Measuring likelihood in cybersecurity

Prof. Pablo Corona-Fraga [1]*, Dr. Vanessa Díaz-Rodriguez [2], Dr. Jesús Manuel Niebla-Zatarain [3], and Dr. Gabriel Sánchez-Pérez [4]

1 PhD Candidate from INFOTEC Center for Research and Innovation in Information and Communication Technologies, México. pablo.coronaf@infotec.mx, Twitter: @pcoronaf. ORCiD: https://orcid.org/0000-0001-5012-9468. Corresponding author*.

2 Legal Research Officer at the National Supreme Court of Justice, Mexico, vdiazr@scjn.gob.mx, https://orcid.org/0000-0002-7186-8785

3 Professor and researcher at the Faculty of Law of Mazatlan of the Autonomous University of Sinaloa, Mexico, j.niebla@uas.edu.mx, (0000-0001-8460-4538) - ORCID

4 Researcher at Instituto Politecnico Nacional, ESIME Culhuacan, Mexico City 04440, Mexico; gasanchezp@ipn.mx, https://orcid.org/0000-0002-4735-205X

Abstract: Cybersecurity risk is commonly measured by impact and probability, the former is objectively measured based on the consequences of using technology to obtain gains or achieve business objectives. The latter has been measured in sectors such as finance or insurance, based on available historical data, and many other fields have applied the same approach. In cybersecurity, as a new discipline, historical data to support an objective measure of probability is not always available, it may not be public and there is no consistent formatting to store and share it, so a new approach is required to measure cybersecurity event incidence.

Through a comprehensive analysis, including methodologies, frameworks, and incident data; considering tactics, techniques, and procedures (TTP) used by attackers, indicators of compromise (IOC), and defense controls, this study proposes a data model describing a cyber-exposure profile that provides an indirect but objective measure for likelihood, including different sources and metrics to update the model if needed.

We further propose a set of practical, quantifiable metrics for risk assessment—enabling cybersecurity practitioners to measure likelihood using a set of proposed variables. This provides organizations with an actionable framework for continuously refining their cybersecurity strategies.

Keywords: Cybersecurity, metrics, risk, data science, likelihood

1. Introduction

Cybersecurity risks are among the main concerns in the current panorama of technological dependency. Attackers are organized and have formal structures to identify and share techniques, therefore their attacks evolve rapidly, and incidents not related to an attacker (such as accidents) can also have severe consequences in organizations. As mentioned by Michał Choraś (2013) [9] after successful attacks on Estonia, Georgia, Iran and companies like Sony, Google, and Facebook, cyberattacks threaten not only information systems but also critical infrastructures, communications, supply chains, and public services as well as all kinds of daily activities.

According to the World Economic Forum Global Risk Report (2022) [50], " 'cybersecurity failure' is among the top 10 risks that have worsened most since the start of the COVID-19 crisis". These risks not only threaten information and communication technology (ICT) but also the interconnected infrastructure that ICT supports in supply chains for food, water, power, manufacturing, transport, education, etc. Therefore, cybersecurity also affects the tangible world, as stated in [11] by Couce-Vieira et al (2020).

Following the guidelines described in ISO/IEC 27005:2022 [27] risk assessment can be conducted by applying different approaches, including measuring the probability and impact of events and applying qualitative or quantitative scales, but in both cases requiring historical data or an expert's opinion. Major cybersecurity incidents, such as the Heartbleed attack and Wannacry ransomware, used 0-day vulnerabilities for which the probability cannot be objectively measured from past events.

Likar and Trček (2012) [32] proposed a methodology for the provision of sustainable information systems security, based on the evolving landscape of cybersecurity threats and attacks; however, this lacks an objective approach to measure risk probability that supports decision-making.

Bechor and Jung (2019) [7] identified 50 scholarly articles related to cybersecurity published between 2012 to 2018. They classified topics of cybersecurity and data science clustered with significant terms and concepts as follows:

- Advanced/Unseen Attack Detection 22.9%
- Contextual Cybersecurity 19.9%
- Cybersecurity Applied Domain 18.5%
- Data-Driven Adversary 11.7%
- Power System in Cybersecurity 7.9%
- Vulnerability Management 19%

However, none of these articles covers cybersecurity from the strategic point of view of measuring cybersecurity risks probability or creating a data model that supports decision-making. We identified that data science is applied differently according to the field of research, and in cybersecurity, data sets with relevant information or the corresponding data sources are not easily available, so this work focuses on creating a data model based on available sources.

A problem that makes it difficult to apply data science to strategic decisions is that cybersecurity incidents information is obtained, processed, and stored in different formats; therefore, having a unified approach for gathering and presenting information is key to developing a solution. Cybersecurity incidents are often underreported owing to legal, regulatory, and reputational concerns, limiting access to comprehensive datasets. Other problems are related to sharing cybersecurity incidents information and solutions, which are addressed in detail by the European Union Agency for Cybersecurity ENISA (2010) Incentives and challenges for cybersecurity information sharing [16], including social, cultural, and organizational issues that prevent organizations from implementing sharing practices.

To address this, frameworks such as the Trusted Automated Exchange of Intelligence Information (TAXII) and Structured Threat Information Expression (STIX) have been developed to facilitate structured threat intelligence sharing among organizations. However, TAXII and STIX are not yet broadly adopted, and their implementation results remain limited owing to interoperability challenges, lack of incentives for data sharing, and technical complexity in integrating them into existing cybersecurity infrastructures. Consequently, while these frameworks offer a potential path forward, their impact remains constrained, requiring further research and broader industry cooperation for effective large-scale implementation.

The objective of this study is to propose a model that enables comparable and reusable cybersecurity risk sources of information to measure its incidence. After reviewing the state of the art and why the current methodologies to assess this incidence are incomplete, we will present this model and the results from applying this to real case scenarios, proposing the use of a cyber exposure profile, which, in contrast to purely probability-based approaches, refers to a structured collection of relevant organizational data such as asset inventories, threat intelligence, system configurations, and security controls—that together capture an organization's real-time susceptibility to specific types of cyber threats. By focusing on these tangible indicators rather than raw historical frequencies, this profile enables a more adaptive, multi-dimensional view of likelihood—one that acknowledges the limited availability and reliability of past data in an ever-changing threat environment.

According to Foroughi and Luksch (2018) [17], data science methodology for cybersecurity projects can be divided into four general steps: problem definition and formulation, data gathering, data analysis, and production. This work was structured by applying the principles and steps presented by Rollins (2015) [40], including elements to be considered for data gathering, data sources available, indicators for data analysis and decision making, and ending with the structure of the data model proposed to generate a cyber security strategy and defining metrics to assess the level of risk, particularly for measuring its incidence. This should be applicable to different types of organizations according to their context, use and dependencies on ICT. In the final section, conclusions are presented on how to use this model, the result of its implementation, and along with some issues that can be the subject of future research.

2. Methodology

To create a data model for measuring cybersecurity risks, we applied the Foundational Methodology for Data Science and steps proposed by Foroughi and Luksch (2018) [17]. The rest of the document is structured according to an adaptation of this methodology, the phases and steps applied are described in Figure 1.

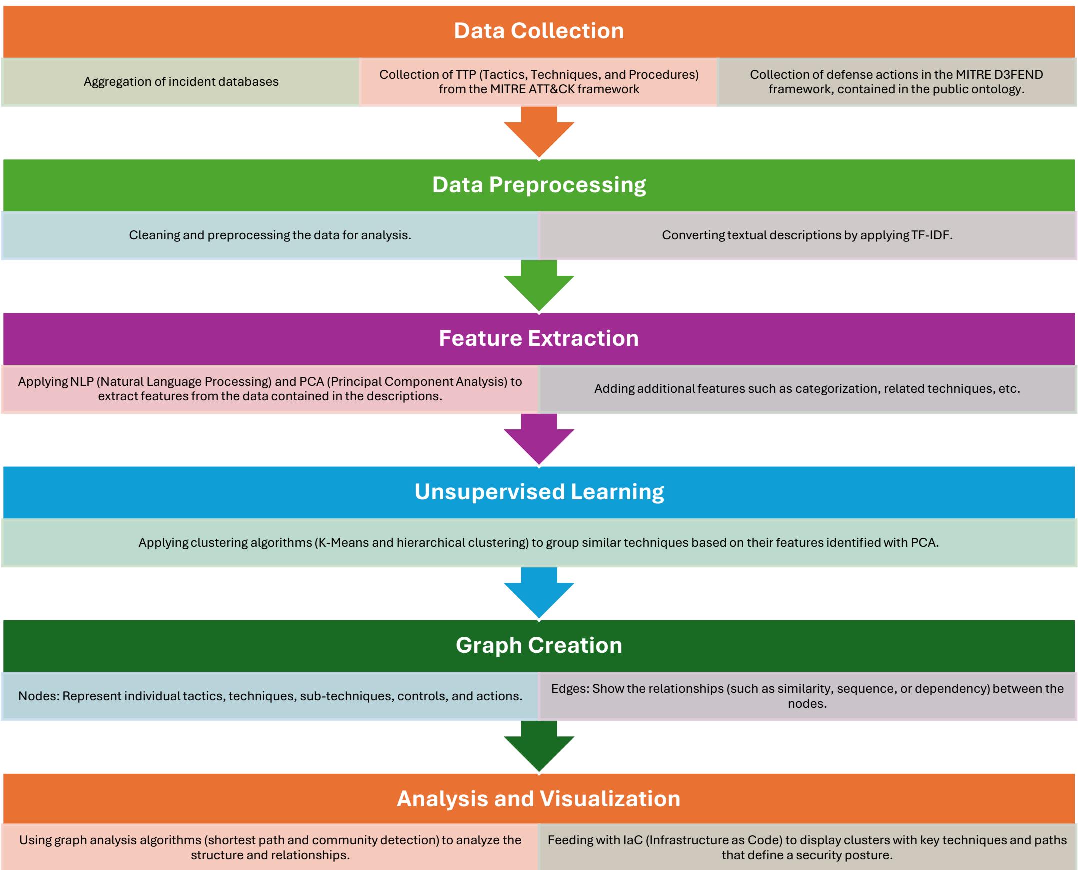

Figure 1 - Step-by-step methodology

2.1 Problem definition and formulation

To understand risk, we need to take into account why we run into that risk, because there is no risk-taking activity that is not taken to search for gains or pursue an objective, even those that are taken for entertainment and leisure purposes, are taken to gain adrenalin, fame, reputation, etc. Therefore, business risks should be measured based on their business impact or the gains pursued by taking risks.

Cybersecurity risks are not an exception, and every time we find a cybersecurity risk, we should track down the business leverage on technology that introduces the risk, generating potential gains for the organization, whether it be revenue, speed, competitiveness, increased geographical coverage, reaching more clients, or having slim and efficient processes. This also must be done the other way round, for every tool or business enabler that helps the organization to achieve its objectives, it is necessary to identify the possible failures or undesired events that can impact the organization or affect the capability to comply with their interested parties' needs and expectations.

The cost-gain of the implementation of cybersecurity strategies, such as the one presented by the National Institute for Standards and Technology (NIST 2014) Cybersecurity Framework [36], was studied by Gordon et al. (2020) [21], providing an approach for cost-gain analysis proposing the Gordon–Loeb (GL) model: (i) the value of the information being protected, (ii) the vulnerability (or probability) of a cybersecurity breach to that information, and (iii) the efficiency of the investments in cybersecurity activities.

In this context risk is measured by two factors: i) impact and ii) probability. Considering that the gains obtained from taking a risk are known, it is important to remember that even if there is a direct relationship between the potential losses and gains obtained, this relationship may not be linear nor symmetric. According to prospect theory by Kahneman and Tversky (1979) [29] losses count more than gains; therefore, this asymmetry should be considered when evaluating cybersecurity risks.

Measuring these gains and losses is paramount to be able to measure risk for assigning the resources required for risk treatment, because of the relationship between the gain obtained and the potential losses that arise from that risk. Agrafiotis et al. (2018) [2] in "A taxonomy of cyber-harms: Defining the consequences of cyber-attacks and understanding how they propagate" define the relationship between cybersecurity events and their impacts, considering the relationship between harm, impact and risk in organizations, insights from criminology and monetizing cyber-incidents to define a taxonomy that identifies impacts to organizations divided in physical/digital, economic, psychological, reputational, social/societal.

The risk level is commonly obtained from a combination of the probability and impact. To determine the level of risk an organization is exposed to, different approaches have been applied; however, a common failure is in calculating the probability of events based on past data because these data can be incomplete, outdated or the underlying risks could be based on 0-day attacks which do not have historic references.

2.2 Analytic Approach

Regarding the incidence, for this study, the concept of likelihood will be used instead of probability, based on the concept of "inverse Probability" defined by R.A. Fisher and explained by Edwards (1984) [14], and then expanded in (1997) [15]. Where they used $P(R|H)$ describing the probability of obtaining results R given the hypothesis H; and defining the likelihood $L(H|R)$, of the hypothesis H given data R for a specific model. Therefore, probability uses a model to find the chance of obtaining results (data), whereas likelihood evaluates how well a given model explains the observed data.

In this regard, the use of Bayesian methods allows for addressing uncertainty and the use of data from past events to create and adjust a suitable model that can be applied to something not readily observable (as a 0-day attack) as stated by Diaconis and Friedman (1991) [13] considering the difficulty of an agent in formulating a prior in these situations.

To determine these models, we integrated real-world datasets with statistical and machine learning techniques based on the work of Sarker et al. (2020) [42], and the proposed directions in Simulation for cybersecurity: state-of-the-art and future directions Kavak et al. (2021) [30] established the following for future directions of research in data science applied to cybersecurity:

i. advanced data collection and access,
ii. generate new theoretical constructs, and
iii. improve the behavioral models for simulation.

In the following sections, these subjects are considered to construct the proposed model.

The first theoretical new construct based on the difference between gains and losses, creates a (a)symmetric relationship that we propose to consider with three components:

- Time symmetry. The moment of the loss in comparison with the gain;
- Gain symmetry. The magnitude of the potential gain obtained, as compared with the magnitude of potential loss;
- Likelihood symmetry: The likelihood of obtaining a gain in comparison with the likelihood of obtaining negative potential consequences.

Different events have different symmetry (or asymmetry) values for each component listed above. This can be exemplified by the flip of a coin for a bet, in which you know at the same time if you win or lose(time symmetry), the amount of money that can be won in this bet is the same as that which can be lost (gain symmetry), and the likelihood of winning is the same as that for losing (likelihood symmetry). Most of the risks we observe daily are not symmetric in one or more of the previous aspects and cybersecurity is no exception.

Likelihood symmetry considers how the consequences of the risks can be materialized, taking into account that there are risks in which a single event is diluted among the number of elements of the system, so the impact of this single incidence can be low in comparison to its impact on the whole system. This is in contrast with events that have a very high impact and are difficult to consider in previous models; therefore, a single incidence can produce a high impact on the whole system.

This was described by Taleb (2007) in Black Swan [46] as a "black swan" event that has no historical and statistical information that can be used to assess the risk level, so the likelihood of the event cannot be measured objectively. For example, in fields such as financial and insurance risks, a single person having an accident or not paying a credit is diluted between the number of people who pay for the service or the credit interest rate. This type of risk has been measured for many years and there is historical information to establish risk criteria to measure the probability, such as age, gender, job, personal and family health conditions, activities, and hobbies.

In cybersecurity, the risks are asymmetric over time because the gain is obtained when a technology is implemented, but the potential negative consequences remain dormant until a threat exploits a vulnerability in the system materializing the risk. In addition, there are asymmetries in gain, because information and communication technology are business enablers providing opportunities, revenue, and other kinds of gains, but if cybersecurity events are materialized, the negative consequences can be bigger than the original gain, considering downtime, fines, penalties or even bankruptcy. The likelihood is asymmetrical and not dilutable with respect the number of elements in the system. This type of event in cybersecurity is called 0-days, for the nonexistence of information before the first event, but once it has an incidence, it is replicated and appears at a high frequency. This proposed concept of non-dilutable risks is characterized by little or no historical information, and a single incident can have very high impacts, as described by a black swan event.

This dilutable risk is the second new theoretical construct we propose in this study, to differentiate risks where the number of elements determines the severity of the impact of a risk, versus those risks that are prone to events with very high impacts from a single element, even if this has low probabilities.

For instance, Heard et al. (2018) [24] in Data Science for Cyber-security mention that: APT attacks often circumvent strong firewalls, which protect value assets, through a technique called pivoting (also known as lateral movement). This allows attackers to expand their access from vulnerable computers to those with higher levels of protection. In their work, they analyzed datasets from different sources and were able to identify user activity states from a Markov-modulated Poisson process (MMPP) suggesting a data-driven decision-making process.

Elvis Pontes et al. (2011) [39] in Applying multi-correlation for improving forecasting in cyber security, use a Two-Step System (TSS) to forecast events in cybersecurity using multi-correlation information from an Intrusion Forecast System (IFS). These works will be the basis for the proposed framework to build a cybersecurity strategy to determine the metrics for likelihood.

2.3 Data collection

Considering the asymmetry mentioned before and the difference between dilutable and non-dilutable risks, the data sources for measuring and monitoring cybersecurity and applying data science should be adequate for the organizational context, in a format that is comparable with other organizations and with previous measurements, and provides information to evaluate the risk of cybersecurity events.

As mentioned above, to calculate the impact an organization can identify the level of dependency on Information and Communication Technology (ICT), and the gain they obtain from the use of these technologies as a business enabler.

However, the probability is often miscalculated by trying to use information from previous events; however, as mentioned, cybersecurity events require a different approach. Liu et al. (2012) [33] proposed a correlation scheme based on rough set theory, using a knowledge base of events regarding network security, with rule generation methods and sets of rules to be matched as an alternative approach.

Wrotniak and Wozniak (2012) [51] used Combined Bayesian Classifiers to address spam detection problems using classifiers for cybersecurity events. On this basis, to improve the accuracy of measuring the odds of occurrence, instead of using classic statistical information to obtain a probability, Bayesian methods are applied to measure the likelihood of cybersecurity risks, including using probabilities to update our beliefs about an event as new information becomes available. The Impact, Stability, Uniqueness, Network, and Factors (ISUNF) formula measures the likelihood of a cyberattack given a specific set of circumstances based on the information on cyber-attacks.

Likelihood = Impact x Stability x Uniqueness x Network x Factors

This study considers these references and proposes measuring cybersecurity risk using an adaptation of cumulative prospect theory [48] and likelihood mathematical representations, as follows:

$$\mathrm{V(x,p)} = \begin{cases} \pi(\mathrm{P(H \mid R)}) \cdot x^{\alpha}, & if\ x \geq 0 (Gains) \\ -\lambda \cdot \pi(\mathrm{P(H \mid R)}) \cdot x^{\beta}, & if\ x < 0 (Losses) \end{cases}$$

Where:

- V(*x,p*) = Perceived risk-adjusted value (subjective cybersecurity impact and probability).
- *x* = Objective impact value (e.g., financial loss from a breach or gain from security investment).
- *p* = Actual probability of the cybersecurity event occurring.
- *π(p)* = Probability weighting function that distorts perceived likelihood:

$$\pi(p) = \frac{p^{\gamma}}{(p\gamma + (1-p)^{\gamma})^{1/\gamma}}$$

- *γ* = 0.61 (empirically determined) affects the probability distortion.
- Small probabilities are overweighted (e.g., rare 0-day attacks seem more threatening).
- High probabilities are underweighted (e.g., unpatched system risks are underestimated).
- *λ*=2.25 = Loss aversion coefficient (cybersecurity losses feel ~2.25x worse than equivalent gains).
- *α,β*=0.88 = Risk sensitivity parameters (diminishing sensitivity for gains/losses).

2.4 Data pre-processing

To determine the variables that increase or decrease risk sensitivity, a comprehensive data collection should consider both online and offline aspects, considering the information technology, communications, human regulatory, and organizational components. Choras (2013) [9] described online aspects including monitoring, intrusion detection and prevention, intrusion tolerance, resilience, reaction, remediation, and testing capabilities. Offline aspects are described by Choras (2013) [9] with examples such as simulations and prediction, visualization, human factors, risk management, vulnerability assessment, development of security policies, and audits.

Collecting and analyzing relevant datasets was one of the first challenges of this work, as detailed by Aldribi et al. (2018) [3]. Identifying appropriate datasets for modelling and evaluation is critical for optimizing cloud security measures. As noted by Aldribi et al. (2020) [4], real-time multivariate statistical change tracking is a relevant tool for detecting complex incident patterns, which supports the proactive nature of new technology security strategies.

Sarker et al. (2020) [42] in “Cybersecurity data science: an overview from a machine learning perspective” suggested a data driven cybersecurity framework based on machine learning techniques considering multiple processing layers. Another relevant contribution of their work is listing a summary of available cybersecurity datasets, their characteristics and the machine learning techniques that can be applied to them.

Also, gfek Real-CyberSecurity-Datasets Github repository [18] lists a comprehensive collection of cybersecurity datasets to be used in machine learning projects and other applications.

2.5 Feature extraction

Not only is the information in real-case scenarios datasets relevant. Understanding the dynamics of hacker collaboration is critical to develop robust cybersecurity strategies. Hausken (2017) [23] explored how cyber attackers share information during successive attacks, which can significantly enhance the efficiency and sophistication of their methods. This has been documented in the ATT&CK framework [6] built by Mitre Corporation.

Pontes et al. [39] defined a system for applying multi-correlation to improve forecasting in cybersecurity, which can be used to process the information obtained and create triggers and alerts based on the input received from "sensors" in the system.

To determine which these "sensors" could be, we used Sarker et al. (2020) [42] and gfek Real-CyberSecurity [18] -Datasets, and applied the method used by Abbiati [1] to merge datasets , resulting in a mega-dataset with more than 3 million relationships, this dataset was difficult to handle and we had problems processing it.

To identiy which datasets were more relevant we took ideas from Heard and Rubin-Delanchy (2018) [24], who cite a series of works and contributions that provide elements for:

- adapting classic data analysis tools through inference for graphs and networks,
- adaptive detection methods for detecting attacks on computer systems,
- sampling statistics for the detection of intruders and anomalies

Based on these, we processed the merged dataset by applying natural language processing and Term Frequency - Inverse Document Frequency (TF-IDF) techniques to the incident descriptions to determine the importance of each word in a description relative to all other descriptions, and extracting the main features, considering additional information such as the related techniques, infrastructure and technologies.

2.6 Unsupervised learning

Using the vectorized textual descriptions obtained by TF-IDF, we applied Principal Component Analysis (PCA) to extract the main characteristics from descriptions, reducing the dimensionality of the terms obtained to be prepared for the clustering process. Applying unsupervised learning for clustering using K-means, we grouped similar incidents based on their descriptions, using the characteristics, identified with PCA.

The merged dataset obtained included incident descriptions, related systems and infrastructure metadata, and relationships with the tactics, techniques and procedures (TTP) in the ATT&CK framework[6]. This is how we identified that the ISOT-CID [26] and HIKARI-2021 Datasets [25] had the strongest relationships with TTP in ATT&CK and the Mordor project security events [47].

Different sizes of K were used until we found a number of clusters that were adequate to describe the dimensions to represent risk likelihood. Manual validation was conducted and features such as related techniques that were missing were added.

The results of the clustering were different groups that were manually labelled, identifying four of them that became the variables for measuring the likelihood of cybersecurity events: exposure, traceability, motivation, and system update. Some sub-variables were also identified, for example, authentication and authorizations that were merged with traceability or number of users and devices that were merged in exposure, which are defined as metrics for the variable in Table 1 below.

The establishment of similar metrics was presented by Breier and Hudec (2011) [8], mainly aligned with the controls contained in the D3FEND framework[12] and ISO/IEC 27002:2022 [27], and later classified according to different organizational levels: Strategic and Leadership, Management and Operational, by Anu (2021) [5]. In particular, the metrics for offline policies that are not self-enforced are defined by Goel et al. (2010) [19] to assess their effectiveness.

2.7 Graph creation

To obtain an operative tool using the multidimensional relationships obtained, the merged dataset was formatted in a graph database using the variables for measuring likelihood we propose from the clusters obtained. The graph is structured as follows:

Nodes (Vertices) represent:

- TTPs (Tactics, Techniques, and Procedures) from the MITRE ATT&CK framework
- Cybersecurity events from merged datasets (ISOT-CID and HIKARI)
- Defensive controls from D3FEND framework [12]
- System components (applications, services, hardware and software assets)

Edges (Links) represent:

- Identified relationships (for example a malicious activity leveraging a specific TTP)
- Control mappings (which D3FEND technique mitigates which ATT&CK tactic)
- Event propagation paths (for example lateral movement between compromised assets)

2.8 Analysis and visualization

The graph network obtained enable us linking the TTP related to the technology used by an organization to build a profile including the technological baseline, that can be described as Infrastructure as code (IaC), and determining the applicable Indicators of Compromise (IOC), thus the resulting model can be used not just to determine which attack is more suitable to certain baseline, but also to classify the type of controls required to address the risks, according to the underlying systems and infrastructure. Then estimating the level of risk, particularly the likelihood, filtering the relevant remediation methods and creating recommendations to build a cybersecurity strategy for organization-specific needs.

For these means, on the obtained the graph we applied the following:

- Let N be the total number of controls in the ATT&CK and D3FEND databases.
- Let Ci be the ith control, where 1 <= i <= N in the list of available controls in ATT&CK and D3FEND.
- Let Ri be the number of references for control Ci in the mapping as a countermeasure to
- cybersecurity events.
- The weight for control Ci can be calculated as follows:

$$\text{Weight(Ci)}= \frac{\text{Ri}}{\sum \text{Rj}}$$

for all j where 1 <= j <= N in the list of available controls in ATT&ACK and D3FEND

To determine the variables, each incident description and metadata were classified according to their direct or indirect impact on the occurrence and severity of cybersecurity incidents using the following criteria:

- Incident frequency: The number of times a specific type of cybersecurity incident appeared in the merged dataset. This variable helps establish a baseline frequency of incidents, which is critical for predicting the likelihood of future events. A higher frequency indicates a greater likelihood of recurrence.
- Incident severity: The impact level of an incident, typically categorized as low, medium, or high, according to the ATT&CK framework. This helps to prioritize incidents and understand the potential damage they can cause. It informs risk mitigation strategies by focusing on high-impact incidents.
- Attack Vector: The method or pathway used by attackers to exploit vulnerabilities and how easily they can be exploited. Identifying common attack vectors enables reinforcement of specific areas of infrastructure and reduces the risk of these attacks.
- Affected Assets: Types of assets (e.g., servers, databases, end-user devices) affected by each incident. Understanding which assets are most frequently targeted helps in allocating resources for better protection and resilience of critical assets.

To create the data model considering those consequences, and avoiding the common probability miscalculation mentioned earlier, an indirect measurement can be applied to assess the likelihood of events using a formula adapted from prospect theory and likelihood.

To determine how each of the variables can be measured, based on the technology adopted by the organization and its context, we identified the following data sources queried through Application Programming Interfaces (APIs):

- Google Safe browsing [20]
- GreyNoise Intelligence [22]
- Metasploit API [35]
- PhishTank [38]
- SecurityTrails [43]
- Shodan.io [44]
- VirusTotal [49]

These data sources can be divided into different categories:

- Those that collect malware samples, addresses and signatures of malicious files, their activities, or their origin can be detected and blocked through antimalware software, firewalls, and email filters.
- Those that provide information on the Indicator of Compromise (IOC) that are symptoms can be identified in a compromised system, according to the actions produced by threat actors and malware, this could be existing communications from or to an address, the modification of a system file, the creation and activities of a user.
- Those that describe tactics, techniques and procedures (TTP) conducted by malicious actors, including mapping and scanning of the environment, identifying vulnerabilities and weak parts of the system, gaining access privileges, moving

laterally, to make persistent changes in the system to avoid detection and removal, identification of valuable assets and information, and the extraction or attack of assets.

3. Resulting model

The resulting data model is built based on the one proposed by Sarket et al. (2020) [42], expanding it with specific data sources relevant for each variable and metric in Table 1, mapping with the TTP and IOC, to determine the remediation actions that will be applicable as controls that build the cybersecurity strategy applicable to the underlying infrastructure and systems for each organization, that we define as the cyber-exposure profile, completing the cycle back from this ICT dependencies to the tactics, techniques and procedures applied by attackers.

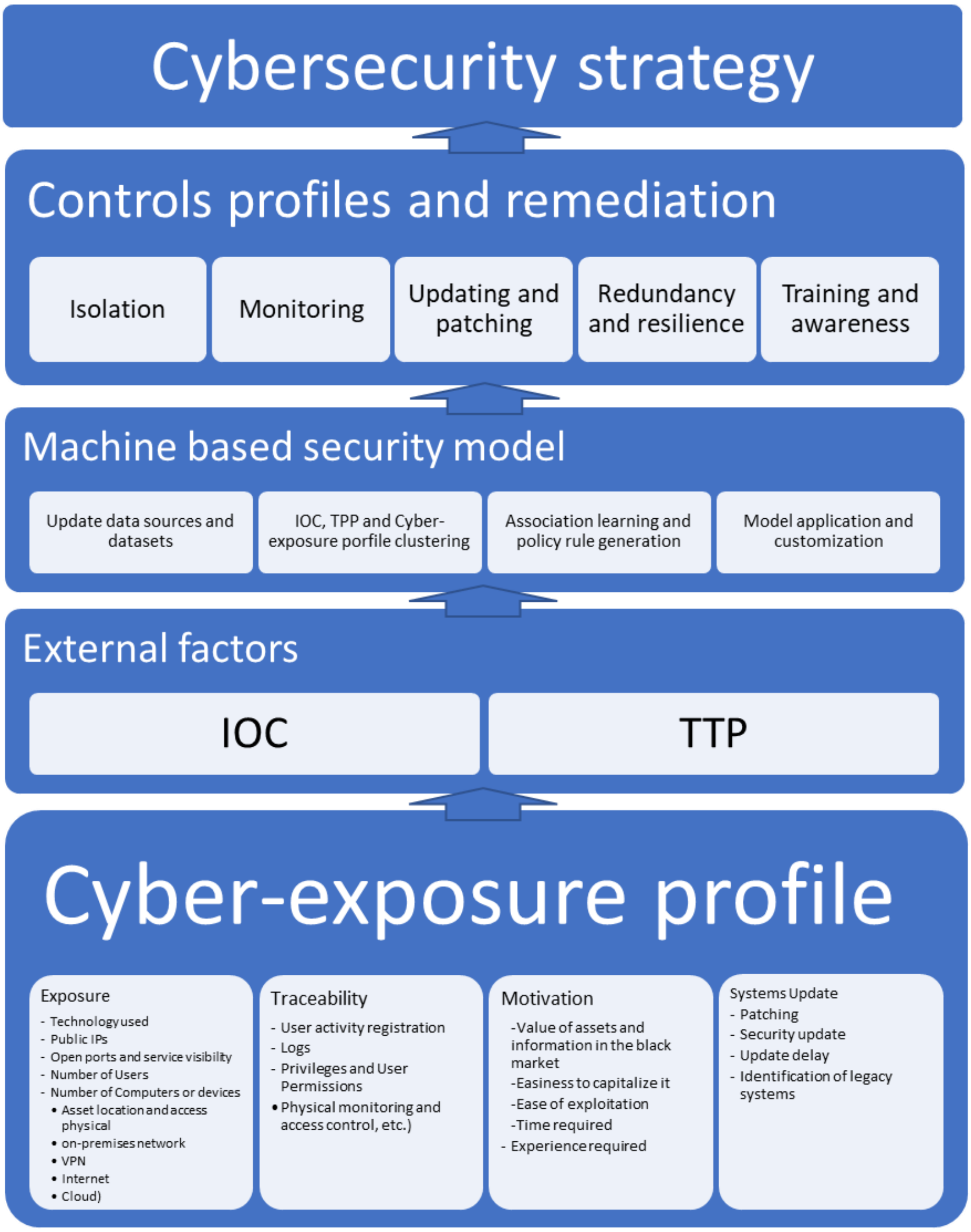

Figure 2 – Cybersecurity strategy data model and cyber-exposure profile

3.1. Metrics for proposed variables

Based on the concept of "sensors" defined by Choras (2013) [9] the resulting model defines four variables proposed to measure the likelihood of events. Table 1 also presents methods for gathering information, tools, and metrics for each of these variables in the cyber-exposure profile, including data sources for those "sensors", associated with each proposed variable.

Table 1. Metrics for proposed variables

Variable	*Unit*	*Gathering method/tool*	*Metric*
Exposure	*Technology used*	*Organization assets inventory* *Nmap [37]*	*# of different technologies used*
	Public IPs	*Nmap* *Shodan*	*# public IPs/# necessary public IPs*
	Open ports and service visibility	*Nmap* *Sqlmap [45]*	*# visible ports/# necessary visible ports*
	Number of Users	*Active directory* */etc/passwd*	*# privileged users /# total users* *# authenticated users/# total users* *#users with shared accounts/#total users*
	Number of Computers or devices	*Nmap* *Active directory*	*# registered computers/#total computers*
	Asset location and access (accessible through the physical, on-premises network, VPN, internet or the cloud)	*Physical inventory*	*Access to assets from location/ business rules requiring access from a location*
Traceability	*User activity registration*	*File and application logs*	*User authentication vs expected authentication*
	Logs	*Operative system logs* *Network logs* *Device logs* *Application logs* *Database logs*	*Automated activities vs expected transactions* *devices online/ expected or registered devices* *Traffic or transactions/expected traffic and transactions volume*
	Privileges and User Permissions	*Active directory* *Operative system* *Application and system configuration*	*User authorization for activities/roles and responsibilities*
	Physical monitoring and access control, etc.)	*CCTV* *Door locks* *Fire alarms*	*Registered behaviour / expected behaviour*

		Temperature sensors *Presence sensors*	
Motivation	*Value of assets and information in the black market*	*Internal information and asset classification* *Brand monitoring in forums and social media*	*Information and assets by their value and business gain*
	Easiness of capitalization	*Type of information and value in the market*	*Information and assets that can be restored from backup* *Information and assets that are made public can harm the organization*
	Ease of exploitation	*Residual risk from implemented controls*	*Effectiveness of controls from Vulnerability analysis, pen-testing* *Updated and patched systems/total number of systems*
	Time required	*Residual risk from implemented controls*	*Updated and patched systems/total number of systems*
	Experience required	*Maturity level of implemented controls*	*Effectiveness of controls from Vulnerability analysis, pen-testing*
Systems update	*Update policy (e.g. n-1 updates, three-month delay, etc.)*	*Systems inventory* *Asset management*	*Updated and patched systems/total number of systems* *Time to delay updates or number of versions from current*
	Test updates	*Measure time and resources to test updates before deployment*	*Time to test updates* *Resources to test update*
	Legacy systems with no updates	*Systems inventory* *Asset management*	*# of systems that cannot be updated* *Complementary controls for those systems*
	Time to update for critical updates	*Measure time to identify and apply critical updates*	*Time to release updates after identification of vulnerability* *Time to implement updates after release*

The metrics proposed in table 1 can be used to parameterize the likelihood formula, resulting in:

$$\text{Likelihood(incident)} = \frac{E^{\alpha} \times M^{\beta}}{T^{\gamma} \times U^{\delta}}),$$

Where:

- E,T,M, and U∈[0,1] are normalized measures of each variable (0 = minimal, 1 = maximal) based on column 4 in Table 1.
- E(Exposure) and M (Motivation) are direct "risk enhancers"; higher values drive the likelihood upward.
- T (Traceability) and U (System Updates) are "risk mitigators"; higher values drive the likelihood downward.
- α,β,γ,δ are tunable exponents that allow one to weight each variable's relative importance based on the organizational context or empirical data.

In practice, this is applied following these steps:

1. Normalize each raw metric for example ratio of privileged users to total users for exposure, fraction of fully patched devices for system updates, so they fall between 0 and 1, according to Table 1 metrics.
2. Estimate or tune α,β,γ, and δ using historical performance, expert judgment, or optimization against observed incidents.
3. Continuously update E,T,M,U (for instance, after new user roles are implemented, software patches are applied or changes in attacker incentives) to recalculate a current snapshot of the likelihood.

4. Results

The model was implemented in Mexican organization of different sectors with the following steps:

1. Define their infrastructure and Systems as IaC.
2. Monitoring cybersecurity incidents and gathering information available from previous incidents (before implementing the model or new controls).
3. Obtain their cyber-exposure profile.
4. Determining the appliable controls in DEF3ND.
5. Implementing the necessary controls for the cybersecurity strategy.
6. Monitoring cybersecurity incidents, particularly the attacks mapped in the model (after implementing the model and new controls).
7. Vulnerability assessment.
8. Evaluating the effectiveness of the cybersecurity strategy implemented.

Only three organizations completed all the steps due to economic and time limitations of the rest, for which the implementation results are presented in Table 2.

Table 2. Implementation evaluation

Evaluation criteria	***Organization A***	***Organization B***	***Organization C***
Baseline (before applying the model)			
Total Devices	*275*	*361*	*437*
Unregistered Devices	*86*	*134*	*369*
Privileged Users	*27*	*19*	*45*
Devices with Logging	*Minimal (below 30%)*	*Minimal (below 30%)*	*Inconsistent (below 50%)*
Post-Implementation (after applying the model)			
Registered Devices	*200*	*250*	*400*
Privileged Users	*9*	*11*	*16*
Devices with Logging	*Comprehensive (over 90%)*	*Comprehensive (over 90%)*	*Standardized (over 80%)*
Outcome			
Reduction in Incidents	*39%*	*47%*	*35%*
Detection & Response Time	*Reduced by 26%*	*Reduced by 11%*	*No significant reduction*
Compliance & Audit Readiness	*Complied with PCI-DSS assessment*	*Significant improvements in compliance vs ISO/IEC 27001*	*No compliance framework implemented*

One of the main advantages of using this model is that each of the units for measuring the variables represents not just a value to measure risk likelihood but also an action that should be taken. For example, if the ratio of registered devices to the total number of devices is low, the action is to register these devices. If there is a large number of privileged users, the action is to question whether there is a need for each of these privileged uses, and to create a log registering the activities for each of them, creating rules to correlate them with other logs, in order to get alerts on abnormal activities. If some devices or systems do not log users' activities, the action is to reduce the number of users or the exposure. These actions are in practice controls to be implemented so this model not only describes metrics but also associates each one with the actions and controls to be used.

4.1 Feedback

To understand the maturity of the controls, a root cause analysis should be conducted to determine for each incident the control that failed and if the objective of the controls related to a specific asset is fulfilled, according to the nature of the breach and the values in Table 1. To enable this analysis, the control attributes in the previous section indicate the nature and objectives of the controls necessary to address the root cause.

For example, if a breach occurs in an information asset, allowing unauthorized access to information, all the controls related to confidentiality, traceability and exposure should be evaluated to determine which applicable control failed to contain the event.

To use an objective evaluation of the effectiveness of the control, the cost associated with incidents is a responsive metric explored by Romanosky (2016) in “Examining the costs and causes of cyber incidents” [41], where they sampled over 12,000 cyber events, including data breaches, security incidents, privacy breaches, and phishing events. The bottom-line cost of cybersecurity events costs firms around 0.4% of their revenues, less than any other loss due to fraud, theft, corruption, or bad debt. Therefore, the cost of implementing controls should be less than the possible damage to the organization by the event it aims to address.

To provide proper feedback on cost-gain, a balance is important to trace the underlying infrastructure and systems that introduce the elements of exposure, traceability, systems update, and motivation to an attacker. Every cybersecurity event and its consequences can be linked to the gain provided to the organization by using that technology in terms of competitiveness, speed, response times, clients, income, etc. which is the value that provides information on how many resources will be assigned to address it.

5. Discussion

The data model resulting from this work enables, based on the information on the technology adopted by each organization, its business context and the current measures they implemented, to measure risk likelihood, and provide information on the necessary actions, the type of controls required to be implemented, and resource allocation to build a cybersecurity strategy.

This study highlights how critical cybersecurity risks are in the context of technological dependency, emphasizing the rapid evolution of attacks and outlining the limitations of existing risk assessment methods, which rely heavily on historical data or expert opinion, underscoring the flaws of these approaches in the cybersecurity domain due to the prevalence of 0-day vulnerabilities.

By leveraging data science principles and processes, and adopting a multidisciplinary approach, this research identifies gaps in current cybersecurity risk measuring strategies and suggests a data-driven framework to objectively evaluate risk likelihood, contributing to the field by providing a structured framework that also facilitates informed decision-making to build a cybersecurity strategy.

Based on the four variables proposed the likelihood can be assessed and the results evaluated directly using the metrics in Table 1, and at the same time, these variables are measurements of risk and indicators of actions that need to be taken..

This study also proposes the use of the cyber-exposure profile based on ATT&CK and D3FEND matrices, although this can be applied to any other set of controls or attack origins, such as Common Vulnerabilities and Exposures (CVE), Common Platform Enumeration (CPE), Common Attack Pattern Enumeration and Classification (CAPEC), and Common Weakness Enumeration (CWE) . This is aligned with what Pablo Corona (2022) [10], in "Practical Guidance for Risk Management", that proposes a model for making decisions on the necessary controls for a cybersecurity strategy considering the events, motivation for an attacker, consequences, and current controls to determine the risk level, which enables decisions on the type of controls needed, where they should be implemented and the resources that should be assigned to that implementation.

Furthermore, the industry sector to which the target organization belongs and the type of information it handles (for example healthcare, finance or manufacturing) should be considered. This can become an extra variable, helping to understand the motivation of the attacker to try to breach a particular organization. This information was not available in the datasets used in this study.

6. Conclusions

Cybersecurity metrics require a consistent and comparable approach to produce repeatable and verifiable results. Thus, using indirect metrics to assess the likelihood of cybersecurity incidents is an important tool for supporting decision-making by measuring the risk likelihood and determining the controls needed to build a cybersecurity protection strategy.

Based on the proposed cyber -exposure profile, an organization can determine the actions to be taken to determine the type of controls needed, where they need to be implemented, and the number of resources based on the gain obtained by using the involved technology.

This research project contributes three elements to define metrics for a cybersecurity strategy:

- Measuring the likelihood of risk in cybersecurity could be difficult using simple statistics because the consequences cannot be diluted, and there is no objective information to measure the probability of events; this study describes asymmetries that are present in cybersecurity events.
- To objectively measure the likelihood of cybersecurity risks, four elements are proposed (exposure, traceability, motivation, and systems update) along with gathering methods and metrics that measure the cybers-exposure profile and help to determine the necessary and applicable controls;
- Measuring the effectiveness and providing feedback for the controls and the entire cybersecurity strategy can be enhanced by using the control attributes in correlation with the applicable TTP, IOC, exposure, traceability, motivation, and system updates.

It is important to highlight that the result of a risk assessment process is not only measuring risk with a value, but also that those values should enable the decision-making process toward the selection of controls, the assets that must be covered or protected and the resources to be assigned, which should be directly aligned to the gain the risk-taking enables and the likelihood of that event.

This research used the Sarket et al. (2020) [42] framework for data models in cybersecurity and applied the Rollins (2015) [40] methodology to define a cybersecurity data science project. Further research can produce automatic gathering methods and algorithms to train machine learning models by applying the proposed framework.

For future research, the application of this data model can be implemented using an automated mechanism to gather further information and build new datasets and using other machine learning techniques such as neural networks for real-time and automatic decision-making, such as those presented by Fei Li et al. (2018) [31] for their work on Intelligent Vehicle Intrusion Detection and Mbona (2022) [34] in "Detecting Zero-day intrusion attacks using semi-supervised machine learning approaches".

Statements and Declarations

No funding was received. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Author Contributions: Conceptualization, Pablo Corona Fraga, Dr. Vanessa Diaz Rodriguez, and Dr. Jesús Manuel Niebla Zatarain; methodology, Pablo Corona Fraga; software, Pablo Corona Fraga; validation, Pablo Corona Fraga, Dr. Vanessa Diaz Rodriguez, and Dr. Jesús Manuel Niebla Zatarain; formal analysis, Dr. Vanessa Diaz Rodriguez, Dr. Gabriel Sánchez-Pérez, and Dr. Jesús Manuel Niebla Zatarain; investigation, Pablo Corona Fraga; resources, Pablo Corona Fraga.; data curation, Pablo Corona Fraga; writing—original draft preparation, Pablo Corona Fraga, Dr. Vanessa Diaz Rodriguez.; writing—review and editing, Dr. Vanessa Diaz Rodriguez, Dr. Gabriel Sánchez-Pérez, and Dr. Jesús Manuel Niebla Zatarain; All authors have read and agreed to the published version of the manuscript.

References

[1] Abbiati G, Ranise S, Schizzerotto A, Siena A. 2021. Merging Datasets of CyberSecurity Incidents for Fun and Insight. *Front Big Data*. https://doi.org/10.3389%2Ffdata.2020.521132

[2] Agrafiotis, I., Nurse, J., Goldsmith, M., Creese, S., Upton, D. 2018. A taxonomy of cyber-harms: Defining the impacts of cyber-attacks and understanding how they propagate, *Journal of Cybersecurity*, Volume 4, Issue 1, tyy006, DOI: https://doi.org/10.1093/cybsec/tyy006

[3] Aldribi A., Traore I., Moa B. 2018. Data sources and datasets for cloud intrusion detection modelling and evaluation. Mishra B., Das H., Dehuri S., Jagadev A. Cloud computing for optimization: foundations, Applications, and challenges. *Studies in Big Data*. (39): 333-366. Springer. DOI: 10.1007/978-3-319-73676-1_13

[4] Aldribi A., Traore I., Moa B., Nwamuo O., 2020. Hypervisor-based cloud intrusion detection through online multivariate statistical change tracking, *Computers & security*, Elsevier, Volume January 88, 2020., DOI: https://doi.org/10.1016/j.cose.2019.101646

[5] Anu, V., 2021, Information security governance metrics: a survey and taxonomy, *Information Security Journal: A Global Perspective*, DOI: https://doi.org/10.1080/19393555.2021.1922786

[6] ATT&CK is a globally accessible knowledge base of adversary tactics and techniques based on real-world observations. Available at https://attack.mitre.org/ consulted 20250515

[7] Bechor T. and Jung, B. 2019. current state and modelling of research topics in cybersecurity and data science, *Computer science*, available at http://www.iiisci.org/journal/CV$/sci/pdfs/ZA483JS19.pdf

[8] Breier, J. and Hudec, L. 2011. Risk Analysis supported by Information Security Metrics, CompSysTech '11, Proceedings of the 12th International Conference on Computer Systems and Technologies, June 2011 Pages 393–398, DOI: https://doi.org/10.1145/2023607.2023673

[9] Choraś, M. 2013. Comprehensive approach to Information sharing for increased Network security and survivability, **Cybernetics and systems**: an international journal, DOI: http://dx.doi.org/10.1080/01969722.2013.818433

[10] Corona-Fraga, P. 2022. *Practical guidance for risk management in the age of cybersecurity*, Kindle Direct Publishing

[11] Couce-Vieira A., Rios Insua, D., Kosgodagan, A. 2020. Assessing and forecasting cybersecurity impacts, decision analysis, DOI: 10.1287/deca.2020.0418

[12] D3FEND framework that encodes a countermeasure knowledge base, and a knowledge graph available at https://d3fend.mitre.org/, consulted 20250515

[13] Diaconis, P. and Zabell, S. 1991. Closed form summation for classical distributions: variations on a theme of De Moivre. *Statistical Science 6* (3), pages 284–302,. DOI: https://doi.org/10.1214/ss/1177011699

[14] Edwards, A. W. F. 1984. *Likelihood*, Cambridge classical science, 8-10.

[15] Edwards, A. W. F. 1997. What did fisher mean by 'inverse probability' in 1912-1922?, *Statistical science 12*, no. 3 (1997): 177–84. Available at http://www.jstor.org/stable/2246368 consulted 20250515.

[16] European Network and Information Security Agency (ENISA). 2010. Incentives and challenges for cyber security information sharing.

[17] Foroughi, F., and Luksch, P. 2018. Data science methodology for cybersecurity projects, DOI: http://dx.doi.org/10.5121/csit.2018.80401

[18] gfek, Real-cybersecurity-datasets, available at https://github.com/gfek/Real-CyberSecurity-Datasets, consulted 20240220

[19] Goel, S., Chengalur-Smith, I. N. 2010. Metrics for characterizing the form of security policies, *The Journal of Strategic Information Systems*, Volume 19, Issue 4, December 2010, Pages 281-295 DOI: https://doi.org/10.1016/j.jsis.2010.10.002

[20] Google safe browsing, https://safebrowsing.google.com/, consulted 20250515

[21] Gordon, L., Loeb, M., Zhou, L. 2020. Integrating cost-gain analysis into the NIST Cybersecurity Framework via the Gordon-Loeb Model, *Journal of Cybersecurity*, Volume 6, Issue 1, 2020, DOI: 10.1093/cybsec/tyaa005

[22] Grey Noise Intelligence, https://www.greynoise.io/, consulted 20250515

[23] Hausken, K. 2017. Information sharing among cyber hackers in successive attacks, *international game theory review* Vol. 19, No. 02, 1750010, DOI: https://doi.org/10.1142/S0219198917500104

[24] Heard, N., Adams, N. and Rubin-Delanchy, P. 2018,.*Data science for cyber-security*, World Scientific. DOI: https://doi.org/10.1142/q0167

[25] HIKARI-2021 Datasets, available at https://github.com/gfek/Real-CyberSecurity-Datasets, consulted 20240220

[26] Information security and object technology, University of Victoria, ISOT-CID dataset, https://onlineacademiccommunity.uvic.ca/isot/2022/11/25/cloud-security-datasets/, consulted 20250515

[27] ISO/IEC 27002:2022, ISO, Information security, cybersecurity and privacy protection — Information security controls

[28] ISO/IEC 27005:2022, ISO, Information security, cybersecurity and privacy protection — Guidance on managing information security risks, ISO

[29] Kahneman, D., & Tversky, A. 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263. https://doi.org/10.2307/1914185

[30]Kavak, H., Padilla, J., Vernon-Bido, D., Diallo, S., Gore, R., Shetty, S. 2021. Simulation for cybersecurity: state of the art and future directions, *Journal of Cybersecurity*, Volume 7, Issue 1, tyab005, DOI: https://doi.org/10.1093/cybsec/tyab005

[31]Li, F., Zhang, H., Wang, J., Liu, Y., Gao, L., Xu, X., Sanin, C., and Szczerbicki, E. 2018. Toward Intelligent Vehicle Intrusion Detection Using the Neural Knowledge DNA, *Cybernetics and Systems*, DOI: https://doi.org/10.1080/01969722.2017.1418788

[32]Likar, B., & Trček, D. 2012. A methodology for provision of sustainable information systems security. *Cybernetics and Systems*, 43, 22 - 33. DOI: https://doi.org/10.1080/01969722.2012.637013

[33]Liu, J., Gu, L., Xu, G., and Niu, X. 2012. A Correlation Analysis Method of Network Security Events Based on Rough Set Theory, Paper presented at the 2012 3rd IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC), Beijing, China, September 21–22. DOI: https://doi.org/10.1109/ICNIDC.2012.6418807

[34]Mbona, I., & Eloff, J.H. 2022. Detecting zero-day intrusion attacks using semi-supervised machine learning approaches. *IEEE Access*, PP, 1-1. DOI: https://doi.org/10.1109/ACCESS.2022.3187116

[35]Metasploit API, https://docs.metasploit.com/api/, consulted 20250515

[36]National Institute for Standards and Technology (NIST). 2014. Framework for improving the critical infrastructure cybersecurity. Version 1.0, 12 February 2014. Available at https://www.nist.gov/system/files/documents/cyberframework/cybersecurity-framework-021214.pdf , consulted 20250515

[37]Nmap ("Network Mapper") is a free and open-source (license) utility for network discovery and security auditing. Many systems and network administrators also find it useful for tasks such as network inventory, managing service upgrade schedules, and monitoring host or service uptime, available at https://nmap.org/, consulted 20250515

[38]PhishTank, https://phishtank.org/, consulted 20250515

[39]Pontes E., Guelfi A. E., Kofuji S. T., Silva A. 2011. Applying multi-correlation for improving forecasting in cyber security, 2011 Sixth International Conference on Digital Information Management, IEEE. DOI: https://doi.org/10.1109/ICDIM.2011.6093323

[40]Rollins, J. B. 2015. Foundational Methodology for Data Science, *IBM big data and analytics hub*, vol. 2017.

[41]Romanosky, S. 2016. Examining the costs and causes of cyber incidents, Journal of Cybersecurity, Volume 2, Issue 2, December 2016, Pages 121–135, DOI: https://doi.org/10.1093/cybsec/tyw001

[42]Sarker, I. H., Kayes, A. S. M., Badsha, S., Alqahtani, H., Watters, P. and Ng, A. 2020. Cybersecurity data science: an overview from machine learning perspective, *Journal of Big Data* volume 7, Article number: 41 (2020) DOI: https://doi.org/10.1186/s40537-020-00318-5

[43]Security Trails, https://securitytrails.com/, consulted 20250515

[44]Shodan.io, https://www.shodan.io/ , consulted 20250515

[45]Sqlmap is an open-source penetration testing tool that automates the process of detecting and exploiting SQL injection flaws and taking over database servers. Available at https://sqlmap.org/consulted 20250515

[46]Taleb, N. N. 2007. *Black Swan*, Penguin Random House.

[47]The Mordor project, https://github.com/UraSecTeam/mordor consulted 20250515

[48]Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323. https://doi.org/10.1007/BF00122574

[49]Virus Total, https://www.virustotal.com/, consulted on 20250515

[50]World Economic Forum, 2022, The global risks report 2022, available at https://www.weforum.org/reports/global-risks-report-2022, consulted on 20250515

[51]Wrotniak, K. and Wozniak, M. 2012. Combined Bayesian classifiers applied to spam filtering problem. Paper presented at the International Joint Conference CISIS’12-ICEUTE´12-SOCO´12 Special Sessions pp 253-260, Ostrava, Czech Republic, September 5–7. DOI: https://doi.org/10.1007/978-3-642-33018-6_26