the j-th business line).

It is possible to demonstrate that the *Criticality Index* estimator is asymptotically normally distributed for $n >= 30$. Moreover, $\hat{I}$ is an unbiased and consistent estimator for I [4].

3 Application

We apply our proposal to real data provided by a telecommunication company that installs telephone exchange systems and offers post-installation technical assistance for upgrading and problem resolution in different event types, that include, in particular, Network communications. The service is offered to a wide range of customers, that are grouped in several business lines.

The main research problem is to estimate for each business line, a measure of cybersecurity risk based on ordinal data, collected by the customer care center, describing the level of severity of cyber attack suffered by customers.

The data are reported in Table 1. Each row shows a business line and each column reports how many times a cyber problem in Network communication has been reported, for levels of severity equal to high (H), medium (M) and low (L).

Table 1 Data for Network communication event type

Business Line	H	M	L
Banking	23	128	4
Computers	3	26	0
Cooperatives	13	93	2
Defence	34	149	7
Health	43	222	8
Hotels	10	108	3
Industry	13	94	7

In Table 2, column 2 and 3, we report the *Criticality Index* estimates and its associated asymptotic confidence interval.

We show that Defence is the business line with the highest level of risk, followed by Health and Banking. Therefore, a mitigation intervention to prevent cyber risk should prioritise the Defence business line, and the customers in that business line.

To evaluate the robustness of our results we compare them with what could be obtained with the approach proposed by [5] in the context of operational risks. We apply their Stochastic Dominance Index (SDI) to our data and their suggested

Table 2 *I* and SDI risk measure estimates and their respective confidence intervals (CI)

Business line	$\hat{I}$	CI	SDI	Bayesian CI
Banking	0.561	0.517-0.606	0.708	0.685-0.728
Computers	0.552	0.473-0.630	0.701	0.654-0.734
Cooperatives	0.551	0.503-0.599	0.701	0.676-0.723
Defence	0.571	0.527-0.616	0.714	0.692-0.735
Health	0.564	0.529-0.599	0.709	0.692-0.726
Hotels	0.529	0.488-0.570	0.686	0.665-0.706
Industry	0.526	0.472-0.580	0.684	0.657-0.711

Bayesian procedure to derive a confidence interval, by means of a Gibbs Sampling algorithm with R=10000 interactions. What obtained is reported in Table 2, column 4 and 5. Note that since $n \simeq 30$ for all the business lines in Table 2, is not strictly necessary to apply a Bayesian approach to derive confidence intervals of our index, we can apply the asymptotic normality. Bayesian confidence intervals may instead be useful for "rare" problem business lines, for which an asymptotic confidence interval is not possible.

By compare the results of Table 2, we observe, obviously, a different values for the two indices since they are calculated in different way: SDI is based on the observed frequencies, whereas our index is based on the observed relative frequencies of the severity variables conditional on the total n of that business line. Moreover, we observe that the SDI values are always higher than the results obtained with our approach. By comparing the results for different BLs, we show that our index and the SDI produce a consistent ranking, indicating similar priorities of intervention. This confirm that our index could be suitable to measure the level of risk, to compare the level of risk for many BLs or even types and that it can be considered as a synthetic priority of intervention indicator.

References

1. Afful-Dadzie, A., and Allen, T.T.: Data-Driven Cyber-Vulnerability Maintenance Policies. Journal of Quality Technology, **46**, 234-250 (2017).
2. Cebula, J.J., and Young, L.R. (2010). A Taxonomy of Operational Cyber Security Risks, Technical Note CMU/SEI-2010-TN-028, Software Engineering Institute, Carnegie Mellon University, 1-34.
3. Edgar, T.W., and Manz, D.O. (2017). Research Methods for Cyber Security, Elsevier.
4. Facchinetti, S., and Osmetti, S.A.: A risk index for ordinal variables and its statistical properties: a priority of intervention indicator in quality control framework. Quality and Reliability Engeneering International, **34**, 265275 (2018).
5. Figini, S., and Giudici, P.: Measuring risk with ordinal variables. Journal of Operational Risk, **8**, 35-43 (2013).
6. Hubbard, D.W., and Seiersen, R.: How to Measure Anything in Cybersecurity Risk. Wiley, New York (2016).