Security Events and Vulnerability Data for Cyber Security Risk Estimation

Luca Allodi[*]

Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, the Netherlands.

Fabio Massacci[†]

Department of Information Engineering and Computer Science,University of Trento, Italy.

Abstract

Current industry standards for estimating cyber security risk are based on qualitative risk matrices as opposed to quantitative risk estimates. In contrast, risk assessment in most other industry sectors aims at deriving quantitative risk estimations (for example Basel II in Finance). This paper presents a model and methodology to leverage on the large amount of data available from the IT infrastructure of an organization's Security Operation Center to quantitatively estimate the probability of attack. Our methodology specifically addresses untargeted attacks delivered by automatic tools that make the vast majority of attacks in the wild against users and organizations. We consider two-stage attacks whereby the attacker first breaches an Internet-facing system, and then escalates the attack to internal systems by exploiting local vulnerabilities on the target. Our methodology factors in the power of the attacker as the number of 'weaponized' vulnerabilities he/she can exploit, and can be adjusted to match the risk appetite of the organization. We illustrate our methodology by using data from a large financial institution, and discuss the significant mismatch between traditional qualitative risk-assessments and our quantitative approach.

1 Introduction

IT systems are affected by a multitude of vulnerabilities that might be exploited by an attacker,[1] and whose exploitation may affect and propagate to other systems in the infrastructure. [2] The quantitative estimation of the risk posed by these vulnerabilities is a critical step towards a more efficient allocation of resources and a more secure overall environment.[3] Indeed, quantitative risk

[*] Contact: `l.allodi@tue.nl`

[†] Contact: `fabio.massacci@unitn.it`

analysis (QRA) is being increasingly adopted in most industry sectors. Apostolakis characterizes the adoption process in three phases:[4] first, best practice adopts only traditional 'safety analysis', with no risk quantification in place; second, the policy maker integrates 'safety analysis' with the additional insights identified by the new quantitative methods; finally, the policy maker trusts the quantitative predictions enough to *relax* the original safety analysis predictions and prioritize quantitative insights. This process can be identified in many industry sectors, including nuclear energy, space, and insurance (e.g. for natural catastrophes). [4, 5] Unfortunately, in the cyber-security risk domain this process has been withheld by technical and organizational difficulties. For example, security information is often difficult to analyze to extract actionable information, as big datasets (e.g. reporting perimeter alarms from an *Intrusion Detection System* [6]) are often of an unstructured nature.

Partially addressing these issues, recent research advancements propose new data-extraction algorithms and models for big data. For example, Khorshidi et. al [7] proposes a technique for the aggregation of qualitative data features with the aim of fostering the risk management activities of complex systems whose data sources may be incomplete or not sufficient for the purpose of the analysis. Similarly, but going in the opposite direction, Susto et al. [8] propose a method for the aggregation of multiple data sources to build models of the data with interpretable regressors.

Other research has gone in the direction of statistically linking security alarm data in IT infrastructures to security events and, ultimately, risk. [9] Yet, the inherent limitations of security monitoring technologies[10] limit the applicability to risk modelling of the data they generate. [11, 12] Similarly, risk assessment procedures prescribed by standards and best practices often fall short in providing quantitative instruments for risk estimation. For example, NIST's Information Security Handbook prescribes the usage of 'risk matrices' to qualitatively estimate the risk associated with a particular event. [13] Yet, it is known that risk matrices may cause risk mis-categorization, and even wrong risk prioritization. [14, 15] This issue is further aggravated by the fact that whilst the estimation of a vulnerability's exploitation technical impact is well understood (e.g. the Common Vulnerability Scoring System - CVSS), [16] current measures for 'exploitation likelihood' have been widely questioned. [17, 18, 19] For example, recent work has shown that patching a hundred vulnerabilities bundled in automated attack tools yields a risk reduction of 40% of attempted attacks in the wild; in contrast, the recommendation of the Payment Card Industry Security Council,[20] which is only based on the qualitative impact indicator of the CVSS, [21] only yields a 4% risk reduction and requires looking at thousands of vulnerabilities. [19] Due to the general permeation of the IT infrastructure in most industrial infrastructures, these issues can potentially impact several industry domains. For example, Basel's banking treaties tie a bank's capital requirements to a quantitative assessment of operational risk (IT risk being an instance) or to a much larger flat allocation if only qualitative measures are used.

The lack of a shared framework for the quantification of risk makes the adoption of sound and comparable measures for risk mitigation currently impossible. Whilst estimating impact is also well understood in practice, [22] the subjectivity of qualitative *attack likelihood* estimations, stemming from the absence of a shared quantitative estimation methodology, [23] has raised many concerns: [24]

existing approximations are known to be often unrealistic or based on implicit and untested assumptions. [23]

To address these shortcomings, in this work we propose a risk estimation model that explicitly *quantifies likelihood of attack* by leveraging on data available to any organization deploying common perimeter defenses such as Intrusion Detection Systems (IDS) and performing periodic vulnerability assessments. [25, 26] Addressing current concerns on the quantification of uncertainties, [27] our methodology factors in the risk model a measure of probability (or, in the definition of Aven, *vulnerability* [27]) that explicitly accounts for the level of knowledge used for the assessment (i.e. the IDS alarms and the technical vulnerabilities discovered on the system) and does not involve 'subjective' assessments otherwise explicitly or implicitly included in current quantitative risk assessment procedures in IT. [23] We illustrate our methodology by using actual data from a large financial institution. Our analysis shows how the qualitative risk assessment methodology based on risk matrices used by cybersecurity standards can lead to widely different estimations of risk from those derived quantitatively from the data. Following Apostolakis, [4] we see our model as a contribution in the transition between 'traditional' risk assessment methods and 'quantitative' risk assessment methods that, in the IT sector, are still lacking behind. [4, 23]

In the rest of the paper we first define the scope of this work (§2), and discuss existing risk assessment procedures in a large financial institution (§3); we then discuss the theoretical limitations of these practices (§4), and illustrate our quantitative methodology to address them (§5). Next, we show how to extract the relevant data from the IT infrastructure (§6), apply the methodology to a real case study of a large financial institution (§7), and conclude the paper (§8).

2 Types of cyber-risk and scope of this work

The applicability of a risk assessment methodology is tied to the nature of the risk it addresses. Broadly speaking, cybersecurity risks are typically generated by a 'human' or 'sentient actor' that initiates the attack.[1] Following Ransbotham and Mitra [28] we can distinguish between two types of attacks by type of initiating actor:

- 'Targeted' attacks typically involve strategic players that may react strategically to the defender's choices. [29, 30] By definition, these events are caused by attacks targeted solely or almost exclusively at specific facilities (such as the cases of the Stuxnet and Duqu malware), and that typically require an extreme level of sophistication. The limitations of PRA's applicability to this type of scenarios is akin to terrorism threats and is well discussed in the literature, [31, 32] as well as its implications for cyber-security. [30] Recent studies estimate that only few cyber-attacks against organizations and consumer system are of this type. [33, 34] Targeted attacks are often referred to in the risk analysis literature as 'Black Swans', [35] or 'Advanced Persistent Threats' (APT) in the computer security literature. [36]

[1] We do not consider here software or hardware failures as part of the cyber-security risk scenario as these can be appropriately studied along the guidelines of traditional safety analysis and QRA in the same fashion that natural risks are studied.

- ‘Untargeted’ attacks are attacks whose targets are not distinguished one from the other by any specific property or characteristic and just happen to be reached by the attack. [28] These attacks are typically launched using automated tools such as exploit kits [37] and are known to drive the vast majority of attacks in the wild. [38] Untargeted attacks have a wide range of potential victims and are known to affect individuals, [37] organizations, [28] and industrial systems alike. [39]

The importance of untargeted attacks in the overall risk scenario has been outlined by recent industry reports; [40, 41, 42] for example, a recent study classifies *Crimeware and Web attacks* as the source of about 70% of the attacks suffered by the Financial sector, [40] a figure in clear accordance with trends previosuly quantified in the literature. [38] Because attacks of this type are largely automated, the attack process typically follows a two-stage mechanism [37] whereby the attacking tool: (1) *attack probing*: the attack ‘probes’ the victim machine, [9, 6] for example to identify if it is vulnerable [43] or if it satisfies some desirable characteristics such as geographic location; [44] (2) *attack delivery*: the attack’s payload (e.g. shellcode, malware, bash scripts,..) is delivered to the target, for example by exfiltrating data or executing otherwise unwarrented actions on the system. [37, 45] Due to the prevalance of untargeted attacks in the overall risk scenario, [46, 47, 38] in this paper we focus on this type of attacks.

3 Cyber Risk Assessment in a Large Financial Institution

(Cyber) Security risk assessment in industry is largely constrained by compliance to regulations and adherence to standards. To illustrate this, Table 1 reports some of the regulations and industry mandated technical standards that must be satisfied by a large financial company, here anonymized as *Company*, that offers integrated services in finance, logistics, and mobile communication with a turnaround of around 24 billion Euro and 150 thousands employees. The *Baseline* illustrates some norms that must be addressed to achieve a minimum compliance whilst the *Perimeter* is the set of affected Services and Applications.

In order to achieve this security baseline a large security infrastructure is needed: a *security operation center* (SOC). There are several best practices to build a SOC[25, 26] and an average SOC can quickly generate several GB of security events per day which can create a significant stress on the human responders. [48] For example, our anonymized and aggregated dataset for Company’s IT infrastructure is over 2GB of data for just a month of processing. Yet, as we shall see, this infrastructure is hardly used by the risk assessment standards. We will return to the key components of a SOC in Section 6 when discussing how to concretely extract data from the infrastructure to feed in our quantitative model.

Once the minimal security measures are in place, the particular risk assessment process to follow might be mandated as well by the regulations. The ISO/IEC 27005 [49] and ISO/IEC 31000 [50] are typical standards used to undertake risk management at the corporate level. The NIST SP 800-30 is an-

Table 1: Security and compliance requirements over diverse perimeters in a financial organization.

Perimeter	Description	Compliance requirements	Security requirements
Privacy	Protection of personal, sensitive and judicial data	National Law and Technical Annexes, Internal Guidelines	Security guidelines, ISO 27001:2013
Financial Data	Protection and tracking of financial transactions, money transfers and financial information	National Law and Technical Annexes, Internal Guidelines	PCI-DSS, Security guidelines for protection of payment systems
Central Bank	Compliance with provisions of management and control issued by CB	National Authority Regulation, National Regulator Terms of Reference	Security guidelines for electronic payments
Traffic Data	EU Communication Directive, Traffic (Phone/Internet) Data Management	Nat. Authority Regulation, Technical Annex to Law, Internal Guidelines	Guidelines for critical infrastructure

other widely used standard for security risk assessment in the US. [51] Other approaches to security risk assessment, with stronger focus on audit, are the COBIT methodology sponsored by the ISACA institute, [52] SABSA used by Accenture, [53] or the COSO Enterprise Risk Management. [54]

While several definitions of risk exist, including concepts such as probability and impact, uncertainty and consequence, and expected consequence [55], in industry standards they all ultimately collapse to the intuitive relation $Risk = Impact \cdot Likelihood$.

To calculate the two parameters and the resulting risk, the default application of an *Information Security Risk Management Process* (ISRM) is basically broken down in the following steps: [49]

1. *Asset and Process Identification* captures the overall enterprise architecture;

2. *Business Impact Analysis* focuses on the information used by each service and the impact of an attack;

3. *Risk Assessment* is then performed in order to identify impact, gaps and current risk levels for all assets;

4. *Security Requirements Identification* addresses those gaps and produces a plateau of security measures for the Service Owner to choose;

5. *Risk Treatment* is performed by the Service Owner on the basis of the

Table 2: Example of Risk Assessment Interviews for ISRM - From [22]

Level	Questions	Time	Unit
Business Information	16	1hrs	Service
Process/People	300	3hrs	Process
Applications	250	3hs	Application
Software components	200	2hrs	Type of Asset
Infrastructure	200	2hrs	Type of Asset
Facilities	100	1hrs	Facility

risk analysis and the business considerations, whilst the ICT Department implements the technical solutions.

The first step of the process generates an enterprise architecture that spans all layers: from Services to Processes, from People to Facilities. For each layer a detailed analysis of impacts and security controls is conducted through several interviews with Service and System Owners. Table 2 illustrates part of this effort for a large financial company [22]

The Business Impact Analysis accounts for the type of data that is processed, the relevance of compliance perimeters, the impact of security compromise (e.g. severity of consequential criminal charges), the economic relevance of the service (e.g. amount of losses due to service downtime), up to the monetary benefit that a competitor might gain.

At Company, these assessments are mapped in five macro categories from C1 (lowest level) to C5 (highest level), to be compliant with the 1-5 *ordinal* levels identified by the ISO standard. At the lower level (C1) we have services that do not manage personal data and are not associated with security perimeters; at the highest (C5) we have services that are fundamental for the company from a business perspective and that are bound to relevant security and compliance perimeters. There might be different mechanisms for doing so that are company dependent. [22]

Likelihood estimations are built by considering the average threat level posed by a vulnerability and the estimated probability of receiving an attack by a certain attacker. For example, attacks of type T may be assigned a certain probability $Pr(AttackType = T)$, and the probability of an attack on a system s is given by $Pr(s \in Attack) = Pr(AttackType = T) \times Severity(v \in s)$, where $Severity(v \in s)$ is a function that considers the severity of the vulnerabilities in the system s. This function may be a transformation of the average or maximum severity. This is what currently prescribed by many standards for Information Risk Management (including PCI-DSS, ISO 27001, NIST 800-30), that suggest to consider a positive correlation between vulnerability severity and likelihood of attack. [56, 57, 17] $Pr(AttackType = T)$ is typically estimated using pre-defined tables (see for example Table 5, defined over Eurocontrol's guidelines).

Other methods focus on using feature extraction to enumerate security events (e.g. as recorded by network sensors) and estimate probability of occurrence from there. This is however known to be a poor indicator of likelihood of attack. [10, 58] As a result, likelihood of attack is typically assessed by means of expert judgment. [23]

To perform the final evaluation both ISO/27001 and NIST 800-30 standards

Table 3: Example of a 3x3 risk matrix.

		Impact:		
		Minor	Severe	Critical
Likelihood:	Rare	Low	Low	Med.
	Frequent	Low	Med.	High
	Certain	Med.	High	High

suggest the use of *risk matrices* as a tool to support such decisions. Table 3 reports a simple example of a 3x3 risk matrix, where the interaction between the *rare*, *frequent*, *certain* likelihood levels and the *minor*, *severe*, *critical* consequence levels, results in a final 3-level risk evaluation from *low* to *high*.

4 Limitations of current risk assessment methodologies

Risk quantification considers the measurement of quantities representing the uncertainty and the consequences of an event. Several definitions of risk exist, [59] and its mathematical form may vary; however risk should be ideally represented as a *cardinal* value resulting from some function transforming uncertainty and consequence assessments in a synthesized risk value. [14] This process is often aided by technological means to measure or estimate those parameters. [4] For example, probabilistic risk assessment is often based on historical data (e.g. measured by sensors such as seismographs, or records of past nuclear incidents) which goes in input to data and risk models that provide the final risk estimate. This same process is applied to IT risk. Several models for cyber-risk quantification exist in the literature. A summary of approaches to cyber-risk estimations is given in Table 4.

Attack trees and graphs aim at quantifying risk of cyber-attacks, both in generality and applied to specific risk scenarios. [2] Attack graphs represent network or system structures, where each node corresponds to a vulnerability and an edge indicates the possibility for an attacker to exploit two vulnerabilities in sequence. Weights on the edges of the graph represent the probability of receiving attacks on the specific *vulnerability chain.* This model has proven to be very successful in the literature with several applications to a number of cases such as industrial control systems [72, 73] and organization networks. [68] While attack graphs are a powerful method to reason over cyber-risk, the model parameters are assumed to be known or measured by other means. [64] Other models enumerating security events or vulnerabilities have also been proposed in the literature. [63, 60] Similarly, vulnerability estimation and assessment methodologies provide a quantitative way of measuring weaknesses and formulate impact estimations. [57, 65] However, the resulting estimations are widely regarded as unrealistic, as recently showed in scientific studies comparing vulnerability exploits with resulting metrics, [17, 19] and industry reports openly criticizing vulnerability models and measures as effective proxies for risk of attack. [74]

More advanced models consider the complex interactions between system vulnerabilities and attacker actions to devise so-called *time-to-compromise* models that provide a general framework to evaluate the probability of successful attack given certain starting preconditions on the system and the attacker. [69, 70] On the same line, but from an economic perspective, other approaches consider the interplay between attacker and defender to define game-theoretic models with the goal of deriving mixed equilibria (whose outcomes are defined in terms of impacts and probabilities) leading to different risks. [66, 75] However, these approaches leave the estimation of "likelihood of event" to the judgment of an expert that either directly sets probabilities of attack, or sets some parameters of the computational or game theoretical model that derives the quantitative probabilities as a function of that input. [68]

Patented methodologies often employ a mixture of system [76, 77] and game-theoretic [78] approaches to evaluate overall system risk. These models build on top of attack graphs and derive risk estimates based on graph traversing [76] or determination of minimal graph cut-sets [79] to block attacks out. Attack probabilities are obtained by employing one (or a combination thereof) of the methods presented in Table 4. [79, 80]

Overall, quantitative methodologies for cyber risk estimations estimate attack likelihoods by assuming an underlying distribution of attacks (e.g. a poissonian distribution) whose parameters are assessed by an 'oracle' (that, for example, estimates the number of expected occurrences λ). This is often a problem of lack of data models that can transform observations in predictions. For example, Cherdantseva et al. report *"In order to deal with the absence of historical data, some PRA methods rely on subjective data such as expert opinion. In some cases, expert opinion is more easily available and may even be more valuable than historical data."* [23] Indeed, even when security data is available (e.g. for network events), it is known to be extremely noisy [10] and fraught with errors of unknown size (see for example the discussion provided in [74]). Hence the need to "*[..] devote more attention to techniques for capturing, formalising and ultimately turning into numeric values expert knowledge*". [23]

This situation is worsened by the current lack of formal procedures to share incident data that may reveal mounting attack trends, or new attacks (something that the recent EU 2013/0027 NIS - Network and Information Security - Directive tries to address; similarly, the US NTIA and the US Department of Commerce's Internet Policy Task Force are currently addressing these issues in a set of call for comments with the Industry. [81]) Official guidelines such as NIST's Information Security Handbook [13] prescribe a qualitative assessment of risk over its quantitative estimation, as it allows the decision maker to operate within easily understood intervals that separate a 'likely' attack from an 'unlikely' attack, and a 'severe' consequence from a 'minor' consequence. Inevitably, moving from a quantitative to a qualitative framework causes some loss in resolution. Worse still, Cox [14, 15] shows that this loss in resolution may cause mis-categorization of risks and misguide the decision maker in believing that a certain risk is qualitatively higher than another, while the opposite is true quantitatively. In the absence of data, a qualitative assessment of probability of attacks is often necessary, although it is well known that 'expert assessments' of highly uncertain events are generally unreliable . [82, 83]

To address this, we propose a quantitative assessment methodology that allows the user to objectively estimate likelihood of (untargeted) attacks against his/her infrastructure.

Our method can be combined with any approach for the quantification of impact. Several studies on the quantification of impact exist, including impact from direct losses, [84, 85] technical impacts, [60] financial impacts, [86] and reputational and operational impacts. [87]

Table 4: Summary of quantitative methodologies for cyber-risk estimation. An overview of related patents is given in the Appendix, Tab. 13.

Ref.	Quantification methodology	Likelihood estimation
[60, 61]	Attack-surfaces measure the entry points that an attacker can exploit to breach the system. Attack surface estimations rely on vulnerability scanners data or vulnerability models (for a practical example, see [62]). Attack surfaces do not explicitly measure risk of attack, but assume that this is proportional to the number of identified entry points.	Probability of attack is assumed to (positively) correlate with attack surface ('attack opportunities'). For example Howard et al. evaluate "targets and enablers, channels and protocols, and access rights" to estimate attack likelihood. [61] The quantitative relation between attack surface and likelihood is however not defined.
[2, 63]	Attack graphs are graphs where a node is a vulnerability exploited by the attacker, and edges represent the occurrence of an attack leading the attacker from one vulnerability to the next. Probabilistic attack-graphs attach probabilities of exploitation as weights to those edges.	Attack graphs do not provide an indication on how to estimate probabilities of attacks. Some methodologies apply a Bayesian approach for the probability estimation based on an initial belief. No indication on the update function or information used for the update is provided. For example, Poolsappasit et al. propose a formal model of belief propagation on top of previous approaches, where prior probabilities are "subjectively assigned by the administrator". [64]
[57, 65]	Vulnerability-oriented methodologies estimate risk of attack by considering the characteristics of the vulnerabilities. For IT systems, several studies suggest the usage of the Common Vulnerability Scoring System (CVSS) metric [21] to make probability estimates. For example, in [64] the authors state "we use the metrics defined in NIST's Common Vulnerability Scoring System (CVSS) to estimate the attack likelihood".	The association between severity and likelihood estimations in the CVSS score is not substantiated by empirical evidence,[19, 17] and is not claimed by the standard itself. [21]
[66, 67]	Game-theoretic models consider the interplay between a strategic attacker and a defender to derive equilibrium points where the resulting mixed strategies are the outcome of the strategic game. The probabilities assigned to each strategy correspond to the probability of an attack following a certain mitigation action by the defender.	The estimation of "likelihood of event" is left to the judgement of an expert that either directly sets probabilities of attack, or sets some parameters of the computational or game theoretical model that derives the quantitative probabilities as a function of that input. [68]
[69, 70]	Time-to-compromise models are also used as proxies to perform risk estimations. Probability quantifications are possible based on the underlying model. For example, Henry and Haimes assume known vulnerability and exploit distributions for the estimation of the time required for the attack to succeed. [70]	The estimation of the probability distribution of attack are expert-based (see for example [71]).

5 A Quantitative Model for Likelihood of Cyber Attacks

The definition of Probabilistic Risk Assessment (PRA) by Ezell et al.[88] links the relation between risk, probability of attack attempt (*Attack*), probability of successful attack (*Comprimise*), and impact (*Impact*):

$$\begin{aligned} \mathit{Risk} = \mathit{Likelihood} \cdot \mathit{Impact} = \\ \mathit{Pr}(\mathit{Compr}|\mathit{Attack}) \cdot \mathit{Pr}(\mathit{Attack}) \cdot \mathit{Impact} \end{aligned} \tag{1}$$

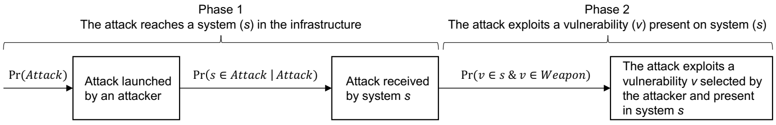
Brown et al. [89] have critiqued this definition, as they deem critical the interpretation of the term $\mathit{Pr}(\mathit{Attack})$, and consequently question the condition on the second term: these probabilities can not be reliably estimated without knowing the 'reason why' an attacker attempts the attack. This also assumes that *Impact* is a deterministic function. In contrast, the operational risk literature treats both Impact and Likelihood as uncertain variables by considering *compound risk distributions* that account for the associated uncertainty. [90] This is for example the standard Basel-II approach used in the banking and financial sectors. These distributions are of the type $\sum_{j=1}^{N} \mathit{Impact}_j$, where Impact_j is the randomly distributed severity of a single loss. Assuming independence between losses and frequency of occurrence (the standard insurance modeling approach used in Basel-II [90, def. 3.5, pp. 98]), one can rewrite Eq.1 as

$$\begin{aligned} \mathit{Risk} = \sum_{z=0}^{\infty} \mathit{Pr}_z(\mathit{Compromise}|\mathit{Attack}) \cdot \\ \mathit{Pr}_z(\mathit{Attack}) \cdot \sum_{j=1}^{z} \mathit{Pr}[\mathit{Impact}_j < x] \end{aligned} \tag{2}$$

where $\mathit{Pr}_z(\cdot)$ is the probability associated with z incidents, x is the total loss, and $\sum_{j=1}^{z} \mathit{Pr}[\mathit{Impact}_j < x]$ is the total probability of loss for z incidents. The risk estimate can then be obtained by deriving the convolution of the probability and impact distributions (for an introductory discussion see [90, Ch. 3.4.3, pp.102]). Whereas several studies on the quantification of impact exist, [84, 85, 60, 86, 87] current risk assessment methodologies lack of a well-specified method for the measurement of attack likelihood.[23]

As discussed in Section 2, in this work we specifically refere to *untargeted* attacks. While the dynamics driving the production of this class of cyber-attacks are not fully specified in the literature yet, [91] their empirical distribution is known or can be obtained from the data. For example, distribution of attacks against web software components seem to follow a log-normal distribution whereby for certain software types 95% of attacks are driven by as little as 5% of the software's vulnerabilities [46]. Hence, the frequency and impact distributions defined in Eq. 2 can be collapsed to a single probability estimate for a given number of events z. We can therefore specify the likelihood of attack for a certain system s in the organization's infrastructure as:

$$\begin{aligned} \mathit{Likelihood}_s = \mathit{Pr}(s \in \mathit{Compromise}|s \in \mathit{Breach}) \cdot \\ \mathit{Pr}(s \in \mathit{Breach}|s \in \mathit{Attack}) \cdot \\ \mathit{Pr}(s \in \mathit{Attack}|\mathit{Attack}) \cdot \mathit{Pr}(\mathit{Attack}) \end{aligned} \tag{3}$$


Phase 1
The attack reaches a system (s) in the infrastructure
Phase 2
The attack exploits a vulnerability (v) present on system (s)
Pr(Attack)
Attack launched by an attacker
Pr(s ∈ Attack | Attack)
Attack received by system s
Pr(v ∈ s & v ∈ Weapon)
The attack exploits a vulnerability v selected by the attacker and present in system s

Phase 1. The infrastructure is attacked with a certain prior probability $Pr(Attack)$ and targets a specific system s with probability $Pr(s \in Attack|Attack)$. The prior probability of attack is evaluated over the whole infrastructure and is therefore immaterial to calculate relative risks for each system. $Pr(s \in Attack|Attack)$ can be obtained by evaluating the exposure of the system to incoming traffic as reported by network sensors. *Phase 2.* Once the attack reaches a system s, the probability of successful compromise is equal to the probability of a vulnerability exploit being included in the attack ($v \in Weapon$) and the match between those and the vulnerabilities on system s ($v \in s$). The resulting probability, $Pr(v \in s \wedge v \in Weapon)$ is therefore the probability of successful attack.

Figure 1: Conceptual representation of the two-phases model and the relationship with its mathematical form.

where $Pr(s \in Attack|Attack)$ is the probability the attack against the organization will materialize as an attack on a particular system type, $Pr(s \in Breach)$ captures the chances of the attack to breach into the system, and $Pr(s \in Compromise)$ is the probability of the final successful compromise.

To compute these components from the technical infrastructure, empirical evidence suggests that if the automated attack tool has incorporated the exploit of a vulnerability present on the system, then success of the attack is almost certain. [92, 37, 47] Therefore the probability of a successful intrusion equals the probability that the appropriate combination of vulnerabilities is present on the attacked system ($v \in s$), and that the attacker actually weaponized v in his or her toolkit ($v \in Weapon$): $Pr(s \in Compromise|s \in Breach) \cdot Pr(s \in Breach|s \in Attack) \approx Pr(v \in s \wedge v \in Weapon|s \in Attack)$. As the presence of a technical vulnerability is an intrinsic property of the configuration, and is not dependent on the attacker selecting the particular system s, we have $Pr(v \in s \wedge v \in Weapon|s \in Attack) = Pr(v \in s \wedge v \in Weapon)$. We can then re-write Eq. 3 as:

$$\begin{aligned} Likelihood_s \approx{} & Pr(v \in s \wedge v \in Weapon) \cdot \\ & Pr(s \in Attack|Attack) \cdot Pr(Attack) \end{aligned} \tag{4}$$

The two distinct attack phases (*attack probing* and *attack delivery*) identified in the computer security literature [37, 28] and introduced in Section 2 naturally emerge from Eq. 4. Figure 1 visualizes a schematic representation of the two-phase attack process and its relation with the mathematical form expressed above.

An agreement on how to measure $Pr(Attack)$ still does not exist. Adopting a purely frequentist approach, $Pr(Attack)$ could be for example the fraction of malicious incoming network packets in which case we would obtain $Pr(Attack) \approx 0$. [10] If we consider the fraction of days with at least an attack we would obtain $Pr(Attack) \approx 1$. [93, 94] From the perspective of prioritizing risk treatments between systems *within* the organization, this value would be mostly immaterial. In this respect qualitative estimates might be appropriate and might as well take into account the motivation of the attackers as advo-

Table 5: Example of attack's likelihood assessment derived from the Eurocontrol Air Traffic Management Risk Tool Kit.

	Likelihood				
	Frequent	Likely	Occasional	Unlikely	Rare
Skills	Automated attack	Semi-automated attack	Attack needs re-engineering	Highly skilled attacker	Insider & skilled
Attack vector	Any	Known public technique	Marketed access	Limited access	Unique access
Profit	High	Significant	Modest	Little	None
Attention	Media coverage world-wide	National interest	Local interest	Little attention from media	No attention
Attack identification	Impossible to detect attack	Unlikely to detect	Likely to detect	Detection almost certain	Obvious
Prosecution	No chance	Little chance	Likely	High chance	Certainty

cated by Brown et al. [89] For example Table 5, derived from the Eurocontrol Air Traffic Management Risk Tool Kit, shows a fine grain classification considering different incentives and dis-incentives for attackers. Quantitative estimates can be obtained empirically by assuming a certain probability function of arrival of attacks (for example Binomial or Poissonian) and deriving the expected value by means of Monte Carlo simulation or equivalent approaches [90]. Naturally, the underlying distribution assumption is of central importance to obtain realistic estimates and is currently an unsolved problem in the cyber-security domain. [23].

In contrast, $P(s \in Attack|Attack)$ and $Pr(v \in s \wedge v \in Weapon)$ are technological measures and can therefore be objectively estimated using tools and procedures commonly available, and often mandated by compliance, in any complex-enough organization as well from the technological assumptions about the power of the attacker.

In absence of complete information on existing exploits (e.g. because the vendor of the security tool did not find the exploit yet), the defender can only assume that the attacker possesses an unknown set of vulnerabilities of size k that he/she may use out of the V_{tot} present in the target infrastructure.[2] Hence, the attack against system s will fail if the attacker has chosen a set of k exploits that does not include any vulnerability among the $V_{tot} - V_s$ that *do not* affect system s. More formally, the attack fails if the attacker chooses a set of attacks of size k from $\binom{V_{tot}-V_s}{k}$ out of the possible $\binom{V_{tot}}{k}$ choices he or she has. We can then estimate the probability of a successful attack for an attacker with

[2] Estimates for k can be derived from the literature. For example, the study of exploits kits as software artefacts by Kotov and Massacci [43] showed that each kit uses on average 11 exploits. Even an allegedly nation-state malware such as Stuxnet with 30 fully automated functionalities (including updating itself and communicating to the remote command and control server, etc.) only exploits 8 vulnerabilities overall. Similarly, the Duqu malware exploited one kernel vulnerability in Windows to breach the system and then exfiltrated and propagated itself in the network.

weaponizing power k as follows.

$$Pr(v \in s \wedge v \in Weapons | AttackerPower = k) = \\ 1 - \frac{\binom{V_{tot}-V_s}{k}}{\binom{V_{tot}}{k}} \approx 1 - \left(1 - \frac{V_s}{V_{tot}-k}\right)^k \tag{5}$$

The latter approximation can be derived by either using Stirling approximation for the binomial coefficient with $V_{tot} \gg V_s > 1$ or by simply drawing k vulnerabilities from V_{tot} with replacement.

This estimation can be refined if it is possible to discriminate between network exploitable vulnerabilities V_n and locally exploitable vulnerabilities V_l, as the former are typically used to establish just a breach of the system and the latter to gain complete control of it. For example, the recent RIG exploit kit only features $k_n = 6$ vulnerabilities[3] and the dropped malware may be executed thanks to some local mis-configuration (e.g. acquiring privileged access to the system, or vulnerabilities in the antivirus software - see Table 7), i.e. $k_l = 2$. In the past, the successful worm Conficker and subsequent iterations exploited a network vulnerability in Windows RPC service (MS08-067), propagated thanks to HTTP pull mechanisms, anonymous network shares and weak network share passwords ($k_n = 4$). Conficker could propagate locally by exploiting removable media content auto-execution ($k_l = 1$).[4] Hence we can break down the ability of attackers to weaponize k_n network facing vulnerabilities and k_l local vulnerabilities as the complement of the probability that the attack fails either because the *network attack* failed (Phase 1), or because the *local attack* failed (Phase 2) even if Phase 1 succeeded. This yields the following equation:

$$\begin{aligned} &Pr(v \in s \wedge v \in Weapons | Power = k_n + k_l) \\ &= 1 - \frac{\binom{Vn-Vn_s}{k_n}}{\binom{Vn}{k_n}} - \left(1 - \frac{\binom{Vn-Vn_s}{k_n}}{\binom{Vn}{k_n}}\right)\frac{\binom{Vl-Vl_s}{k_l}}{\binom{Vl}{k_l}} \\ &\approx\ 1 - \left(1 - \frac{Vn_s}{Vn-k_n}\right)^{k_n} - \left(1 - \frac{Vl_s}{Vl-k_l}\right)^{k_l} \\ &\quad + \left(1 - \frac{Vn_s}{Vn-k_n}\right)^{k_n}\left(1 - \frac{Vl_s}{Vl-k_l}\right)^{k_l} \end{aligned} \tag{6}$$

where $\frac{\binom{Vn-Vn_s}{k_n}}{\binom{Vn}{k_n}}$ denotes the probability that the attack fails in Phase 1, and $(1 - \frac{\binom{Vn-Vn_s}{k_n}}{\binom{Vn}{k_n}})\frac{\binom{Vl-Vl_s}{k_l}}{\binom{Vl}{k_l}}$ denotes the probability of failure in Phase 2 given a successful Phase 1 attack. As the attacker needs to reach at least one network vulnerability to breach the system and one local vulnerability at least to avoid detection, we assume $k_n \geq 1$ and $k_l \geq 1$.

To estimate the probability that an automated attack is actually directed towards a particular system s we leverage on the information gathered by perimeter sensors such as border firewalls or intrusion prevention systems (IPSs) and

[3] See for example the technical analysis at `http://www.kahusecurity.com/2014/rig-exploit-pack/`, *last visited June 2017.*

[4] See technical report at `https://www.icann.org/en/system/files/files/conficker-summary-review-07may10-en.pdf`, *last visited June 2017.*

intrusion detection systems (IDSs) which log unwanted or anomalous incoming traffic toward the organization.

Unfortunately, IDS technology is known to have a relatively low true detection rate and therefore can not be considered to be directly related to successful attacks [10] albeit vulnerability and port scans detected by IDS are known to be followed by attacks [95, 96]. Still, the relative distribution of alerts per system can give us a practical proxy measure for the probability of an attack being directed towards a system (type) s, given that an attack happened.

$$Pr(s \in \mathit{Attack}|\mathit{Attack}) \approx \frac{|\mathit{Alerts}_s|}{|\mathit{Alerts}|} \tag{7}$$

At this point we have all necessary information to calculate the quantitative value of $\mathit{Likelihood}_s$ for system s by using Equation (4) with the values from Equation (7), and the values for Equation (5) or (6) can extracted from the IT system infrastructure; overall, we consider:

1. *IPS/IDS Alerts*, both global and system specific, can be obtained from IPS and IDS data and can be used to evaluate system exposure to malicious traffic (Eq. 7).

2. *Vulnerabilities*, both global V_{tot} and system specific V_s (possibly broken down into network or locally exploitable), can be obtained from internal pen-testing scans or from VAs and can be used to evaluate a system's vulnerability to attacks (Eq. 5, 6).

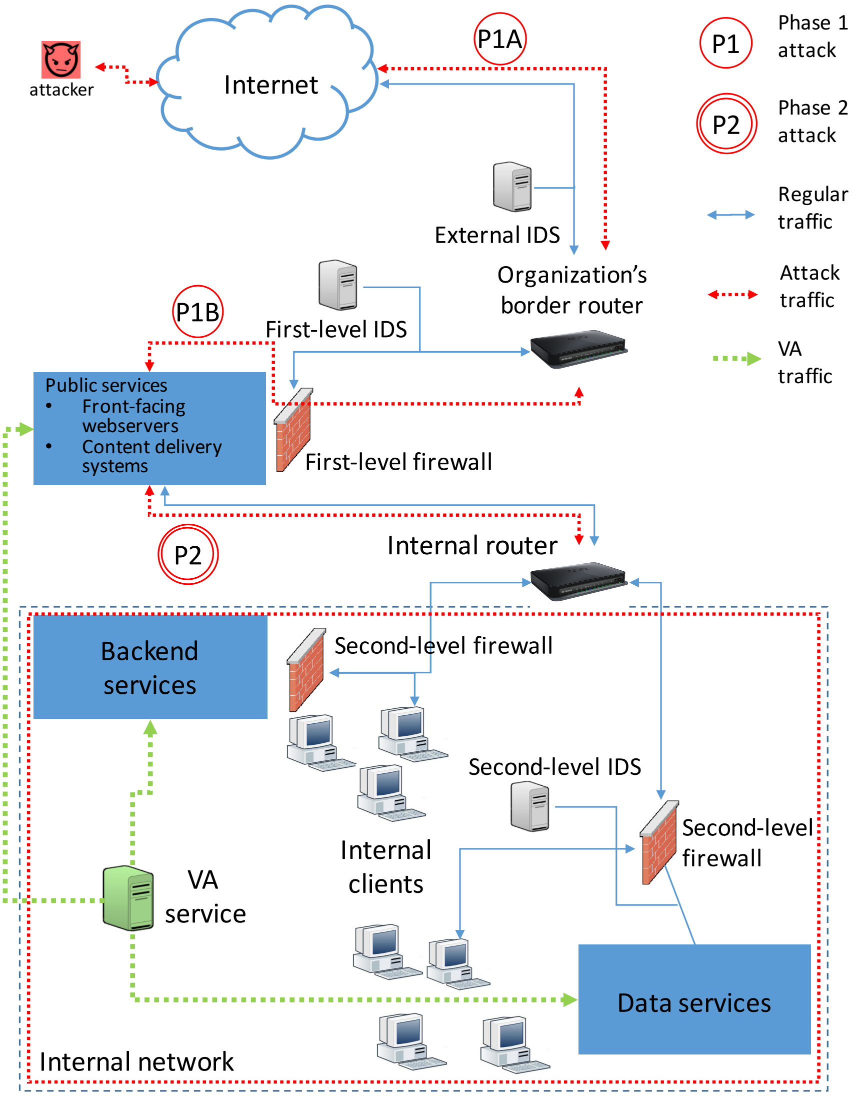
3. *Power of the attacker k* derived from analysis of malware and exploit kits in the wild or simply procured from intelligence services by security companies (Eq. 5, 6).

This data is generated by a cyclic process mandated by several IT security management standards such as ISO 27001 and 31000. The VA process is typically run periodically by network scanning tools, (a requirement defined by compliance, see for example Req. 11 of PCI-DSS [20]). IDSs and IPSs are constantly monitoring the network traffic and the underlying infrastructure generates periodic reports. [25] We illustrate how to concretely obtain this data from the technical infrastructure in the next section.

6 Infrastructural tools to evaluate system risk

Figure 2 exemplifies the considered attack dynamics on a typical network infrastructure with the key components of a security operations infrastructure. [26] The first entry barrier for an attacker is the border router or some first-level firewall that filters incoming traffic. The external IDS monitors all incoming traffic, including traffic that will be filtered by the first-level firewall and the border router. More specific filtering and traffic monitoring can be implemented at the second and lower levels of the network (e.g. to detect attacks against specific services).

The attacker may then initiate a probing phase of the network ($P1A$) in an attempt to obtain some information about the organization's infrastructure. Because these network scans often follow specific patterns, they are likely to be



Firewalls and IDSs are replicated at different levels in the network to monitor traffic towards specific subnets or allow for easier correlation between network and system events (e.g. the external IDS does not typically see the internal network as it sits outside). Further, a Vulnerability Assessment (VA) service is in place that scans configurations for known vulnerabilities in the organization's systems (scans represented by green arrows). An attacker has initially to probe the network (*P1A*) to collect information necessary to deliver the attack, e.g. operating system used by internet-facing systems or a spear-phished employee invoking an exploit kit (*P1B*). The attacker can further propagate the attack to internal systems, e.g. drop some self-propagating malware, or use the internal network interface of public services to bypass the first-level firewall (*P*2).

Figure 2: Schematics of a typical network infrastructure and attack scenario.

Table 6: Example of IDS alarms.

Time	Src IP	Dst IP	Src Port	Dst Port	Description	Status
3/3/16 11:50:01	ip_1	ip_2	39033	https/443	HTTP S Apache ClearText DoS	Detected event
7/3/16 20:17:20	ip_3	ip_4	58171	84	UDP Port Scan	Detected attack (vuln not scanned recently)
24/3/16 21:55:02	ip_5	ip_6	http/80	27710	Script Evasive Concatenation	Detected event
30/3/16 09:50:39	ip_7	ip_8	http/80	40231	JavaScript Packer Delta	Attack failure (blocked by appliance)

reported by perimetral sensors such as IDSs or firewalls. More in general, firewalls and IDSs will report all 'suspicious' incoming network packets ($P1B$), e.g. matching a signature for a known attack. Further, alarms reported by perimeter sensor technologies often report additional information such as detected threat, source and destination of the network packet, and time of report. This can be used to infer the exposure of an internal system to external threats. Table 6 reports an example of entries for an IDS alarm log. In this example, on the 7th of March 2016 an UDP port scan was detected and reported. Similarly, on the 30th of March an attack from a service acting on HTTP port 80 was blocked by the firewall. The alarms generated by perimetral sensors can give the assessor an objective proxy measure of the *exposure to attacks* of a certain network, subnetwork or network component. [95]

The second phase of the attack ($P2$) consists in the execution of malicious behaviour on the breached network host. This can be in the form of a malware software (e.g. a rootkit or ransomware), or system commands that may exfiltrate data or contact other systems (e.g. bypassing the first-level firewall in Fig. 2). This may be allowed by some mis-configuration (e.g. ineffective application filtering) or by some software vulnerability that the attack can exploit to freely operate on the system.

Table 7 reports an example of a VA report. Each entry is a vulnerability detected on an internal system in the organization, and indicates (where available) the type of affected service, and a description of the threat and its severity. For example, an attacking tool reaching system D could run system commands with admin privileges.

To link a Phase 1 network attack with its Phase 2 effects, it is necessary to link a system's exposure to network attacks with the system's vulnerabilities. In several configurations this may not be straightforward. This is because an outward-facing network interface reachable from the Internet masks the real, private IP addresses of the internal systems. This 'translation' from public to private addresses is usually performed by border routers (e.g. by Network Address Translators). Depending on the position of the IPS/IDS on the network topology, it may therefore be not possible to immediately map an alarm generated externally (e.g. by the external IDS in Fig. 2) with the vulnerable system.

The analyst can however reconstruct this translation process by using the routing table of the border router that the organization controls. The information necessary for the translation is typically of the form `Dst IP, Dst Port,`

Table 7: Example of a vulnerability assessment report for four systems.

Sys ID	Type	Service	Severity (1-3)	Vulnerability description
A	Local	-	3	A code execution vulnerability is present in some versions of Oracle Java SE and Java for Business.
B	Network	ftp (21/tcp)	2	A listening not necessary service has been detected on the host.
C	Network	snmp (161/udp)	2	A SNMP community name is set to the default value (e.g. public or private).
D	Local	-	2	Detected presence of user with administrative privileges
E	Local	-	2	A vulnerability exists in the scanning functionality in McAfee products that may allow malware to bypass scans.

Table 8: Example of information necessary to map perimeter logs with VA logs (if sensor is outside of the network's border).

Dst IP	Dst Port	Real Dst IP	Sys. ID	Real Dst Port
ip_2	ftp/21	ip_a	B	ftp/21
ip_4	MySQL /3306	ip_b	...	custom/555
ip_6	http/80	ip_c	...	http-alt/8080
ip_8	snmp/161	ip_d	D	snmp/161

`Real Dst IP, Real Dst Port`. Table 8 reports an example of information needed to map incoming traffic toward a certain public IP ($ip_{\{2,4,6,8\}}$) and port with the correct real destination.

If this information is not immediately available or is difficult to gather, it is still possible to approximate it by looking at classes of functionally similar systems rather than individual systems. For example, in some cases systems might be multiplexed for load balancing to multiple, identical virtual machines. It is thus reasonable to assume that whichever duplexed system the attack was directed to, it would be affected by the same vulnerabilities as any other system of the same type. For example, the probability of an infection ($Pr(s \in Compromise|s \in Attacks)$) directed toward a web server (http/80) would likely be the same regardless of the actual system it reaches, as most webservers would be configured very similarly or identically. In this scenario we can link the destination port of the incoming traffic (`Dest Port` in Table 6) with the service available on the system scanned by the VA (`Type: Network` in Table 7). Vulnerabilities not associated to a network service are local vulnerabilities.

Table 9 gives a summary overview of the information used to estimate the probabilities of attack and infection and the relevant infrastructural technology needed to gather that information.

Table 9: Summary of probability estimates and of the relevant information used for the estimation.

Estimate	Technology	Alert info	Topology-dependent
$P(Attack)$	None	Security expert's perception of attack likelihood	No
$P(s \in Attack \mid Attack)$	IDS	Suspicious traffic directed toward an internal or external system	Yes. Outward-facing sensors may not be directly correlated with internal systems. Sensors with specialised functionalities (e.g. detect attacks against SQL) may be weighted differently than non-specialised sensors.
	Firewalls/ IPS	Unwanted traffic or application protocol usage detected.	Yes. Depending on position in the network, firewalls may provide more detailed traffic analysis (e.g. static vs stateful filtering vs application filtering).
$Pr(v \in s \mid s \in Attack)$	VA	A vulnerability on the system exist that could lead to data exfiltration or other security breaches.	No.

7 Application example: the case of a large financial organization

We now consider the application of this methodology to the case of Company, an anonymized large financial organization. This analysis reports the example of the organization's Intrusion Detection System and Vulnerability Assessment data for the month of March 2016. The IDS data reports 10^6 fired alarms, generated from more than 5000 unique IP addresses. The VA reports data relative to 376 unique systems inside the organization.

To minimize the disclosure of potentially sensitive information, we report normalized quantities for each system and aggregate IDS and VA data by service type (the coarsest granularity described in Section 6). We aggregate network services to nine service types. Table 14 in the Appendix describes the nine categories. To map each TCP port to a service type category as previously described, we extended the Common Ports table provided by Microsoft[5]. Table 10 reports the distribution of the unique services running on systems in the organization (a system may run more than one service). At a first approximation we map one IP address to one (either physical of virtual) system. The organization's infrastructure has roughly the same number of Chat, Http, Infrastr, and RemCtrl services. Share, Mail, SQL, services are the second

[5] Port Assignments for Commonly-Used Services `https://msdn.microsoft.com/en-us/library/cc959833.aspx`, *last visited June 2017*

Table 10: Number of service types on unique systems.

System service type	Occurrences
Auth	2.59%
Chat	14.92%
Http	15.44%
Infrastr	12.63%
Mail	5.32%
RemCtrl	16.22%
RPC	2.42%
Share	7.66%
SQL	6.66%
Other	16.13%
tot.	**100.00%**

Table 11: Service exposure by relative frequency of IDS alarms per system type

System Type	IDS Alert Frequency
RPC	8.54%
SQL	0.39%
Mail	0.47%
Http	89.63%
Infrastr	0.39%
RemCtrl	0.26%
Chat	0.15%
Share	0.03%
Auth	0.14%

most common; RPC, SQL, and Auth are the fewest. Uncategorized services (Other) are about 16% of the total. Other are services that rely on the UDP network protocol and/or use non-default port numbers. As UDP typically supports most TCP applications (Domain Name resolution Services being an exception), we do not count UDP services as these are already likely accounted for as TCP services.

7.1 Data analysis

In Table 11 we report the exposure to network attacks of defined service types. As expected, the most suspicious network activity is directed toward Http services. They are typically outward-facing and therefore more easily identifiable by the attacker. RPC services account for the second largest fraction of suspicious incoming traffic; these services allow remote systems (e.g. owned by the attacker) to interact with procedures implemented locally, and potentially execute arbitrary, privileged actions on the target systems. The remaining services receive substantially fewer network requests raising alarms. Among these, SQL, Mail, Infrastr and RemCtrl services are cumulatively exposed to about 1% of the incoming traffic. SQL and Mail services are services that typically need to be exposed to external network traffic for their normal functionality. In contrast, Infrastr and RemCtrl services are typically inward-facing, meaning that it is more difficult for an attacker to reach them from outside the network.

Table 12: Overall distribution of vulnerabilities per system type.

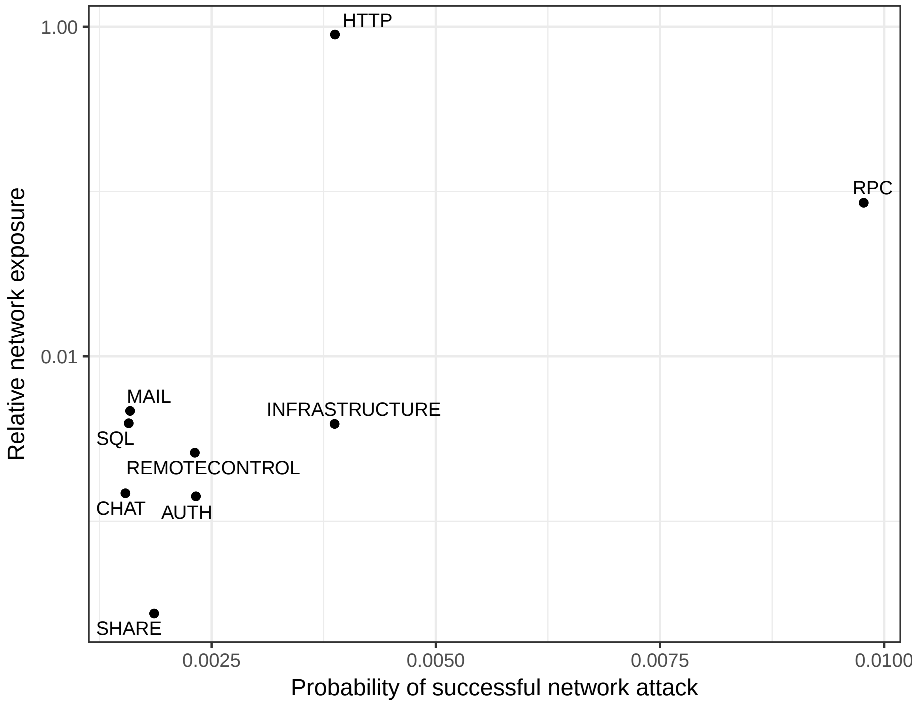
Sys. type	% of vulnerabilities
AUTH	0.46%
CHAT	1.73%
HTTP	4.52%
INFRASTR	3.69%
MAIL	0.64%
REMCTRL	2.83%
RPC	1.79%
SHARE	1.07%
SQL	0.79%
LOCAL	73.88%
OTHER	8.6%

Yet, potentially malicious traffic toward these resources reaches the external border of the network.

Table 12 reports the overall distribution of vulnerabilities per system type. The category LOCAL indicates vulnerabilities that do not belong to network-reachable systems, and that can therefore be only exploited by an attacker that has already gained local access to the system. Overall, the largest fraction of vulnerabilities are of this type. Hence, the attack surface exposed only locally to the attacker is, on average, higher than the network attack surface. This is desirable as a reduced network attack surface minimizes the likelihood of breach, and indicates the employment of good practices by the organization. Among network vulnerabilities, HTTP and INFRASTR services share the highest fraction.

For illustrative purposes, we now consider an attacker that launches a spear-phishing attack against an employee, and is capable of attacking $k_n = 8$ network vulnerabilities [43], and relies on the employee's administrative privileges and antivirus misconfiguration to run the malware locally ($k_l = 2$). The attacking tool has therefore power $k = 10$. The exposure of system services to Phase 1 attacks can be visualized by plotting the relation between exposure to malicious network traffic and network vulnerabilities that can be remotely exploited. Figure 3 plot this relationship. Each dot represents a service type, and its position on the graph indicates the relation between network alarms to which it is exposed (vertical axis, logarithmic) and the probability that an attacker of network power $k = 8$ successfully breaches the system. The farther toward the top right corner a service type is, the higher the associated likelihood according to Eq. (7). Services on the bottom left of the plot have low exposure and low probability of first breach. For example, RPC services suffer from the highest probability of a network breach ($\approx 1\%$), but are exposed to an order of magnitude less malicious connections than HTTP services. In contrast, HTTP services receive the largest fraction of malicious traffic, but face a relatively small probability of breach.

We now consider the Phase 2 risk of a propagation or escalation of the attack that impacts the organization. The success probability of this attack phase is proportional to the number of additional local vulnerabilities that the attacker can exploit on the breached system. This depends on the specific configuration of the attacked system. Figure 4 reports the probability of a successful network attack (x-axis) versus the probability of a successful local breach subsequent to



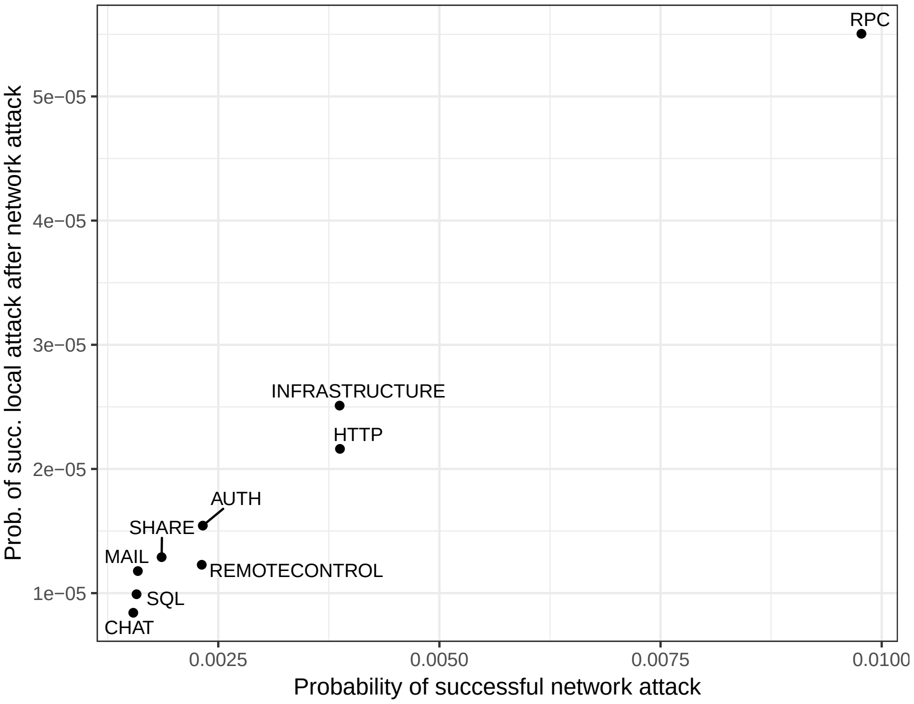
Service exposure to malicious network traffic (y-axis, log scale) against probability of a successful network attack for an attacker with network power $k_n = 8$: $Pr(v_n \in s \wedge v_n \in Weapons|Power = k_n = 8)$ (Phase 1). Each dot represents a service type. The higher the network exposure of the service type and the higher the probability of success, the higher the chance of a successful network attack.

Figure 3: Exposure and success probability of network attacks.

the network attack (y-axis). RPC systems show the highest overall probability of breach ($\approx$ 0.005%, i.e. one out of twenty thousand automated attack is expected to be successful), as opposed to Chat systems for which the probability of breach is a fifth of RPC's.

Further, we can visualize the discrepancy between the proposed risk estimation and current practice. Figure 5 shows a bubble plot of the classification yielded by the quantitative proposed methodology (y-axis), and the classification yielded by the qualitative risk assessment currently employed in the organization (based on ISO 27001:2013) on the x-axis as described in Section 3. Circle size is proportional to the number of systems in the service type. Estimations along the diagonal are similar for both methods. Chat, SQL, and RemCtrl systems are classified similarly by both the qualitative and quantitative methodology. Auth, Mail, and Share services are assigned the highest risk level of all services by the qualitative approach. On the contrary, the proposed quantitative risk estimation assigns a minimum risk level to these services. Similarly, RPC services are qualitatively assigned a 'low risk' profile, while quantitatively they are assigned the highest risk level among all services. This inversion in the assigned risk level is coherent with what previously predicted by Cox, [14] and may lead to a systematic mis-allocation of resources.[6]

[6] This analysis does not report the final 'risk estimate' as this depends on the impact of an



Probability of a network breach (x axis) for an attacker with up to $k_n = 8$ weaponized network vulnerabilities (Phase 1 attack)versus the probability of a local breach (y axis) which exploits up $k_l = 2$ local vulnerabilities (Phase 2 attack). Each dot represents a service type in the organization. Highly-critical service types on the top-right of the graph likely require action.

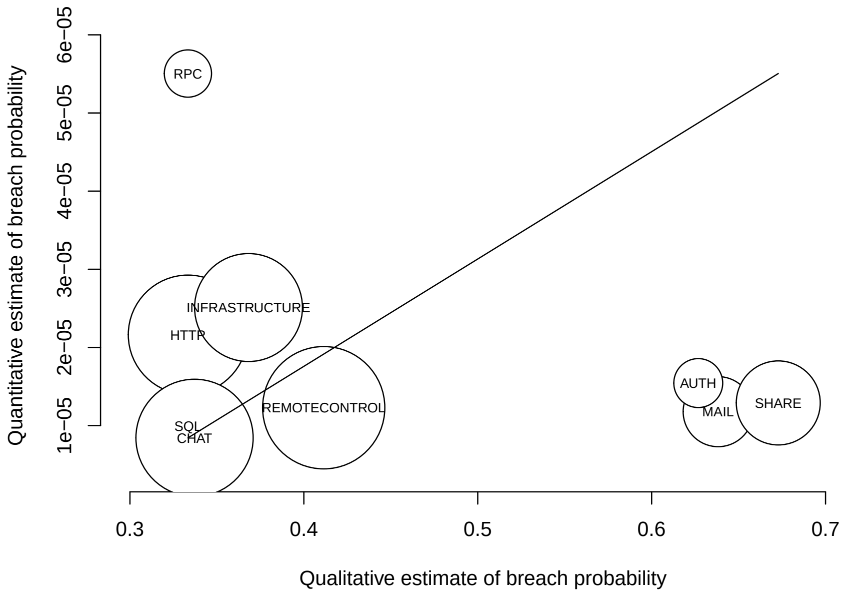
Figure 4: Probability of breach by service type.

Figure 6 reports a contour plot of the the final probability of attack $Pr(v \in s \wedge v \in Weapons|Power = k_n + k_l)$ for varying attack powers. The values for v_n and v_l are those of the proposed case study and are here distributed uniformly among system types. Displayed values are therefore only realistic, and do not represent the real risk profile of Company. The power k of the attacker may be set by the organization depending on its risk appetite. Even a powerful attacker with twelve network vulnerabilities and seven local vulnerabilities achieves an average probability of success of 2%. More specific assessments can be run for each service type by considering the known levels of v_n and v_l for the specific system as opposed to their average.

7.2 Application of the model to other scenarios

Organizations can be very different one from the other; for example, many large organizations may have WANs spanning the whole globe and others may have much simpler structures localized in a single city or even building. The generality of the presented model allows our methodology to be applied to a large number of real settings by capturing the main aspects of an attack. For exam-

attack on the systems. However this is immaterial for the discussion at hand as all systems will be multiplied by the same value for both quantitative and qualitative probability estimates.



Circle size is proportional to the number of distinct systems in each service type. The quantitative risk estimation is calculated for an attacker with power $k_l = 2$ and $k_n = 8$. The diagonal represents situations whereby the proposed and the current risk metrics yield the same categorization of risk in the proposed case study. Auth, Mail and Share services probability estimates are significantly overestimated by the qualitative approach. RPC probability of breach is significantly underestimated.

Figure 5: Bubble plot of probability of breach estimations per service type.

ple, the network structure exemplified in Figure 2 can be extended to include additional subnets or virtual lans (e.g. controlled by managed switches). Still, the effects of different network structures is encoded naturally in the meaning of IDS and security event data recorded by the sensors: if security control measures are in place to limit access to a system, this will necessarily be reflected in a low rate (or null rate) of alarms towards that system, i.e. in our model, $Pr(s \in attack|attack) \approx 0 \rightarrow Risk_s \approx 0$. Combining network effects can also be useful; however, in practice most IDS data is aggregated (e.g. on a periodic basis by the managing infrastructure).

Integration with other risk models Our model can be integrated in any system-level risk assessment methodology following the directions indicated by Haimes [97] (the limitations of which are well discussed by Aven [27]). For example, the estimation of likelihood of attack from our model can be used to weight a system's attack surface by considering its exposure to external attacks. Similarly, attack graphs incorporating network topologies can use computed probabilities as the prior distribution in place of the expert assessment. Expert judgment can then be used to update those probabilities with additional evidence or considerations (e.g. how difficult a vulnerability exploitation is).

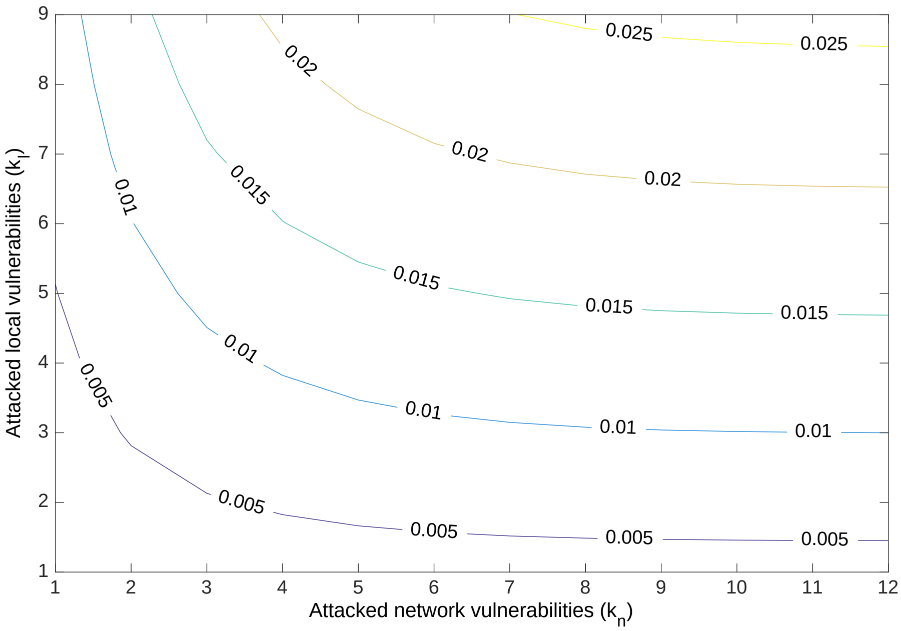


Figure 6: Probability of attack at varying the number of k_l local and k_n network weaponized vulnerabilities

Black swan events Our model does not refer to rare security events. There are other techniques that can be applied to black swan scenarios. Our approach could be generalized by considering the adversarial nature of the attacker, following the guidelines proposed in the literature. [29] The study of these events requires the joint analysis of the defender and attacker's strategies, and need consider the information asymmetries that exist between the two players in the definition of the equilibrium conditions. For example, the 'unknown unknowns' of a black swan attack (e.g. a 0-day vulnerability known by the attacker but not the defender) require the distinction between uncertainty and probability that is still part of the debate in the risk analysis literature. [59] This may require the investigation of attacker motives (e.g. to appropriately define an utility function), and the definition of the conditions under which the attacker can be regarded as *rational*. For example, very powerful attackers such as governments may not aim at maximising a specific utility function, as opposed to using cyber-attacks to (un)balance political relations between countries or other political and economical forces. [98]

7.3 Limitations

The example application proposed in Section 7 should not be interpreted as evidence that the quantitative methodology proposed in this paper is a better approximation of reality than traditional qualitative assessments. Unfortunately this would require ground truth data that does not typically exist or is very limited in nature. This limitation is intrinsic to any risk assessment methodology and is one of the problems behind the transition, historically, from traditional risk assessment and quantitative risk assessment. [4]

Differently, our model results (as depicted in Figure 5) highlight the difference between a replicable and objective criteria for risk quantification, and a qualitative measurement based on expert judgement, and stresses the resulting relative distances in risk levels. This directly informs the process of introducing quantitative methods for risk assessment in standard practices as it already happened in the Nuclear Energy sector in the seventies and the space industry later on. Apostolakis provides and excellent discussion on this. [4]

8 Conclusions

The current industry adoption of qualitative risk matrices, recommended by world-wide standards such as ISO 27001:2013 and NIST 800-30 and related 'handbooks', makes cyber-risk assessment essentially a function of the expert's belief that a particular attack will happen in the future. Assessments based solely on expert opinions are known to be biased [82] and may lead to hard to compare risk assessment. [55] Further, qualitative evaluations of risk using risk matrices can lead to decisions that are systematically opposite to those indicated by a quantitative measure of risk. [14, 15] This may lead to suboptimal resource allocation. [99, 100]

Our methodology proposes a quantitative way to evaluate likelihood of untargeted attacks. This mitigates the 'loss of resolution' caused by the employment of risk matrices discussed by Cox [14], and provides *comparable* risk measures between systems and organizations. Most importantly, this measure is generated by technical data that all medium-large organizations already have in their infrastructure. This data is currently often used in an unstructured way to either generate automatic reports on vulnerability severity, or to try to traceback known incidents. Our methodology proposes to correlate this data to measure on one side the exposure of a system to potential attacks, and on the other the opportunities that a successful attack has to breach a vulnerable system and escalate to the infrastructure. By enabling users in performing objective estimations of risk, our methodology makes a step forward toward the establishment of comparable measures for security. [101, 102]

To compute an organization absolute risk there is still the need to estimate the probability that the organization is attacked $Pr(Attack)$. For the purpose of prioritization within the organization [99] this is not necessary as all systems would be subject to the same value. Its full individual assessment might also not be necessary for *untargeted* attacks, which are the focus of this paper, so that one might calculate its value by sector (e.g. financial institutions or small enterprises). The calculation of $Pr(Attack)$ is instead necessary to achieve *comparable measures of risk* and to provide a baseline to assess the risk of *targeted* attacks. This requires the definition of models that jointly evaluate attacker's and defender's strategies: [89]several independent studies showed that most attacks are driven by a handful of vulnerabilities only, suggesting that attackers *choose* vulnerabilities to exploit as opposed to launch attacks drawn randomly from a pool of exploits for all vulnerabilities. [47, 103, 46] Capturing these aspects may require to integrate socio-economic models to evaluate attacker's incentives in marketing or buying a new vulnerability [91, 102] or choosing a target. [89] We consider these aspects for future work.

Acknowledgements

This work has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 285223 (SECONOMICS) and from the Italian PRIN Project TENACE.

References

[1] Van Goethem T, Chen P, Nikiforakis N, Desmet L, Joosen W. Large-scale security analysis of the web: Challenges and findings. In: Trust and Trustworthy Computing. Springer; 2014. p. 110–126.

[2] Wang L, Islam T, Long T, Singhal A, Jajodia S. An Attack Graph-Based Probabilistic Security Metric. In: Proceedings of the 22nd IFIP WG 11.3 Working Conference on Data and Applications Security. vol. 5094 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg; 2008. p. 283–296.

[3] Chen Py, Kataria G, Krishnan R. Correlated failures, diversification, and information security risk management. MIS Quaterly-Management Information Systems. 2011;35(2):397–422.

[4] Apostolakis GE. How useful is quantitative risk assessment? Risk Analysis. 2004;24(3):515–520.

[5] Lomnitz C. Global tectonics and earthquake risk. vol. 5. Elsevier; 2013.

[6] Cavusoglu H, Mishra B, Raghunathan S. The value of intrusion detection systems in information technology security architecture. Information Systems Research. 2005;16(1):28–46.

[7] Khorshidi HA, Gunawan I, Ibrahim MY. Data-Driven System Reliability and Failure Behavior Modeling Using FMECA. IEEE Transactions on Industrial Informatics. 2016;12(3):1253–1260.

[8] Susto GA, Schirru A, Pampuri S, McLoone S. Supervised aggregative feature extraction for big data time series regression. IEEE Transactions on Industrial Informatics. 2016;12(3):1243–1252.

[9] Valeur F, Vigna G, Kruegel C, Kemmerer RA. Comprehensive approach to intrusion detection alert correlation. IEEE Transactions on dependable and secure computing. 2004;1(3):146–169.

[10] Axelsson S. The base-rate fallacy and the difficulty of intrusion detection. ACM Transactions on Information and System Security (TISSEC). 2000;3(3):186–205.

[11] Pauli D. Register T, editor. Hackers tear shreds off Verizon's data breach report top 10 bug list. The Register; 2016. [online] `http://www.theregister.co.uk/2016/05/12/verizon_dbir_criticised/`. Available from: `\url{http://www.theregister.co.uk/2016/05/12/verizon_dbir_criticised/}` [cited Nov 2016].

[12] Martin B. OSVDB, editor. A Note on the Verizon DBIR 2016 Vulnerabilities Claims. OSVDB; 2016. [online] https://blog.osvdb.org/2016/04/27/a-note-on-the-verizon-dbir-2016-vulnerabilities-claims/. Available from: https://blog.osvdb.org/2016/04/27/a-note-on-the-verizon-dbir-2016-vulnerabilities-claims/ [cited Nov 2016].

[13] Bowen P, Hash J, Wilson M. Information security handbook: a guide for managers. NIST; 2006.

[14] Anthony Tony Cox L. What's wrong with risk matrices? Risk analysis. 2008;28(2):497–512.

[15] Duijm NJ. Recommendations on the use and design of risk matrices. Safety Science. 2015;76:21–31.

[16] Team CS. Common Vulnerability Scoring System v3.0: Specification Document. First.org; 2015.

[17] Bozorgi M, Saul LK, Savage S, Voelker GM. Beyond Heuristics: Learning to Classify Vulnerabilities and Predict Exploits. In: Proceedings of the 16th ACM International Conference on Knowledge Discovery and Data Mining; 2010. p. 105–114.

[18] Houmb SH, Franqueira VN, Engum EA. Quantifying security risk level from CVSS estimates of frequency and impact. Journal of Systems and Software. 2010;83(9):1622–1634.

[19] Allodi L, Massacci F. Comparing vulnerability severity and exploits using case-control studies. ACM Transaction on Information and System Security (TISSEC). 2014 8;17(1).

[20] Council PCISS. PCI-DSS, editor. PCI Data Security Standard (DSS): Requirements and Security Assessment Procedures. PCI-DSS; 2010. Available from: https://www.pcisecuritystandards.org/documents/pci_dss_v2.pdf.

[21] org F. Common Vulnerability Scoring System v3.0: Specification Document. FIRST, Available at http://www.first.org/cvss; 2015.

[22] Giacalone M, Mammoliti R, Massacci F, Paci F, Perugino R, Selli C. Security triage: A report of a lean security requirements methodology for cost-effective security analysis. In: Proc. of ACM/IEE ESEM'14; 2014. p. 25–27.

[23] Cherdantseva Y, Burnap P, Blyth A, Eden P, Jones K, Soulsby H, et al. A review of cyber security risk assessment methods for SCADA systems. computers & security. 2016;56:1–27.

[24] Kunreuther H. Risk Analysis and Risk Management in an Uncertain World. Risk Analysis. 2002;22(4):655–664. Available from: http://dx.doi.org/10.1111/0272-4332.00057.

[25] Kelley D, Moritz R. Best Practices for Building a Security Operations Center. Information Systems Security. 2006;14(6):27–32.

[26] Jacobs P, Arnab A, Irwin B. Classification of Security Operation Centers. In: 2013 Information Security for South Africa; 2013. p. 1–7.

[27] Aven T. On some recent definitions and analysis frameworks for risk, vulnerability, and resilience. Risk Analysis. 2011;31(4):515–522.

[28] Ransbotham S, Mitra S. Choice and Chance: A Conceptual Model of Paths to Information Security Compromise. Information Systems Research. 2009;20.

[29] Rios Insua D, Rios J, Banks D. Adversarial risk analysis. Journal of the American Statistical Association. 2009;104(486):841–854.

[30] Rao NS, Poole SW, Ma CY, He F, Zhuang J, Yau DK. Defense of Cyber Infrastructures Against Cyber-Physical Attacks Using Game-Theoretic Models. Risk Analysis. 2015;.

[31] Merrick J, Parnell GS. A comparative analysis of PRA and intelligent adversary methods for counterterrorism risk management. Risk Analysis. 2011;31(9):1488–1510.

[32] Brown GG, Cox Jr LAT. How probabilistic risk assessment can mislead terrorism risk analysts. Risk Analysis. 2011;31(2):196–204.

[33] Verizon. Verizon, editor. 2016 Data Breach Investigation Report. Verizon; 2016. [online] http://www.verizonenterprise.com/verizon-insights-lab/dbir/2016/. Available from: http://www.verizonenterprise.com/verizon-insights-lab/dbir/2016/ [cited Nov 2016].

[34] Bilge L, Dumitras T. Before we knew it: an empirical study of zero-day attacks in the real world. In: Proceedings of the 19th ACM Conference on Computer and Communications Security (CCS'12). ACM; 2012. p. 833–844.

[35] Paté-Cornell E. On "Black Swans" and "Perfect Storms": risk analysis and management when statistics are not enough. Risk analysis. 2012;32(11):1823–1833.

[36] Tankard C. Advanced persistent threats and how to monitor and deter them. Network security. 2011;2011(8):16–19.

[37] Grier C, Ballard L, Caballero J, Chachra N, Dietrich CJ, Levchenko K, et al. Manufacturing compromise: the emergence of exploit-as-a-service. In: Proceedings of the 19th ACM Conference on Computer and Communications Security (CCS'12). ACM; 2012. p. 821–832.

[38] Provos N, Mavrommatis P, Rajab MA, Monrose F. All your iFRAMEs point to Us. In: Proceedings of the 17th USENIX Security Symposium; 2008. p. 1–15.

[39] Nicholson A, Webber S, Dyer S, Patel T, Janicke H. SCADA security in the light of Cyber-Warfare. Computers & Security. 2012;31(4):418–436.

[40] Verizon. Verizon, editor. 2015 Data Breach Investigations Report. Verizon; 2015. [online] http://www.verizon.com/about/news/2015-data-breach-report-info/. Available from: http://www.verizon.com/about/news/2015-data-breach-report-info/ [cited Nov 2016].

[41] Symantec. Symantec Corporation Internet Security Threat Report 2013. 2012 Trends; 2013. 18.

[42] TrendMicro. TrendMicro, editor. Exploit Kit. TrendMicro; 2016. [online] http://www.trendmicro.com/vinfo/us/security/definition/exploit-kit. Available from: http://www.trendmicro.com/vinfo/us/security/definition/exploit-kit [cited Nov 2016].

[43] Kotov V, Massacci F. Anatomy of Exploit Kits. Preliminary Analysis of Exploit Kits as Software Artefacts. In: Proc. of ESSoS 2013; 2013. p. 181–196.

[44] Sophos. Sophos, editor. Location-based threats: How cybercriminals target you based on where you live. Sophos; 2016. [online] https://blogs.sophos.com/2016/05/03/location-based-ransomware-threat-research/. Available from: https://blogs.sophos.com/2016/05/03/location-based-ransomware-threat-research/ [cited Nov 2016].

[45] Erickson J. Hacking: the art of exploitation. No Starch Press; 2008.

[46] Allodi L. The Heavy Tails of Vulnerability Exploitation. In: Proceedings of the 2015 Engineering Secure Software and Systems Conference (ESSoS'15); 2015. p. 133–148.

[47] Nayak K, Marino D, Efstathopoulos P, Dumitraş T. Some Vulnerabilities Are Different Than Others. In: Proceedings of the 17th International Symposium on Research in Attacks, Intrusions and Defenses. Springer; 2014. p. 426–446.

[48] Bhatt S, Manadhata PK, Zomlot L. The Operational Role of Security Information and Event Management Systems. IEEE Security Privacy. 2014 9;12(5):35–41.

[49] ISO/IEC. 27005:2011–Information Technology–Security Techniques–Information Security Management Systems–Requirements. ISO/IEC; 2011.

[50] ISO/IEC. 31000:2009 – Risk Management. ISO/IEC; 2009.

[51] Stoneburner G, Goguen A, Feringa A. Risk management guide for information technology systems. NIST; 2002.

[52] ISACA. COBIT 5: A Business Framework for the Governance and Management of Enterprise IT. ISACA; 2012.

[53] Sherwood J, Clark A, Lynas D. Enterprise security architecture: a business-driven approach. Backbeat Books; 2005.

[54] Curtis P, Carey M. Risk assessment in practice. Committee of Sponsoring Organizations of the Treadway Commission; 2012. Available on www.coso.org.

[55] Aven T, Cox LA. National and Global Risk Studies: How Can the Field of Risk Analysis Contribute? Risk Analysis. 2016;36(2):186–190. Available from: http://dx.doi.org/10.1111/risa.12584.

[56] Quinn SD, Scarfone KA, Barrett M, Johnson CS. SP 800-117. Guide to Adopting and Using the Security Content Automation Protocol (SCAP) Version 1.0. NIST; 2010.

[57] Naaliel M, Joao D, Henrique M. Security Benchmarks for Web Serving Systems. In: Proceedings of the 25th IEEE International Symposium on Software Reliability Engineering (ISSRE'14); 2014. p. 1–12.

[58] Patcha A, Park JM. An overview of anomaly detection techniques: Existing solutions and latest technological trends. Computer networks. 2007;51(12):3448–3470.

[59] Aven T, Zio E. Foundational issues in risk assessment and risk management. Risk Analysis. 2014;34(7):1164–1172.

[60] Manadhata PK, Wing JM. An Attack Surface Metric. IEEE Transactions on Software Engineering. 2011;37:371–386.

[61] Howard M, Pincus J, Wing JM. Measuring Relative Attack Surfaces. Computer Security in the 21st Century. 2005;p. 109–137.

[62] Manadhata PK, Wing JM, Flynn MA, McQueen MA. Measuring the Attack Surfaces of Two FTP Daemons. In: Proceedings of the 2nd Workshop on Quality of Protection; 2006. p. 3–10.

[63] Baiardi F, Telmon C, Sgandurra D. Hierarchical, model-based risk management of critical infrastructures. Reliability Engineering & System Safety. 2009;94(9):1403–1415.

[64] Poolsappasit N, Dewri R, Ray I. Dynamic security risk management using Bayesian attack graphs. IEEE Transactions on Dependable and Secure Computing. 2012;9(1):61–74.

[65] Ten CW, Manimaran G, Liu CC. Cybersecurity for critical infrastructures: attack and defense modeling. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans. 2010;40(4):853–865.

[66] Serra E, Jajodia S, Pugliese A, Rullo A, Subrahmanian VS. Pareto-Optimal Adversarial Defense of Enterprise Systems. ACM Trans Inf Syst Secur. 2015 Mar;17(3):11:1–11:39. Available from: http://doi.acm.org/10.1145/2699907.

[67] Hewett R, Rudrapattana S, Kijsanayothin P. Cyber-security analysis of smart grid SCADA systems with game models. In: Proceedings of the 9th Annual Cyber and Information Security Research Conference. ACM; 2014. p. 109–112.

[68] Dewri R, Poolsappasit N, Ray I, Whitley D. Optimal Security Hardening Using Multi-objective Optimization on Attack Tree Models of Networks. In: Proceedings of the 14th ACM Conference on Computer and Communications Security. CCS '07. ACM; 2007. p. 204–213. Available from: `http://doi.acm.org/10.1145/1315245.1315272`.

[69] Holm H. A Large-Scale Study of the Time Required to Compromise a Computer System. IEEE Transactions on Dependable and Secure Computing. 2014;11(1):2–15.

[70] Henry MH, Haimes YY. A comprehensive network security risk model for process control networks. Risk analysis. 2009;29(2):223–248.

[71] McQueen MA, Boyer WF, Flynn MA, Beitel GA. Quantitative Cyber Risk Reduction Estimation Methodology for a Small SCADA Control System. Proceedings of the 39nd Hawaii International Conference on System Sciences. 2006;9:226.

[72] Byres EJ, Franz M, Miller D. The use of attack trees in assessing vulnerabilities in SCADA systems. In: Proceedings of the international infrastructure survivability workshop. Citeseer; 2004. p. 3–10.

[73] Ten CW, Liu CC, Govindarasu M. Vulnerability Assessment of Cybersecurity for SCADA Systems Using Attack Trees. In: Power Engineering Society General Meeting, 2007. IEEE; 2007. p. 1–8.

[74] Christey S, Martin B. BlackHat, editor. Buying into the bias: why vulnerability statistics suck. BlackHat; 2013. Black Hat Conference, `https://www.blackhat.com/us-13/archives.html\#Martin`. Available from: `\url{https://www.blackhat.com/us-13/archives.html\#Martin}`.

[75] Grossklags J, Christin N, Chuang J. Secure or insure?: a game-theoretic analysis of information security games. In: Proceedings of the 17th international conference on World Wide Web. ACM; 2008. p. 209–218.

[76] Beresnevichiene Y, Tsifountidis F. Network security risk assessment. Google Patents; 2014. US Patent 8,650,637.

[77] Kang PY, Sim WT, Kim WH. Security risk evaluation method for effective threat management. Google Patents; 2007. US Patent App. 11/941,193.

[78] Tambe M, Paruchuri P, Ordóñez F, Kraus S, Pearce J, Marecki J. Agent security via approximate solvers. Google Patents; 2012. US Patent 8,195,490.

[79] Wiemer D, Robert JM, McFarlane B, Gustave C, Chow S, Tang J. Application of cut-sets to network interdependency security risk assessment. Google Patents; 2005. US Patent App. 11/232,004.

[80] Cohen G, Meiseles M, Reshef E. System and method for risk detection and analysis in a computer network. Google Patents; 2005. US Patent 6,952,779.

[81] NTIA. NTIA, editor. NTIA vulnerability disclosure call for comments. NTIA; 2016. [online] https://www.ntia.doc.gov/files/ntia/publications/fr_meeting_vulnerability_disclosure_msp_04082016.pdf. Available from: https://www.ntia.doc.gov/files/ntia/publications/fr_meeting_vulnerability_disclosure_msp_04082016.pdf [cited Dec 2016].

[82] Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases. Science. 1974;185(4157):1124–1131. Available from: http://science.sciencemag.org/content/185/4157/1124.

[83] Cox Jr LAT. Some limitations of "Risk= Threat× Vulnerability× Consequence" for risk analysis of terrorist attacks. Risk Analysis. 2008;28(6):1749–1761.

[84] Romanosky S. Examining the costs and causes of cyber incidents. Journal of Cybersecurity. 2016;p. tyw001.

[85] Romanosky S, Hoffman D, Acquisti A. Empirical analysis of data breach litigation. Journal of Empirical Legal Studies. 2014;11(1):74–104.

[86] Yayla AA, Hu Q. The impact of information security events on the stock value of firms: The effect of contingency factors. Journal of Information Technology. 2011;26(1):60–77.

[87] Davis G, Garcia A, Zhang W. Empirical Analysis of the Effects of Cyber Security Incidents. Risk Analysis. 2009;29(9):1304–1316. Available from: http://dx.doi.org/10.1111/j.1539-6924.2009.01245.x.

[88] Ezell BC, Bennett SP, Von Winterfeldt D, Sokolowski J, Collins AJ. Probabilistic Risk Analysis and Terrorism Risk. Risk Analysis. 2010;30(4):575–589. Available from: http://dx.doi.org/10.1111/j.1539-6924.2010.01401.x.

[89] Brown GG, Cox LAT Jr. How Probabilistic Risk Assessment Can Mislead Terrorism Risk Analysts. Risk Analysis. 2011;31(2):196–204. Available from: http://dx.doi.org/10.1111/j.1539-6924.2010.01492.x.

[90] Franzetti C. Operational risk modelling and management. CRC Press; 2016.

[91] Allodi L, Massacci F, Williams J. The Work-Averse Cyber Attacker Model. Evidence from two million attack signatures. In: Published in WEIS 2017. Available at https://ssrn.com/abstract=2862299; 2017. p. 1–35.

[92] Allodi L, Kotov V, Massacci F. MalwareLab: Experimentation with Cybercrime attack tools. In: Proceedings of the 2013 6th Workshop on Cybersecurity Security and Test; 2013. p. 1–8.

[93] Dumitras T, Efstathopoulos P. Ask WINE: are we safer today? evaluating operating system security through big data analysis. In: Proceeding of the 2012 USENIX Workshop on Large-Scale Exploits and Emergent Threats. LEET'12; 2012. p. 11–11.

[94] Baker W, Howard M, Hutton A, Hylender CD. 2012 Data Breach Investigation Report. Verizon; 2012.

[95] Panjwani S, Tan S, Jarrin KM, Cukier M. An experimental evaluation to determine if port scans are precursors to an attack. In: Dependable Systems and Networks, 2005. DSN 2005. Proceedings. International Conference on. IEEE; 2005. p. 602–611.

[96] Harris B, Hunt R. TCP/IP security threats and attack methods. Computer Communications. 1999;22(10):885–897.

[97] Haimes YY. On the complex definition of risk: A systems-based approach. Risk analysis. 2009;29(12):1647–1654.

[98] Brito J, Watkins T. Loving the Cyber Bomb-The Dangers of Threat Inflation in Cybersecurity Policy. Harv Nat'l Sec J. 2011;3:39.

[99] Smith S, Winchester D, Bunker D, Jamieson R. Circuits of Power: A Study of Mandated Compliance to an Information Systems Security "De Jure" Standard in a Government Organization. MIS Quarterly. 2010;34(3):463–486.

[100] Boehmer W. Appraisal of the Effectiveness and Efficiency of an Information Security Management System Based on ISO 27001. 2010 Fourth International Conference on Emerging Security Information, Systems and Technologies. 2008;0:224–231.

[101] Mellado D, Fernández-Medina E, Piattini M. A comparison of software design security metrics. In: Proceedings of the Fourth European Conference on Software Architecture. ECSA '10. ACM; 2010. p. 236–242. Available from: http://doi.acm.org/10.1145/1842752.1842797.

[102] Anderson R. Security Economics–A Personal Perspective. In: Proceedings of the 28th Annual Computer Security Applications Conference; 2012. p. 139–144.

[103] Nappa A, Johnson R, Bilge L, Caballero J, Dumitras T. The Attack of the Clones: A Study of the Impact of Shared Code on Vulnerability Patching. In: Proceedings of the 36th IEEE Symposium on Security and Privacy; 2015. p. 692–708.

[104] Ning P, Xu D. Learning attack strategies from intrusion alerts. In: Proceedings of the 10th ACM conference on Computer and communications security. ACM; 2003. p. 200–209.

[105] Kieyzun A, Guo PJ, Jayaraman K, Ernst MD. Automatic creation of SQL injection and cross-site scripting attacks. In: Software Engineering, 2009. ICSE 2009. IEEE 31st International Conference on. IEEE; 2009. p. 199–209.

[106] Ornaghi A, Valleri M. Man in the middle attacks Demos. Blackhat 2003. 2003;19.

A

Table 13: Review of patents for cybersecurity risk assessment.

Reference	Patent name	Contribution
US 20050193430 A1	Method for risk detection and analysis in a computer network	Automated method for the assesment of system risk based on attack graphs and vulnerability assessment. The method approximates likelihood of attack as a function of number of steps in the attack graph to reach the target.
US 20120203590 A1	Technology risk assessment, forecasting, and prioritization	”Environmental scoring of security risks for prioritization. No specific measure of likelihood is defined, whereas the relative distribution of severities makes for the final prioritization index.”
US 20070067845 A1	Application of cut-sets to network interdependency security risk assessment	”Risk assessment of interconnected systems in a network. Likelihood measures are calculated o the relative complexity of developing an exploit, in terms of technical equipement or knowledge required.”
US 7752125 B1	Automated enterprise risk assessment	Risk assessment methodology based on the collection of information related to risk factors that influence overall risk level of the system. No specific measure or metric for likelihood of occurrence is provided.
US 8195490 B2	Agent security via approximate solvers	General framework for the evaluation of threat realization with unknown adversaries. Probability distributions are assumed to be known.
US 8402546 B2	Estimating and visualizing security risk in information technology systems	Discretization of security risks for single or multiple networked systems. Probabilities are defined as probability of losses and are proxied as the “fidelity” of the security assessment.
US 8650637 B2	Network security risk assessment	Simulation-driven approach to risk assessmenet whereby threats are characterized by the interaction between a vulnerability and a remote website. Probabilities are results of the simulation process based on different browsing profiles as the proportion of infectious malware sites time the proportion of malware type instances.
US 20050144480 A1	Method of risk analysis in an automatic intrusion response system	”The method comprises data from an introsion response system to evaluate overall risk of attack. Probabilities of attacker are not explicitly definied, whereas the frequency of an attack is used among other parameters to evaluate final risk levels.”
US 20060064740 A1	Network threat risk assessment tool	Method that provides the user with an overview of the risk levels of the system. The method uses multiple sources to gather intelligence on probabilities of “pervasive” attacs. The overall threat score is computed as a linear combination of a “probability score” and other metrics computed by the method.
US 20090024663 A1	Techniques for Information Security Assessment	The method identifies factors for the evaluation of a final risk score. No specific definition of probability of attack is provided. Frequency of malware infections and other historical parameters are referenced as “informative” for the process.
US 8539586 B2	Method for evaluating system risk	Method for the evaluation of system risk related to a threat and a vulnerability. Probability of event is computed as a function of past events. An aggregate estimation of risk is given by weighting the risk probabilities relative to the affected system.
US 20090106843 A1	Security risk evaluation method for effective threat management	”Threat evaluation method that account for impact degree of attack, asset value, and frequency of attack. No specific definition of probability of attack is provided.”

Table 14: Service type categories and relative service and port examples.

Service type	Service description	Service ex.	Port ex.
RPC	Technology that allows programs to 'call' and execute procedures and functionalities on remote systems. An attacker can exploit this technology to remotely access resources local to the victim [104].	RPC client, rpc-rstatd (32778/tcp)	1500, 2500, 32786
SQL	Service that allows interaction with SQL database servers. An attacker may misuse the SQL language to interact with the underlying database and possibly exfiltrate or modify data without system authorization [105].	SQL session, sqlnet	139,66
Mail	Network services that allow resolution of email addresses and forward messages from one mail infrastructure to another. An attacker may compromise the exchange protocol to read or modify messages without the knowledge of either the receiver or the sender [61].	IMAP, SMTP	143, 25
Http	Network services responding to http(s) traffic. An attacker may interact with the remote server to modify some content on the webpage (e.g. store malicious scripts returned with server's content), or exploit some configuration vulnerability to read and potentially modify the traffic (e.g. connection downgrade) [61, 38].	http, http-admin, apache server	80, 8080, 9090
Infrastr	Set of services enabling functionalities internal to the company. An attacker may exploit service misconfiguration to interact with it remotely and gain privileged access to otherwise protected resources [61].	NNTP, LDAP	119, 389
RemCtrl	Set of services and protocols used to remotely control an operating system. The attacker may send arbitrary commands to the remote system by breaching the service (e.g. encryption downgrade) [106].	SSH, Telnet	22, 23
Chat	Services that enable user communication. This is typical of complex and distributed organization environments where employees may need to communicate at a distance. An attacker may obtain sensitive information by breaching these systems [61].	IRC, MSN	531, 569
Share	Services used to share content (e.g. documents) in a network. An attacker may gain access to sensitive information and possibly modify or delete it [61].	FTP, AFP	21, 548
Auth	Services that enable remote authentication. An attacker may exploit software and configuration weaknesses to gain privileged access to network and local resources [106].	Kerberos, login	543, 513
Other	Other unidentified network services, e.g. operating on non-standard ports or on protocol other than TCP. The organization may limit the fraction of Other systems by surveying the network systems.	-	udp 1434