RESEARCH ARTICLE

MAGIC: A Method for Assessing Cyber Incidents Occurrence

MASSIMO BATTAGLIONI[ID], **(Member, IEEE), GIULIA RAFAIANI**[ID], **(Student Member, IEEE),**
FRANCO CHIARALUCE[ID], **(Senior Member, IEEE), AND MARCO BALDI**[ID], **(Senior Member, IEEE)**
Department of Information Engineering, Università Politecnica delle Marche, 60131 Ancona, Italy

Corresponding author: Massimo Battaglioni (m.battaglioni@univpm.it)

ABSTRACT The assessment of cyber risk plays a crucial role for cybersecurity management, and has become a compulsory task for certain types of companies and organizations. This makes the demand for reliable cyber risk assessment tools continuously increasing, especially concerning quantitative tools based on statistical approaches. Probabilistic cyber risk assessment methods, however, follow the general paradigm of probabilistic risk assessment, which requires the magnitude and the likelihood of incidents as inputs. Unfortunately, for cyber incidents, the likelihood of occurrence is hard to estimate based on historical and publicly available data; so, expert evaluations are commonly used, which however leave space to subjectivity. In this paper, we propose a novel probabilistic model, called MAGIC (Method for AssessinG cyber Incidents oCcurrence), to compute the likelihood of occurrence of a cyber incident, based on the evaluation of the cyber posture of the target organization. This allows deriving tailor-made inputs for probabilistic risk assessment methods, like HTMA (How To Measure Anything in cybersecurity risk), FAIR (Factor Analysis of Information Risk) and others, thus considerably reducing the margin of subjectivity in the assessment of cyber risk. We corroborate our approach through a qualitative and a quantitative comparison with several existing methods.

INDEX TERMS Cyber incident, cyber risk, FAIR, HTMA, probabilistic risk assessment.

I. INTRODUCTION

The massive exploitation of data and information systems in companies and organizations is motivating an increasing attention to cybersecurity and its management. One of the main pillars upon which cybersecurity management relies is cyber risk assessment, for which a plethora of standards and models exist. Risk does not have a unique definition, but according to the NIST (National Institute of Standards and Technology) [1], it is *a measure of the extent to which an entity is threatened by a potential circumstance or event, and is typically a function of: (i) the adverse impacts that would arise if the circumstance or event occurs; and (ii) the likelihood of occurrence*. A universal method for cyber risk assessment does not exist, but international standards provide general guidelines which should be followed when designing

The associate editor coordinating the review of this manuscript and approving it for publication was Fabrizio Marozzo[ID].

a cyber risk assessment method. For example, the ISO (International Standard Organization) unpacks risk assessment into risk identification, analysis and evaluation [2], [3]. In the risk identification phase, the critical services are identified and the threats and vulnerabilities they could face are determined. Risk analysis, instead, is needed to determine the likelihood of occurrence and the impact of these threats [3], [4]. Finally, in the risk evaluation phase the obtained results are compared with some pre-established risk acceptance criteria [2], [3]. In this paper we focus on the second phase: risk analysis and likelihood estimation.

Existing cyber risk assessment approaches can be grouped into quantitative and qualitative methods. In the former ones, the risk analysis phase is carried out through numerical evaluations, mostly based on probability theory, as in classical PRA (Probabilistic Risk Assessment) [5]. Widely speaking, these approaches have some potential advantages, like their robustness, reproducibility, and comparability of results.

While it seems generally reasonable to rely on past events for risk analysis, in the specific case of cyber incidents these data may be unavailable or incomplete and, thus, probabilistic methods must resort to the help of experts, which makes them barely practical in real-case scenarios [1].

Qualitative methods, in turn, exploit a nonnumerical approach and, therefore, they are often simpler to implement and interpret. However, their results are scarcely reproducible and comparable and, above all, they have an intrinsic nature of subjectivity.

A. CONTRIBUTION

To the best of our knowledge, a probabilistic method enabling the computation of the likelihood of occurrence of cyber incidents, based on the posture of the organization rather than on subjective expert estimates, is currently missing. To fill this gap, we propose a method called MAGIC (Method for AssessinG cyber Incidents oCcurrence), which is a probabilistic model to quantitatively estimate the likelihood of occurrence of a cyber incident for a specific organization, starting from qualitative assessment approaches based on questionnaires. Owing to its own nature, MAGIC should be seen as a tool to be used in conjunction with existing probabilistic cyber risk assessment methods, to make their results more reproducible and less subjective.

For this purpose, we consider four indicators representing the *awareness* of the employees, the *maturity*, the *complexity*, and the *attractiveness* of the target organization. These aspects are commonly qualitatively assessed through questionnaires, so we estimate the values of the corresponding indicators starting from questionnaires concerning the target organization, provided by one or more assessors. Then, the four indexes are combined, using a probabilistic quantitative approach, in order to find the likelihood of occurrence of a given type of cyber incident. In particular, we consider two scenarios: in the first scenario, we obtain as output the estimated frequency of occurrence of cyber incidents in a given time period (and the associated likelihood); in the second scenario, we obtain the probability that the organization will face exactly one successful cyber incident in a given time period. Indeed, in the latter case we assume that, after one cyber incident, the organization will change its posture and thus a new assessment needs to be performed.

MAGIC provides inputs that can be used in conjunction with classical PRA methods such as HTMA (How To Measure Anything in cybersecurity risk) [6], FAIR (Factor Analysis of Information Risk) [7], as well as many others. In the former method, it is assumed that the frequency of adverse events follows a log-normal distribution, whose mean and variance is not directly related to the organization posture. Similarly, in the FAIR method, for a given adverse event, an expert is needed to estimate the minimum, maximum and most likely values for its frequency of occurrence. In the case of cyber events, these processes often become subjective, due to the lack of reliable historical data concerning each specific type of cyber incident. Our target is to reduce such a

subjectivity as much as possible, by following and extending the approach in [8], [9].

In short, according to the proposed method, the main effort required from the target organization is to provide information regarding the state of their technological systems and protection measures: expert evaluations are no longer required and the relations among all the components of the considered infrastructure do not need to be figured out. We consider questionnaires based on international cybersecurity standards and frameworks, which are comprehensive and widely recognized, but MAGIC is general and can be applied with other types of questionnaires.

A basic version of our model has been introduced in [8]. Even though the approach proposed in this paper has some similarities with that in [8], since the final goal of both of them is to estimate the likelihood of occurrence of cyber incidents based on the organization posture, there are some profound differences between them. The most important ones are described next:

- the approach in [8] does not take the awareness of the employees into account as a determinant factor for risk assessment, which is instead considered in the model we propose;
- we provide the set of controls to be used for complexity assessment;
- the approach in [8] stands as an independent method for probabilistic assessment, whereas MAGIC is a transversal tool to provide tailor-made inputs for existing probabilistic cyber risk assessment methods; owing to this, we are also able to report a larger number of numerical results;
- in this paper, the final likelihood is computed differently from that in [8], where the probability that the organization faces a certain number of cyber attacks is not considered in the computation. Through the novel approach, instead, the likelihood of a cyber incident can be easily converted into a frequency, comparable to that obtained through expert evaluations in existing methods;
- differently from [8], in this paper we also consider the practical scenario in which the target organization does not immediately realize that a cyber incident has occurred.

B. PAPER ORGANIZATION

The paper is organized as follows. In Section II we provide a high-level description of many related works. In Section III we recall the basic functioning of some probabilistic cyber risk assessment methods. In Section IV the components of the cyber incident occurrence model we propose are discussed. In Section V we show how our approach can be combined with known probabilistic methods in order to overcome their main limitations when dealing with cyber risks. In Section VI we present some numerical results, aimed at validating the effectiveness of the proposed approach. In Section VII we provide both a quantitative and a qualitative comparison of

our method with other approaches. Finally, Section VIII concludes the paper.

II. RELATED WORKS

A plethora of approaches for cyber risk assessment can be found in existing literature. Besides those proposed by national and international organizations (like the mentioned ISO/IEC (International Standard Organization/International Electrotechnical Commission) 27005:2018 [3] and NIST SP (Special Publication) 800-30 [1]), others have been introduced by public and private organizations, like EBIOS (Expression des Besoins et Identification des Objectifs de Sécurité) [10], CRAMM (Central computer and telecommunications agency Risk Analysis and Management Method) [11], OCTAVE (Operationally Critical Threat, Asset, and Vulnerability Evaluation) [12], MEHARI (MEthod for Harmonized Analysis of RIsk) [13], MAGERIT (Risk Analysis and Management Methodology for Information Systems) [14], IRAM2 (Information Risk AssessMent version 2) [15], IT-Grundschutz [16], and CORAS [17]. An extensive and critical literature review can be found in [18], [19]. A lot of attention has been devoted to solving the problem of estimating the likelihood of occurrence of a threat and the corresponding impact. For example, several methods have been proposed using different techniques like Bayesian networks [20], attack path graphs [21], fuzzy logic [22], probabilistic model checking [23], vulnerability assessments [24], Monte Carlo simulations [6], [7], and others.

Next we provide a brief description of some of the aforementioned methods, highlighting the differences with MAGIC as well as the possible common aspects. A deeper comparison is then carried out in Section VII, after the description of our method.

- **ISO/IEC 27005:2018** [3]: is an international framework for managing information risks. It describes all processes for risk management, including risk assessment. In this case, the estimation of the likelihood can be performed in a qualitative, quantitative, or hybrid way. However, the ISO/IEC 27005 standard only provides guidelines for doing it, without describing any specific practical method.
- **NIST SP 800-30** [1]: is a guide for conducting risk assessment. In order to determine the likelihood of occurrence of a security incident, this method identifies all the potential vulnerabilities and the probability of their exploitation. The likelihood is then described using a qualitative or semi-quantitative scale. Such an approach is opposed to the numerical one we propose, since it leaves space to subjectivity.
- **EBIOS** [10]: is a scenario-based approach for risk management that relies on the establishment of a strong link among different stakeholders. It uses a modular approach for identifying risk causes. However, all the phases considered in this method, including risk assessment, are the result of security debates among the team and, therefore, they are subjective.

- **CRAMM** [11]: is a qualitative method for risk assessment. It measures risks as the product of asset, threat, and vulnerability values. It uses trained experts. Threats and vulnerabilities are not exhaustively assessed, but the assessor can choose among different predefined threat/asset and threat/impact combinations. Relying on structured questionnaires, and/or on the expertise of the assessor, the method determines the likelihood of threats and vulnerabilities; however, differently from MAGIC, it does not directly link those likelihoods with the likelihood of occurrence of a cyber incident.
- **OCTAVE** [12]: is an asset-driven method for assessing information security risks. While the original approach is designed for large organizations, the OCTAVE-S [25] and the more recent OCTAVE-Allegro [26] versions can be also applied to small and medium enterprises. This method firstly identifies all the assets, and then focuses on the critical ones. For each of them, it determines the related threats, and qualitatively labels their likelihood of occurrence as "Low", "Medium" or "High".
- **MEHARI** [13]: is a risk management model with the aim of helping the implementation of ISO/IEC 27005. It performs the risk assessment phase through an audit. This phase includes the identifications of assets, threats and vulnerabilities. The likelihood of occurrence of a threat is qualitatively described using a four levels scale.
- **MAGERIT** [14]: is an asset-oriented risk management model. Using this model, security professionals evaluate the assets and all the possible threats. Then, they describe the likelihoods of occurrence of those threats using a numerical scale with a limited number of levels. The likelihood is usually evaluated relying on the annual rate of occurrence of any specific threat for each specific asset, which is substantially estimated through historical data.
- **IRAM2** [15]: is a qualitative threat-driven method for information risk assessment and treatment. The likelihood of success of a threat is estimated using lookup tables having as input two numerical values describing the strength of the threat and that of the security controls implemented by the organization. Then, using this method, the residual likelihood is evaluated (also, through a lookup table) as a combination of likelihood of success and likelihood that an attacker will try to cause an incident based on the considered threat.
- **IT-Grundschutz** [16]: is a qualitative method for identifying and assessing security incidents. Using this method, some qualified staff has to identify all possible threats. Then, for each of them this method evaluates the frequency of occurrence using a qualitative scale. This frequency is finally combined with the impact through a risk matrix.
- **CORAS** [17]: is a method for security risk analysis. All the risk assessment phases are performed through structured brainstorming, where people with different backgrounds and competences collaborate to identify

assets, vulnerabilities, threats, and the related likelihoods of occurrence. The likelihood assessment can be done both in a qualitative or a quantitative way; however, the evaluation is strictly related to the subjectivity of the assessors.

- **LiSRA** [21]: is a risk assessment method taking a bidimensional input; on one dimension domain-specific information is required from an expert, while the other dimension is filled by the user, according to the practices of the considered organization. Then, the risk is conventionally obtained as the combination of the probability to have successful attacks and their impact. The probability of success is computed based on attack trees but, differently from our method, the latter need to be entirely computed by external experts. Therefore, MAGIC might be seen, with some adaptations, as an alternative method allowing to bypass the need of experts, rather than a completely different approach.

- **Risk analysis based on fuzzy decision theory** [22]: the first step of this approach is to identify an expert; then, a taxonomy of events and scenarios has to be defined (second step). Finally, the expert builds a matrix with potential accidents on the rows and possible scenarios on the columns: each entry of the matrix has to be filled with a probability that the accident takes place in a certain scenario. Also in this case, differently from our method, external expert estimates are used. Starting from this matrix, the fuzzy decision theory is then applied, which returns the expected value of the considered option. Also in this case, MAGIC is not in contrast with this method, but can be seen as a variation of it. In fact, in MAGIC the a-priori probabilities can be numerically derived from the posture of the organization, rather than estimated by an expert.

- **CVSS-based risk assessment**: the CVSS (Common Vulnerability Scoring System) [27] gives scores to threats exploiting vulnerabilities on the basis of three categories. One can combine this approach with attack graphs [24] in order to derive the vulnerabilities starting from known threat sources, or to Bayesian decision networks [20], in which the attacks are modeled starting from correlated alerts. The probabilities that the considered threats exploit certain vulnerabilities are computed from the scores associated to each of them, according to CVSS. In particular, CVSS 2.0 provides for the use of three types of metrics: base, temporal and environmental. Base metrics represent the intrinsic characteristics of threats exploiting vulnerabilities that are constant over time and user environments; temporal metrics represent the features of threats exploiting vulnerabilities that change over time but not over user environments; environmental metrics are based on the characteristics of threats exploiting vulnerabilities that are unique to a particular user's environment. These methods may be directly comparable to MAGIC, in that the probabilities derive from numerical assessments, even though

a subjective component is still required. In particular, in the basic version of CVSS-based risk assessment methods, the probability of occurrence of the *i*-th item in the CVE (Common Vulnerabilities and Exposures) list can be computed as a product of some CVSS metrics, *i.e.,*

$$L_i = \text{AV} \times \text{AC} \times \text{Au} \times \text{E} \times \text{RC},$$

where AV is the access vector metric, AC is the access complexity metric, Au is the authentication metric, E is the exploitability metric and RC is the report confidence metric. The exact numerical values of all these metrics need to be chosen in a range of prefixed values, which makes also this numerical approach subjective.

There are several reasons that make many of the aforementioned approaches difficult to apply in real case scenarios. In fact, as explained above, they make risk assessment result in a long process requiring the availability of a significant amount of data. Moreover, the application of these methods often requires the help of an expert assessor external to the organization and able to provide quantitative measures of the cyber risk, which makes them cumbersome and exposed to subjectivity. MAGIC, instead, allows computing quantitative parameters starting from simple questionnaires, which makes the risk assessment process straightforward and less exposed to subjectivity.

III. PROBABILISTIC CYBER RISK ASSESSMENT: PRELIMINARIES

In this section we describe two state-of-the-art probabilistic cyber risk assessment methods based on Monte Carlo simulations, *i.e.,* HTMA [6] and FAIR [7].

A. PROBABILISTIC RISK ASSESSMENT USING HTMA

In a nutshell, the HTMA method for cyber risk assessment [6] is based on the following steps:

- definition of the potential cyber threats;
- estimation of the likelihood of occurrence and impact of each event;
- Monte Carlo simulation for generating the scenarios;
- results interpretation.

While the likelihood of occurrence of each threat is given by a single value, the impact is associated to a 90% confidence interval, identified by a lower and an upper bound. All these three numerical values are to be determined by external experts.

The Monte Carlo simulation, in each scenario, works as follows:

1) for any threat *i*, a real number *r* is generated by sampling uniformly at random the range between 0 and 1, boundaries included, denoted as [0, 1]. If $r < L_i$, where L_i is the likelihood of the *i*-th threat, it is assumed that the event has happened, and vice versa;

2) the impact of events which did not occur is 0, whereas the impact of occurred events is obtained by randomly sampling a log-normal distribution of the impacts,

obtained according to the boundaries of the given 90% confidence interval;

3) all the impacts of occurred events are summed, in order to obtain an estimate of the total annual risk.

The results obtained with the Monte Carlo simulation are used to construct the LEC (Loss Exceedance Curve), which corresponds to the graphical representation of the complementary cumulative distribution function of the annualized loss expectancy.

B. PROBABILISTIC RISK ASSESSMENT USING FAIR

The FAIR methodology [7] can be described through the following four steps:

- definition of the scenario under exam and its decomposition into sub-scenarios;
- estimation of the parameters for any sub-scenario;
- generation of the frameworks through Monte Carlo simulations;
- results interpretation.

More generally, FAIR defines an ontology for the risk, which is summarized in Fig. 1. The ontology describes how the assessment of risk can be obtained; in particular, it considers all the risk factors that contribute to the evaluation of the risk and all the relationships between them. Risk factors can be measured and estimated; then, it is possible to calculate risk using the mathematical expressions of the relationships among factors. In essence, the risk is computed as a combination of LEF (Loss Event Frequency) and Loss Magnitude. In order to facilitate the process of risk evaluation, these two factors can be individually decomposed in other factors that, in their turn, could be further decomposed, as well. This way, the user can assess the risk considering a specific layer of the ontology, according to which factors he is able to estimate. The LEF depends on several factors; the most relevant ones for our model are redefined next: the TEF (Threat Event Frequency) is defined as the frequency with which, in a given time period, the attacker tries to breach the organization; the Vulnerability is the probability of success of any of these breaches. The Loss Magnitude, instead, is the sum of the losses caused by a certain primary threat event and of the losses caused by its side-effects (such as, for example, the reaction of secondary stakeholders), which are included in the concept of secondary risk.

Conventionally, the outputs of the FAIR approach are scatterplots with LEF and Loss Magnitude on the *x*-axis and *y*-axis, respectively, and/or tables summarizing various results of the Monte Carlo simulation.

IV. CYBER INCIDENT MODEL

In this section we introduce the quantitative approach we propose for estimating the likelihood of occurrence of a cyber incident, called MAGIC. First of all, we provide some basic definitions:

- **Cyber threat** [28]: any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation),

organizational assets, or individuals through an information system via unauthorized access, destruction, disclosure, modification of information, and/or denial of service. We will simply refer to cyber threats as "threats" in the following.

- **Cyber attack**: any realized attempt to partially or totally disclose, expose and/or compromise data by a malicious entity. We will simply refer to cyber attacks as "attacks" in the following.
- **Cyber incident**: any cyber attack which had success.

The proposed model, with all the key parameters, which we will discuss next, is schematized in Fig. 2. We need to distinguish between two types of threats: those coming from external threat agents and those happening as a consequence of human misbehavior, *i.e.*, non-malicious threats. In Fig. 2, the dashed arrows are referred to the former scenario, the dotted arrows describe the latter, whereas the solid arrows are valid for both scenarios. The model components in black are described in detail in the following subsections, while those in blue are not object of our theoretical analysis, but they are taken into account in Section VI, where numerical results are provided.

A. AWARENESS OF THE EMPLOYEES

The awareness of the employees can be defined as their level of consciousness about the cybersecurity risks. It can be directly related to the training programs the organization supplies to its workers.

In order to practically evaluate the awareness, it is possible to consider a list of best practices devoted to this issue, as that in the NIST SP 800-53 [28], or to rely on specific subsets of controls proposed in cybersecurity frameworks; for example, Control 14 in the CIS (Center for Internet Security) Controls [29] deals with security awareness and training programs. The evaluation of the awareness parameter can be carried out in different ways. For example, the assessor can determine if every identified control is fully implemented by the organization or not, simply associating to it a *Yes* or *No* answer. Another approach can be to use a scale in order to determine at which degree each considered control is implemented and, therefore, assigning better ratings as the level of implementation completeness of the control increases. With both these approaches, an *N/A* option needs to be included; this should be used when one or more controls are considered to be not applicable in the context under examination. Then, according to the results of the assessment, a numerical score should be assigned to every control. For example, using the first of the approaches stated above, a score equal to 1 may be assigned to all the controls with *Yes* as an answer, while 0 may be assigned to the controls answered with *No*. Instead, when using the second approach, one may consider as many scores as the number of the possible implementation levels. In other words, if the scale used for the implementation evaluation has 5 possible levels, the score that can be assigned to each control can be, for example, a number between 0 and 4. But, obviously, other choices are possible. In the following,
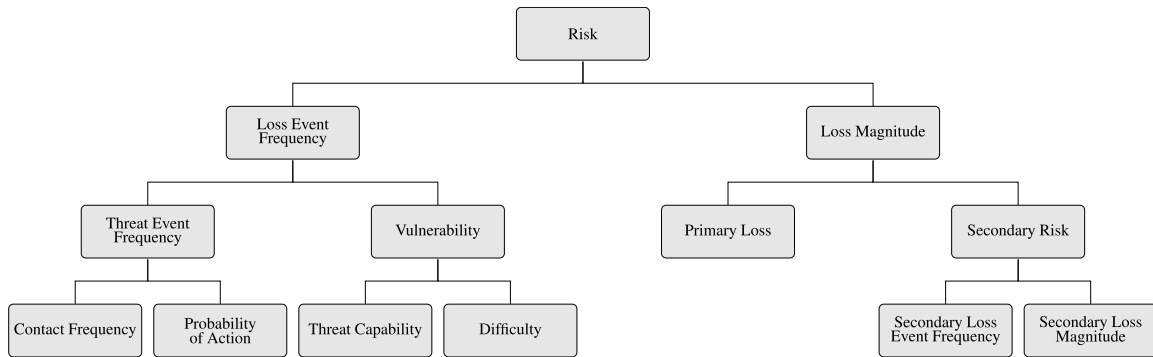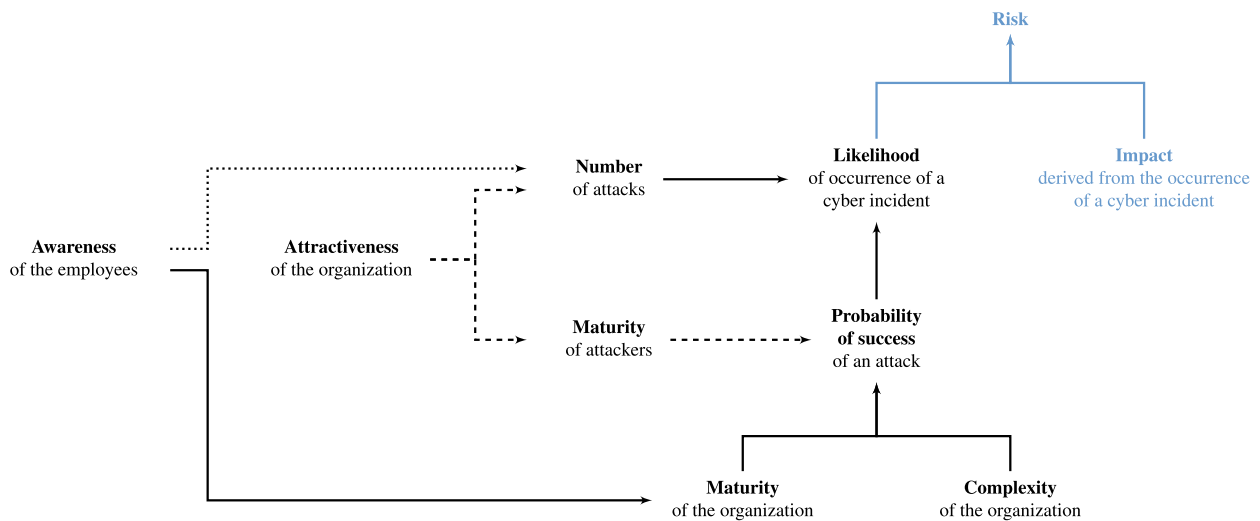
FIGURE 1. Ontology of the FAIR model.

FIGURE 2. Relations among the components of MAGIC.

we denote by s_{max} the maximum score value (so, $s_{max} = 1$ or $s_{max} = 4$ in the mentioned examples). Moreover, a weight may be assigned to every control. The weights are given mainly according to the aspects that the organization wants to stress out, so they are not mandatory. In fact, an organization may want to focus its analysis on some specific training strategies, while it may not be interested in other ones; in this case, it can assign higher (or lower) weights to the controls that it considers more crucial (or less relevant). Finally, a weighted average is calculated in order to obtain a final value for the awareness. Note that the *N/A* controls should not be counted in the average.

In this paper, we consider the awareness index as a number between 0 and 10. Let E be the total number of considered controls, $s_i \in [0, s_{max}]$ and $a_i \geq 0$ the score and the weight associated to the i-th control, respectively; the awareness index is computed as

$$\text{Awareness Index} = \text{AI} = \frac{\sum_{i=1}^{E} s_i \times a_i}{\sum_{i=1}^{E} a_i} \times \frac{10}{s_{max}}. \quad (1)$$

Notice that, as shown in Fig. 2, the awareness directly influences the maturity of the organization, formally defined

in Section IV-B. This holds for both malicious and non-malicious threats. Moreover, when specifically dealing with non-malicious threats, the awareness is inversely related to the number of threats. In other words, if employees are well-trained about cyber risks, we can assume that the organization will suffer fewer potential threats, since the employees are less prone to causing cyber incidents.

B. MATURITY OF THE ORGANIZATION
The maturity of an organization can be defined as the level of implementation of all practices and procedures that the organization executes in order to reduce the risk of receiving cyber attacks which may cause security breaches, data leakage, denial of service, and so on. As mentioned in Section IV-A, awareness concurs to determine maturity (being one of its most relevant components) but other issues need to be taken into account, too. The evaluation of this further part of the maturity can be done by assessing the organization compliance to the security controls proposed by one or more cybersecurity frameworks. The frameworks chosen as the reference ones will influence the area of application and, therefore, the kind of risk the organization is going

to evaluate. An organization, in fact, may be interested in assessing the risk of suffering breaches, or the risk of losing, as a consequence of an attack, data confidentiality and/or integrity and/or availability, or may be interested in assessing the risk related to a combination of them both. The security frameworks that can be used as reference are, for example, the one proposed by CIS in [29] for assessing cybersecurity risk, or the set of controls introduced by the ENISA (European Union Agency for Cybersecurity) in [30] for assessing data protection risk, or the NIST Cybersecurity Framework [31] for assessing both cybersecurity and data protection risks. Therefore, first of all, the organization should choose the kind of risk it intends to assess and, according to the choice, select one or multiple appropriate security frameworks as reference. The next step is to evaluate if and at which level all the security controls listed in the reference framework are implemented within the organization. Using the controls of cybersecurity frameworks to assess the maturity will result in an accurate picture of the actual cyber posture of the organization. In fact, since the security frameworks are usually considered as a set of best practices, their use will lead to a robust and complete mapping of the area of interest. Moreover, since the security frameworks are continuously updated, the maturity can be dynamically assessed, simply repeating the evaluation as soon as new versions of the frameworks are released.

Because of the need to combine the mentioned different aspects, the maturity index evaluation is performed through a two-step procedure. The first step is equivalent to that described in the previous section for the evaluation of the awareness index. Clearly, the controls of the chosen framework about cybersecurity awareness and training programs must be excluded from the list used for this part of the evaluation. Thus, after answering to all the remaining controls, a score is associated to any answer. The *N/A* option should be included as well. Also in this case, it is possible to assign a weight to each control, according to its relevance for the kind of assessment the organization wants to perform. Let T be the total number of considered controls, $s'_i \in [0, s'_{\max}]$ and $a'_i \geq 0$ respectively the score and the weight associated to the i-th control; a weighted average is computed as follows

$$M = \frac{\sum_{i=1}^{T} s'_i \times a'_i}{\sum_{i=1}^{T} a'_i} \times \frac{10}{s'_{\max}}. \quad (2)$$

Then, in the second step of the procedure, the maturity index is eventually obtained as a weighted average of the awareness index and M. In this case, AI and M are weighted by $E/(E+T)$ and $T/(E+T)$, respectively. We finally obtain

$$\text{Maturity Index} = \frac{\text{AI} \times E + \text{M} \times T}{E + T}. \quad (3)$$

We observe that, according to this definition, the maturity index is a real number ranging between 0 and 10.

Note that the maturity of the organization will directly influence the probability of success of an attack. In fact, a higher maturity means that the organization has addressed more attention to cybersecurity practices and, therefore, the probability for an attack attempt to be successful will decrease, and vice versa.

C. COMPLEXITY OF THE ORGANIZATION

The complexity of an organization can be defined as the measurement of the intricacy of its technological infrastructure and of how the processes, the activities, and the services are managed. The concept that the risk does not only depend on the maturity of the organization, but also on its complexity is introduced in [29], where three IGs (Implementation Groups) are defined. Following that approach, every organization should identify itself, mainly according to its dimension, in one of the proposed IGs and, therefore, should implement a specific subset of security controls. This is due to the fact that smaller organizations are usually expected to be less exposed to threats, when compared to larger organizations. However, the dimension should not be considered as the only discriminating parameter. For this reason, starting from the controls proposed in [32] for the evaluation of the inherent risk, we have identified a set of punctual and specific controls useful to assess the inherent complexity of an organization.[1] Coherent with the above considerations, the controls we address do not consider only the dimension of the organization, but they try to identify all possible critical points of hardware, software, networks, and facilities. More precisely, we have grouped all the considered controls into five categories: Networks and Infrastructure, IP (Internet Protocol) network technologies, Applications, Services and IT (Information Technology) department. Therefore, beside the number of employees, the assessment includes the number and the characteristics of the components (physical and software systems) and their interconnections, the number of services and their characteristics, and the entropy of the IT system management. In order to facilitate the assessment, for each control we have identified five possible guided answers. The answers are associated to *Very Low*, *Low*, *Medium*, *High*, and *Very High* complexity.

In order to practically assess the complexity, the process is equivalent to the maturity assessment. After evaluating all the controls, a score is associated to every evaluation. The *N/A* option should be included as well. Also in this case, it is possible to assign a weight to each control, according to its relevance for the kind of assessment the organization wants to perform. Then, a weighted average is calculated in order to obtain a complexity index for each one of the five categories considered above. Similarly to the previous indexes, we consider the complexity index as a number between 0 and 10. Let C_j be the number of controls included in the category j, $s''_{i,j} \in [0, s''_{j,\max}]$ and $a''_{i,j} \geq 0$ respectively the score and the weight associated to the i-th control of the j-th category; the

---

[1] The set of controls considered in this paper for the complexity assessment can be found at https://github.com/secomms/cyber-risk-assessment.

complexity index of the category j is computed as

$$\text{CI}(j) = \frac{\sum_{i=1}^{C_j} s_{i,j}'' \times a_{i,j}''}{\sum_{i=1}^{C_j} a_{i,j}''} \times \frac{10}{s_{j,\max}''}. \quad (4)$$

Finally, a global *complexity index* for the organization is computed as the weighted average of the five complexity indexes computed before; in this case, the weight associated to each category is simply computed as the number of controls included in that category divided by the total number of controls, i.e., $b_j = C_j/(\sum_{j=1}^{5} C_j)$. Then, we have

$$\text{Complexity Index} = \sum_{j=1}^{5} (\text{CI}(j) \times b_j). \quad (5)$$

We observe that the same approach, with a weighted average for each category and a final weighted average for the index computation, could be used for the maturity assessment when the security controls chosen for the evaluation are divided into categories.

Note that the complexity of the organization will inversely influence the probability of success of an attack. In fact, a higher complexity means that the technological infrastructure of the organization is more intricate and, therefore, it is more prone to being successfully targeted. In other words, the probability for an attack attempt to be successful will increase for increasing complexities, and vice versa, according to the law that will be described in Section IV-F.

D. ATTRACTIVENESS OF THE ORGANIZATION

The attractiveness of an organization can be defined as the level of interest the organization causes in potential attackers. The attractiveness depends on several factors, as the type of business, the kind and the amount of data the organization manages, etc. The insertion of this parameter in our model is due to the assumption that cyber criminals will likely attack organizations from which they can obtain larger profits. The dimension of the organization does not affect its attractiveness. In fact, for example, small organizations may operate in critical environments and/or may process a large amount of sensitive data, making them more attractive to cyber criminals if compared to larger organizations with a limited technological value.

In order to practically estimate the attractiveness, we have examined the data proposed in cybersecurity reports like [33]. In particular, we have considered the number of attacks any type of organization has suffered in one year, with respect to the total number of attacks analyzed in the report. According to the given data, we have classified the type of organization as reported in Table 1.

Therefore, during the assessment, the organization should simply identify itself in one of the proposed types of business, and the attractiveness will be assigned consequently.

Note that the attractiveness will influence both the number of attacks and the maturity of attackers. In fact, we assume that a highly attractive organization will likely suffer more

TABLE 1. Classification of the organizations as a function of the percentage π of attacks received with respect to the total number of attacks.

Percentage of attacks	Type of organization
$\pi < 1.25\%$	*Very lowly* attractive
$1.25\% \leq \pi < 2.5\%$	*Lowly* attractive
$2.5\% \leq \pi < 5\%$	*Averagely* attractive
$5\% \leq \pi < 10\%$	*Highly* attractive
$\pi \geq 10\%$	*Very highly* attractive

and more structured attacks if compared to a lowly attractive organization.

E. MATURITY OF ATTACKERS

Given the definition of cyber attack proposed at the beginning of this section, let us devise the attacker model. As mentioned above, we should distinguish between attacks coming from malicious attackers, and threats deriving from the lack of awareness of the employees. For the probabilistic model we devise, however, there is no need to mathematically distinguish between these cases. Thus, for the sake of simplicity, in the rest of the paper we will also refer to non-malicious threats as "attacks", keeping in mind that they come from non-malicious employees, who assume the role of unaware attackers in the model.

In general, the organization might be targeted by multiple attackers, but we assume that they cannot conduct more than one attack in the same time slot Δt. The value Δt can be chosen sufficiently small, so that attacks performed in close periods of time can be distinguished. Moreover, under this assumption, the identity of the attacker does not play a relevant role and, therefore, we will generically refer to "the attacker". Another assumption that we make, in order to keep the analysis feasible, is that different attack attempts are not correlated. In other words, each attack attempt does not depend on previous attack attempts, and its outcome does not influence future attack attempts. Clearly, this hypothesis may not always be verified and, in the scenario where the assumptions are too optimistic, our model provides a lower bound to the likelihood of the adverse events, rather than an estimate.

As shown in Fig. 2, the attractiveness is related to the maturity of the attackers: attractive organizations will more likely face more structured attacks, and vice versa. The maturity of the attackers, in its turn, influences the probability of success of an attack, described in the next section.

F. PROBABILITY OF SUCCESS OF AN ATTACK

Obviously (and luckily) attacks are not always successful and, therefore, more than one attempt may be needed before breaking through the organization defenses. The single attack attempt is thus associated to a probability of success. In particular, the computation of the probability of success of a single attack takes into account the three key indexes described above. Considering the maturity of the organization as a variable x, we need a function that decreases when x increases,

since we expect more mature organizations to be more resistant to cyber attacks, and vice versa. Furthermore, we do not expect this trend to be linear, since slight improvements of the maturity of a very immature organization may not be sufficient to significantly decrease the probability of success of an attack and, similarly, very mature organizations should not need to further improve significantly their posture, since the probability of an attacker breaching them may already be low enough. A function that may fit well this scenario is the logistic function [34]

$$f(x) = \frac{K}{1 + e^{-B(x-t)}}, \quad (6)$$

where K is the *saturation level*, i.e., the upper horizontal asymptote that limits the curve's maximum value; t is the *midpoint*, i.e., the value of x corresponding to half the saturation level and B is the *growth rate*, i.e., the steepness of the curve. In order to obtain the aforementioned trend, B must be chosen negative. According to (6), $f(x)$ has a lower horizontal asymptote equal to 0.

It is possible to extend (6), by considering a lower asymptote different from 0 and a non-symmetric shape, thus obtaining the so-called *generalized logistic function* [35]

$$f(x) = A + \frac{K - A}{(1 + Q \times e^{-B(x-x_0)})^{1/\nu}}, \quad (7)$$

where A is now the *lower asymptote*, Q is a variable, related to $f(0)$, that influences the inflection point, $f(x_0) = A + \frac{K-A}{(1+Q)^{1/\nu}}$ and $\nu > 0$ determines the asymmetry of the curve. Notice that, by choosing $Q = \nu = 1$, x_0 corresponds to the point at which the curve is at its midpoint and has the maximum slope.

By using the generalized logistic function, with $Q = \nu = 1$, to express the probability of success of a single attack $P_s(x)$, (sometimes P_s, for notation simplicity, in the following) we have

$$P_s(x) = A + \frac{K - A}{1 + e^{-B(x-x_0)}}. \quad (8)$$

By increasing x_0, $P_s(x)$ shifts toward right. So, we associate the value of the complexity of the organization to x_0, to take into account that, for the same value of x, more complex organizations are expected to face attacks with a larger probability of success.

According to the discussion in Section IV-B, the maturity x takes values in the range from 0 to 10. Moreover, we assume that the probability of success cannot reach the extreme and thus unrealistic values of 1 and 0. So, we set the maximum value U at $x = 0$ and the minimum value L at $x = 10$. Consequently, K and A depend on x_0 and can be easily obtained by solving the system with $f(0) = U$ and $f(10) = L$. An example of curves for different values of x_0 is shown in Fig. 3.

Indeed, as mentioned above, a more significant value of the probability of success can be obtained by taking into account the maturity of the attackers, the latter playing the role of
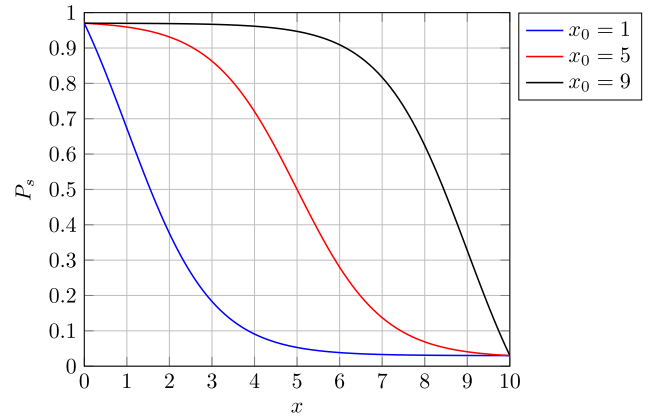

FIGURE 3. Probability of success of a single attack (P_S) for different values of x_0. $B = -1$, $U = 0.97$, $L = 0.03$.

weight coefficient $0 < w \leq 1$. Explicitly this means that the final value of the probability of success of a single attack results in

$$p^* = p_s(x) = wP_s(x). \quad (9)$$

A possible choice for the values of w is given in Table 2, showing how more attractive organizations suffer attacks which are more likely to be successful because of the higher maturity of the attackers. In this case, we assume that a linear law describes well the expected behaviour. We observe that, in case of non-malicious threats, w is always equal to 1, since attractiveness does not play any role.

TABLE 2. Possible values of w as a function of the attractiveness of the organization.

Attractiveness	Very Low	Low	Medium	High	Very High
w	0.6	0.7	0.8	0.9	1

G. NUMBER OF ATTACKS

Let us consider a time period containing t time slots, each of duration Δt. An attacker (malicious or non-malicious, according to the discussion in Section IV-F) may perform an attack attempt or not in any time slot, with a certain probability. Therefore, at most t attack attempts will be suffered by the organization in the considered time period. Let us define two discrete random variables:

- $\mathcal{A}$: represents the outcome of the attacker choice in a time slot. Either he performs an attack attempt ($\mathcal{A} = 1$), or he does not ($\mathcal{A} = 0$). Therefore, $\mathcal{A}$ is a binary random variable and follows a Bernoulli distribution.
- $\mathcal{N}$: represents the number of attack attempts suffered by the organization in the considered time period. Its realization n respects the condition $0 \leq n \leq t$.

We assume that $\mathcal{N}$ follows a binomial distribution.[2] The PMF (probability mass function) associated to the binomial

---

[2]An alternative, which might be considered when n_{avg} is significantly smaller than t, is the Poisson distribution, whose probability mass function is $\Pr(\mathcal{N} = n) = \frac{\lambda^n}{n!} e^{-\lambda}$, where $\lambda = n_{avg}$.

distribution is therefore

$$\Pr(\mathcal{N} = n) = \binom{t}{n} \Pr(\mathcal{A} = 1)^n \Pr(\mathcal{A} = 0)^{(t-n)}$$

$$= \binom{t}{n} \left(\frac{n_{\text{avg}}}{t}\right)^n \left(1 - \frac{n_{\text{avg}}}{t}\right)^{(t-n)}. \quad (10)$$

In (10) we have denoted by n_{avg} the average number of attack attempts suffered by the organization in the considered time period, so that, reminding the properties of the binomial distribution, $\Pr(\mathcal{A} = 1) = \frac{n_{\text{avg}}}{t} = 1 - \Pr(\mathcal{A} = 0)$. A possible choice is to pick n_{avg} as the number of attack attempts suffered by the organization in a previous time period having the same length of the considered one.

Then, wishing to calculate the probability that the organization suffers at most (or at least) n attack attempts, we have

$$\Pr(\mathcal{N} \leq n) = \sum_{k=0}^{n} \Pr(\mathcal{N} = n)$$

$$\Pr(\mathcal{N} \geq n) = \sum_{k=n}^{t} \Pr(\mathcal{N} = n),$$

respectively.

H. LIKELIHOOD OF OCCURRENCE OF CYBER INCIDENTS

In the previous section, we have analyzed the probability that the organization suffers a certain number of attack attempts. Even though this represents an interesting information for the stakeholders, it is more relevant to estimate the likelihood of these attacks being successful. In fact, in the definition of risk [1] the likelihood of a successful attack, *i.e.,* a cyber incident, is multiplied by its impact, and the impact of non-successful attacks is obviously 0. We therefore introduce a further random variable $\mathcal{S}$, which models the number of cyber incidents faced by the organization in the considered period of time. We consider two scenarios:

- the likelihood of facing s cyber incidents, after at most t attempts by the attacker. In this approach, it is assumed that the organization does not promptly realize when it is successfully attacked, and, therefore, does not take the appropriate countermeasure to change posture;
- the likelihood of facing exactly one cyber incident, after at most t attack attempts by the attacker. In this more realistic approach, it is assumed that the organization immediately notices the breach, and tries to improve its posture.

1) ORGANIZATIONS WHICH DO NOT CHANGE POSTURE

In this scenario, having assumed that the organization does not react immediately after experiencing one or more cyber incidents, we can model the attacks as Bernoulli experiments, with success probability given by (9). We can actually assume that also the outcome of the single attack attempt is a continuous random variable $\mathcal{P}$ with realization p. In particular, taking into account the uncertainty inherent in the problem (where the maturity index, for example, results from the opinion

of some experts or from the assessment of questionnaires) we assume that $\mathcal{P}$ follows a PERT (Project Evaluation and Review Techniques) distribution, with $p_m = p_s(\min\{x + q, 10\})$, $p_M = p_s(\max\{x - q, 0\})$, for some arbitrary value of q, and $p^* = p_s(x)$ respectively the minimum, maximum, and most likely values. In the following, we have considered $q = 1$. Therefore, the probability that s out of $n \leq t$ trials are successful is given by

$$\Pr(\mathcal{S} = s | \mathcal{N} = n) = \int_{p_m}^{p_M} \binom{n}{s} p^s (1-p)^{(n-s)} f_{\mathcal{P}}(p) \mathrm{d}p,$$

where

$$f_{\mathcal{P}}(p) = \frac{(p - p_m)^{\alpha-1}(p_M - p)^{\beta-1}}{B(\alpha, \beta)(p_M - p_m)^{\alpha+\beta-1}},$$

being $\alpha = 1 + 4\frac{p^* - p_m}{p_M - p_m}$ and $\beta = 1 + 4\frac{p_M - p^*}{p_M - p_m}$, while $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$ is the Beta function. For the likelihood, we finally obtain

$$L^{(\text{NC})}(s) = \Pr(\mathcal{S} = s)$$

$$= \sum_{n=1}^{t} \Pr(\mathcal{S} = s | \mathcal{N} = n) \Pr(\mathcal{N} = n). \quad (11)$$

2) ORGANIZATIONS WHICH CHANGE POSTURE

In this scenario, we assume that the organization runs into a single cyber incident, detects it immediately and decides to promptly take countermeasures, changing its posture. Therefore, a new cyber risk assessment should be performed after one successful attack, since the initial conditions, and the value of the maturity index above all, should have been changed. In this case, we consider a cumulative geometric distribution, *i.e.,*

$$\Pr(\mathcal{S} = 1 | \mathcal{N} = n) = \int_{p_m}^{p_M} \sum_{k=1}^{n} p(1-p)^{(k-1)} f_{\mathcal{P}}(p) \mathrm{d}p.$$

Then, the likelihood of occurrence of a single cyber incident can be obtained by considering all possible values of n, *i.e.,*

$$L^{(\text{C})} = \Pr(\mathcal{S} = 1) = \sum_{n=1}^{t} \Pr(\mathcal{S} = 1 | \mathcal{N} = n) \Pr(\mathcal{N} = n). \quad (12)$$

V. CYBER INCIDENT MODEL IN HTMA AND FAIR

In this section we show how the values of the likelihood (or frequency) computed in Sections IV-H1 and IV-H2 can be plugged into HTMA and FAIR, minimizing the subjectivity of the assessor and, rather, allowing the computation of risk according to the posture of the considered organization.

A. USE OF MAGIC WITH HTMA

According to the description in Section III-A, HTMA does not take into account the possibility that a threat happens more than once in a year and, therefore, we must consider the likelihood that the organization faces the threat just once. This means that the scenario described in Section IV-H2 must be applied. As for the monetary impact of these threats, they

are not object of investigation in this paper and, therefore, we rely on available results.

In particular, MAGIC can be used to estimate the likelihood of occurrence of the threats considered in the Monte Carlo simulation (point 1. in Section III-A). In fact, instead of relying on the expertise of the assessor, it is possible to employ the likelihood obtained through our cyber incident model (12) as input to the HTMA method. However, we need to associate a likelihood to every single cyber threat in the considered list (and not a single, general likelihood value) and, therefore, the model must be adapted. Explicitly the procedure works as follows. We start by defining a list of threats. This list is then combined with the list of considered controls through a table containing weight coefficients $\omega_{i,j}$ as shown in Table 3, where τ_j and γ_i denote the j-th threat and the i-th control, respectively.

TABLE 3. Weight coefficients for likelihood assessment in HTMA.

	τ_1	τ_2	$\cdots$	τ_n
γ_1	$\omega_{1,1}$	$\omega_{1,2}$	$\cdots$	$\omega_{1,n}$
γ_2	$\omega_{2,1}$	$\omega_{2,2}$	$\cdots$	$\omega_{2,n}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
γ_m	$\omega_{m,1}$	$\omega_{m,1}$	$\cdots$	$\omega_{m,n}$

First of all, one should answer the question "if the control γ_i is not implemented, does the risk relative to the threat τ_j increase?". If not, then $\omega_{i,j}$ is 0, otherwise a non-zero value can be assigned to $\omega_{i,j}$, according to some predetermined rule. This way, for each threat we can select a subset of controls that includes only those controls that have a non-zero weight ω for that specific threat. Notice that the subsets are not necessarily disjoint. We can compute a maturity index for each of these subsets and obtain the likelihood of occurrence of a cyber incident caused by the threat τ_j, noted by L_j, following the procedure presented in the previous sections.

B. USE OF MAGIC WITH FAIR
Referring to the FAIR ontology in Fig. 1, the blocks that are directly influenced by our model are the TEF, the Vulnerability and the LEF (both for the primary and the secondary threat event). Clearly, these blocks will also indirectly influence the upper layers of the ontology which, however, also depend on blocks which are not object of our analysis. We remind that the TEF is defined as the frequency with which, in a given time period, the attacker tries to breach the organization. In Section IV-G, we have devised a model that returns the probability that the organization faces a given number of attacks. As discussed below, the link between number of attacks and frequency of the attacks is immediate, once the reference time period has been defined. The Vulnerability, defined as the probability of success of any of these breaches, assumes the same meaning of the probability of success of the attacks, defined in Section IV-F and denoted as p^*. Finally, the LEF is obtained by combining the other two parameters; it is defined as the frequency with which the attacker

succeeds, causing a monetary loss to the organization. Also this parameter has been covered by our analysis, especially that in Section IV-H1, where the likelihood that the organization is breached a given number of times is computed through (11). Also in this case, the connection between likelihood and frequency is straightforward.

In the original FAIR approach, it is assumed that the LEF follows a PERT distribution, whose parameters must be computed based on experts' estimates. Using MAGIC, we maintain the Monte Carlo approach, as already done for the HTMA approach, but we sample the LEF from the distribution of $\mathcal{S}$ (i.e., $\Pr(\mathcal{S} = s)$) reported in (11), which is computed taking into account both the TEF and the Vulnerability. The random sampling can be performed very easily, since it is referred to a discrete distribution: we can associate a probability to each allowed value of s and sample according to these probabilities. Notice that, once s is obtained by random sampling, the LEF in the considered framework can be simply computed as LEF $= \frac{s}{t}$.

VI. NUMERICAL RESULTS
In this section we provide some numerical results, which aim to validate the effectiveness of MAGIC.

A. CASE STUDY WITH HTMA
In order to take into account a realistic scenario, we have considered the list of threats given in [9, Table 1] (originally taken from [36]), and here reported, for the sake of clarity, in Table 4. Furthermore, as a case study, we have analyzed a dummy organization in the "Healthcare" sector. We have simulated a posture assessment, finding, for each threat, a maturity index as in Table 4. As for the attractiveness, according to the discussion in Section IV-D, we have found that organizations in the healthcare sector are very highly attractive. Indeed, this is not surprising, and well-known in the literature. Finally, we have considered a range of values for the complexity index, from 4.5 to 7.5 with step 0.5. We have also considered different values of n_{avg}, ranging from 2 to 5; we have assumed that, in each simulation, n_{avg} is the same for all threats (for the sake of comparison between different threats). As in [8], we choose $B = -2$, $U = 0.97$ and $L = 0.03$ as parameters of the generalized logistic function. Looking at Table 4 we notice that we are talking about an attractive organization, which however has some serious maturity shortcomings, which may lead to weaknesses against cyber threats.

The parameters of the PERT distribution, along with the likelihood of occurrence of each threat in a year ($t = 365$ and $\Delta t = 1$, assuming that the attacker performs at most one attack a day), computed by (12), are also shown in Table 4 for $n_{\text{avg}} = 4$ and complexity index 5. As expected, due to the relatively low values of the maturity indexes associated to each threat, the organization has a relatively high probability of suffering a cyber incident due to them. Still, threats corresponding to lower maturity indexes are less likely

TABLE 4. Parameters of the proposed model, as inputs of HTMA for the considered scenario, when the complexity index is 5 and $n_{avg} = 4$.

ID	Threat	Maturity index	p^*	p_m	p_M	L^C
1	Malware	4.3	0.50	0.28	0.72	0.86
2	Web-based attacks	5.6	0.23	0.11	0.43	0.61
3	Denial of services	3.6	0.66	0.43	0.83	0.92
4	Malicious insiders	1.9	0.90	0.79	0.95	0.97
5	Phishing and social engineering	3.6	0.66	0.43	0.83	0.92
6	Malicious code	6.0	0.17	0.08	0.34	0.52
7	Stolen devices	4.8	0.38	0.19	0.62	0.78
8	Ransomware	5.1	0.32	0.16	0.55	0.72
9	Botnets	4.3	0.50	0.28	0.72	0.86

to be suffered than threats corresponding to higher maturity indexes.

Regarding the monetary impacts, we have considered those in [9, Table 2] and also reported in Table 5, for the sake of completeness. In order to keep a uniform notation with the rest of the paper, we have applied a currency exchange from dollars (used in [9]) to Euros. Notice that, assuming that all threats happen and the associated cyber incidents cause the largest possible loss (given by the upper value of the corresponding interval), we obtain an upper bound for the total loss which, in the proposed example, is a value between 11.5 and 12 million Euro (precisely, 11.8361 million Euro).

TABLE 5. Impact range of the considered threats, in million Euro.

ID	Threat	Impact range
1	Malware	[2.1360, 2.3941]
2	Web-based attacks	[1.8156, 2.0381]
3	Denial of services	[1.4151, 1.5842]
4	Malicious insiders	[1.2816, 1.4329]
5	Phishing and social engineering	[1.1748, 1.3172]
6	Malicious code	[1.1659, 1.2994]
7	Stolen devices	[0.77875, 0.87576]
8	Ransomware	[0.48060, 0.53845]
9	Botnets	[0.31684, 0.35600]

As usual, we express the final output of the model as a LEC, summarizing the probability that the loss faced by the organization will be greater than or equal to a certain value, obtained with a Monte Carlo simulation with 10, 000 trials.

We show the LECs, obtained using the aforementioned values of the complexity and considering $n_{avg} = 4$, in Fig. 4. Notice that, coherent with our analysis, more complex organizations are more likely to suffer cyber attacks and, therefore, they will be subject to larger monetary losses.

Finally, we have shown in Fig. 5 the LECs for complexity index equal to 5 and different values of n_{avg}. Also in this case, as expected, the losses increase for increasing values of n_{avg}, which we have chosen as the number of attack attempts suffered in the previous year, relatively to each threat.

B. CASE STUDY WITH FAIR

We have considered the same scenario described in [7, Chapter 8]. For the sake of simplicity, we have assumed that the Loss Magnitude is only caused by primary losses, thus ignoring secondary risk. Clearly, this assumption can be removed without changing the rationale of the model.
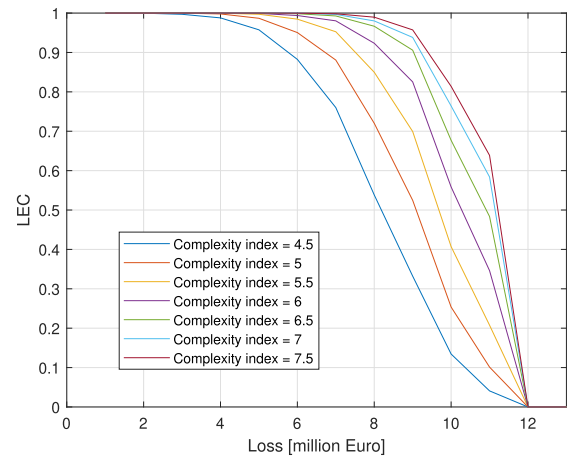

FIGURE 4. LEC for the considered (very highly attractive) organization, for different values of the complexity index and $n_{avg} = 4$, using the HTMA method combined with MAGIC.
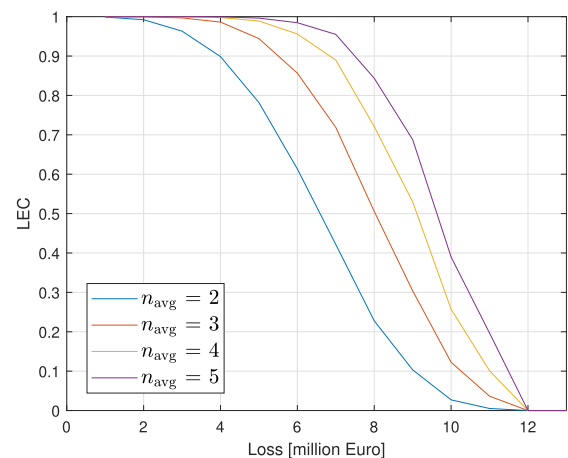

FIGURE 5. LEC for the considered (very highly attractive) organization, for different values of n_{avg} and complexity index equal to 5, using the HTMA method combined with MAGIC.

Concerning the Loss Magnitude, we have not changed the minimum, maximum, and most likely values (and neither we have changed the confidence level) for the considered categories in [7, Chapter 8]: productivity, response, replacement, fines & judgements, reputation and competitive advantage. The only categories having non-zero impact values are response and replacement and, for the sake of completeness,

we have reported them in Table 6. Also in this case, we have considered a dummy organization in the "Healthcare" sector. Simulating a different posture assessment from the previous section, we have found a complexity index equal to 5.2 and a maturity index equal to 6.9. Being in the healthcare sector, the organization is considered again very highly attractive. The values of B, U and L are once again set as -2, 0.97 and 0.03, respectively. The analysis is carried out on daily intervals within a year, i.e., $t = 365$ and $\Delta t = 1$, and we have considered $n_{\text{avg}} = 84$. Notice that this value is significantly larger than those adopted for the examples in Section VI-A, just in view of enlarging the ensemble of considered scenarios.

TABLE 6. Impact range of the considered event, in €.

Category	Minimum	Maximum	Most Likely	Confidence
Response	2, 750	8, 250	22, 000	20
Replacement	20, 000	30, 000	50, 000	20

Running the Monte Carlo simulation, we obtain the Loss Magnitudes in Fig. 6. Each point corresponds to a different simulation. The red circle in the figure represents the mean value (for both axes). Defining the k-th percentile as the score below which the $k\%$ of the scores fall in the given distribution, we have found that the 10-th percentile and the 90-th percentile, for this example, are approximately 18, 000 € and 58, 000 €, respectively. The results of the Monte Carlo simulation are also summarized in Table 7 where, besides the minimum, the maximum and the mean, also the mode (that is the most frequent value occurred in the Monte Carlo simulation, differently from the most likely value, that is the result of the statistical inference for the same quantity) is reported.
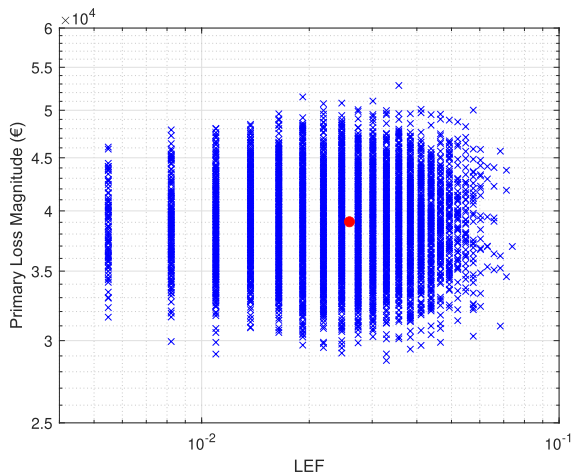
FIGURE 6. Loss magnitude for the considered organization, which is very highly attractive, when the complexity index is 5.2 and the maturity index is 6.9, obtained through the FAIR method.

Clearly, fixed t, a smaller value of Δt increases the number of possible attacks per year (which is now 365), thus increasing the number of possible values of the LEF, in its turn.

TABLE 7. Summary of the results of the Monte Carlo simulation, when the complexity index is 5.2 and the maturity index is 6.9.

	Minimum	Mean	Mode	Maximum
Primary Loss Events per year	1	9.5	8	27
Primary Loss Magnitude (€)	28, 709	39, 045	32, 948	52, 822
Total Loss Exposure (€)	35, 369	369, 260	295, 520	1, 141, 600

In order to evaluate the impact of different values of the maturity index, complexity index and n_{avg}, we have fixed two of these parameters and let the other change, obtaining Figs. 7, 8 and 9, respectively. The results are coherent with our analysis: on the one hand, for increasing values of the complexity and of n_{avg}, the total loss exposure increases; on the other hand, the total loss exposure decreases when the organization has increasing values of the maturity.
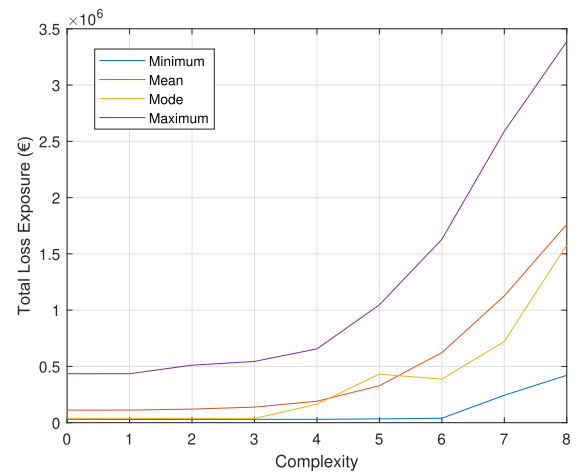
FIGURE 7. Minimum, mode, mean and maximum of the total loss exposure, for different values of the complexity index.
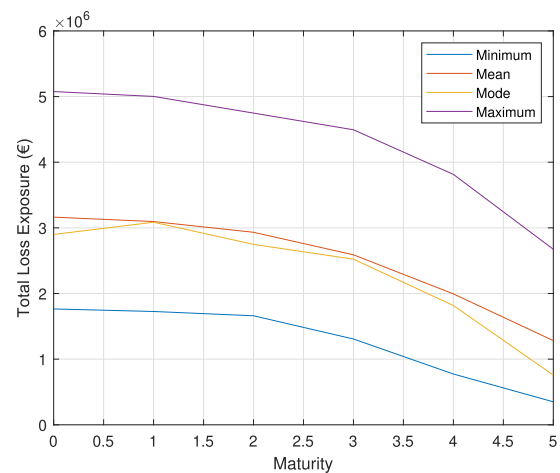
FIGURE 8. Minimum, mode, mean and maximum of the total loss exposure, for different values of the maturity index.

VII. COMPARISON AND DISCUSSION
In this section we perform both a quantitative and a qualitative comparison of MAGIC with other methods.
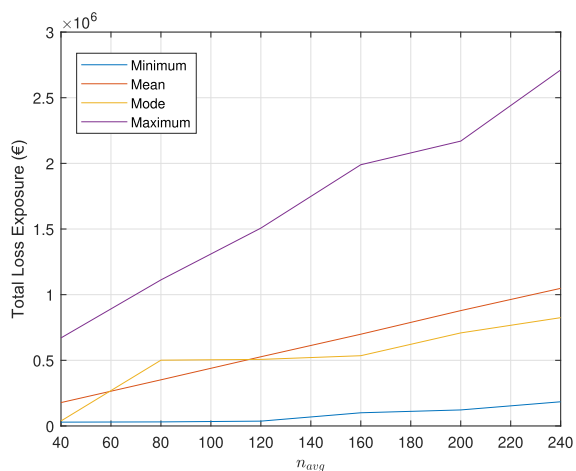
FIGURE 9. Minimum, mode, mean and maximum of the total loss exposure, for different values of n_{avg}.

TABLE 8. Low level metrics for likelihood assessment using CVSS.

Metric	CVSS	Metric evaluation	Numerical value
Access Vector (AV)	Base	Local	0.4
		Adjacent	0.6
		Network	1
Authentication (Au)	Base	Multiple	0.5
		Single	0.55
		None	1
Access Complexity (AC)	Base	High	0.5
		Medium	0.75
		Low	1
Exploitability (E)	Temporal	Unproven	0.85
		Proof of concept	0.9
		Functional	0.95
		High	1
		Not defined	1
Report Confidence (RC)	Temporal	Unconfirmed	0.9
		Uncorroborated	0.9
		Confirmed	1
		Not defined	1

TABLE 9. Likelihood of occurrence of the considered threats, obtained using different methods.

ID	Threat	L^{C}	L^{CVSS}	L^{EXPERT}
1	Malware	0.86	0.90	0.80
2	Web-based attacks	0.61	0.64	0.75
3	Denial of services	0.92	0.95	0.90
4	Malicious insiders	0.97	0.15	0.55
5	Phishing and social engineering	0.92	0.90	0.95
6	Malicious code	0.52	0.81	0.70
7	Stolen devices	0.78	0.09	0.45
8	Ransomware	0.72	0.99	0.85
9	Botnets	0.86	0.86	0.80

A. QUANTITATIVE COMPARISON

Let us consider the same organization in the "Healthcare" sector as in Section VI-A, for which a posture assessment has returned a complexity index equal to 5. The maturity index associated to each of the considered threats is given in Table 4 (third column). Owing to its sector, the organization is considered very highly attractive.

We have computed the likelihood of occurrence of cyber incidents caused by the same threats, by using the basic version of CVSS-based risk assessment method, relying on the fundamentals described in [27]. The low level metrics we use for the comparison are those given in [27], which we also report in Table 8 for the sake of completeness. The assignment of the numerical values to these CVSS low level metrics, which is the possibly subjective phase of this approach, has been performed by means of a brainstorming session, based on the experience of the authors and on available evidences. It must be stressed that some assignments are not subjective: for example, the Access Vector of stolen devices or malicious insiders is objectively "Local". However, this is neither true for all metrics, nor for all threats.

Lastly, a team of certified external experts, taking as inputs the status of the organization used for the posture assessment, in order to find the maturity, complexity and attractiveness indexes, has provided estimates for the likelihood of occurrence of cyber incidents deriving from the nine considered threats. This expert-based approach is used in many modern cyber-risk assessment methods, such as HTMA, FAIR, MAGERIT, CRAMM, and many others.

The results obtained through MAGIC and the aforementioned approaches are compared in Table 9, where we denote the likelihoods resulting from the newly proposed method, the CVSS-based method and the expert estimates as L^{C}, L^{CVSS} and L^{EXPERT}, respectively.

We notice that, in several cases, the three considered methods provide comparable results. However, in some cases, the numerical methods (MAGIC and the CVSS-based) give similar results. We argue this is due to the incapability of the

CVSS-based method to catch that different organizations may react differently to the same threats, if the state of their organizational infrastructure is dissimilar. In fact, the CVSS-based method considers only marginally and indirectly the degree of compliance of the organization to the best practices for threat prevention, whereas MAGIC takes it into account in the preliminary posture assessment. In contrast, on the one hand, the external experts evaluation seems very balanced in most situations but, on the other hand, as already asserted, it is subjective, in that a different team of experts might return significantly different results. Moreover, the involvement of external experts is costly, it is time consuming, and requires a significant amount of data to be processed, which may not be available.

B. QUALITATIVE COMPARISON

In this section we provide a qualitative comparison of several cyber risk assessment and/or management methods. The considered approaches, the evaluated metrics and the results of the assessment are shown in Table 10.

In short, we notice that among the most relevant existing risk assessment/management methods, most of them require the subjective estimates of calibrated or external experts. Clearly, the extensive use of external experts (described in Section VII-A) restricts the range of the organizations that can afford it to those in the medium and large scale, making the associated methods barely applicable to small organizations. Some methods, such as MEHARI or MAGERIT,

TABLE 10. Comparison between the considered methods. Legend: RM=risk management; RA=risk assessment; LA=likelihood assessment.

Method	Scope	Likelihood Assessment	Assessors	Target Organizations
ISO/IEC 27005:2018	RM	Guidelines only	External experts	All
NIST SP 800-30	RA	Guidelines only	External experts	Governmental*
EBIOS	RM	Qualitative	Team of different stakeholders	All
CRAMM	RA	Qualitative	External experts	Large scale*
OCTAVE	RM	Qualitative	Internal experts	All
MEHARI	RM	Qualitative	Internal experts	All
MAGERIT	RM	Numerical	Calibrated experts	Public Administration*
IRAM2	RM	Qualitative	Internal experts	All
IT-Grundschutz	RM	Qualitative	Internal and external experts	All
CORAS	RA	Qualitative or numerical	Internal and external experts	Medium and large scale
HTMA	RA	Numerical	Calibrated experts	Medium and large scale
FAIR	RA	Numerical	Calibrated experts	Medium and large scale
LiSRA	RA	Numerical	External experts	All
Fuzzy-logic based	LA	Numerical	External experts	Medium and large scale
CVSS-based	LA	Numerical	Internal experts	All
MAGIC	**LA**	**Numerical**	**Internal experts**	**All**

* Can be extended to other types of organizations.

rely on internal experts, like our method, but they implement only qualitative approaches. Therefore, they do not provide a numerical estimate of the likelihood of occurrence of cyber incidents, leading to results which might need an involved interpretation and may not be easily reproducible. Even in the CVSS-based methods, where the likelihood is numerically estimated based on internal experts analyses, the process is still subjective, since there is no automatic and universal way to assign scores to the threats exploiting the considered vulnerabilities. Moreover, none of the considered methods directly relate the likelihood of occurrence of cyber incidents to the posture of the organization. MAGIC tries to overcome all these shortcomings, by providing a numerical approach for the likelihood assessment, mostly relying on the simple initial task of some internal experts compiling questionnaires about the status of the target organization.

VIII. CONCLUSION
Most existing cyber risk assessment methods are affected by subjectivity or require a high level of expertise from the organizations, which makes them barely practical in real-case scenarios. We have tackled this issue by proposing a novel probabilistic method, called MAGIC that, based on the organization posture, estimates the likelihood (or the frequency) of occurrence of a cyber incident or, more generally, of a list of cyber incidents. Through numerical simulations we have shown how MAGIC can be plugged into statistical cyber risk assessment methods, like HTMA and FAIR, to eventually assess the risk as the likelihood of occurrence of cyber incidents combined with their impact. We have performed both qualitative and quantitative comparisons with alternative approaches, showing that our method is reliable and general, in the sense that it can provide inputs to other risk assessment and/or risk management methods. Moreover, MAGIC tries to catch all the advantages of existing likelihood assessment methods. In conclusion, due to its simplicity and rather low computational demand, we argue that MAGIC can be applied to any type of organization (small, medium, or large), without any loss of generality.

REFERENCES
[1] *Special Publication 800–30 Revision 1 Information Security: Guide for Conducting Risk Assessments*, National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA, Sep. 2012.

[2] *Risk Management Guidelines*, Standard ISO 31000:2018, International Organization for Standardization (ISO), Feb. 2018.

[3] *Information Technology Security Techniques Information Security Risk Management*, Standard ISO/IEC 27005:2018, International Organization for Standardization (ISO), Jul. 2019.

[4] P. Shamala, R. Ahmad, and M. Yusoff, "A conceptual framework of info structure for information security risk assessment (ISRA)," *J. Inf. Secur. Appl.*, vol. 18, no. 1, pp. 45–52, Jul. 2013.

[5] M. Stamatelatos, "Probabilistic risk assessment: What is it and why is it worth performing it?" NASA Office of Safety and Mission Assurance, Washington, DC, USA, Tech. Rep., May 2000. [Online]. Available: https://copswiki.org/w/pub/Common/M1922/pra%20-%20Probabilistic %20Risk%20Assessment%20-%20What%20is%20it%20and%20why %20is%20it%20work

[6] D. W. Hubbard and R. Seiersen, *How to Measure Anything Cybersecurity Risk*. Hoboken, NJ, USA: Wiley, 2016.

[7] J. Freund and J. Jones, *Measuring and Managing Information Risk: A FAIR Approach*. Oxford, U.K.: Butterworth-Heinemann, 2014.

[8] G. Rafaiani, M. Battaglioni, M. Baldi, F. Chiaraluce, G. Libertini, L. Spalazzi, and G. Cancellieri, "A functional approach to cyber risk assessment," in *Proc. AEIT Int. Annu. Conf. (AEIT)*, Oct. 2021, pp. 1–6.

[9] P. Santini, G. Gottardi, M. Baldi, and F. Chiaraluce, "A data-driven approach to cyber risk assessment," *Secur. Commun. Netw.*, vol. 2019, pp. 1–8, Sep. 2019.

[10] National Cybersecurity Agency of France (ANSSI). *EBIOS Risk Manager*. Accessed: Jun. 8, 2022. [Online]. Available: https://www.ssi.gouv.fr/ en/guide/ebios-risk-manager-the-method

[11] Z. Yazar, "A qualitative risk analysis and management tool—CRAMM," *SANS InfoSec Reading Room White Paper*, vol. 11, pp. 12–32, Apr. 2002.

[12] C. Alberts, A. Dorofee, J. Stevens, and C. Woody, "Introduction to the OCTAVE approach," Carnegie-Mellon Univ. Pittsburgh PA Software Eng. Inst., Pittsburgh, PA, USA, Tech. Rep. ADA634134, 2003. https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=51546

[13] CLUSIF (Club de la Sécurité de l'Information Français). *MEHARI (MEthod for Harmonized Analysis of RIsk)*. Accessed: Jun. 8, 2022. [Online]. Available: https://www.enisa.europa. eu/topics/threat-risk-management/risk-manageme%nt/current-risk/risk-management-inventory/rm-ra-methods/m_mehari.html

[14] Ministerio de Hacienda y Administraciones Publicas (Spanish Ministry of Finance and Public Administrations). *MAGERIT—Version 3.0 Methodology for Information Systems Risk Analysis and Management, Book I—The Method*. [Online]. Available: https://administracionelectronica.gob.es/pae_Home/dam/jcr:80b16a91-75b1%-432d-ab23-844a12aab5fc/MAGERIT_v_3_book_1_method_PDF_NIPO_630-14-162-0.pdf

[15] Information Security Forum. *Information Risk Assessment Methodology 2 (IRAM2)*. Accessed: Jun. 8, 2022. [Online]. Available: https://www.securityforum.org/solutions-and-insights/information-risk-a%ssessmentmethodology-iram2/

[16] Bundesamt Für Sicherheit in der Informationstechnik (BSI). *BSI Standard 100-3–Risk Analysis Based on IT-Grundschutz*. Accessed: Jun. 8, 2022. [Online]. Available: https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Grundschutz/Interna%tional/bsi-standard-2002_en_pdf.html

[17] SourceForge. *The CORAS Method*. Accessed: Jun. 8, 2022. [Online]. Available: http://coras.sourceforge.net/

[18] O. Giuca, T. Popescu, A. Popescu, G. Prostean, and D. Popescu, "A survey of cybersecurity risk management frameworks," in *Soft Computing Applications. SOFA 2018* (Advances in Intelligent Systems and Computing), vol. 1221, V. Balas, L. Jain, M. Balas, and S. Shahbazova, Eds. Cham, Switzerland: Springer, 2021.

[19] D. Gritzalis, G. Iseppi, A. Mylonas, and V. Stavrou, "Exiting the risk assessment maze: A meta-survey," *ACM Comput. Surv.*, vol. 51, no. 1, pp. 1–30, Jan. 2019.

[20] M. Khosravi-Farmad and A. Ghaemi-Bafghi, "Bayesian decision network-based security risk management framework," *J. Netw. Syst. Manage.*, vol. 28, no. 4, pp. 1794–1819, Oct. 2020.

[21] C. Schmitz and S. Pape, "LiSRA: Lightweight security risk assessment for decision support in information security," *Comput. Secur.*, vol. 90, Mar. 2020, Art. no. 101656.

[22] A. P. H. Gusmão, L. C. E. Silva, M. M. Silva, T. Poleto, and A. P. C. S. Costa, "Information security risk analysis model using fuzzy decision theory," *Int. J. Inf. Manage.*, vol. 36, no. 1, pp. 25–34, 2016.

[23] A. Handa, S. Mukhopadhyay, S. Mallick, N. Kumar, S. K. Shukla, R. L. Minz, S. P. Nagarmat, and R. Rakesh, "Cyber risk assessment of networked cyber assets using probabilistic model checking," in *Proc. IEEE Conf. Inf. Commun. Technol.*, Dec. 2019, pp. 1–6.

[24] M. U. Aksu, M. H. Dilek, E. I. Tatli, K. Bicakci, H. I. Dirik, M. U. Demirezen, and T. Aykir, "A quantitative CVSS-based cyber security risk assessment methodology for IT systems," in *Proc. Int. Carnahan Conf. Secur. Technol. (ICCST)*, Oct. 2017, pp. 1–8.

[25] C. Alberts, A. Dorofee, J. Stevens, and C. Woody, *OCTAVE-S Implementation Guide, Version 1.0. Manuelélectronique*. Pittsburg, PA, USA: Software Engineering Institute, Carbegie Mellon Univ., 2005.

[26] R. Caralli, J. Stevens, L. Young, and W. Wilson, "Introducing OCTAVE allegro: Improving the information security risk assessment process," Softw. Eng. Inst., Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU SEI-2007-TR-012, 2007, [Online]. Available: http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=8419

[27] P. Mell, K. Scarfone, and S. Romanosky. (2007). *A Complete Guide to the Common Vulnerability Scoring System Version 2.0*. [Online]. Available: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=51198

[28] *Special Publication 800-53a Revision 5 Assessing Security and Privacy Controls in Information Systems and Organizations*, National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA, Jan. 2022.

[29] *CIS Controls V.8*, Center of Internet Security (CIS), East Greenbush, NY, USA, May 2021.

[30] *Handbook on Security of Personal Data Processing*, European Union Agency For Network and Information Security (ENISA), Athens, Greece, Dec. 2017.

[31] *Framework for Improving Critical Infrastructure Cybersecurity Version 1.1*, National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA, Apr. 2018.

[32] *Cybersecurity Assessment Tool*, Federal Financial Institutions Examination Council (FFIEC), Richmond, VA, USA, May 2017.

[33] *Rapporto Clusit 2021 Sulla Sicurezza ICT in Italia*, Dipartimento di Informatica Giovanni Degli Antoni, Univ. degli Studi di Milano, Milano, Italy, Oct. 2021.

[34] D. Fekedulegn, M. M. Siurtain, and J. Colbert, "Parameter estimation of nonlinear growth models in forestry," *Silva Fennica*, vol. 33, no. 4, pp. 327–336, 1999.

[35] P. Verboon and G.-J.-Y. Peters, "Applying the generalized logistic model in single case designs: Modeling treatment-induced shifts," *Behav. Modification*, vol. 44, no. 1, pp. 27–48, Jan. 2020.

[36] "Cost of cyber crime study," Ponemon Inst. LLC, Traverse City, MI, USA, Tech. Rep., 2017. [Online]. Available: https://www.accenture.com/_acnmedia/pdf-62/accenture-2017costcybercrime-us-final.pdf

MASSIMO BATTAGLIONI (Member, IEEE) was born in Macerata, Italy, in 1991. He received the Laurea degree in electronic engineering, the Laurea Magistrale degree *(summa cum laude)* in electronic engineering, and the Ph.D. degree in information engineering from Marche Polytechnic University, Ancona, Italy, in 2013, 2015, and 2019, respectively. In 2017, he was a Visiting Student with the Electrical and Information Technology Department, LTH, Lund University, Sweden. In 2018, he has been a Visiting Student with the Klipsch School of Electrical and Computer Engineering, NMSU, Las Cruces, NM, USA; and the School of Electrical and Electronic Engineering, University College Dublin, Ireland. Since 2019, he has been a Postdoctoral Researcher with the Department of Information Engineering, Marche Polytechnic University. His research interests include cyber risk assessment and coding techniques for communications reliability and cryptography, with particular attention to LDPC codes and their application to cryptography and blockchain. He also serves as an Editor for IEEE COMMUNICATIONS LETTERS.

GIULIA RAFAIANI (Student Member, IEEE) was born in San Severino Marche, Italy, in 1994. She received the Laurea degree in biomedical engineering and the Laurea Magistrale degree *(summa cum laude)* in biomedical engineering from Marche Polytechnic University, Ancona, Italy, in 2017 and 2020, respectively, where she is currently pursuing the Ph.D. degree with the Department of Information Engineering. Her research interests include cyber risk assessment and security of distributed ledger technologies.

FRANCO CHIARALUCE (Senior Member, IEEE) was born in Ancona, Italy, in 1960. He received the Laurea degree *(summa cum laude)* in electronic engineering from the University of Ancona, in 1985. Since 1987, he has been with the Department of Electronics and Automatics, University of Ancona. He is currently a Full Professor in telecommunications with Marche Polytechnic University, Ancona. He has coauthored more than 300 scientific articles and three books and holds three patents. On his research topics, he collaborates with national and international companies. His main research interests include various aspects of communication systems theory and design, with a special emphasis on error correcting codes, cryptography, and physical layer security. He is a member of the IEICE.

MARCO BALDI (Senior Member, IEEE) received the Laurea degree (Hons.) in electronics engineering and the Ph.D. degree in electronics, computer, and telecommunications engineering from Marche Polytechnic University, Ancona, Italy. Since 2019, he has been an Associate Professor with the Department of Information Engineering, Marche Polytechnic University, where he also coordinates the local node of the CINI Cybersecurity National Laboratory and takes part in the Research and Service Center for Privacy and Cybersecurity (CRiSPY). He has coauthored more than 150 scientific articles and one book, and holds four patents. His research interests include coding and cryptography for information security and reliability. He is a member of AEIT, CINI, CNIT, IEEE Communications Society, and IEEE Information Theory Society. He serves as a Senior Associate Editor for IEEE COMMUNICATIONS LETTERS and an Associate Editor for the *EURASIP Journal on Wireless Communications and Networking* and the *Information* journal (MDPI).

• • •