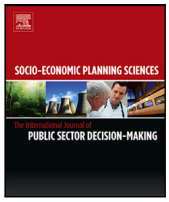
Socio-Economic Planning Sciences 87 (2023) 101584

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

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



Network models for cyber attacks evaluation

Silvia Facchinetti [a,1], Silvia Angela Osmetti [a,1], Claudia Tarantola [b,*,1]

[a] Department of Statistical sciences, Università Cattolica del Sacro Cuore, Milano, Italy
[b] Department of Economics and Management, University of Pavia, Pavia, Italy

ARTICLE INFO

Keywords:
Bayesian Network
Cyber risk
DAG
Random Forest
Social Network

ABSTRACT

The significant recent growth in digitization has been accompanied by a rapid increase in cyber attacks affecting all sectors. Thus, it is fundamental to make a correct assessment of the risk to suffer a cyber attack and of the resulting damage. Quantitative loss data are rarely available, while it is possible to obtain a qualitative evaluation on an ordinal scale of the gravity of an attack from experts of the sector. In this paper, we discuss how network models can be useful instruments for the evaluation of the risk associated to a cyber attack. In particular, we consider Bayesian Networks, Random Forests and Social Networks to study different aspects of the examined problem. Along with the description of the methodology, we examine a real set of data regarding serious cyber attacks occurred worldwide before and during the pandemic due to Covid-19. In the analysis, we also investigate how the Covid-19 period had an impact on the cyber risk landscape in terms of frequency and gravity of the observed attacks.

1. Introduction

Over the past decade, the frequency and impact of cyber attacks have increased and no sector can consider itself completely sheltered from these types of events [1]. Attackers modify techniques, targets and tools at a very high pace, leading to a threat landscape in continuous evolution. Therefore, cyber risk evaluation has become an important area of interest for standard-setting bodies and international cooperation.

Cyber risk can be viewed as a particular type of operational risk that arises from external or internal attacks compromising a computer database or network, or from transactions on the Internet [2]. Following the definition provided by the Basel Committee on Banking Supervision (BCBS) operational risk can be defined as the risk of loss resulting from inadequate or failed internal processes, people, and systems or from external events [3]. For more details on operational risk modelling see e.g., [4–7] and references herein. As discussed in [8,9] cyber risk differs from "traditional" operational risk in various aspects. In cyber attacks, confidentiality of data may be impaired, as may be the integrity and/or availability of data and information. Cyber attacks propagate throughout the system at a significantly faster pace than other types of risk via the existing technological linkages. Moreover, they are associated with significant uncertainty and are not constrained by geographical boundaries. Finally, cybercriminals execute a deliberate attack to damage a system, unlike the "traditional" operational risks often associated with accidental failures. We refer to [10,11] for major details on cyber attack taxonomy and definition.

Cyber risk is a research topic that has attracted considerable academic, industry and government attention over the last years. Fields studying cyber risk include computer science, behavioural science, economics, technological sector, management science, law, and political science [12,13]. Unfortunately, to date, cyber risk research has been piecemeal and uncoordinated mostly due to the cross-disciplinary characteristics of cyber risk, this implies the decentralization of the study across different academic sectors.

In addition, cyber loss data are very difficult to obtain since institutions are not often willing to disclose them as their reputation and security could be at risk. Quantitative models for cyber risk assessment are limited, and the lack of a shared framework makes the adoption of comparable measures for risk mitigation very difficult [14]. Nevertheless, it is possible to obtain a qualitative evaluation of the level of severity of an attack, expressed on a rating scale. In this way, while not knowing the actual magnitude of cyber attacks, we can identify the most dangerous types of attacks. Currently, there is no internationally recognized standard classification of the gravity of cyber attacks. As described in Section 2, we consider a classification provided by *Hackmanac*, a society collaborating with Clusit (Associazione Italiana per la Sicurezza Informatica - https://clusit.it), the principal Italian authority in the field of cyber security.

* Corresponding author.
E-mail address: claudia.tarantola@unipv.it (C. Tarantola).
[1] Contributed equally to the work.

https://doi.org/10.1016/j.seps.2023.101584
Received 20 December 2022; Received in revised form 24 February 2023; Accepted 20 March 2023
Available online 21 March 2023

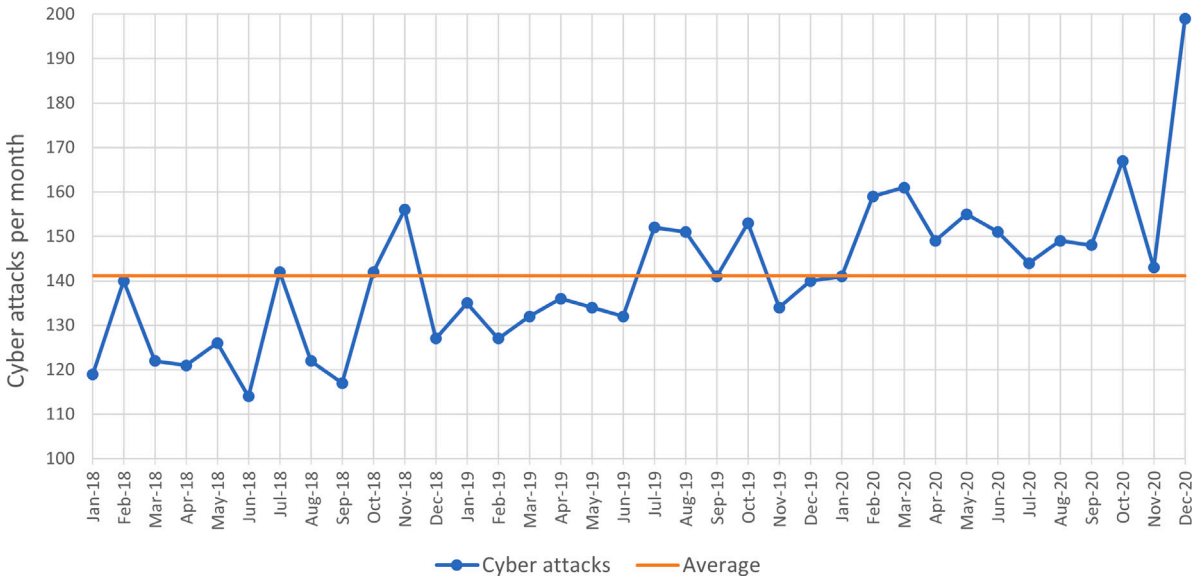


Fig. 1. Trend of cyber attacks per month during the period 2018–2020. Figure produced with `Excel`.

In this paper, we focus on three different types of network models for cyber risk evaluation: Bayesian Networks (BNs), Random Forests (RFs) and Social Networks (SNs). These three types of networks allow different aspects of cyber risk assessment issues to be investigated. In particular, BNs and RFs are predictive models that allow us to assess the severity levels of cyber attacks for different configurations of their features (e.g., type of attack, target). BNs are used to provide a visual representation of the relationship among the examined variables and to evaluate alternative risk configurations. While RFs are used to detect the relevant factors that influence the severity of an attack. On the other hand, SNs are models of visualization and analysis. They are used for representing and analysing the links and relationships between subjects impacted by a cyber attack. In this paper, we consider them to measure the interconnection among the targets of cyber attacks and to monitor the diffusion of attacks.

For our analysis, we examine a dataset that includes information on serious cyber attacks occurred worldwide in the years 2018–2020. Our data refers to years that encompass the period of Covid-19 pandemic,[2] hence it turns natural to investigate the Covid-19 effect on cyber risk assessment. It is well known that the lifestyle change due to the Covid-19 pandemic has caused significant modifications in the everyday life. We witnessed an increase in the use of technology, and an expansion of the attack surface for cybercriminals, leading to a relevant rise in the number of cyber attacks and their severity. Cybercriminals have taken advantage of the collective unease, as well as extreme hardship, experienced by certain sectors to target their victims.

The plan of the paper is as follows. Section 2 contains the analysis and description of the considered cyber attacks data. In Sections 3–5 we show how BNs, RFs and SNs can be useful tools for cyber risk evaluation before and during the pandemic period. We end the paper with some concluding remarks.

2. Cyber attacks data

We consider a dataset containing information on more than 5,000 relevant cyber attacks occurred worldwide from 2018 to 2020. It was provided to us by *Hackmanac*, a society based in Dubai that monitors the evolution of real global cyber threat by the analysis and classification of several open sources (such as national and international newspapers, web articles and press releases). *Hackmanac* collaborates with Clusit, performing the analysis for their bi-annual report, see [17]. More details on this society can be found at the webpage https://hackmanac.com/.

In the following part we present a brief description of the examined data. For each cyber attack we consider the following variables: **Date**, **Attacker**, **Attack Technique**, **Target**, **Continent** (where the attack took place) and **Severity** (ordinal variable describing the gravity of the attack); see Table A.1 in Appendix for a short description.

The number of cyber attacks has surged in the examined period: 1,554 attacks occurred in 2018; 1,667 in 2019; and 1,867 in 2020, with a growth of approximately 17% between 2018 and 2020.[3] Fig. 1 reports the trend of cyber attacks per month from January 2018 to December 2020 together with the average number of attacks (about 141 attacks).

The geographical distribution of the cyber attacks of the period 2018–2020 is provided in Fig. 2; more than 100 countries around the world have been affected by cyber attacks in the period 2018–2020; values expressed in a percentage scale. We notice that the majority of the attacks are directed to `America` (46%), `Europe`(13.21%) and `Multiple Continents` (27.36%). As reported in [17], by comparing the trend over the years, there is a overall increasing trend for `America` and `Europe`. On the contrary, attacks against geographically distributed targets (`Multiple Continents`) are decreasing on a global scale.

Overall, for the whole considered period, the main attacker is `Cybercrime` (4,119 attacks), the mostly used technique is `Malware` (2,110 attacks), the majority of the attacks are directed to `Multiple Targets` (1,132 attacks), the most affected continent is `America` (2,338 attacks) and a great part of the attack are of `High` severity (2,028 attacks).

[2] Covid-19 is a type of coronavirus (together with MERSnCoV and SARS-nCoV) that can spread to humans [15]. At present, it represents one of the most serious worldwide emergencies, potentially able to change the lifestyle of people, and to destroy whole economies. The first case of Covid-19 was reported on 31 December 2019 in Wuhan, China. About 45 days after the first detection, the epidemics started to affect several other countries and has now spread worldwide [16]. In March 2020, the World Health Organization declared the Covid-19 outbreak a pandemic.

[3] Note that the examined data represent a situation less critical than the actual one, since many attacks may not be disclosed, or may be disclosed at a much later date.

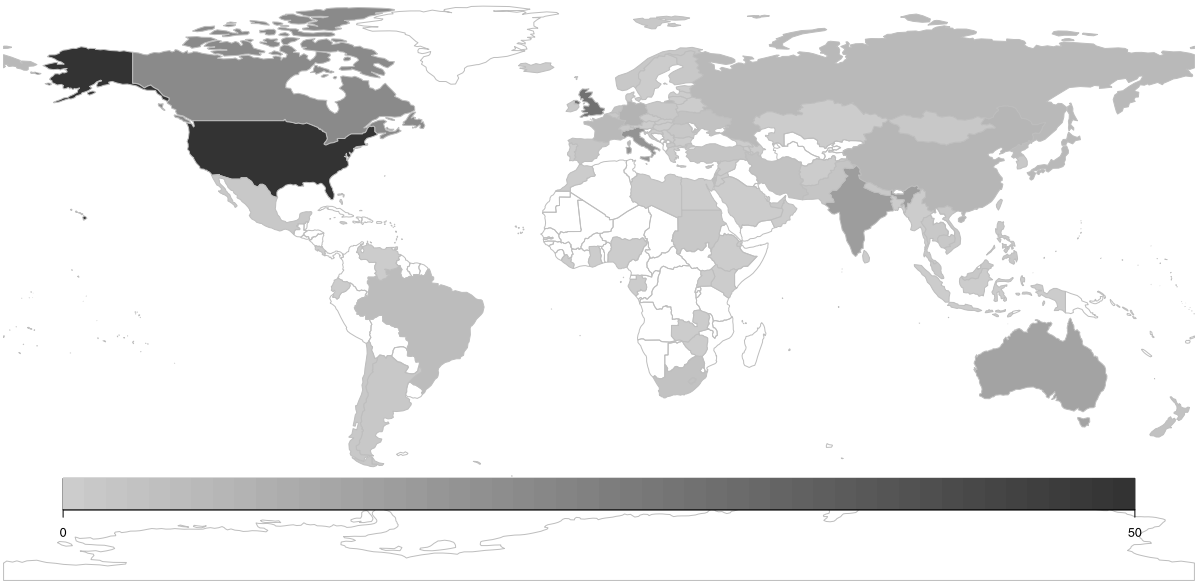


Fig. 2. Map of the percentage frequency of attacks map in the period 2018–2020. Graph produced with '**geomap**' package of R.

The classification of attack techniques is derived from MITRE ATT&CK (Adversarial Tactics, Techniques & Common Knowledge, https://attack.mitre.org), ENISA's Threat Taxonomy (European Union Agency for Cyber security, https://www.enisa.europa.eu), and Open Threat Taxonomy (https://www.auditscripts.com/free-resources/open-threat-taxonomy/).

Targets of cyber attacks are classified according to the International Standard of Industrial Classification (ISIC) of All Economic Activities (https://ilostat.ilo.org/resources/concepts-and-definitions/classification-economic-activities/).

The classification of the attackers came instead by the experience of the researchers collaborating with *Hackmanac* in the field of cyber risk and considers the main actors and motivations of the attack.

Finally, the researchers of the *Hackmanac* developed an ordinal classification of cyber risk severity (`Low`, `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 and economic impact on the targets. More precisely, the geopolitical impact is considered relevant if governmental or national security institutions are involved; the social impact is based on the number of individuals involved; the economic impact is measured in terms of the amount of estimated damages.

The distribution of the **Severity** of cyber attacks is reported in the clustered line chart of Fig. 3. The situation changes from 2018 to 2020: while the number of `Medium` level attacks slightly increased (+19%), that of `Low` (+343%) and `Critical` (+96%) impact attacks significantly increased. Conversely, the number of `High` severity attacks decreased (−20%).

In the following Sections (3, 4 and 5), we present three different types of networks (BNs, RFs and SNs) and show how they can be used to study different aspects of the problem under investigation. Each of the following sections start with a brief theoretical introduction and then focuses on the application of the proposed methodology to the analysis of cyber risk data.

3. Bayesian network analysis

In this section, we show how BN can be used not only to provide a pictorial representation of the relationship among the examined variables, but also to investigate alternative risk scenarios. After a brief introduction on the main notation and terminology of BN, we illustrate the application of this methodology to cyber risk data. We conclude this part with an analysis of attacks directed towards `Healthcare` sector during the Covid-19 pandemic.

3.1. BN: Notation and terminology

A BN is a multivariate statistical model that uses a Directed Acyclic Graph (DAG) to describe the dependence structure among a set of variables. A DAG is a directed graph $D = (V; E)$, where $V = \{v_i; i = , \dots, |V|\}$ is the set of nodes and E is the set of directed edges (v_i, v_j) connecting a pair of nodes. Nodes represent a set of random variables $\mathbf{X} = (X_1, \dots, X_{|\mathbf{X}|})$, and directed edges (depicted as arrows) indicate dependencies among the corresponding variables. In a DAG no directed cycle is present, this means that it is not possible to start from a node and, following the directions of the arrows, return to the starting one.

If an arrow points from node v_i to node v_j then v_i is called parent of v_j; the set of parents of node v_j is denoted as $pa(v_j)$, or equivalently as $pa(X_j)$. The absence of an edge between two nodes may indicate marginal/conditional independence between the corresponding variables.

For example, the following graph configurations $v_i \to v_j \to v_\ell$ and $v_i \leftarrow v_j \to v_\ell$ indicate that variables X_i and X_ℓ are independent given X_j. On the other hand, a graph configuration such as $v_i \to v_j \leftarrow v_\ell$ indicates that variables X_i and X_ℓ are dependent given X_j.

Exploiting the graphical structure, we can easily factorize the joint probability distribution as follows

$$p(X_1, \dots, X_{|\mathbf{X}|}) = \prod_{i \in V} p\Big(X_i | pa(X_i)\Big),$$

where $p\big(X_i|pa(X_i)\big)$ denotes the conditional probability distribution of variable X_i given its parents set. If node X_i, has no parents $p\big(X_i|pa(X_i) = \emptyset\big) = p(X_i)$, with $p(X_i)$ marginal distribution of X_i. For more details on BN semantics and properties, see e.g. [18] and [19]. For examples of applications of BNs to education, banking, data integration and official statistics, see [20–23], respectively.

3.2. Application of BN to cyber attacks data

The graphical structure could be either specified by using expert knowledge or learned from the data. In our analysis, we follow the second approach, and use the PC structure learning algorithm implemented by the software `GeNIe` (https://www.bayesfusion.com/genie/), see [24] for the original formulation of the PC algorithm. The PC algorithm is one of the earliest and the most popular constrained based structural algorithms. It carries out a series of independence tests and construct a graph which satisfies the discovered independence

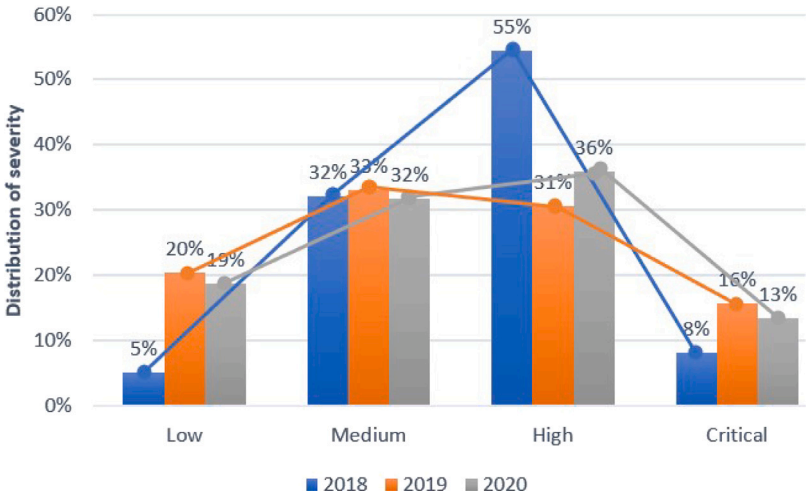


Fig. 3. **Severity** distribution of cyber attacks during the period 2018–2020. Graph produced with `Excel`.

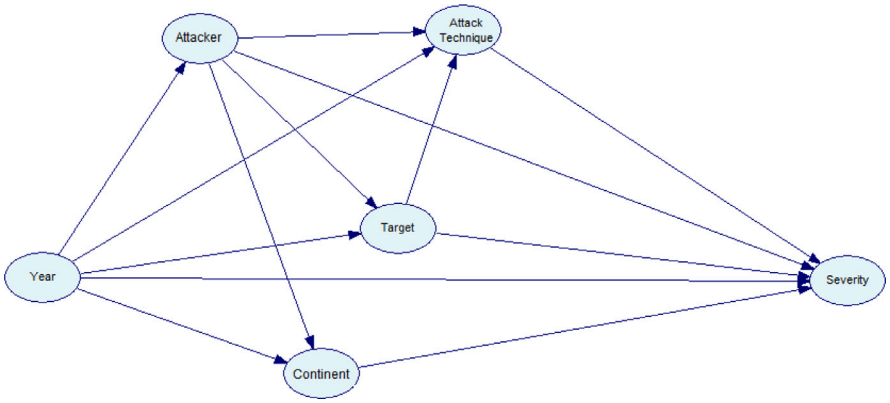


Fig. 4. Selected network structure for cyber attacks data. Graph produced with `GeNIe`.

statements. Furthermore, differently from the other structural learning algorithm available in `GeNIe`, the version of the PC algorithm implemented in `GeNIe` allows the user to choose among independence equivalent models, the one most suitable for the analysed problem (i.e. choosing the direction/presence of specific edges).

The network in Fig. 4 represents the selected dependence structure for the cyber data obtained setting variable **Severity** as a response variable. While, Fig. 5 displays the marginal probability tables estimated from the data. We notice that all explicative variables are directly connected to **Severity**, with variable **Target** playing a central role in the network. Indeed, **Target** is directly influenced by variables **Year** and **Attacker** and influence in turn the **Severity** and **Attack Technique**. As discussed in Section 2, the majority of cyber attacks are `Cybercrime` based on a `Malware` among the alternative techniques.

The network can be used to evaluate in a mouse-click time alternative risk scenarios. For example, Fig. 6 represents the situation in which only `Critical` attacks are considered. It turns out that these types of attacks are mostly directed to specific targets. Indeed, if we examine the marginal distribution of variable **Target** we notice a decrease in the percentage of attacks directed to `Multiple Targets` (13% in contrast to the previous value of 22%) while other types of attacks such as `Government` and `Financial Insurance` become more frequent (18% versus 13% and 13% versus 8%, respectively).

Furthermore, the network can be easily used to investigate the threat landscape of each single target, taking also in consideration the different periods. In fact, as pointed out in the Clusit report of March 2021 [17] it is important to provide a specific risk scenario of each single target.

As an exemplification, Fig. 7 reports the situation for year 2020. Focusing only on attacks directed to the `Government` sector, we obtain the representation in Fig. 8. We notice that while the proportion of attacks carried out for cybercriminal purposes against this sector is significantly lower than the general one of 2020 (61% versus 81%) the component attributable to `Espionage/Sabotage` is almost duplicated (26% versus 14%), while the attacks made for purpose of `Hacktivism` and for `Information Warfare` purposes are almost tripled.

Taking into consideration only attacks directed towards the `Government` sector, we notice that the percentage of attacks made for `Hacktivism` and for `Information Warfare` purposes are stable over time. On the other hand, we notice that `Cybercrime` attacks increases in 2019 (67%) versus 2018 (9%); the reverse is observed for `Expionage/Sabotage` (26% in 2018 versus 19% in 2019).

Replicating the previous analysis for the other categories of target (results are not reported here for lack of space), we observe significant differences among the alternative scenarios. This indicates that, each category of target, has its own particular threat landscape from which it must protect itself, implementing a specific defensive strategy.

3.3. Analysis of Covid-19 effect

We conclude this section with the analysis of the `Healthcare`, a sector that has been significantly impacted by Covid-19 related attacks during 2020 (see [17] pp. 31–35), see Fig. 9.

The `Healthcare` sector was affected mainly by attacks for cybercriminal purposes, in particular extortion (ransomware) and theft of

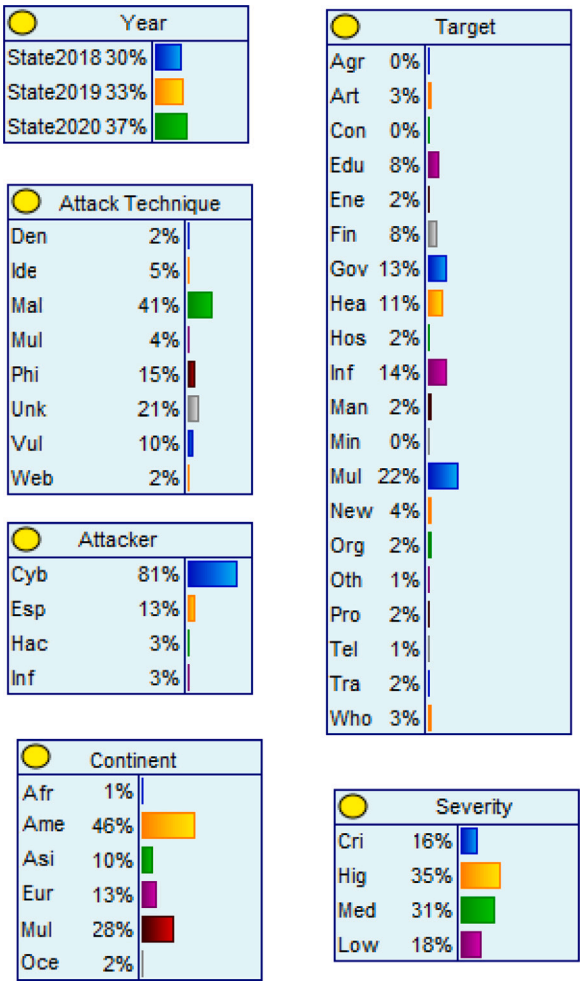


Fig. 5. Estimated marginal probabilities (percentage values) for the cyber attack network. Figure produced with `GeNIe`.

personal data, to be used to carry out further attacks. The `Espionage` component goes from 2% in 2019 to 4% in 2020 (1% in 2018), with a sensible increase in risk evaluation (61% of `Critical` and `High` severity attacks with respect to 48%, the global value of the period), see Fig. 10. This is due to the intelligence activities that interested the category (with particular interest in the development of vaccines for the Covid-19) during the year.

4. Random forest

In this section, we show how RF can be a useful instrument for cyber risk evaluation. Firstly, we present a brief description of the model, and we refer to [25,26] for more details. Then, we concentrate on the application of RF to cyber risk data.

4.1. RF: Notation and terminology

RF model is a supervised machine learning algorithm that makes predictions by averaging outcomes from a collection of uncorrelated Decision Trees (DTs). The algorithm works as follows. Firstly, it generates T different training subsets from the considered data set by using a bootstrap sampling approach; then T decision trees are built by training these subsets. Finally, a random forest is constructed from these DTs.

More specifically, DTs allow investigating interaction effects of explicative variables on the response one considering a suitable partition of the space of the explicative variables in non-overlapping regions with similar response values. In each region, the prediction of the value of the response variable is the mean or the mode of the value assumed by the observations classified in that region. The set of splitting rules used to define the regions can be graphically depicted as a tree, hence the name DT methods. The tree is a particular type of DAG where nodes are connected, that is, there is a path connecting each node to any other one.

DT is constructed splitting a data set into different subsets according to certain cutoff values of the explicative variables. It can be implemented by using a binary recursive splitting algorithm. At the first step, the original data set is split in two subsets according to a certain cut-off value of the selected explicative variable. This leads to a partition that allows achieving the best fit in terms of predictive accuracy; see [26] for details. In the following steps, this process is repeated on each derived subset recursively until a suitable stopping rule is satisfied, see [27].

The nodes of the tree represent steps of the splitting process. The root node represents the initial situation where the full data set is considered; decision nodes denote the criteria for subsequently classifying the observations in subclasses corresponding to the different regions; leaf nodes (terminal nodes) that represent the final regions. All nodes have a parent node, except the root one. A forest is then obtained as a collection of trees.

The term "Random" Forest derives from the fact that, at every step of the trees' growth process, the algorithm searches the most important explicative variable for the splitting within a random subset of all p explicative variables.

More precisely, for each node, a random sample of $q \approx \sqrt{p}$ variables is selected from the full set of p input explicative variables. This procedure ensures low correlation among DTs and improves predictive performance of the model, see [26].

4.2. Application of RF to cyber attacks data

In this section, we apply RF to data of cyber attacks over the years 2018–2019 in order to detect the relevant variables that influence the severity of an attack. For the implementation of the model we used the package **'randomForest'** of R program, [28].

First, we implement a RF model by considering all the available nominal explicative variables (complete model) and we evaluate the importance of these variables to predict the level of severity of cyber attacks.

The importance of each variable is computed by considering the mean decrease in Gini coefficient. It measures how each variable contributes to the purity of the nodes (averaged over all trees). A node is pure if it contains observations from a single class of the response variable. The higher the purity of the node, the lower its classification error. See [26,28] for the definition and the implementation of this measure, respectively.

The higher the mean decrease in Gini coefficient, the higher the importance of the variable in the model. Fig. 11 show the importance plot for the cyber attacks data. The variables are presented in decreasing order of importance.

The results indicate that **Target** and **Attack Technique** are the two most impactful variables in the prediction of response across all the trees considered in the RF.

We also apply the RF by considering only one explicative variable at a time. In this way, we can evaluate the importance of the categories of each variable in the prediction of the attack severity. More specifically, we implement p univariate RFs with a unique explicative variable i, with $i = 1, 2, \dots, p$. For each category of the explicative variable i we compute the importance measure. We repeat the procedure for each explicative variable i, for $i = 1, 2, \dots, p$. The graphs of Fig. 12 report the results of the importance measures for the categories of each variable used in the p models.

First, we analyse **Target** and **Attack Technique** that are the two most important variables resulting from the complete model (Fig. 11).

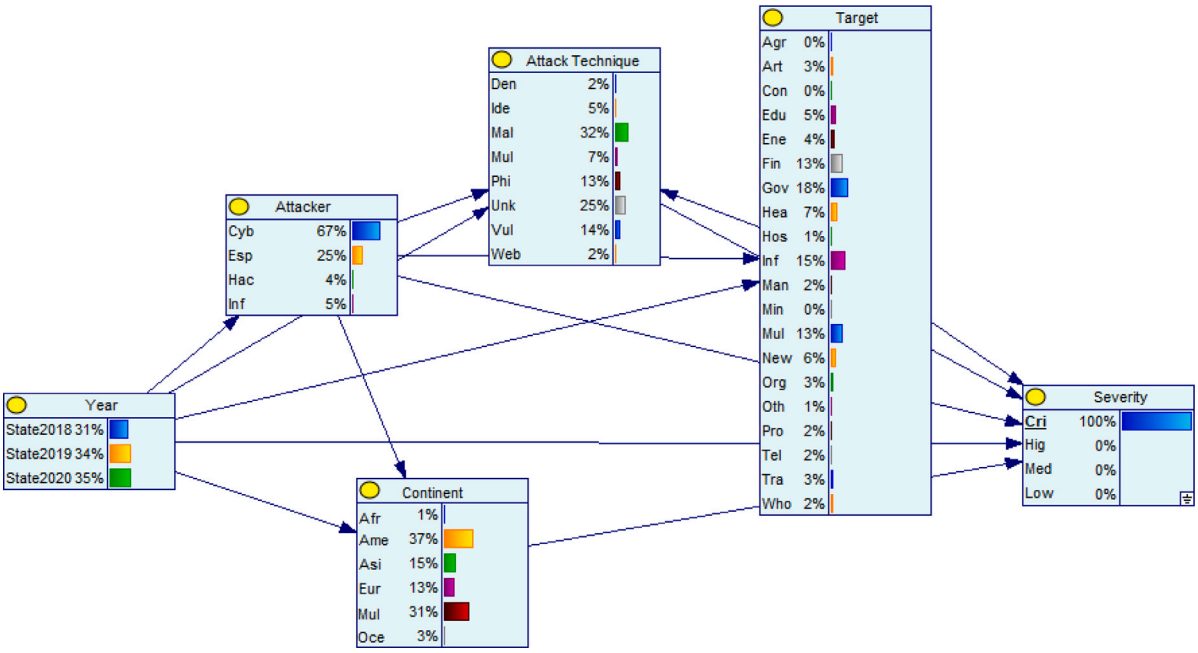


Fig. 6. Critical attacks evaluations. Graph produced with GeNIe.

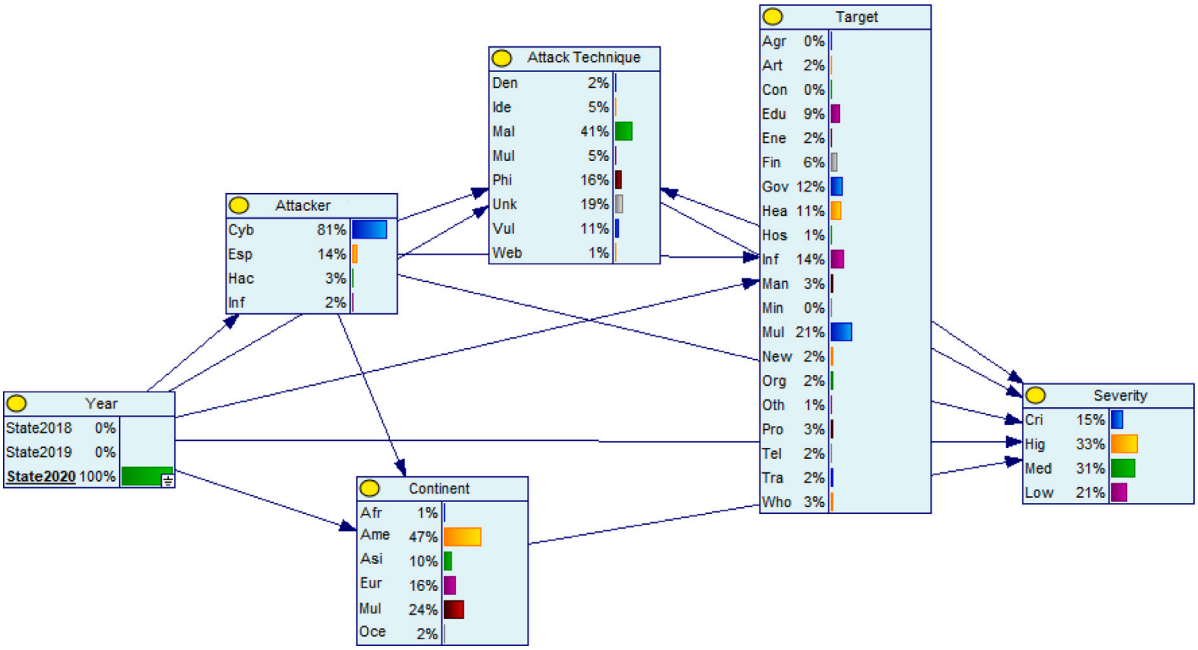


Fig. 7. Cyber attack network for year 2020. Graph produced with GeNIe.

From Fig. 12 we see that the most relevant target categories are Multiple Targets, Education, and Health, while the most important attack techniques are Unknown, Phishing Social Engineering and Multiple Techniques.

Moreover, for variable **Attacker**, the most influential type of attacks are Espionage and Cybercrime. Finally, for the variable **Continent**, we see that the most important categories are Multiple Continent and America.

4.3. Analysis of Covid-19 effect

To evaluate if the Covid-19 pandemic affects the importance of the variables in the prediction of the attack severity levels, we apply the RF to attacks observed in 2020 (year dominated by Covid-19). Fig. 13 reports the importance measures of the explicative variables for a complete RF implemented on the data by considering all the explicative variables. By comparing the results of Fig. 13 with the ones reported in Fig. 11, we note that the ranking of the explicative variables, based on their importance measure, is not changed during the Covid-19 pandemic. **Target** remains the most relevant variable with the highest value of the importance measure. However, we see a decrease in the Gini scores for variables **Continent**, **Attacker** and **Attack Technique** during the pandemic.

Finally, we apply univariate RFs on data of 2020 by considering only one explicative variable at a time. The aim is to evaluate the

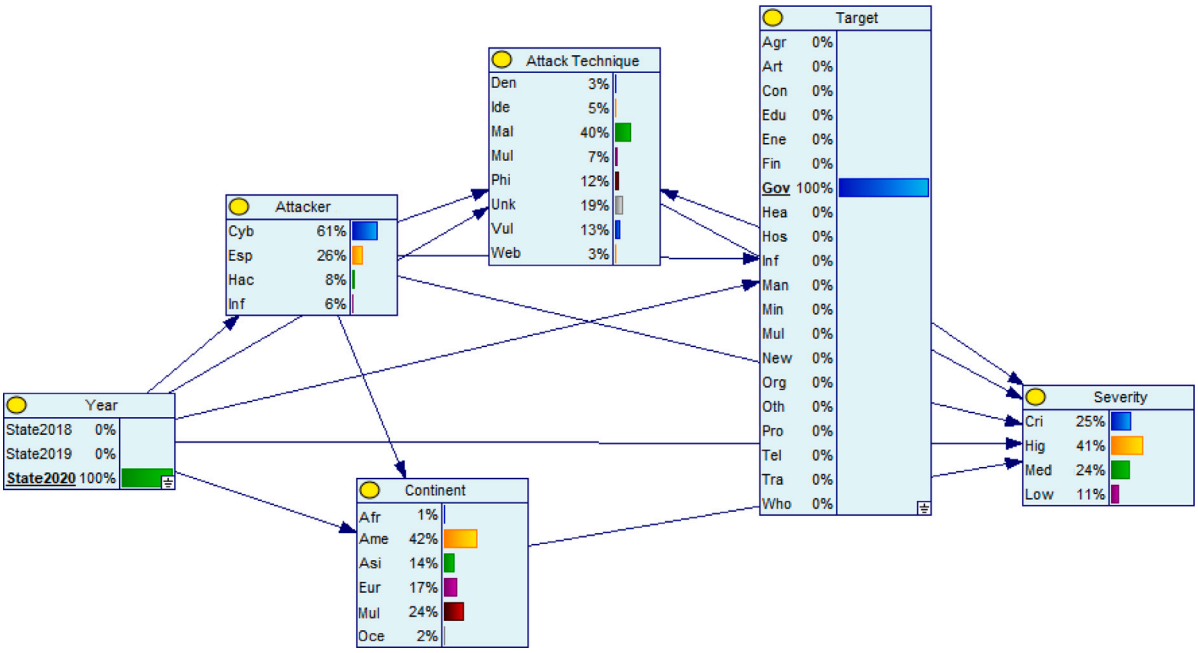


Fig. 8. Cyber attack network towards the `Government` sector in 2020. Graph produced with `GeNIe`.

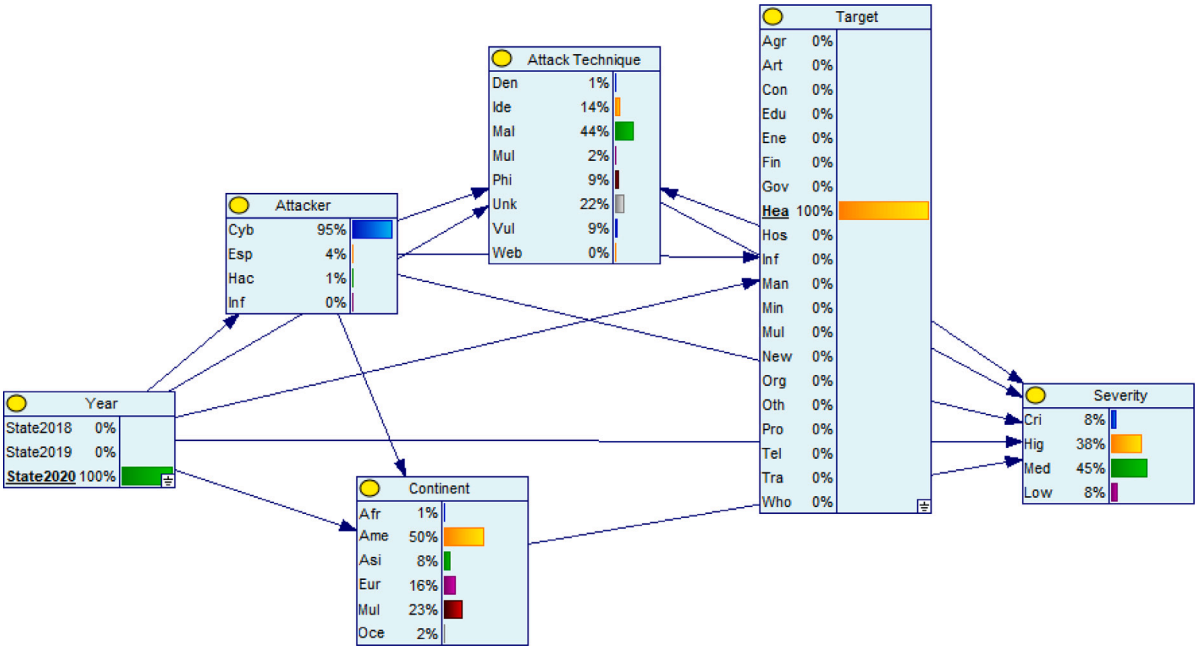


Fig. 9. Cyber attack network towards the `Healthcare` sector in 2020. Graph produced with `GeNIe`.

change in the importance ranking of the categories of each variable during the Covid-19 pandemic. The results are reported in Fig. 14.

By comparing the results obtained during the pandemic with the ones obtained before and reported in Fig. 12, we see a change in the importance of the raking of the categories of some variables. For **Attacker** and **Attack Technique**, we note a change in the top category: before the pandemic the most influential type of attacker in the prediction of **Severity** was `Espionage`, while during the Covid-19 pandemic `Cybercrime` becomes the most important category. For the variable **Attack Technique** the category `Phishing` moves to the top of the ranking.

We also analyse the variables **Continent** and **Target**. From the figure, we do not see any change in the top positions of the ranking of the categories for both variables. `Multiple Continent` and `Multiple Targets` are the most important categories in the prediction of the severity level. However, the value of the Gini measure for both categories increased during the Covid-19 pandemic in comparison to the previous period.

5. Social network analysis

In this section, we show how SN analysis and its centrality measures can be used to measure the interconnection among the targets and to

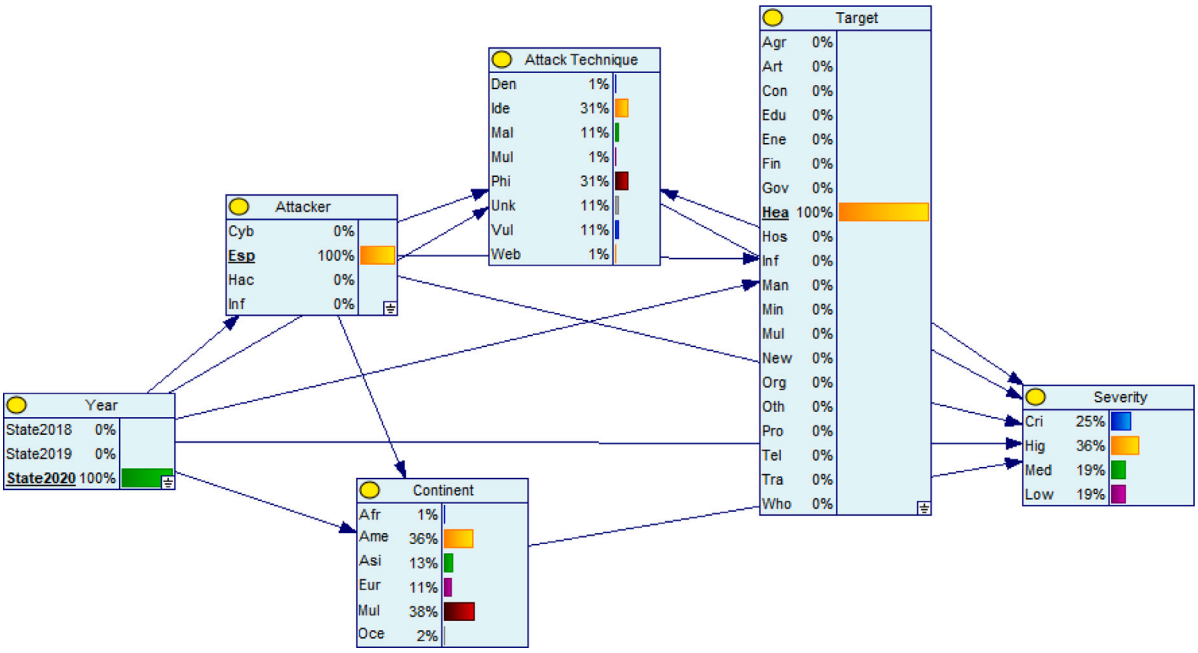


Fig. 10. Espionage cyber attack network towards the Healthcare sector in 2020. Graph produced with GeNIe.

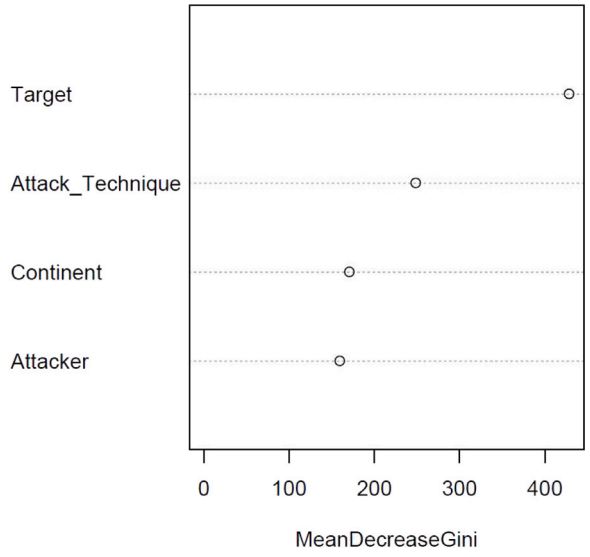


Fig. 11. The variable importance plot for cyber data over the period 2018–2019. The values are expressed relative to the maximum. Figure produced with the package **'randomForest'** of R.

monitor the diffusion of cyber attacks. We start the following part with a brief introduction to the main notation and terminology of SN, for more details see e.g. [29]. Subsequently, we illustrate the application to cyber attacks data, and we evaluate if the Covid-19 pandemic affects the relationships between targets of cyber attacks.

5.1. SN: Notation and terminology

SN analysis studies the behaviour of different actors (i.e. individuals or objects), the pattern of their relationships and the interactions between them using a graph.

A SN is a weighted graph $G = (V, E, \mathbf{W})$, where $V = \{v_i; i = , \dots, |V|\}$ is the set of nodes (actors), E is a set of edges and $\mathbf{W}$ is the weighted adjacency matrix. If weight w_{ij} is greater than 0, then the corresponding nodes v_i and v_j are connected. All edges are undirected; if $(v_i, v_j) \in E$ then $(v_j, v_i) \in E$. An edge between two nodes indicates that there is a connection between the corresponding actors. Weights $w_{ij} \in \mathbf{W}$ are used to measure the strength of such a connection. In a graph, the width of each edge is proportional to the corresponding weight, so that a high weight indicates a stronger connection between two nodes.

To study how actors are connected across the whole network and identify key elements, we rely on centrality measures, see e.g. [30]. Many centrality measures have been found over the years and used in different contexts; for a survey on centrality measures in social networks, see e.g. [31]. Here we focus on closeness and betweenness centrality measures belonging to the shortest path category of measures [32]. A path from $v_i \in V$ and $v_j \in V$ is defined as an alternating sequence of nodes and edges, beginning with v_i and ending with v_j, such that each edge connects its preceding with its succeeding node. The shortest path among $v_i \in V$ and $v_j \in V$ is the one with the minimal distance $d^*(i, j)$ between the nodes. The minimal distance is given by

$$d^*(i,j) = \min\{d(i,j)\} = \min\left\{ \frac{1}{|w_{12}|} + \cdots + \frac{1}{|w_{(h-1)h}|} + \cdots + \frac{1}{|w_{(|V-1|-1)|V|}|} \right\},$$

where the weight $w_{(h-1)h}$ associated to the edge (v_{h-1}, v_h) corresponds to the partial correlation between the variables associated with nodes v_{h-1} and v_h.

The closeness centrality measure c_i scores each node v_i considering its "proximity" to all other nodes in the network, see [33,34]. It is proportional to the inverse of the sum of the shortest path distances between the examined node and all other nodes in the network. Mathematically, for node v_i is defined as

$$c_i = \frac{1}{\sum_{i \neq j} d^*(i,j)}.$$

The more central a node is, the lower its total distance to all other nodes.

A node has a high closeness centrality if the information from this node can reach other nodes quickly and therefore, it can interact efficiently with other nodes in the network.

It is possible to have more than one shortest path between a pair of nodes v_i and v_j. Hence, one can use the betweenness centrality measure that indicates the number of times a node lies on the shortest path

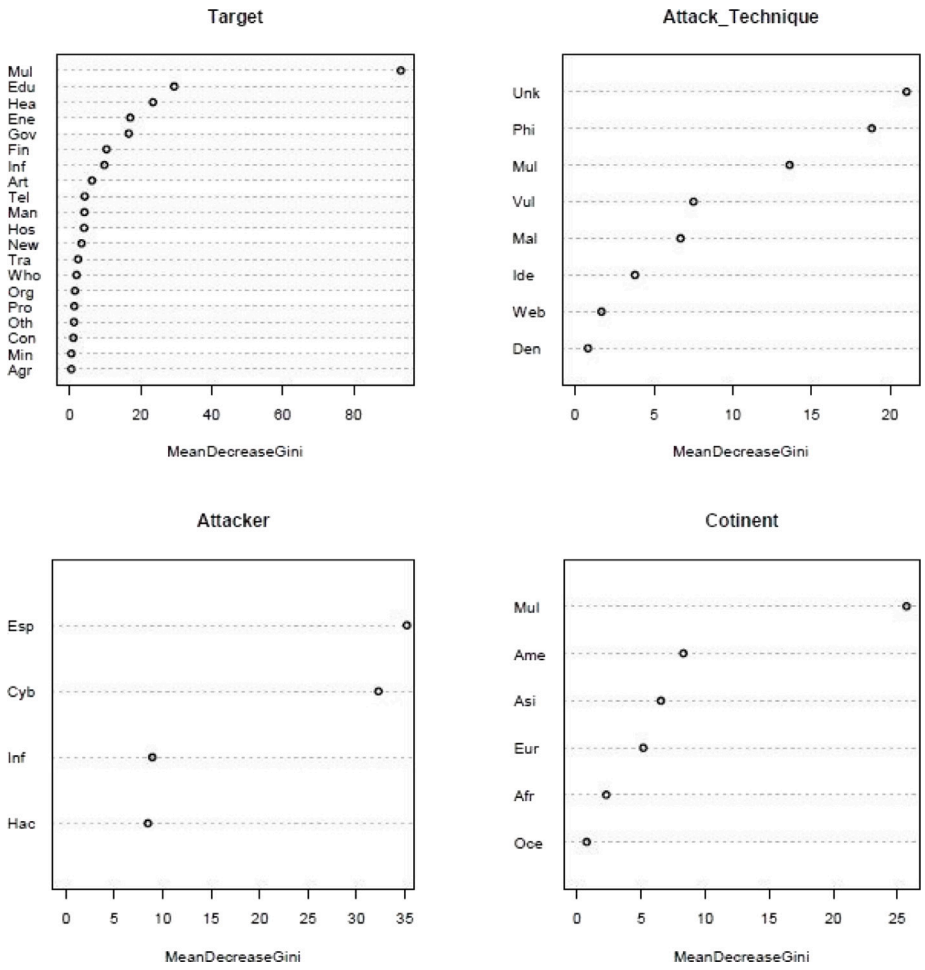


Fig. 12. The importance plots for the categories of each explicative variable of the cyber data over the period 2018–2019. The values are expressed relative to the maximum. Figure produced with the package **'randomForest'** of R.

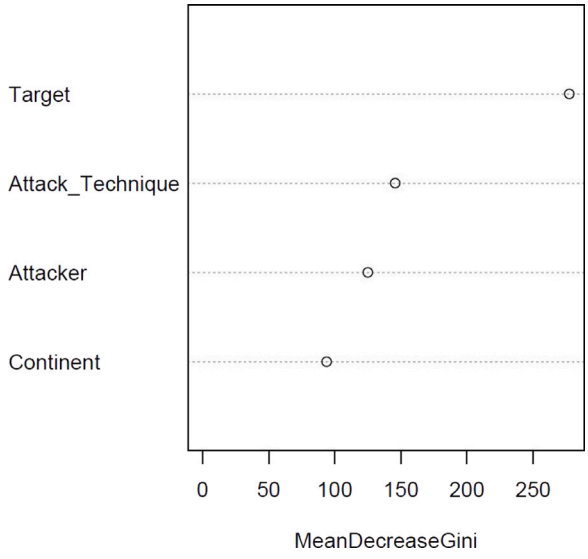


Fig. 13. The variable importance plot for cyber data during the Covid-19 pandemic. The values are expressed relative to the maximum. Figure produced with the package **'randomForest'** of R.

between other nodes; see [35,36]. It is a measure of how important each node might be to effective communication within the network. Mathematically, if n_{hj} denotes the number of all shortest paths between v_h and v_j in the network and $n_{hj}(i)$ is the number of shortest paths connecting v_h and v_j that pass through node v_i, the betweenness centrality measure b_i of the node v_i is defined as

$$b_i = \sum_{j \neq h \neq i} \frac{n_{hj}(i)}{n_{hj}}.$$

A node with a high betweenness centrality can influence the information flow between not directly connected nodes. Removing a node of high betweenness will lengthen the paths connecting several other nodes, rendering communication between them less efficient.

5.2. Application of SN analysis to cyber attacks data

In this section, we illustrate how SN analysis may help to visualize interactions among targets (nodes v_h) of cyber attacks and to identify the more important and the isolated ones. For the implementation, we used the package **'qgraph'** of R, [37].

To build a network connecting the targets of cyber attacks we consider the Criticality index proposed in [38,39]. It is a normalized index that can be used to provide an indication of vulnerability of the targets of cyber attacks: the higher is the value of this index, the higher is the vulnerability of the corresponding target. The Criticality index is based on the relative cumulative frequencies $\hat{F}_k$ of cyber attacks

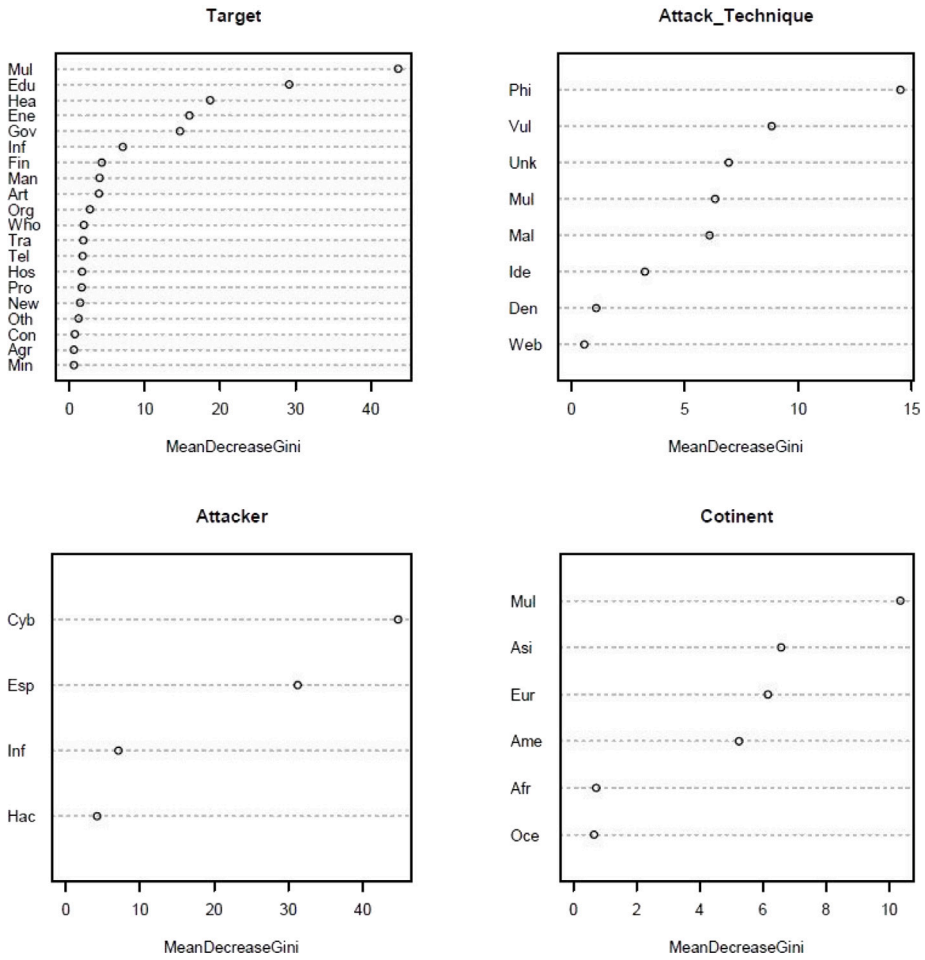


Fig. 14. The importance plot for the categories of each explicative variable of the cyber data during the Covid pandemic. The values are expressed relative to the maximum. Figure produced with the package **'randomForest'** of R.

suffered by a target, for $k = 1, 2..., K$ increasing levels of severity. Given a target v_h and for a specific week, this index is calculated as

$$\hat{I}^{v_h} = 1 - \frac{\sum_{k=1}^{K} \hat{F}_k - 1}{K - 1}.$$

In the considered setting $K = 4$ and the values of k from 1 to 4 correspond to `Low`, `Medium`, `High` and `Critical` severity, respectively.

For each target, we compute the weekly Criticality index time series, and we derive the absolute partial correlation matrix among them. We then connect via an edge targets presenting a non zero absolute partial correlation; the corresponding network is shown in Fig. 15. The greyscale and the width of the edges show how strong the correlation is. A high partial correlation value indicates a strong dependence in terms of vulnerability of the two considered targets. Otherwise, a low partial correlation value highlights that the vulnerabilities of the two targets do not influence each other.

In order to identify the targets that influence more the entire network, we rely on closeness and betweenness centrality measures described in the previous section.

If a target is strongly correlated (positively or negatively) with many others, it is also highly interconnected to them in the network. An attack inflicted on this target could have an indirect effect also on the interconnected ones. Otherwise, there is a low effect in case of weak correlation (correlation coefficient near zero) among the targets in the network.

Fig. 16 shows the centrality plot of the examined network. The centrality plot is a standard representation for these types of measures;

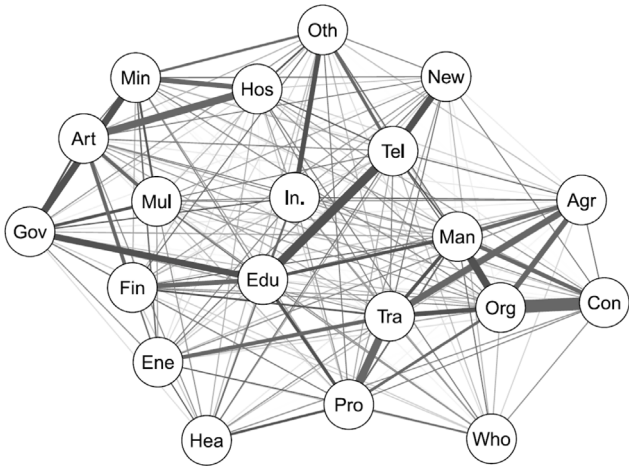


Fig. 15. Network structure among the targets of cyber attacks. Figure produced with the package **'qgraph'** of R.

see e.g. [40]. The top three targets, ranked by centrality measures, are reported in Table 1.

Considering the relation among the two examined centrality measures, we notice that they are positively related, with high values of Pearson's (0.88951), Sperman's (0.92962) and Kendall's (0.81250) indices.

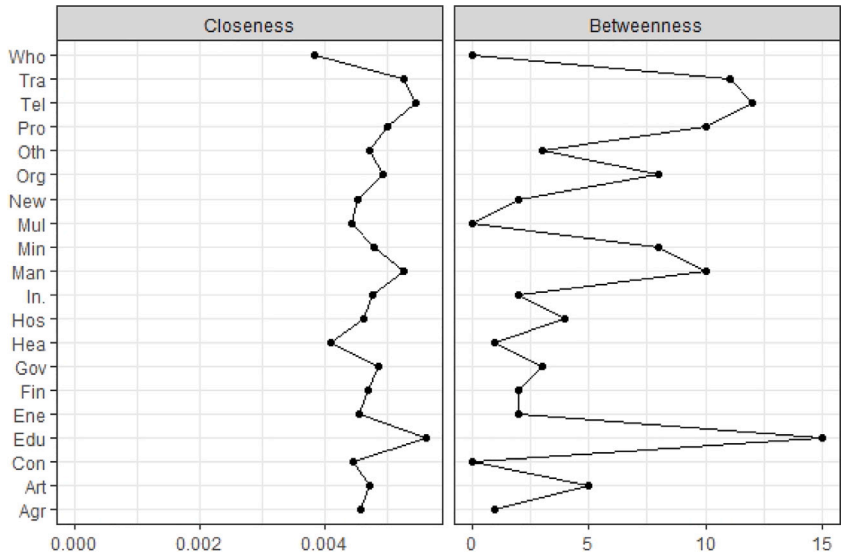


Fig. 16. Closeness and betweenness centrality measures. Figures produced with the package **'qgraph'** of R.

Table 1
Top three targets ranks.

Rank	Closeness	Betweenness
1	Education	Education
2	Telecommunications	Telecommunications
3	Manufacturing	Transportation Storage

Each centrality measure captures a different aspect: the closeness centrality evaluates targets focusing on how strong it affects other targets; the betweenness centrality evaluates targets with more possibility to contact the other targets. Education, which is ranked first according to both centralities (followed by Telecommunications), is considered the centre of the network.

5.3. Analysis of Covid-19 effect

To evaluate if the Covid-19 pandemic affects the interconnection among the targets of cyber attacks, we apply the procedure described in Section 5.2 separately for years 2018–2019 and 2020 (year dominated by Covid-19).

Fig. 17 shows the absolute partial correlation coefficients for years 2018–2019 (left) vs 2020 (right) among the targets of cyber attacks.

Firstly, we highlight that a new target that had not been considered by cybercriminals in previous years was hit in 2020: Agriculture Forestry Fishing.

The greyscale and the dimension of the balls in Fig. 17 show how strong the correlation between the targets is. A comparison between the two matrices point out higher values of the absolute partial correlation after the start of Covid-19 pandemic, thus a stronger dependence in terms of vulnerability of the targets of cyber attacks. For example, in 2018–2019 the absolute partial correlation between Financial Insurance and Government Military Law Enforcement was equal to 0.02881, while in 2020 it rises to 0.54517.

Observing the values of the closeness and betweenness centrality reported in Fig. 18, we note that the interconnectedness between targets also changes after the start of the pandemic.

Observing closeness centrality, we note an increase in the speed with which one target affects others. In fact, the mean value passes from 0.00542 in 2018–2019 to 0.00913 in 2020. While before the beginning of pandemic Manufacturing was ranked first according to this measure, in 2020 it ranked 16th. The target higher interconnected in 2020 cyber network is Financial Insurance.

Table A.1
Description of variables of the cyber risk dataset.

Variable	Category
Attacker (4 categories)	Cybercrime (Cyb) Espionage/Sabotage (Esp) Hacktivism (Hac) Information Warfare (Inf)
Attack Technique (8 categories)	Denial of Service (Den) Identity Theft Account Cracking (Ide) Malware (Mal) Multiple Techniques (Mul) Phishing Social Engineering (Phi) Unknown (Unk) Vulnerabilities (Vul) Web attack (Web)
Target (20 categories)	Agriculture Forestry Fishing (Agr) Arts Entertainment (Art) Construction (Con) Education (Edu) Energy Utilities (Ene) Financial Insurance (Fin) Government Military Law Enforcement (Gov) Healthcare (Hea) Hospitability (Hos) Information Communication Technology (Inf) Manufacturing (Man) Mining Quarrying (Min) Multiple Targets (Mul) News Multimedia (New) Organizations (Org) Other Services (Oth) Professional Scientific Technical (Pro) Telecommunications (Tel) Transportation Storage (Tra) Wholesale Retail (Who)
Continent (6 categories)	Africa (Afr) America (Ame) Asia (Asi) Australia/Oceania (Aus) Europe (Eur) Multiple Continents (Mul)
Severity (4 categories)	Low (Low) Medium (Med) High (Hig) Critical (Cri)

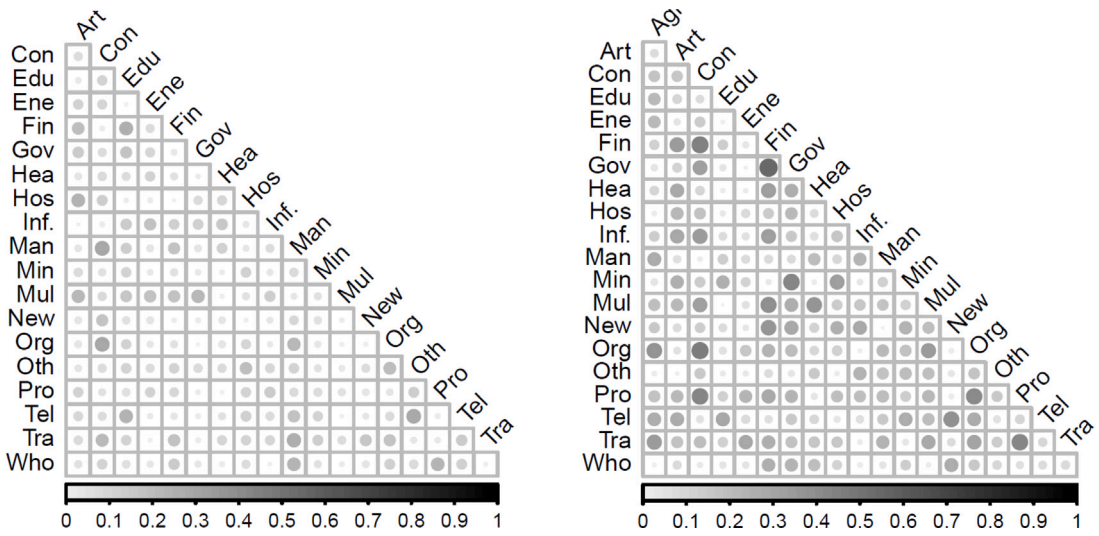


Fig. 17. Triangular part of the absolute partial correlation matrix (without the main diagonal): 2018–2019 (left) vs 2020 (right). Figures produced with the package '**corrplot**' of R.

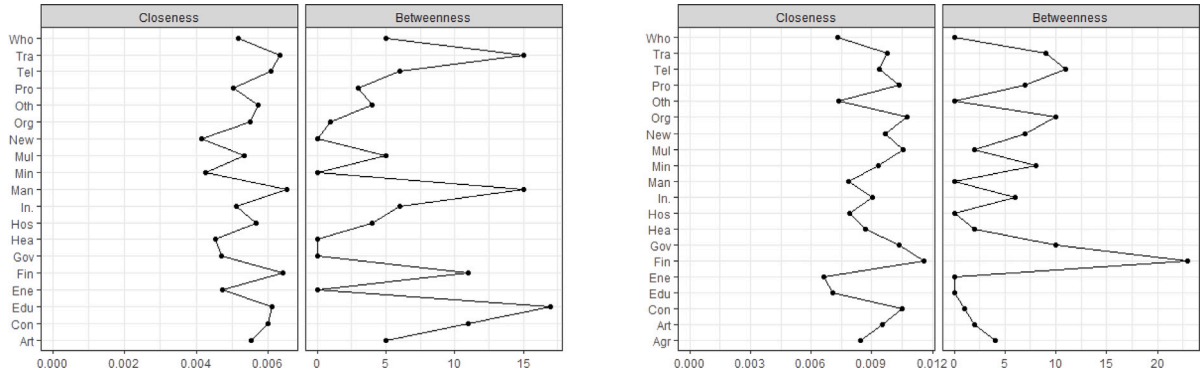


Fig. 18. Closeness and betweenness centrality measures: 2018–2019 (left) vs 2020 (right). Figure produced with the package '**qgraph**' of R.

If we referred to betweenness centrality, the total number of times a target lies on the shortest path between other targets is almost the same (108 in 2018–2019 and 102 in 2020), but the interconnections considerably change. In fact, while before the beginning of pandemic `Education` was ranked first, in 2020 its betweenness was equal to zero on an equal footing with `Energy Utilities`, `Hospitability`, `Manufacturing`, `Other Services` and `Wholesale Retail`. `Financial Insurance` is the target with more possibility to contact the others in 2020.

6. Conclusions

In this paper, we have shown how BNs, RFs and SNs are suitable instruments to evaluate cyber risk. An application of these models to real data of cyber attacks observed before and during Covid-19 pandemic allowed us to analyse several aspects of the risk assessment. Bayesian Networks provided a visual representation of the relationship among the variables that influenced the severity of an attack. We notice that all variables have a direct influence on the **Severity** with **Target** playing a central role in the network. The network permitted us to evaluate alternative risk configurations and evaluate in a real-time alternative risk landscape. Random Forest detected the relevant factors that influence the severity of the attacks. According to the considered importance measures, **Target** and **Attack Technique** are the two most important variables in the prediction of response. The most important target categories are `Multiple Targets`, `Education` and `Health`, while the most important attack techniques are `Unknown`, `Phishing Social Engineering` and `Multiple Techniques`. Finally, the implemented Social Network among the targets of cyber attacks and its centrality measures allowed us to evaluate the interconnection among victims and the presence of a possible contagion effect. From the analysis, `Education` was considered the centre of the network since it ranked first according to the examined centrality measures (followed by `Telecommunications`). We also investigated the risk assessment during the Covid-19 period. The results obtained applying the three models suggest that there is a slight effect of the pandemic on cyber risk evaluation.

The networks models discussed here can be tools in service of practitioners and regulators for setting properly cyber security policy considering the specific characteristic of the observed attacks. The results emerging from the application of these three types of networks can guide supervisory authority and professionals in defining their risk profile and then studying appropriate improving actions to reduce it and prevent reputational and monetary damage.

CRediT authorship contribution statement

Silvia Facchinetti: Conceptualization, Design, Analysis, Drafting, Revision of the manuscript. **Silvia Angela Osmetti:** Conceptualization, Design, Analysis, Drafting, Revision of the manuscript. **Claudia Tarantola:** Conceptualization, Design, Analysis, Drafting, Revision of the manuscript.

Data availability

The authors do not have permission to share data.

Acknowledgements

We thank the AE and the two referees for helpful comments and suggestions. This work acknowledges research support by COST Action CA19130 'Fintech and Artificial Intelligence in Finance - Towards a transparent financial industry' (FinAI), supported by COST (European Cooperation in Science and Technology). We also thank the experts of the Hackmanac Project for sharing the dataset.

Appendix

In Table A.1 a description of the variables **Attackers**, **Attack Technique**, **Target**, **Continent**, **Severity** is reported.

References

[1] IBM Security. Cost of a data breach report. 2020.
[2] Moore R. Cybercrime: Investigating high-technology computer crime. Cleveland: Anderson Publishing; 2005.
[3] Risk Management Group. The 2002 loss data collection exercise for operational risk: Summary of the data collected. Report to the bank for international settlements. 2003, Available at: http://www.bis.org.
[4] De Luca G, Carità D, Martinelli F. Statistical analysis of operational risk data. Springer International Publishing; 2020.
[5] Dalla Valle L, Giudici P. A Bayesian approach to estimate the marginal loss distributions in operational risk management. Comput Statist Data Anal 2008;52:3107–27.
[6] Jarrow RA. Operational risk. J Bank Financ 2008;32:870–9.
[7] Moosa IA. Operational risk management. New York: Palgrave Macmillan; 2007.
[8] Curti F, Gerlach J, Kazinnik S, Lee MJ, Mihov A. Cyber risk definition and classification for financial risk management. 2019.
[9] Kashyap A, Wetherilt A. Some principles for regulating cyber risk. AEA Pap Proc 2019;109:482–7.
[10] Edgar TW, Manz DO. Research methods for cyber security. Elsevier; 2017.
[11] Hartwig RP, Wilkinson C. Cyber risks: The growing threat. USA: Insurance Information Institute; 2014, p. 1–27.
[12] Falco G, Eling M, Jablanski D, Miller V, Gordon L, Wang S, et al. A research agenda for cyber risk and cyber insurance, Conference: Workshop on the economics of information security (WEIS), Boston, MA. 2019.
[13] Ramirez R, Choucri N. Improving interdisciplinary communication with standardized cyber security terminology: A literature review. IEEE Access 2016;4:2216–43.
[14] Allodi L, Massacci F. Security events and vulnerability data for cybersecurity risk estimation. Risk Anal 2017;37:1606–27.
[15] Hui DS, Azhar IE, Madani TA, Ntoumi F, Kock R, Dar O, et al. The continuing 2019-nCoV epidemic threat of novel coronavirus to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. Int J Infect Dis 2020;91:264–6.
[16] World Health Organization (WHO). Coronavirus disease 2019 (COVIDC-19) situation report - 35. WHO; 2020.
[17] Antonielli A, Arsene L, Barletta VS, Butti G, Caivano D, Ciardi N, et al. Rapporto clusit 2021 sulla sicurezza ICT in Italia, clusit. 2021.
[18] Jensen FV. An introduction to Bayesian networks. London: UCL Press; 1996.
[19] Kenett RS. Bayesian networks: Theory, applications and sensitivity issues. Encyclopedia with semantic computing. Singapore: World Scientific press; 2017.
[20] Pietro LD, Mugion RG, Musella F, Renzi MF, Vicard P. Reconciling internal and external performance in a holistic approach: A Bayesian network model in higher education. Expert Syst Appl 2015;42:2691–702.
[21] Tarantola C, Vicard P, Ntzoufras I. Monitoring and improving Greek banking services using Bayesian networks: An analysis of mystery shopping data. Expert Syst Appl 2012;39:10103–11.
[22] Dalla Valle L, Kenett R. Social media big data integration: a new approach based on calibration. Expert Syst Appl 2018;111:76–90.
[23] Marella D, Vicard P. Object-oriented bayesian networks for modelling the respondent measurement error. Comm Statist Theory Methods 2013;42(19):3463–77.
[24] Spirtes P, Glymour C, Scheines R. Causation, prediction, and search. Springer verlag lectures in statistics, 1993.
[25] Breiman L. Random forests. Mach Learn 2001;45:5–32.
[26] James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning with applications in R. New York: Springer; 2017.
[27] Genuer R, Poggi JM. Random forests with R. 1st ed. Cham: Springer; 2020.
[28] Liaw A, Wiener M. Package 'RandomForest'. Breiman and Cutler's random forests for classification and regression. R package version 4.6-14. 2018.
[29] Wasserman S, Faust K. Social network analysis: Methods and applications. Cambridge University Press; 1996.
[30] Freeman LC. Centrality in social networks: Conceptual clarification. Social Networks 1979;1:215–39.
[31] Kousik D, Sovan S, Madhumangal P. Study on centrality measures in social networks: a survey. Soc Netw Anal Min 2018;8:1–11.
[32] Opsahla T, Agneessensb F, Skvoretzc J. Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks 2010;32:245–51.
[33] Bavelas A. A mathematical model for group structures. Appl Anthropol 1948;7:16–30.
[34] Sabidussi G. The centrality index of a graph. Psychometrika 1966;31:581–603.
[35] Freeman LC. A set of measures of centrality based on betweenness. Sociometry 1977;40:35–41.
[36] Shaw ME. Group structure and the behavior of individuals in small groups. J Psychol 1954;38:139–49.
[37] Epskamp S, Costantini G, Haslbeck J, Isvoranu A, Cramer AOJ, Waldorp LJ, et al. Qgraph: Graph plotting methods, psychometric data visualization and graphical model estimation. R package version 1.6.9. 2021.
[38] Facchinetti S, Giudici P, Osmetti SA. Cyber risk measurement with ordinal data. Stat Methods Appl 2019. http://dx.doi.org/10.1007/s10260-019-00470-0.
[39] Facchinetti S, Osmetti SA. A risk index for ordinal variables and its statistical properties: a priority of intervention indicator in quality control framework. Qual Reliab Eng Int 2018;34:265–75.
[40] Dalege J, Borsboom D, van Harreveld F, van der Maas H. Network analysis on attitudes: A brief tutorial. Soc Psychol Pers Sci 2017;8(5):528–37.

Silvia Facchinetti holds a degree in Statistics from Università Cattolica del Sacro Cuore, Milano, Italy, and a Ph.D. in Statistics from University of Milano-Bicocca, Italy. She is a researcher in the Department of Statistical Science, Università Cattolica del Sacro Cuore, Milano, Italy. She is a member of the Italian Statistical Society and author of publications in the area of methodological and applied statistics: nonparametric statistics, statistical quality control, and statistical models for operational risk. Silvia teaches undergraduate courses in statistics and statistical methods for finance and insurance. She has taken part in various research programs, both national and international.

Silvia Angela Osmetti holds a degree in Statistics from Università Cattolica del Sacro Cuore, Milano, Italy, and a Ph.D. in Statistics from University of Milano-Bicocca, Italy. She is associate professor of Statistics at the Department of Statistical Science, Università Cattolica del Sacro Cuore, Milano, Italy. She has taken part in various research programs, both national and international. She is a member of the Italian Statistical Society and author of publications in the area of methodological and applied statistics: statistical methods for ordinal data, copulas and pair copulas, optimal design of experiment, and statistical models for operational risk.

Claudia Tarantola is associate professor of Statistics at the Department of Economics and Management of the University of Pavia, Italy. Her research interests revolve mainly around multivariate statistical models for categorical data (in a frequentist and Bayesian setting) and in particular to graphical models. Her research interests also include Markov Chain Monte Carlo methods and financial risk models. She has taken part in various research programs, both national and international. From May 2018 until May 2022 she was the President of the Hermes University Network (http://hermes-universities.eu/). She is member of the Managing Committee of the Cost action "FinTech and Artificial Intelligence in Finance-Towards a Transparent Financial Industry (CA19130).