A statistical approach for assessing cyber risk via ordered response models

Claudia Tarantola
DEM University of Pavia and Hermes Universities Network

based on recent papers with

Silvia Facchinetti (UNICATT), Silvia Osmetti (UNICATT) and Maria Iannario (UNINA)

"*Recent Security Advances in the Finance Sector*", January 14, 2021

Department of Economics & Management

Cyber risks

Cyber risks can be defined as "an operational risks emerging from the use of Information and Communication Technologies (ICT) systems that compromises the confidentiality, availability, or the integrity of data or services"

(Kopp et al, 2017 and Cebula and Young, 2010)

- Among operational risks related to ICT systems, cyber risks are gaining increasing importance.
- According to World Economic Forum Global Risks Report (2020), cyber risks are consolidating their position alongside environmental risks in the high-impact/high-likelihood quadrant of the Global Risks Landscape.

Cyber risk data

- Although cyber security is a field with a growing amount of research in many contexts (e.g. computer science, law, and business management), there is a limited number of statistical papers addressing the problem; this is mainly due to the lack of quantitative data available.
- Public and commercial datasets exist but they are incomplete, have different coverage and use different definitions of cyber attacks.
- Data are often collected on an ordinal scale using information provided by experts of the sectors.

Clusit data

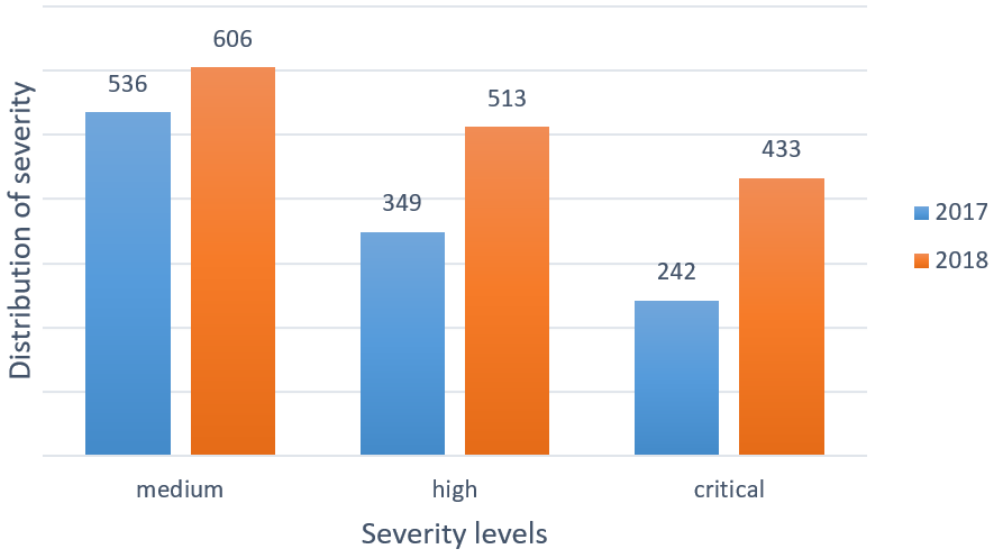
We consider a set of data collected by the researchers of the *Hackmanac Project* and described by **Clusit** in the 2019 Annual Report on ICT Security in Italy.
Clusit is the largest and most respected Italian association in the field of cyber security.
It was created in 2000 at the University of Milan.

Clusit activity

- Since 2012 Clusit publishes on biannual base a report on "serious" cyber attacks occurred worldwide in the previous year. In the period 2011-2018 Clusit collected 8417 observations.
- The current classification criteria used to identify a "serious" cyber attack is different from the one of the early years.
- In 2017 the experts of the "*Hackmanac Project*" developed for Clusit an ordinal classification of cyber risk severity (medium, high, critical) on the base of their expertise. The aspects that determine the risk assessment of each attack are multiple, and include the geopolitical, social, economic, image and cost/opportunity impact on the victims.

Severity levels 2017-2018

Severity is a measure of the impact of a cyber attack



Data source: Clusit 2019 Annual Report on ICT Security in Italy.

The data

We consider for our analysis 2679 cyber attacks that occurred Worldwide in 2017-2018.

In addition to the severity levels we have information on these categorical variables:

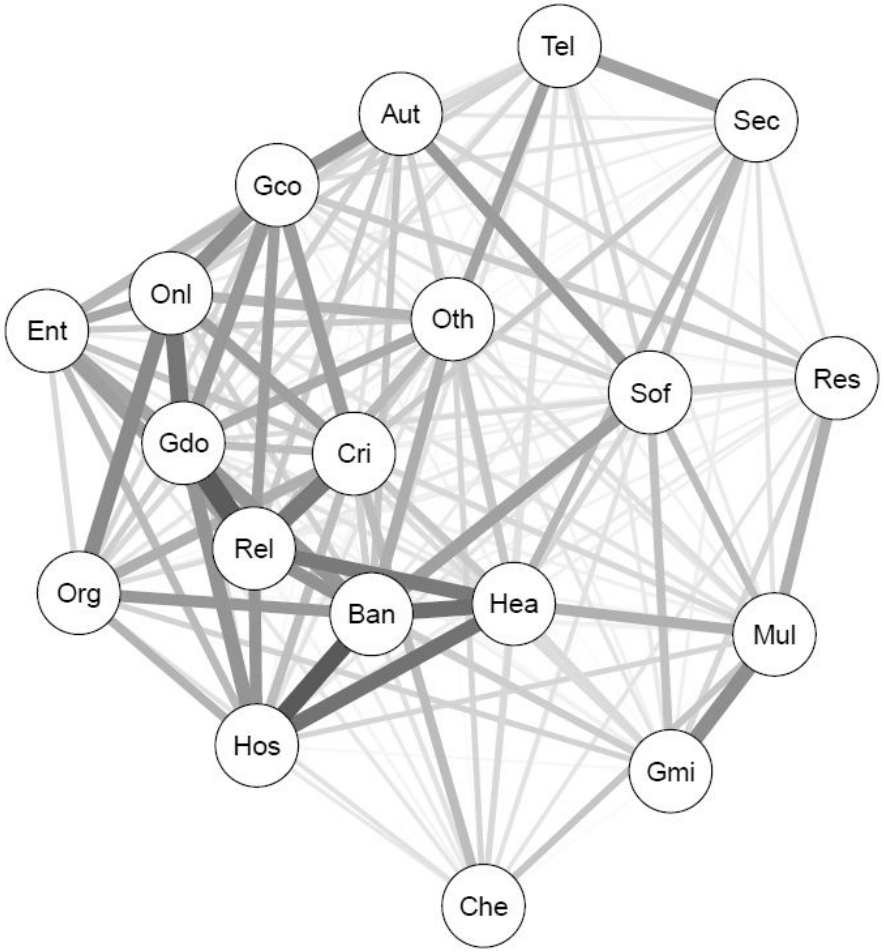
- *Type of Attack*
- *Attack Technique*
- *Continent*
- *Victim*

The data-continue

- *Type of Attack* coded in **4 categories**:
  ▷ Cybercrime, Hactivism, Espionage/Sabotage, Information Warfare; baseline->Cybercrime
- *Attack Technique* coded in **5 categories**:
  ▷ 0-day, Multiple Threats, Trivial Threats, SQL Injection, Unknown; baseline->SQL Injection
- *Continent* coded in **6 categories**:
  ▷ Africa, America, Asia, Australia/Oceania, Europe, Multiple Continents; baseline->Multiple Continents
- *Victim* coded in **19 categories**:
  ▷ Automotive, Chemical/Medical, Banking/Finance, Critical Infrastructures, Entertainment/News, GDO/Retail, Gov-Mil-LEAs - Intelligence, Gov.Contractors/Consulting, Health, Hospitability, Multiple Targets, Online Services/Cloud, Organization/ONG, Others, Religion, Research- Education, Security Industry, Software/Hardware Vendor, Telco.

Variable Closeness

A measure of diffusion of the effects of the attacks obtained via the use of a network structure is incorporated into the model.
The nodes of the network corresponds to the victims of cyber attacks, while the edges indicate the strength of their connections.



Victim

Automotive (Aut)
Banking/Finance (Ban)
Chemical/Medical (Che)
Critical Infrastructures (Cri)
Entertainment/News (Ent)
GDO/Retail (Gdo)
Gov-Mil-LEAs-Intelligence (Gmi)
Gov.Contractors/Consulting (Gco)
Health (Hea)
Hospitability (Hos)
Multiple Targets (Mul)
Online Services/Cloud (Onl)
Organization/ONG (Org)
Others (Oth)
Religion (Rel)
Research-Education (Res)
Security Industry (Sec)
Software/Hardware Vendor (Sof)
Telco (Tel)

Variable Closeness (continue)

The network is constructed as follows.

- For each victim we calculated the weekly time series of the Criticality Index, Facchinetti et al. (2019). This index is based on the frequencies of cyber attacks suffered by the victims for the different levels of gravity. It can be used to provide an indication of vulnerability of the victims of cyber attacks.
- We estimated the network based on the partial correlation matrix among the previous series.
- Variable Closeness scores each node on the basis of its distance to all the other nodes in the network and the strength of the connection (it is bases on the network centrality measure proposed by Opsahla et al., 2010).

Ordered response model

Due to the ordinal nature of the severity variable, it turns out natural to consider ordinal response models.

In particular, we applied the following models:

- Cumulative link models under the proportional odds assumption (POM), see e.g. Agresti (2010).
- Combination of Uncertainty and Preference of the respondents (CUP) models, Tuds et al. (2017).

POM

$$link[P(S_i \leq s)] = \alpha_s - \mathbf{x}_i\boldsymbol{\beta} \qquad s = 1, 2, \quad i = 1, \ldots, n$$

- $\mathbf{x}$ is the vector of covariates
- α_s is the intercept
- β is the vector of the regression parameters
- "link" is a suitable link function

The larger the value of $\mathbf{x}\boldsymbol{\beta}$, the higher the probability to obtain an elevated level of cyber risk.

CUP

$$P(R_i = s) = \pi P_M(S_i = s|\boldsymbol{x}_i) + (1 - \pi_i)P(U_i = s) \quad s = 1, 2, \quad i = 1, \ldots, n$$

The Prefence part $P_M(S_i = s|\boldsymbol{x})$ is obtained via a cumulative link model. The uncertainty parameter π may also depend on the same/different set of covariates. The uncertainty component accounts for the difficulty of the respondent to express a rating regarding a specific object/item.

Model parameters interpretation

- For both **POM** and **CUP** models, the use of a **nonlinear link function** leads to model parameters not as simple to interpret as slopes and correlations in ordinary linear regression models.
- The model effect parameters, related to measures such as odds ratios, may not be easily understood or can even be misinterpreted. This enhances the need to introduce simpler ways to interpret the effects of the covariates.
- Following Agresti and Tarantola (2018) and Iannario and Tarantola (2021a) we evaluate the effect of each explanatory variable, on the preference and on the uncertainty component, using the so-called **Marginal Effect (ME)** measures.

Marginal effects

A ME shows how a variation in a covariate affects the outcome distribution, holding all the other variables constant.

MEs are computed differently for continous and categorical covariates:

▷ the ME of a quantitative variable for level s of the Severity is the partial derivative of $P(S = s)$ with respect to the examined variable, holding all the other variables constant.

▷ the ME for level l of a categorical variable (discrete change) measures the change in the probability $P(S = s)$ when the level changes to the baseline one, holding all the other variables constant.

Marginal effects

Alternative types of marginal effects:

- **Average Marginal Effect** (AME): is obtained by calculating the ME of a specific covariate for each observation in the sample, and then averaging across all values
- **Marginal Effects at the Mean** (MEM): it is computed by setting all the other covariates equal to their mean value
- **Marginal Effects at Representative values** (MER): it is computed by choosing representative values of the other covariates

Marginal effects

Among these summary measures, we use the **AME**:

- it averages the effects across all cases observed in the sample and thus can be interpreted as the sample average of the MEs

It has behavior reminiscent of effects in **ordinary linear models**:

- it is roughly stable when we add an explanatory variable that is uncorrelated with the variable whose effect we are describing

Note that, as any explanatory variable increases, cumulative link models imply monotononicity only in the extreme-category probabilities. Hence, it is sensible to consider MEs only for extreme category of the response scale.

POM: Alternative link functions

We consider the following link functions

Link	logit	probit	log-log	cloglog	cauchy	bgeva ($\tau = -0.25$)
Res. Dev.	5538.333	5541.049	5540.789	5565.718	5577.085	5733.813
AIC	5562.333	5565.049	5564.789	5589.718	5601.085	5757.813

Note that *logit*, *probit* and *log-log* cumulative link models outperform the others, presenting similar values of residual deviance (Res. Dev.) and Akaike Information Criteria (AIC). They also provide coherent parameter estimations.

In the following, we only present the results for the logit link that is the most commonly used and easier to interpret.

POM: Empirical results

- Type of attack
- Attack technique
- Continent
- Closeness

POM: Empirical results

```
--------------------------------------------------------------------------------
ocAME_CAT(logit.m) # new function available from the authors

$ME.1 (medium severity)
                          effect     std.error     z.value     p.value
T-Espionage/Sabotage      -0.234         0.037      -6.420       0.000
T-Information Warfare     -0.238         0.043      -5.521       0.000
A-Trivial Threats          0.425         0.162       2.619       0.009
A-Unknown                  0.426         0.163       2.612       0.009
C-America                 -0.083         0.086      -0.968       0.333
C-Asia                    -0.194         0.031      -6.293       0.000
C-Australia/Oceania       -0.263         0.064      -4.136       0.000
C-Europe                  -0.163         0.027      -6.055       0.000
Closeness                 -0.217         0.099      -2.186       0.029

$ME.3 (critical severity)
                          effect     std.error     z.value     p.value
T-Espionage/Sabotage       0.213         0.045       4.696       0.000
T-Information Warfare      0.177         0.032       5.493       0.000
A-Trivial Threats         -0.316         0.121      -2.621       0.009
A-Unknown                 -0.317         0.121      -2.614       0.009
C-America                  0.050         0.056       0.895       0.371
C-Asia                     0.144         0.023       6.186       0.000
C-Australia/Oceania        0.196         0.048       4.097       0.000
C-Europe                   0.122         0.020       5.936       0.000
Closeness                  0.161         0.074       2.180       0.029
--------------------------------------------------------------------------------
```

POM: Empirical results

▷Type of Attack

- Although Cybercrime attacks (baseline level) represent a huge percentage of the total number of attacks occurred in the period 2017-2018 (77.98%), in terms of gravity they are nowadays classified as minor risks. This is confirmed by the analysis of MEs.

- Espionage/Sabotage and Information Warfare are the type of attacks that are more likely than Cybercrime to generate a critical level of severity. According to our experts this can be explained by the fact that this type of attacks are mainly used to steal important geopolitical and economical information, and the available countermeasures are at the moment ineffective.

POM: Empirical results

▷Attack technique

- SQL Injection (the baseline attack technique) is the one influencing more a critical level of severity. All MEs are negative.
- Even if SQL Injection is a low frequency attack technique (0.30%), its impact on critical severity is quite relevant. The SQL injection attack consists of reading and modifying sensitive data, performing unauthorized operations as an administrator on a database, retrieving the contents of a given system, and in some cases commanding the operating system. Al these issues can cause serious problems to the victim.

POM: Empirical results

▷ Continent

- Critical severity attacks directed at individual continents show a positive effect (the ME measures are positive), while attacks of medium severity are effective against multiple continents (the ME measures are negative).
- Attack targeted against individual continent are more effective than the ones directed to more continents in which the effect is dispersed (the baseline level is multiple continent).

POM: Empirical results

▷ Closeness

- The quantitative variable Closeness has a significant impact on the attack severity level (positive ME) and its inclusion in the model improves the predictive performance.

- If a cyber attack hits a victim strongly connected in the network, its effect propagate rapidly through the network structure (affecting other connected victims) and may enlarge its gravity. Strongly connected victims should work together to prevent critical attacks and hopefully collaborate for the development of common security protocols.

Concluding remarks

- There is no internationally recognized standard classification of the gravity of a cyber attack.
- A different classification could lead to different results.
- There is the necessity to introduce a standardized classification of cyber risk levels that can be adopted worldwide.

References

- Agresti, A., and Tarantola, C. (2018). Simple Ways to Interpret Effects in Modeling Ordinal Categorical Data, Statistica Neerlandica, 72: 210-223.
- 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.
- Facchinetti, S., Giudici, P. and Osmetti, S.A. (2020). Cyber risk measurement with ordinal data. *Stat Methods Appl*, **29**, 173-185.
- Facchinetti, S., Osmetti, S.A. and Tarantola, C. (2021). A statistical approach for assessing cyber risk via ordered response models, submitted.
- Kopp, E., Kaffenberger, L., and Wilson, C. (2017). Cyber Risk, Market Failures, and Financial Stability, IMF Working Paper WP/17/185, 1-35.
- Iannario, M., Tarantola, C. (2021a) How to interpret the effect of covariates on the extreme categories in rating data models. *Sociological methods & research*, in press.
- Iannario, M., Tarantola, C. (2021b). Effect measures for group comparisons in a two-component mixture model: a cyber risk analysis. In *Data Analysis and Knowledge Organization*, Springer book series Studies in Classification, in press.
- Tutz, G., Schneider, M., Iannario, M., Piccolo, D. (2018). Mixture Models for Ordinal Responses with a Flexible Uncertainty Component. *Journal of Applied Statistics*, **46**, 1-20.
- World Economic Forum (2020). *Global Risk Report 2019*, 15th Edition.