a cyberSecurity Platform for vIrtualiseD 5G cybEr Range services

D5.4: An empirical decision support framework for econometric analysis of cyber risk and investment

Grant Agreement number:	833685
Project acronym:	SPIDER
Project title:	a cyberSecurity Platform for vIrtualiseD 5G cybEr Range services
Start date of the project:	01/07/2019
Duration of the project:	36 months
Type of Action:	Innovation Action (IA)
Project Coordinator:	Name: Pier Luigi Polvanesi Phone: +39 010 600 2662 e-mail: pierluigi.polvanesi@ericsson.com

Due Date of Delivery:	31/10/2021
Actual Date of Delivery:	31/10/2021
Work Package:	WP5 – Economics of 5G Security
Type of the Deliverable:	Report
Dissemination level:	Public
Main Editors:	Michail Chronopoulos (CITY)
Version:	1

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833685.

List of Authors, Contributors, Reviewers

Name	Role	Organisation
Maria Tsiodra	Author	CITY
Michail Chronopoulos	Co-Author	CITY
Matthias Ghering	Contributor	CLS
Irene Karapistoli	Contributor	CLS
Antonio Álvarez	Contributor	ATOS
Jorge Martinez Olmo	Contributor	ATOS
Panagiotis Gouvas	Contributor	UBITECH
Maurizio Giribaldi	Contributor	INFO
Manos Athanatos	Reviewer	FORTH
Nikolaos Petroulakis	Reviewer	FORTH
George Amponis	Reviewer	K3Y

History of Changes

Version	Date	Change History	Author	Organisation
0.0	12.04.2021	Initial version	Maria Tsiodra and Michail Chronopoulos	CITY
0.1	26.06.2021	Updated version.	Maria Tsiodra and Michail Chronopoulos	CITY
0.2	01.10.2021	Updated version.	Maria Tsiodra and Michail Chronopoulos	CITY
0.3	15.10.2021	Updated version. Contributions to Section 9.	Matthias Ghering	CLS
0.4	15.10.2021	Cleaned-up version submitted for internal review.	Maria Tsiodra and Michail Chronopoulos	CITY
0.5	27.10.2021	Updated version following the internal review.	Maria Tsiodra and Michail Chronopoulos	CITY
0.6	28.10.2021	Updated version following the internal review.	Maria Tsiodra and Michail Chronopoulos	CITY
0.7	29.10.2021	Updated version. Contribution to Section 9.	Maria Tsiodra and Michail Chronopoulos	UBITECH
0.8	29.10.2021	Updated version. Contribution to Section 9.	Maria Tsiodra and Michail Chronopoulos	UBITECH
0.9	30.10.2021	Updated version following the internal review.	Maria Tsiodra and Michail Chronopoulos	CITY
1	31.10.2021	Final version.	Maria Tsiodra and Michail Chronopoulos	CITY

The content of this deliverable is PUBLIC and must be handled according to SPIDER Consortium Agreement.

Glossary

Acronym	Explanation
APILA	Asset Pricing and Impact Loss Analysis
APT	Advanced Persistent Threat
CIC	Cybersecurity Investment Component
CIS	Centre for Internet Security
CRAE	Continuous Risk Assessment Engine
CVaR	Condition Valua at Risk
DCF	Discounted cash Flow
GDPR	General Data Protection Regulation
PV	Present Value
NPV	Net Present Value
VaR	Value at Risk

Disclaimer

The information, documentation and figures available in this deliverable are written by the SPIDER Consortium partners under EC co-financing (Call: H2020-SU-DS-2018, Project ID: 833685) and do not necessarily reflect the view of the European Commission.

The information in this document is provided "as is", and no guarantee or warranty is given that the information is fit for any particular purpose. The reader uses the information at his/her sole risk and liability.

Contents

List of Figures

List of Tables

1 EXECUTIVE SUMMARY

Controlling cyber risk is a critical, albeit notoriously complex task of cyber security management not only due to the uncertainties associated with a cyber attack and the resulting risk exposure for an organisation, but also due to the availability of scarce resources for investment in mitigation measures. In this report, we explore how the optimal set of mitigation measures, a.k.a. controls, may be driven by different optimisation objectives, in order to subsequently quantifying the risk exposure that each set of controls entails. Specifically, we first adopt a set covering formulation to determine the least number of controls required to cover all vulnerabilities. Subsequently, we use a Knapsack mathematical optimisation to identify the set of controls that minimise the expected impact of a cyber attack. Note that the latter approach builds upon D5.3 [2], where we evaluate the cost that a firm incurs as a result of a cyber security breach that progresses in phases, assuming that both the duration of an attack phase and the associated cost are random variables. In both the set covering and the Knapsack formulation, the optimisation is designed to account for various relevant constraints, such as the limited availability of financial resources and the desired efficacy of the controls. Thus, apart from identifying the optimal set of controls in each case, these models may be used to demonstrate not only how different objectives can be achieved, but also which approach is more effective in mitigating cyber risk.

2 INTRODUCTION

2.1 PURPOSE AND SCOPE

Building upon the Asset Pricing and Impact Loss Analysis (APILA) framework of D5.3, this report aims to demonstrate how the application potential of traditional mathematical programming techniques may be extended within the context of cyber security to facilitate the selection of mitigation measures subject to different constraints. Specifically, the scope of D5.4 is to apply optimisation methods based on the set covering and the knapsack formulation in order to address the following two questions:

i. What is the minimum number of controls that offer a baseline coverage of the network's vulnerabilities.

ii. What is the set of controls that minimises the present value (PV) of the impact of a cyberattack.

Note that the first question only utilises information regarding the coverage of the network's vulnerabilities by different controls and their cost, and, thus, it does not directly address the risk of a cyberattack. However, its contribution is in providing a point of reference in terms of comparing the residual risk under different optimisation objectives. Indeed, by gauging the risk reduction following the implementation of the controls proposed by the set covering problem, we can compare it with the reduction of risk associated with the second question. The latter utilises the output of the risk assessment framework developed in D5.3 to formulate objective functions that can be optimised to produce a set of controls that minimise the PV of the impact of a cyberattack. Thus, the scope of D5.4 is to analyse how the selection of mitigation measures is driven by different optimisation objectives and gauge the resulting risk exposure in each case in order to derive managerial insights.

2.2 MOTIVATION

Breakthroughs and advancements in the area of computer information systems have improved the operational efficiency of critical infrastructures, but have also rendered these substantially more vulnerable. The risk exposure and financial consequences that cyber attacks entail for an organisation can be demonstrated through a range of examples. Among the most recent breaches is that at Marriott that revealed personal details of approximately 5.2 million hotel guests. The breach at Twitter allowed fraudulent tweets about Bitcoin generating Bitcoin worth more than $\$100,000$, while the Solarwinds hack managed to compromise multiple government systems along with many fortune 500 companies, globally. The latter, also resulted in an 8% fall in the share price of FireEye after it disclosed information

about the cyber attack[1], and is expected to cost cyber insurers $\$90$ million for incident response and forensic services[2].

The aforementioned examples demonstrate how cyber security is a critical defensive manoeuvre as well as a strategic decision that may increase an organisation's competitive advantage. Furthermore, they emphasise the increasing need for developing economic models to assess cyber risk and deriving insights on how to invest in measures to mitigate them. However, while controlling cyber risk is the cornerstone of information security management, the uncertainties associated with cyber attacks, the resulting risk exposure, and the availability of scarce resources for investment in mitigation measures, make it challenging for organisations to assess and control cyber risk. Indeed, while cyber security models for optimising the selection of mitigation measures have evolved substantially from standard to multi-objective, bi-level models that facilitate strategic interactions, these have been developed mainly within a deterministic context that ignores key uncertainties of cyber attacks, and, consequently, they do not provide a formal assessment of cyber risk.

2.3 RELATION TO OTHER WORK IN THE PROJECT

A diagrammatic overview of the connection of the different tasks within WP5 as well as between WP5 and the general SPIDER platform has already been presented in D5.3 via Figure 1, which is also included below for ease of reference. Specifically, the scope of Figure 1 is to indicate: i. how the SPIDER simulated/emulated infrastructure as well as the SPIDER platform infrastructure provide data to be utilised within WP5; ii. the nature of the information and the way that this information is passed from one WP5 module to another; and iii. how the output of WP5 is reported to different SPIDER visualisation components. Note that, unlike in D5.3, here we highlight T5.4 in order to emphasise the key task associated with and underlying D5.4.

Furthermore, as T5.3 and T5.4 form part of the methodologies underlying the Cybersecurity Investment Component (CIC) that is developed in T5.5 [3], T5.4 relies on the same input as T5.3. More specifically, the quantitative analysis of the continuous risk assessment engine (CRAE) in T5.1 [1] receives input from the SPIDER Simulated/Emulated Infrastructure and provides improved information on asset values to the CIC. Also, as indicated in Figure 1, the CIC also receives information on attackers, assets and controls directly from the SPIDER Simulated/Emulated Infrastructure. Thus, together with the information from T5.1, the CIC calls the risk assessment and optimisation models developed in

[1] https://www.cnbc.com/2020/12/08/fireeye-shares-fall-after-security-company-discloses-cyberattack.html
[2] https://www.isaca.org/resources/news-and-trends/industry-news/2020/top-cyberattacks-of-2020-and-how-to-build-cyberresiliency

T5.3 and T5.4, respectively.

Also, note that T5.4 builds upon and extends T5.3 by integrating the output of the APILA framework within models for optimal selection of mitigation measures. Hence, the required input for T5.4 is provided partly by T5.3, in the form of the risk metrics to be optimised. However, T5.4 also requires information regarding controls, their efficacies, and limitations of financial resources in relation to the budget constraint. This input is not required in T5.3, as its scope does not extend beyond the assessment of cyber risk. Both the set of controls and the associated efficacies are provided the SPIDER platform, while constraints are provided by the SPIDER visualisation components.

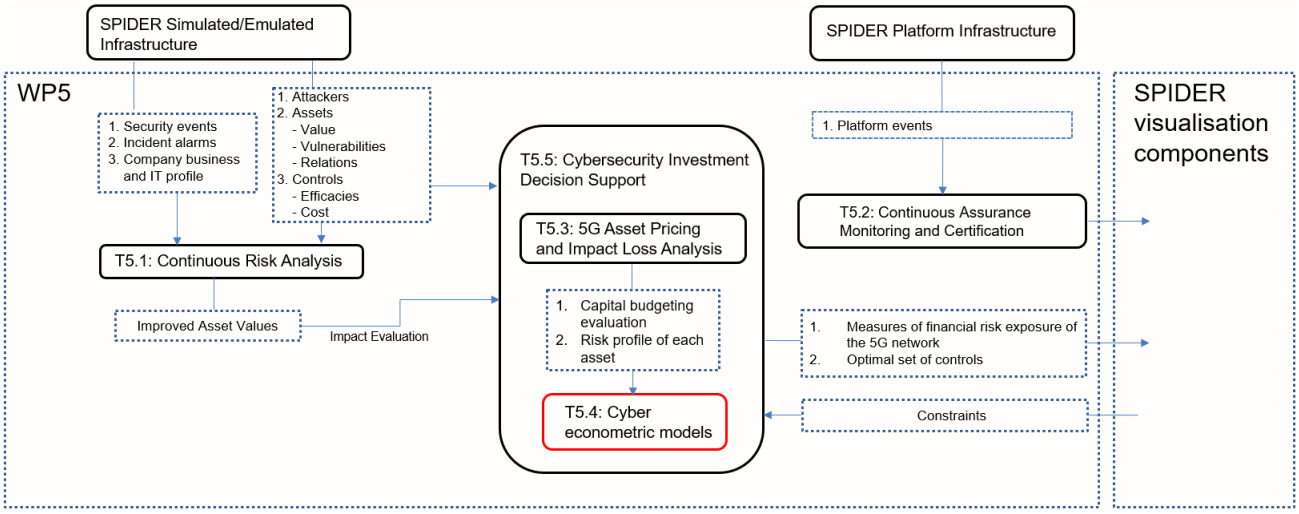


Figure 1: WP5 reference architecture.

2.4 STRUCTURE OF THE DOCUMENT

The report is organised as follows: In Section 3, we present the methodology we follow in order to produce D5.4, and, in Section 4, we provide an overview of the key novelties of this deliverable. Next, Section 5 discusses some related work in order to emphasise the contribution of D5.4 relative to the existing literature, while Section 6 gives a brief overview of the work developed in D5.4.

The control optimisation framework is presented in Section 7, where we begin with a description of the assumptions and the notation, and then proceed with a detailed treatment of the methodology in Section 8. For ease of exposition, we present the key elements of the APILA framework in order maintain a link with D5.3 and to facilitate a smooth transition to the development of the optimisation models.

The first optimisation model is presented in Section 8.2.1, where we cast the optimal selection of mitigation measures as a set covering problem. The set covering problem is a representative combinatorial optimization problem with many practical applications. Here, we adopt this formulation in order to determine the minimum number of controls that offer a baseline coverage of the network's vulnerabilities subject to efficacies and budget constraints. The second optimisation model is presented in Section 8.2.2, where we adopt a Knapsack formulation in order to address the question of how to select a subset of controls in order to minimise the expected PV of the impact of a cyber attack, which has already been determined in D5.3.

A comprehensive demonstration of the optimisation models will be presented upon completion of their integration within the CIC. Nevertheless, a brief overview is presented in Section 9 in terms of: i. a sample 5G network topology that is used to illustrate possible paths of the propagation of an attack, and, in turn, the 5G context within which the optimisation models will be carried out; and ii. the relevant input in terms of the list of available controls for a given set of vulnerabilities. Finally, Section 10 concludes the report by providing a summary of the key finding as well as an outline of the next steps in relation to the integration of D5.4 within the CIC.

3 METHODOLOGY FOLLOWED TO PRODUCE THIS DELIVERABLE

3.1 DATA COLLECTION

Note that, since D5.4 builds upon the APILA framework developed in D5.3, there is a substantial overlap in terms of the data utilised in each case. Therefore, detailed references to the data utilised in D5.3 will be omitted in order to emphasise the input required for the risk control framework of D5.4, and we begin the description of the methodology with an overview of the data on which it is based. The latter, can be classified into two main categories, i.e. qualitative and quantitative data.

3.1.1 QUALITATIVE DATA

Qualitative data pertain to various network characteristics, such as the assets of the network and the vulnerabilities within each asset. In D5.3, we make use of such data in order to produce risk metrics that can now be utilised within D5.4. In addition, since the objective here is to develop a decision-support framework that facilitates the selection of mitigation measures, a critical input is a comprehensive description of mitigation measures (controls) and how they map to different vulnerabilities. Security controls are safeguards or countermeasures to avoid, detect, counteract, or minimize security risks to physical property, information, computer systems, or other assets. In the field of information security, such controls protect the confidentiality, integrity and availability of information.[3]

3.1.2 QUANTITATIVE DATA

The key functionality of the risk control framework is to propose an optimal set of mitigation measures, in the light of scarce financial resources, taking into account their efficacy. Consequently, this requires information about the efficacy of each mitigation measure and the available budget. The efficacy of a cyber security control is defined as the probability of preventing the event that an attack is successful in compromising a vulnerability. Consequently, the efficacy of a given control is a number within the interval $[0, 1]$, although in many cases the implementation of a control means that a vulnerability is patched, i.e. protected with certainty. Data on the efficacies as well as their cost and the available budget will be provided by the SPIDER Simulated/Emulated infrastructure.

[3] https://en.wikipedia.org/wiki/Security_controls

3.2 RISK CONTROL FRAMEWORK

The goal of the risk control framework is to provide decision support regarding the selection of cyber security controls constrained by a financial budget. The objective of this selection is to find the combination of controls (also referred to as *security package*) that consists of: i. the minimum number of controls required to patch all vulnerabilities (Approach 1); and ii the controls that minimise the expected PV of the impact (Approach 2). Regarding Approach 1, the security package is derived as a solution to a set covering problem in order to account for the potential interaction among security packages that may overlap in terms of the vulnerabilities they cover. Regarding Approach 2, we use the risk assessment part of D5.3 along with the improvement of the expected impact when deploying security packages. We then derive the security package, which maximises the improvement of the expected PV of the impact subject to a financial budget. The same function may also take into account the cost of each security package, which is subtracted by the improvement. In this way, this framework optimises the return on security investment, which is defined as the ratio (benefit of security - cost of security) / cost of security [9].

4 INNOVATION

By extending the application of standard optimisation methods, namely set covering and knapsack, within the area of cyber security, the novelty of D5.4 is twofold: first, by adopting a set covering formulation, we obtain a benchmark solution in terms of an optimal set of controls that offers a baseline coverage of the network's vulnerabilities. Second, we adopt a knapsack formulation that directly integrates the expected PV of the impact of a cyber attack obtained via the APILA framework of D5.3. Specifically, the latter approach integrates the discounted cash flow (DCF) method for assessing the cyber risk associated with a multi-phase attack within a model for optimising the selection of cyber risk mitigation measures. Consequently, key novel aspects of the APILA framework, i.e. uncertainties regarding the impact and the duration of each phase of a cyber attack, are accounted for and integrated within an optimisation model, which, in turn, facilitates more informed decisions for investment in cyber security controls. Additionally, by developing two optimisation models for optimal selection of cyber security controls, we enable a comparison of different optimisation objectives in terms of the reduction in cyber risk following the implementation of the solutions they propose.

Furthermore, the novelty of D5.4 is that it facilitates important directions for future work. Indeed, the potential to utilise different risk metrics within the proposed optimisation models enables the development of a decision-support frameworks that account for risk preferences. In turn, this implies that the insights obtained via these optimisation models can reflect the implications of attitudes towards risk for investment in cyber security. Hence, although we demonstrate here the application potential of optimisation models in terms of minimising the expected PV of the impact of a cyber attack, other risk metrics could also be implemented within these models. Thus, the risk control framework of D5.3 is suitable for both risk-neutral and risk-averse decision makers.

Also, the optimisation methods proposed in D5.4 may be adopted within a framework that considers the strategic aspect of cyber security interactions. The objective would be to capture the interaction between the defender and the attacker as a non-cooperative game to determine the best responses against strategic attackers. From an application point of view, future work could extend the analysis to obtain recommendations against known threat actors by including adversarial information from CAPEC and MITRE ATT&CK framework.

Finally, within the context of public policy, governments advise organisations to get cyber security compliance certifications to demonstrate compliance with prescribed guidelines. For example, the UK government demands organisations to get certified for Cyber Essentials[4], which is a government-

[4] https://www.ncsc.gov.uk/cyberessentials/overview

backed scheme aimed to protect organisations from a range of most common cyber attacks. To be certified, organisations have to satisfy a list of requirements that cover five technical control themes: firewall, secure configuration, user access control, malware protection, and security update management. The set covering approach could be a comparable method used to identify controls to meet these requirements. However, this is a basic method that provides ideas on how to invest in cyber security and does not deal with cyber risk minimisation. On the other hand, the Knapsack optimisation method could be used to overcome the inefficacy of the set covering method to identify controls that minimise cyber risk.

5 BASELINE KNOWLEDGE

A strand of the cyber security economics literature draws on the theory of investment and project valuation under uncertainty [12, 8], with the main objective to derive the expected value of an investment in cyber security controls along with the investment threshold price and the probability of investment within a given time horizon [14]. This methodology, also known as *real options*, addresses the problem of investment under uncertainty while reflecting the value from embedded managerial discretion. For example, Gordon *et al.* [20] extend the framework of [19] by showing that information-sharing regarding vulnerabilities can decrease uncertainty about risks, and, in turn, the value of deferment options. More recently, Benaroch [8] develop a real options model to cast the cyber security investment problem as one of selecting a subset of uncertainty-reducing mitigation measures, whose availability is controlled by decision-makers and their size is log-normally distributed. The novelty of his work is to improve the efficiency of cyber security investments by balancing the costs of mitigation against their incremental uncertainty-reduction impact on cyber security loss expectancy. In the same line of work, Chronopoulos *et al.* [10] analyse how uncertainty over the cost of a cyber attack and the arrival of a control impacts the optimal time of investment in cyber security.

Although the optimal time of investment in cyber security controls is an important problem, especially considering the intensity and irreversibility of this capital expenditure as well as the various underlying uncertainties, the main limitation of the aforementioned literature is two-fold: First, decisions for mitigation of threats and protection of a network must be taken promptly, and, therefore, the value of waiting, which real options theory emphasises, is not as pronounced as it is in other industries, e.g. pharmaceutical, research and development, and energy. Second, real options models can be used to derive the expected value of an investment opportunity along with the investment threshold price, but they do not quantify the degree to which risk is hedged. The latter problem often fits within a security planning process, in terms of optimal selection of countermeasures [27, 28]. However, such models are typically deterministic, as they ignore key uncertainties underlying the nature of attacks for which the selection of countermeasures is designed [15, 22]. Consequently, the implications of uncertainty underlying key aspects of a cyber attack for the financial exposure of an organisation's assets and the choice of mitigation measures remains an important open research question.

Examples of models for optimal selection of cyber security controls include Leskovec *et al*. [25], who consider the general problem of detecting outbreaks in networks and demonstrate the application potential of their optimization model within the context of detection of contaminants or malicious ideas in a physical or social network, respectively. Also, Zhuo and Solak [39] propose a stochastic program-

ming model to optimize a firm's cybersecurity budget in an investment portfolio taking into account the uncertainty in the effectiveness of the countermeasures. The problem of optimal policy development for cyber vulnerability maintenance is presented in Afful-Dadzie and Allen [5], who propose multiple Markov decision processes to tackle data scarcity when developing IT network security maintenance policies. Furthermore, Nagurney *et al*. [26] propose a supply chain game theory framework consisting of retailers and consumers for managing vulnerability in electronic Internet transactions. While retailers compete noncooperatively in order to maximize their expected profits by determining their optimal product transactions as well as cybersecurity investments in the presence of network vulnerability, consumers reveal their preferences via demand price functions that depend on the product demand and on the average level of security in the supply chain network.

In the same line of work, Smeraldi and Malacaria [37] investigate the challenge of how to spend a security budget optimally and propose methods, such as optimisation algorithms, combinatorial optimisation and classical knapsack problem, that can deal with overlapping safeguards that exhibit nonlinear relationships. Similarly, Fielder *et al*. [15] propose a methodology for investing in CIS controls, considering a single value for a vulnerability, and a number of implementation levels for each control. The latter represent the information security levels proposed in the seminal work on the economics of information security by Gordon and Loeb [18]. Also, [28] extends the methodology proposed in [15] to obtain an optimal set of controls to protect various employee groups of a healthcare organisation from social engineering attacks.

Additionally, [38] cast the problem of optimal selection of controls as a set covering problem, whereby they first solve a deterministic version to analyse the incentive to implement complementary mitigations to reduce supply chain vulnerabilities. Subsequently, they extend the deterministic version to allow for limitations on the choice as well as uncertainty over the efficacy of the different controls. Building upon [6], [22] develop a game-theoretic framework, whereby the defender chooses a security plan seeking to minimise its security risk, while the attacker aims to maximise it via the most effective attack path. This is modelled as a min-max optimisation problem, where the maximisation problem is the attacker's, and the minimisation problem is the defender's, keeping in mind the reaction of the attacker.

A limitation of the aforementioned optimisation models is that they have been developed mainly within a deterministic context that ignores key uncertainties of cyber attacks, and, as a result, they do not provide a formal assessment of cyber risk. In finance, the risk exposure of a project is often measured by its Value at Risk (VaR), which is the minimum project value for a given confidence level during a specified time horizon, and by its Conditional VaR (CVaR), which is the expected value of the project

given that it is less than the VaR. Such risk measures have been recently applied to the cyber world and examples of empirical models that utilise risk measures, such as VaR and CVaR, for assessing risk and for analysing investment decisions in cyber security controls include [16], who propose an approach to estimate both the VaR and the Tail VaR using information on data breaches obtained from [31]. Also, Ekelund and Iskoujina [13] demonstrate how to find the optimal investment level in protecting an organisation's assets. Their framework combines theory and empirical findings, and proposes a new approach to determining optimal security investment levels. Via a case study on an international financial organisation, they demonstrate that the optimal security investment levels can be found through computer simulation of VaR using historical incident data. Specifically, by combining various scenarios, they plot the convex graph of the risk cost function, whereby the minimum of the graph indicates the optimal invest level for an asset. Other examples in the same line of work include [24, 29, 30, 32, 35].

Despite their novelty and contribution, a limitation of the aforementioned risk assessment and optimisation models is that they overlook the serial nature of a cyber attack and key uncertainties, such as the time it takes to exploit a vulnerability and the cost that the system incurs once a vulnerability is compromised. Hence, to assist in the anticipation and control of the financial impact of cyber attacks [33, 17], our work builds upon the literature on the valuation of serial projects to develop quantification tools for assessing the risk associated with a security breach that progresses in phases. For example, Creemers [11] studies the net present value (NPV) of a serial project, whereby a cash flow may be incurred at the start of each phase, a payoff is obtained at the end of the project, while the duration of each phase is a random variable with a general distribution function. The novelty of this work is that it derives an exact closed-form expression for the moments of the NPV of a project as well as a closed-form approximation of the distribution of the project's NPV. Therefore, we adopt and extend [11] within the context of cyber security in order to derive risk measures and quantify the risk exposure that a security breach entails so that we can subsequently utilise the risk measure within objective functions for optimal selection of mitigation measures.

Specifically, having derived the analytical expressions for the distribution of the PV of the impact of a cyber attack for each one of its phases, we will now develop methods for optimising the selection of controls. The latter is achieved by minimising the PV of the expected impact, yet the methodology we develop could be adopted to optimise any risk measure that can be derived from the distribution of the PV, e.g. VaR and CVaR. By casting the optimal selection of mitigation measures as the solution to a Knapsack and a set covering problem, we develop a framework for analysing the following set of testable hypothesis:

i. The Knapsack formulation provides solutions that not only are more affordable but also entail lower risk than those proposed by the set covering formulation.

ii. Greater investment intensity does not necessarily result in an analogous reduction of risk, which, in turn, implies that the rate of risk reduction decreases beyond a certain level of investment intensity.

iii. The VaR corresponding to the solutions obtained via the Knapsack formulation is lower compared to the VaR corresponding to the solutions obtained via the Set Covering formulation. This would further emphasise how the former approach is more suitable for providing solutions for controlling risk.

Thus, recommendations from the optimisation models proposed in this report can assist organisations to determine effective cyber security strategies aligned with the guidelines advocated by the UK government to protect small businesses against cyber attacks.

6 WORK DEVELOPED

The work developed in this report is twofold. First, we develop a framework for integrating the risk assessment methodology of D5.3 within objective functions to be optimised with respect to the choice of mitigation measures. Second, we develop two optimisation models for selection of mitigation measures in order to analyse and compare different objectives in terms of their respective proposed mitigation measures and the resulting risk exposure for the network. Note that via the integration of the risk assessment methodology of D5.3, the selection of mitigation measures is driven by objectives that reflect various critical uncertainties underlying a cyber attack. Consequently, the novelty of this risk-based approach to mitigating cyber risk is that it has the potential to facilitate optimal investment decisions that account for attitudes towards risk.

Additionally, in terms of the optimisation models themselves, the work developed in this report consists of two main parts that reflect the models for optimal selection of mitigation measures. Specifically, we develop two models in order to analyse and compare the reduction of cyber risk under different optimisation objectives.

i. The first model is based on a set covering formulation, and its objective is to determine the *minimum number of controls* that offer a baseline coverage of the network's vulnerabilities based on a budget constraint and the desired level of efficacy resulting from each patch.

ii. The second model is based on the Knapsack problem, which is a problem in combinatorial optimization. In its general version, given a set of items, each with a weight and a value, we want to determine the number of each item to include in a collection so that the total weight is less than or equal to a given limit and the total value is as large as possible.

7 CONTROL OPTIMISATION FRAMEWORK

7.1 SYSTEM MODEL

In this section, we present a summary of the framework for controlling cyber risk by adopting a mathematical programming approach that couples the risk assessment framework of D5.3 with optimisation of mitigation measures. In summary, this section discusses: i. the underlying system model with an organisation that wishes to protect its systems (Defender) and hackers who target the organisation (Attacker) and ii. how patching a percentage of these vulnerabilities, given a limited budget, leads to the challenge of optimally allocating cyber security controls, which may naturally overlap in terms of vulnerabilities they patch. We assume that the Defender's infrastructure consists of a number of systems and networks, referred to as assets, which the Defender aims to protect from the Attacker. Each asset $i \in \mathbb{N}$ has a set of vulnerabilities $\mathcal{V}_i = \{v_{i1}, v_{i2}, \ldots, v_{im_i}\}$ that the Attacker may exploit.

7.2 EXPECTED IMPACT

The objective function that will be optimised in Section 8.2.2 builds upon the expected PV of the impact of the cyber attack, as this is determined in [2]. Recall that the expected PV of the impact from exploitation of asset i is denoted by K_i, which is defined in (1).

$$K_i = A_i \cdot \langle R_i, S_i \rangle. \tag{1}$$

More specifically, A_i denotes the value of asset i, while $R_i = (r_{i1}, r_{i2}, \ldots, r_{im_i})$ and $S_i = (s_{i1}, s_{i2}, \ldots, s_{im_i})$. Note that r_{ij} is the likelihood of the Attacker attempting to exploit vulnerability v_{ij} and expresses the degree of attractiveness of a vulnerability to the Attacker, while s_{ij} is the probability of the same vulnerability being successfully breached. Thus, the likelihood of occurrence of an attack is expressed as $\langle R_i, S_i \rangle$, i.e. the inner product between R_i and S_i [34].

7.3 PRESENT VALUE OF THE EXPECTED IMPACT

For an attack phase i, T_i represents the time required to exploit a vulnerability v_{ij} in $\mathcal{V}_i$, and we refer to this as the hardness of exploiting a vulnerability. We assume that T_i follows a general distribution function denoted by $\Psi_{T_i}(\cdot)$, as shown in Figure 2. Assuming that an attack consists of a number of phases, each of them compromising an asset, we compute the duration of the attack (W_k) as the sum of the exploitation times required to compromise an asset in each phase, i.e. $W_k = \sum_{i=1}^{k} T_i, 1 \leq k \leq n$.

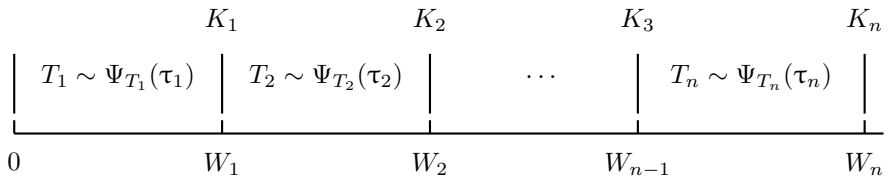


Figure 2: Sequential, multi-phase security breach.

To realise the expected impact that can materialise in the future, we determine the distribution of the PV of the expected impact associated with the attack. The aggregated expected impact (Z_n) over n attack phases in described in (2), where U_i denotes the PV of K_i, ρ denotes the discount rate, and W_i denotes the duration of the attack until phase i.

$$Z_n = \underbrace{K_1 e^{-\rho W_1}}_{U_1} + \underbrace{K_2 e^{-\rho W_2}}_{U_2} + \cdots + \underbrace{K_n e^{-\rho W_n}}_{U_n}, \tag{2}$$

As discussed in [2], the PV is used to introduce the concept of discounting into the calculation of the current value of the impact of a cyber attack, thereby supporting effective decision-making [11]. From a modelling standpoint, the contribution of this is to facilitate the development of risk measures that can be used to gauge the financial risk exposure of the Defender [21], and, subsequently, to enable the development of optimisation objectives for selections of mitigation measures.

7.4 CYBER RISK CONTROL

After gauging the potential risk exposure associated with each vulnerability, the risk control framework subsequently focuses on optimising the coverage of vulnerabilities in each asset by determining the appropriate Security Package. The latter refers to the set of controls that minimise the expected PV of the impact from an attack. This is done by patching asset vulnerabilities, thereby reducing an asset's attack surface or by increasing the effort required in successfully breaching the asset. The risk control framework specifically considers that the implementation of a control will mitigate the expected impact of an attack by reducing the probability of the latter being successful. We denote by $\mathcal{C} = \{C_1, C_2, \ldots, C_g\}$ the set of available controls and by E_{ijl} the efficacy of control C_l against vulnerability v_{ij}, where $l \in \{1, 2, \ldots, g\}$. Intuitively, E_{ijl} reflects the degree of protection offered by control C_l for a vulnerability v_{ij}. The effect of each control is formulated as $\widehat{S}_i = S_i \cdot \varepsilon_j$, where ε_j reflects the residual risk, i.e. $\varepsilon_j = \prod_l (1 - E_{ijl})$.

As the implementation of cyber security controls is not cost-free, the associated *direct* and *indirect* costs must be considered by the Defender. According to [9], the former refers to the acquisition, deployment, and maintenance costs of this control, while the latter can be anything else that inflicts loss to the Defender, such as slowing down essential processes due to incompatibility of controls and training employees to get acquainted with the new controls. For the sake of brevity, we assume that each control C_l comes with a set of costs Ξ_l associated with each level of implementation, inclusive of the direct and indirect costs. Table 1 presents a summary of the notation used in this report, and, specifically, complements the notation introduced in D5.3 with the notation relevant to the optimisation models. Note also that, although the optimisation models proposed in Section 8.2 are designed to account for the level of implementation of a given control, their application may be adjusted in simpler contexts, where these features may not be pertinent due to availability of data.

Symbol	Description
i	phase of attack or an asset ($i = 1, 2, \ldots, n$)
$\mathcal{V}_i$	set of vulnerabilities in asset i, $\mathcal{V}_i = \{v_{i1}, v_{i2}, \ldots, v_{im_i}\}$
$\mathcal{C}$	set of cyber security controls, $\mathcal{C} = \{C_1, C_2, \ldots, C_g\}$
$\mathcal{L}_l$	set of levels of control C_l, $\mathcal{L}_l = \{L_1, L_2, \ldots, L_h\}$, where $l = 1, 2, \ldots, g$
Ξ_l	set of cost of each level of control C_l, $\Xi_l = \{\xi_{l1}, \xi_{l2}, \cdots, \xi_{lh}\}$
$\mathcal{E}_{ijl}$	set of efficacy of each level of control C_l on vulnerability $v_{ij}, \mathcal{E}_{ijl} = \{E_{ijl1}, E_{ijl2}, \cdots, E_{ijlh}\}$
v_{ij}	vulnerability within asset i ($j = 1, 2, \ldots, m_i$)
r_{ij}	Probability of vulnerability v_{ij} being targeted (attack occurrence)
s_{ij}	Probability of vulnerability v_{ij} being compromised when attacked (success rate)
$x_{l\ell}$	Indicates whether level L_ℓ, $\ell = 1, 2, \ldots, h$, of control C_l is selected
y_{jl}	Indicates whether vulnerability v_{ij} is covered by C_l
A_i	Value of asset i
T_i	Time required to exploit a weakness in asset i
λ_i	Rate parameter for attack phase i
W_k	Total duration of the attack until phase k, $\sum_{i=1}^{k} T_i$ where $1 \leq k \leq n$
K_i	Expected impact from exploiting asset i
U_i	PV of the expected impact for attack phase i
Z_n	Aggregated expected impact for the first n attack phases, $\sum_{i=1}^{k} U_i$ where $1 \leq k \leq n$

Table 1: List of symbols.

8 ANALYSIS

8.1 RISK ASSESSMENT

For ease of exposition and to facilitate a smooth transition to the risk control framework, which utilises and builds upon the risk metrics developed in D5.3, we begin with a brief overview of the APILA framework. Note that the scope of the APILA framework is to assess the PV of the expected impact from a breach, and, to achieve this, we perform a phase-wise analysis of a cyber attack to determine: i. the distribution of the PV of the expected impact at each phase $i = 1, 2, \ldots, n$; ii. the distribution of the PV of the aggregated expected impact (Z_n) over n phases; and, iii. risk measures to gauge the financial risk exposure following the multi-staged attack. The main outcome of the APILA framework can be summarised via the result of the arbitrary n-th phase. Recall that $W_n \sim \text{Hypo}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and the cdf and pdf of W_n is described in (3) and (4), respectively [36].

$$\Phi_{W_n}(w_n) = 1 - \sum_{i=1}^{n} e^{-\lambda_i w_n} \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \tag{3}$$

$$\phi_{W_n}(w_n) = \sum_{i=1}^{n} \lambda_i e^{-\lambda_1 w_n} \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \tag{4}$$

Next, the cdf and pdf of U_n is described in (5) and (6), respectively.

$$\Theta_{U_n}(u_n) = \sum_{i=1}^{n} \left(\frac{u_n}{K_n}\right)^{\frac{\lambda_i}{\rho}} \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \tag{5}$$

$$\theta_{U_n}(u_n) = \sum_{i=1}^{n} \frac{1}{\rho K_n} \left(\frac{u_n}{K_n}\right)^{\frac{\lambda_i}{\rho}} \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \tag{6}$$

Having derived the distribution of U_n, we can now derive a wide range of risk measures to gauge the risk exposure associated with this arbitrary attack phase. In (7), we indicate the analytical expression of the mean, which we will use to formulate the optimisation objective.

$$\mathbb{E}[U_n] = K_n \sum_{i=1}^{n} \frac{\lambda_i}{\lambda_i + \rho} \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \tag{7}$$

Note also that since we have already derived the distribution of the impact for each phase, as expressed by the mean of present values U_i for all attack phases, the expected PV of Z_n is given by (8).

$$Z_n = \sum_{i=1}^{n} U_i \quad \Rightarrow \quad \mathbb{E}[Z_n] = \sum_{i=1}^{n} \mathbb{E}[U_i] = \sum_{g=1}^{n} K_g \sum_{i=1}^{g} \frac{\lambda_i}{\lambda_i + \rho} \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \tag{8}$$

8.2 RISK CONTROL

The optimisation models developed in T5.4 aim to provide decision support regarding the selection of cyber controls subject to different constraints, e.g. budget or efficacy constraints. Specifically, the objective of this selection is to find:

i. the minimum number of controls required to patch all vulnerabilities or

ii. the controls that reduce the expected PV of the impact optimally.

The first objective is addressed by adopting a set covering formulation, which also accounts for the potential interaction among security packages that may overlap in terms of the vulnerabilities they cover. For the second objective, we adopt a knapsack formulation that integrates the risk assessment part of the APILA framework developed in D5.3 together with the improvement of the expected impact following the implementation of security packages. Subsequently, we derive the security package for which the improvement of the expected PV of the impact subject to a financial budget is maximised.

8.2.1 SET COVERING PROBLEM

Here, we cast the problem of optimal selection of mitigation measures as a set covering problem [7], which is a classical question in combinatorics, computer science, operations research, and complexity theory. In its basic implementation, it involves a set of elements $\mathcal{U} = \{1, 2, 3, \ldots, M\}$ and a collection of subsets $\mathcal{S}_n, n \in \mathbb{N}$, whose union equals $\mathcal{U}$. The objective of the set covering problem is to identify the smallest sub-collection of subsets $\mathcal{S}_n$, so that their union equals $\mathcal{U}$. A basic illustration of this problem is provided in Figure 3, where each node denotes an element of $\mathcal{U}$ and each block represents a subset $\mathcal{S}_n$ of $\mathcal{U}$. Note that this reflect the basic implementation of this problem, which may be further refined depending on the intersection of the different subsets and the presence of constraints.

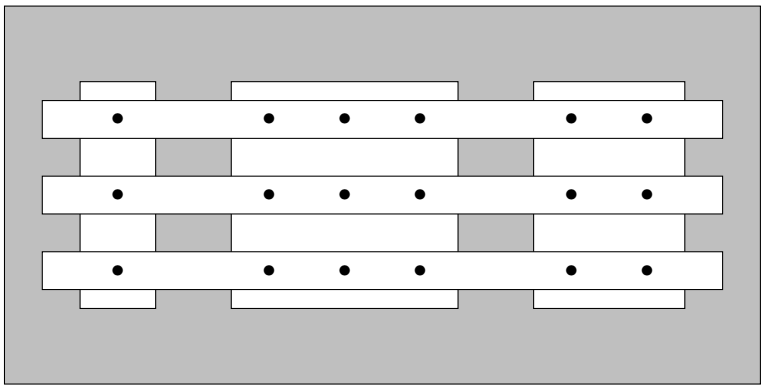


Figure 3: Set cover problem.

In the area of cyber security, the objective is to determine the *minimum number of controls* that offer a baseline coverage of the network's vulnerabilities. The formulation of the problem may also account for a budget constraint and the desired level of efficacy resulting from each patch. In terms of context, each node in Figure 4 represents a specific vulnerability and the set of all vulnerabilities is denoted by $\mathcal{U}$. However, as in D5.3, we assume that the total number of vulnerabilities within a network can be expressed in terms of each asset. Therefore, we denote by V_i the subset of vulnerabilities corresponding to asset $i = 1, 2, \ldots, n$. Also, we denote the set of controls by $\mathcal{C} = \{C_1, C_2, \ldots, C_g\}$ and assume that each control can protect a given subset of vulnerabilities at least partially or even offer protection that extends to another subset of vulnerabilities.

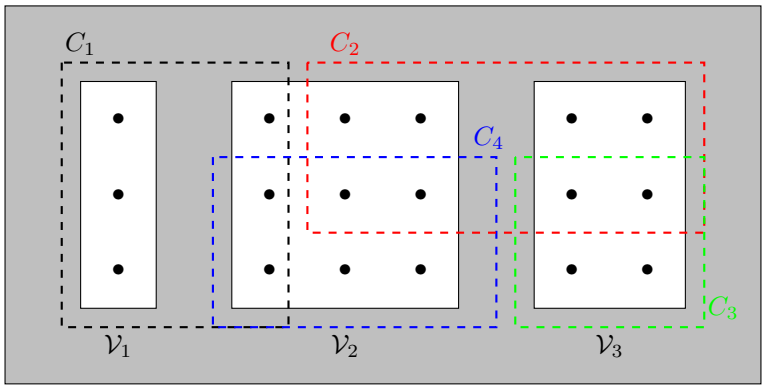


Figure 4: Cybersecurity set cover problem.

This optimisation is described in (9), where $x_{l\ell}$ is a binary variable indicating whether a specific level ℓ of a security control C_l is applied. Constraint (10) ensures that each vulnerability is covered by at least one control, while (11) is the budget constraint. Finally, constraint (12) ensures that the choice of controls offers a minimum coverage of all vulnerabilities at a desired level of efficacy, $\hat{e}$.

$$\min \sum_{l=1}^{g} x_{l\ell} \tag{9}$$

$$s.t. \sum_{l:E_{ijl\ell}>0} x_{l\ell} \geq 1, \;\; x_{l\ell} \in \{0,1\}, \;\; \forall i,j \in \mathbb{N} \tag{10}$$

$$\sum_{l=1}^{g} x_{l\ell}\xi_{l\ell} \leq B \tag{11}$$

$$E_{ijl\ell} > \hat{e}, \;\; \forall i,j \in \mathbb{N}, \;\; \hat{e} \in (0,1) \tag{12}$$

To facilitate the implementation of the set covering optimisation model, we include below the rel-

evant algorithm. Note that Approach 1 is appropriate when the underlying controls are related to patching vulnerabilities, as the degree of their effectiveness is 1, meaning that either the vulnerability is patched or not. In its basic implementation (Algorithm 1), the set covering problem does not account for the expected cost of the security breach. Indeed, the solution obtained via (9)-(12) ensures the minimisation of the number of controls, but does not consider whether the proposed controls minimise the expected impact of the attack or how they affect the associated risk. Furthermore, when we study preventative controls (e.g. firewalls), we must take into account their degree of effectiveness (henceforth referred to as Control Efficacy), of control against a vulnerability.

In a quantitative approach, the control efficacy falls within the interval $(0, 1)$. Algorithm 2 presents the set covering implementation with cost and control efficacy constraints. As stated in the formulation of the problem, the constraints affect the choice of controls. The Defender has a fixed Budget that can be utilised for implementing controls. A control is selected based on the number of vulnerabilities it covers and its cost. The selection is aimed at maximising the vulnerability coverage given the budget constraint. The selection process is terminated when: i. Budget is depleted; ii. there is no more controls to select. The only difference between Algorithm 1 and Algorithm 2 is the consideration of the control efficacy in line 4 which allows the selection of controls with efficacy of at least $\hat{e}$.

Algorithm 1: Set Cover Problem with Cost Constraint

Input: $\mathcal{V}_i, \mathcal{C}, \Xi_l$
Output: Minimum set of controls within the budget

1 **Function** `SetCoverCost`$(\mathcal{V}_i, \mathcal{C}, \Xi_l)$**:**
2 **for** $C_l \, in \, \mathcal{C}$ **do**
3 $price \leftarrow \xi_{l\ell}/len(C_l \cap \mathcal{V}_i)$
4 **if** $price < cost$ **then**
5 $cost \leftarrow price$
6 $cover \leftarrow C_l$
7 **return** $(cover, cost)$
8 **while** *len*$(\mathcal{V}_i) \,! = 0$ *and* $Budget \,! = 0$ **do**
9 $(cover, cost) \leftarrow$ `SetCoverCost`$(\mathcal{V}_i, \mathcal{C}, \Xi_l)$
10 $\mathcal{V}_i \leftarrow \mathcal{V}_i - cover$
11 $Budget \leftarrow Budget - cost$

Algorithm 2: Set Cover Problem with Cost and Control Efficacy Constraints

Input: $\mathcal{V}_i, \mathcal{C}, \Xi_l, \mathcal{E}_{ijl}, \hat{e}$
Output: Minimum set of controls with budget and efficacy bound

1 **Function** `SetCoverEfficacy`$(\mathcal{V}_i, \mathcal{C}, \Xi_l, \mathcal{E}_{ijl}, \hat{e})$**:**
2 **for** $C_l \, in \, \mathcal{C}$ **do**
3 $price \leftarrow \xi_{l\ell}/len(C_l \cap \mathcal{V})$
4 **if** $price < cost$ *and* $E_{ijl\ell} > \hat{e}$ **then**
5 $cost \leftarrow price$
6 $cover \leftarrow C_l$
7 **return** $(cover, cost)$
8 **while** *len*$(\mathcal{V}_i) \,! = 0$ *and* $Budget \,! = 0$ **do**
9 $(cover, cost) \leftarrow$ `SetCoverEfficacy`$(\mathcal{V}_i, \mathcal{C}, \Xi_l)$
10 $\mathcal{V}_i \leftarrow \mathcal{V}_i - cover$
11 $Budget \leftarrow Budget - cost$

8.2.2 KNAPSACK FORMULATION

Unlike Section 8.2.1, the objective here is to select the controls to invest in so that the expected PV of the impact, as expressed in (8), is minimised. The challenge of optimal budget allocation in cyber security can be addressed through combinatorial optimisation [15]. Besides minimising the expected value of a security breach, any rational Defender would explore ways to minimise the investments in security controls. Based on the selection of a control at a particular level, indicated through $x_{l\ell} \in \{0, 1\}$, the probability of exposure can be expressed as $\varepsilon_j = \prod_{l\in\mathcal{C},\ell\in\mathcal{L}_l} (1 - x_{l\ell}E_{ijl\ell})$. Notice how ε_j is a strictly decreasing function of $x_{l\ell}$, thereby indicating that the inclusion of a control will reduce the likelihood of exposure. T5.4 uses 0-1 Knapsack to obtain the optimal security package, which minimises the PV of the expected impact of a cyber attack given a budget constraint. Thus, T5.4 solves the following Knapsack problem:

$$\max_{\vec{x}} \min_{\mathbb{E}[Z_i]} \sum_{i=1}^{n} \left\{ \left\{ \prod_{l=1}^{g} \prod_{\ell=1}^{h} (1 - x_{l\ell}E_{ijl\ell}) \right\} \cdot \mathbb{E}[Z_i] \right\}, \quad \forall i, j \in \mathbb{N} \tag{13}$$

$$s.t. \quad \sum_{l=1}^{g} \sum_{\ell=1}^{h} x_{l\ell}\xi_{l\ell} \leq B \tag{14}$$

$$\sum_{\ell=1}^{h} x_{l\ell} = 1, \quad x_{l\ell} \in \{0, 1\}, \quad \forall l = 1, \cdots, g. \tag{15}$$

To facilitate the integration of the Knapsack optimisation model within the CIC, an outline of the implementation steps is indicated in Algorithm 3. Note that the optimal efficacy matrix O is constructed iteratively for all the cost values within the Budget, and for each value the problem is solved considering all the available levels of control within that cost. The optimal aggregated efficacy value $O[l, cost]$ depends on the control level selected for the i-th cost. For a detailed analysis of 0-1 Knapsack optimisation using dynamic programming refer to [23].

Algorithm 3: Dynamic Programming based 0-1 Knapsack Optimisation

Input: $\mathcal{V}_i, \mathcal{C}, \mathcal{L}_l, \Xi_l, \mathcal{E}_{ijl}$
Output: Optimal set of controls and total cost
Function `KnapsackOptimisation`$(\mathcal{V}_i, \mathcal{C}, \mathcal{L}_l, \Xi_l, \mathcal{E}_{ijl})$**:**
  for C_l *in* $\mathcal{C}$ **do**
    for $cost$ *in* $Budget$ **do**
      $O[C_l, cost] \leftarrow O[C_l - 1, cost]$
      for ℓ *in* $\mathcal{L}_l$ **do**
        if $cost \geq \xi_{l\ell}$ **then**
          $O[C_l, cost] \leftarrow \max\{O[C_l, cost], (O[C_l - 1, cost - \xi_{l\ell}] + (1 - E_{ijl\ell})\}$

9 APPLICATION TO 5G NETWORKS

A detailed demonstration of D5.4 will be presented once its integration within the CIC is complete. Hence, here we provide a high-level overview of the key steps underlying the demonstration of D5.4, in a way that also illustrates the connection to the APILA framework.

i. The optimisation begins with the estimation of the impact of a cyber attack based on a sample network topology. As in D5.3, we consider a 5G network that consists of three assets, each with two vulnerabilities, and a possible propagation of a cyber attack across the different vulnerabilities is illustrated via Figure 5. Although this is a partial view of a 5G network, different variations may be produced via the more general illustration of network of assets indicated in Figure 11.

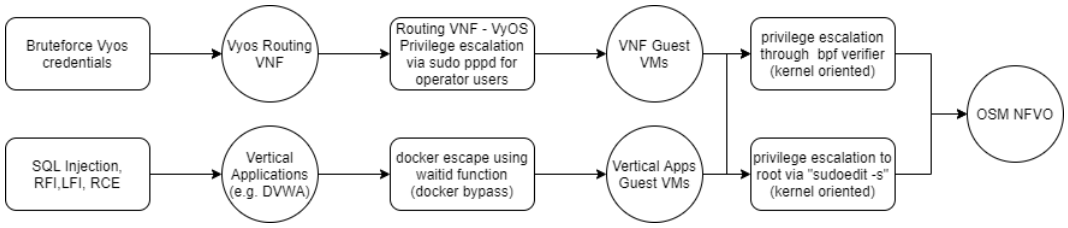


Figure 5: Attack graph representing the attacker actions in a vulnerable artificial 5G infrastructure.

ii. Next, we continue by considering how the selection of controls can be carried out to address the objectives presented in Section 8.2. The controls that may be utilised to safeguard the vulnerabilities indicated in the above attack graph are described in Table 9.

Table 2: Controls and efficacies.

Vulnerability		Control	
V1	Privilege escalation to root via "sudoedit -s" (kernel oriented)	C1	Update 'sudo' to 1.9.8p2
V2	Privilege escalation through bpf verifier (kernel oriented)	C2	Apply Linux Kernel 5.6.1
V3	Docker escape using waitid function (docker bypass)	C3	
V4	Remote code execution using OpenSLP (hypervisor takeover from external entity) (host to guest)	C4	Upgrade VMWareESXI to 7.0
V5	EPYC escape (guest to host escape for KVM)	C5	Apply Linux Kernel 5.11.12
V6	Routing VNF - VyOS Privilege escalation via sudo pppd for operator users	C6	Patch VyOS to 1.1.9
V7	"Attacking the SDN Interface of an OSS - SQL injection in the component database(SQLite) without authenticating to the controller or SDNInterfaceapp"	C7	
V8	WiFi - dictionary attack	C8	– use WPA → low mitigation – use WPA and complex password → medium mitigation – use WPA and complex password and rotating policy → high mitigation (not full)
V9	Remote Code Execution	C9	Update to blazar-dashboard to 6.0
V10	RRC IMSI catcher	C10	
V11	UE Denial Service (OpenLTE)	C11	
V12	SBA null policy	C12	Deactivate Null Policies
V13	SBA key retrival (non integrity)	C13	
V14	Unencrypted PWS	C14	
V15	SQL Injection, RFI,LFI	C15	Install WAF
V16	weak credentials, bruteforce attack	C16	Enforce corporate policies fo strong credentials

10 CONCLUSIONS

Efficient cyber security risk management relies on managerial strategies that are responsive to various uncertainties associated with cyber attacks. The need for such strategies becomes particularly pronounced considering the critical impact cyber attacks may have on organisations and the limited time to make executive decisions. Hence, risk management within the area of cyber security is a considerably delicate task. The presence of uncertainties raises the incentive to postpone decisions and, in turn, the value of waiting, which is often a luxury that cannot be afforded. In this report, we built upon D5.3, where we take into account the serial nature of the cyber attack and the uncertainty in the time required to exploit a vulnerability and develop a decision-support framework to evaluate the risk exposure of an organisation, and propose an optimal set of mitigation measures. This framework is designed to provide insights to the following three testable hypothesis.

i. The Knapsack optimisation methodology not only provides solutions that entail greater reduction of expected impact and lower cost when compared to the solutions of the set covering approach. This means that, by implementing the suggested solution, we achieve a better reduction of expected impact and better chances to gain a positive return on the security investment.

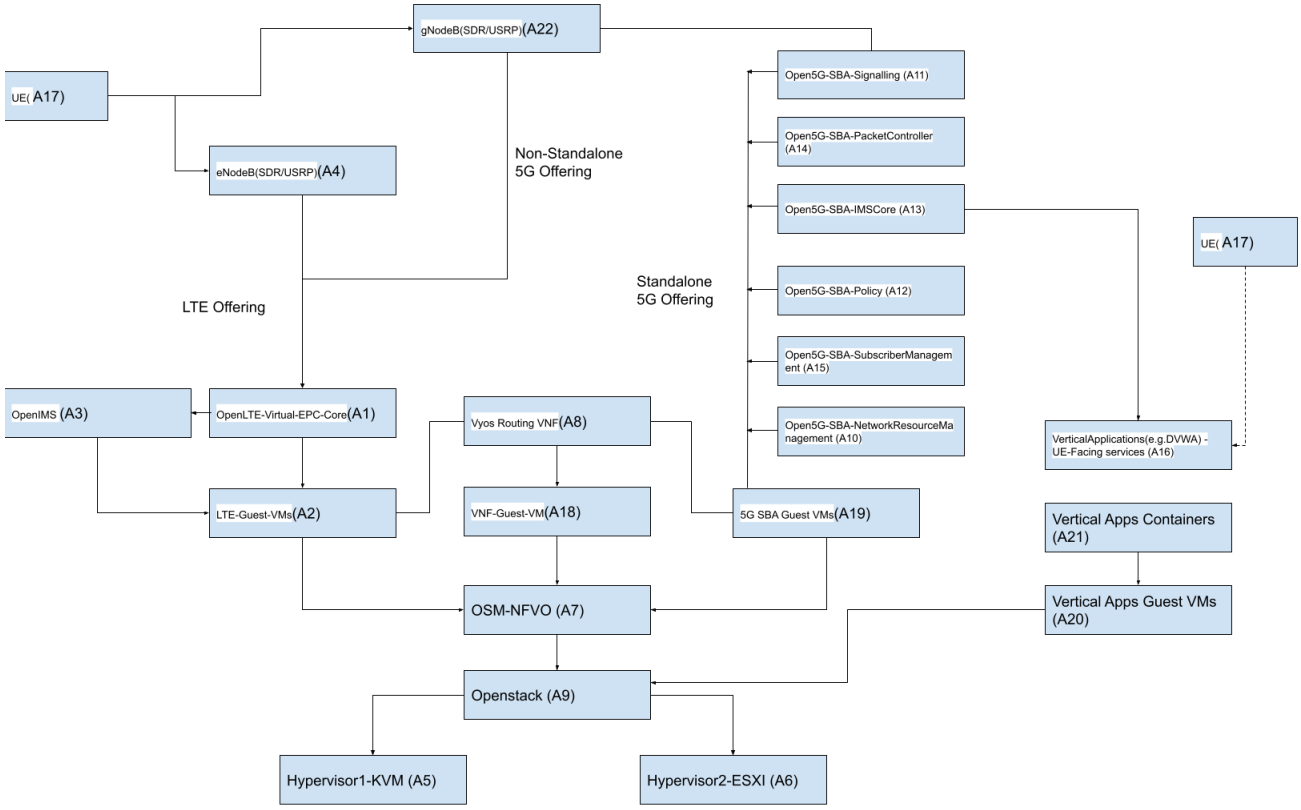
ii. Investing more in security does not necessarily lead to a better reduction of risk. This would comply with the acclaimed Gordon-Loeb [18], who highlight that the rate of risk reduction decreases after a point of security investment. With the decreasing rate of risk reduction, there is a diminishing return on each additional control implemented thereafter, which means that after a certain threshold it is costly for the organisation to mitigate additional risk.

iii. The VaR corresponding to the solutions obtained via the Knapsack formulation is in most cases lower compared to the VaR corresponding to the solutions obtained via the set covering formulation. This would further emphasise how the former approach is more suitable for providing solutions for controlling risk.

Testing the aforementioned hypothesis, is the objective of the next steps within the context of SPIDER. This requires i. input regarding a comprehensive list of available controls as well as the efficacy and cost of each control; and ii. the integration of the optimisation models proposed in D5.4 within the CIC.

11 APPENDIX

Based on the scenarios described in WP2, a general asset graph is illustrated in Figure 11. This may be used to revise the network topology proposed in this report.



ID	Name	Description	Vulnerabilities	Value
A1	OpenLTE-Virtual-EPC-Core	Implements the Radio Resource Control signalling	V10,V11	M
A2	LTE-Guest-VMs	Linux-based VMs that are used for hosting virtualized LTE components	V1,V2	M
A3	OpenIMS	An open source implementation of LTE-IMS		M
A4	eNodeB(SDR/USRP)	The attachment interface of LTE	V15	M
A5	Hypervisor1-KVM	Raw compute resources based on KVM hypervisor that are used during slicing	V5	M
A6	Hypervisor2-ESXI	Raw compute resources based on ESXI hypervisor that are used during slicing	V4	M
A7	OSM NFVO	Open Source Mano - Network Function Virtualization Orchestrator	V15	M
A8	Vyos Routing VNF	Vyos Open source routing VNF used during slice generation	V6	H
A9	Openstack VIM	The base Virtualized Infrastructure Manager	V9	M
A10	Open5G-SBA-NetworkResourceManagement	An open source reference implementation of the Network Resource Management components of the 5G SBA		M
A11	Open5G-SBA-Signalling	An open source reference implementation of the Signalling components of the 5G SBA	V12	M
A12	Open5G-SBA-Policy	An open source reference implementation of the Policy enforcement components of the 5G SBA		M
A13	Open5G-SBA-IMSCore	An open source reference implementation of the IMS components of the 5G SBA	V15	M
A14	Open5G-SBA-PacketController	An open source reference implementation of the Packet controller components of the 5G SBA		M
A15	Open5G-SBA-SubscriberManagement	An open source reference implementation of the subscriber management components of the 5G SBA	V13	H
A16	Vertical Applications (e.g. DVWA)	It represents any (layer-7) vulnerable component that offers IP-based services to UEs	Layer-7-Vulns	M
A17	UE	The user equipment		M
A18	VNF Guest VMs	Linux-based VMs that are used for hosting virtualized VNFs	V1,V2	M
A19	5G SBA Guest VMs	Linux-based VMs that are used for hosting virtualized SBA components	V1,V2	M
A20	Vertical Apps Guest VMs	Linux-based VMs that are used for hosting vertical application containers	V1,V2	M
A21	Vertical Apps Containers	The containers that isolate the vertical applications	V3	M
A22	GNodeB(SDR/USRP)	The attachment interface of 5G (in a standalone version)	V8,V14	M

Figure 6: Indicative list of assets.

ID	Name	Formal Identifier (if existing)
V1	privilege escalation to root via "sudoedit -s" (kernel oriented)	CVE-2021-3156
V2	privilege escalation through bpf verifier (kernel oriented)	CVE-2020-8835
V3	docker escape using waitid function (docker bypass)	CVE-2017-5123
V4	remote code execution using OpenSLP (hypervisor takeover from external entity) (host to guest)	CVE-2020-3992
V5	EPYC escape (guest to host escape for KVM)	CVE-2021-29657
V6	Routing VNF - VyOS Privilege escalation via sudo pppd for operator users	CVE-2018-18556
V7	Attacking the SDN Interface of an OSS - SQL injection in the component database(SQLite) without authenticating to the controller or SDNInterfaceapp	CVE-2018-1132
V8	WiFi - dictionary attack	CVE-2019-12941, CVE-1999-1152, CVE-2001-1291, CVE-2001-0395, CVE-2001-1339, CVE-2002-0628
V9	Remote Code Execution	CVE-2020-26943
V10	RRC IMSI catcher	
V11	UE Denial Service (OpenLTE)	
V12	SBA null policy	
V13	SBA key retrival (non integrity)	
V14	Unencrypted PWS	
Layer-7-Vulns	SQL Injection, RFI,LFI	
V15	weak credentials, bruteforce attack	

Figure 7: Indicative list of vulnerabilities.

References

[1] D5.1 - Continuous risk analysis: models and assessment engine - initial version.

[2] D5.3 - Asset pricing and impact loss analysis: An empirical framework.

[3] D5.7 - SPIDER cybersecurity investment component – final version.

[4] Grant Agreement NUMBER 833685 - SPIDER.

[5] Anthony Afful-Dadzie and Theodore T. Allen. Data-driven cyber-vulnerability maintenance policies. *Journal of Quality Technology*, 46(3):234–250, 2014.

[6] Hussain M.J. Almohri, Layne T. Watson, Danfeng Yao, and Xinming Ou. Security optimization of dynamic networks with probabilistic graph modeling and linear programming. *IEEE Transactions on Dependable and Secure Computing*, 13(4):474 – 487, 2016.

[7] John E Beasley and Paul C Chu. A genetic algorithm for the set covering problem. *European journal of operational research*, 94(2):392–404, 1996.

[8] Michel Benaroch. Real options models for proactive uncertainty-reducing mitigations and applications in cybersecurity investment decision making. *Information Systems Research*, 29(2):315–340, 2018.

[9] Rainer Böhme. Security metrics and security investment models. In *International Workshop on Security*, pages 10–24. Springer, 2010.

[10] Michail Chronopoulos, Emmanouil Panaousis, and Jens Grossklags. An options approach to cybersecurity investment. *IEEE Access*, 6:12175–12186, 2017.

[11] Stefan Creemers. Moments and distribution of the net present value of a serial project. *European Journal of Operational Research*, 267(3):835–848, 2018.

[12] Avinash K Dixit, Robert K Dixit, and Robert S Pindyck. *Investment under uncertainty*. Princeton university press, 1994.

[13] Stale Ekelund and Zilia Iskoujina. Cybersecurity economics – balancing operational security spending. *Information Technology & People*, 32(5):1318–1342, 2019.

[14] Alison Etheridge and Martin Baxter. *A course in financial calculus*. Cambridge University Press, 2002.

[15] Andrew Fielder, Emmanouil Panaousis, Pasquale Malacaria, Chris Hankin, and Fabrizio Smeraldi. Decision support approaches for cyber security investment. *Decision Support Systems*, 86:13–23, 2016.

[16] Maria Francesca Carfora and Albina Orlando. Quantile based risk measures in cyber security. In *2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, pages 1–4, 2019.

[17] Daniel Geer, Kevin Soo Hoo, and Jaquith Andrew. Information security: Why the future belongs to the quants. *IEEE Security & Privacy*, 1(2):32–40, 2003.

[18] Lawrence A Gordon and Martin P Loeb. The economics of information security investment. *ACM Transactions on Information and System Security (TISSEC)*, 5(4):438–457, 2002.

[19] Lawrence A Gordon, Martin P Loeb, and William Lucyshyn. Information security expenditures and real options: A wait-and-see approach. *Computer Security Journal*, 19(2), 2003.

[20] Lawrence A Gordon, Martin P Loeb, and William Lucyshyn. The impact of information sharing on cybersecurity underinvestment: A real options perspective. *Journal of Accounting and Public Policy*, 34(2):509–519, 2015.

[21] Lawrence A Gordon, Martin P Loeb, William Lucyshyn, and Lei Zhou. Increasing cybersecurity investments in private sector firms. *Journal of Cybersecurity*, 1(1):3–17, 2015.

[22] MHR Khouzani, Zhengliang Liu, and Pasquale Malacaria. Scalable min-max multi-objective cyber-security optimisation over probabilistic attack graphs. *European Journal of Operational Research*, 278(3):894–903, 2019.

[23] Satish Kumar, Arnab Sarkar, and Arijit Sur. A resource allocation framework for adaptive video streaming over lte. *Journal of Network and Computer Applications*, 97:126–139, 2017.

[24] Yee Wei Law, Tansu Alpcan, and Marimuthu Palaniswami. Security games for risk minimization in automatic generation control. *IEEE Transactions on Power Systems*, 30(1):223–232, 2015.

[25] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, page 420–429, New York, NY, USA, 2007. Association for Computing Machinery.

[26] Anna Nagurney, Ladimer S. Nagurney, and Shivani Shukla. *A Supply Chain Game Theory Framework for Cybersecurity Investments Under Network Vulnerability*, pages 381–398. 2015.

[27] Pantaleone Nespoli, Dimitrios Papamartzivanos, Félix Gómez Mármol, and Georgios Kambourakis. Optimal countermeasures selection against cyber attacks: A comprehensive survey on reaction frameworks. *IEEE Communications Surveys & Tutorials*, 20(2):1361–1396, 2017.

[28] Sakshyam Panda, Emmanouil Panaousis, George Loukas, and Christos Laoudias. Optimizing investments in cyber hygiene for protecting healthcare users. In *From Lambda Calculus to Cybersecurity Through Program Analysis*, pages 268–291. Springer, 2020.

[29] Chen Peng, Maochao Xu, Shouhuai Xu, and Taizhong Hu. Modeling and predicting extreme cyber attack rates via marked point processes. *Journal of Applied Statistics*, 44(14):2534–2563, 2017.

[30] Chen Peng, Maochao Xu, Shouhuai Xu, and Taizhong Hu. Modeling multivariate cybersecurity risks. *Journal of Applied Statistics*, 45(15):2718–2740, 2018.

[31] Privacy Rights Clearing House. Chronology of data breaches, 2019.

[32] Terry R Rakes, Jason K Deane, and Loren Paul Rees. It security planning under uncertainty for high-impact events. *Omega*, 40(1):79–88, 2012.

[33] T Mercuri Rebecca. Analyzing security costs. *Communications of the ACM*, 46(6):15–18, 2003.

[34] Ronald S Ross. Guide for conducting risk assessments. Technical report, 2012.

[35] Tadeusz Sawik. Selection of optimal countermeasure portfolio in it security planning. *Decision Support Systems*, 55(1):156–164, 2013.

[36] M Sheldon, Ross. *Introduction to Probability Models*. Academic press, 2010.

[37] Fabrizio Smeraldi and Pasquale Malacaria. How to spend it: optimal investment for cyber security. In *Proceedings of the 1st International Workshop on Agents and CyberSecurity*, pages 1–4, 2014.

[38] Kaiyue Zheng, Laura A Albert, James R Luedtke, and Eli Towle. A budgeted maximum multiple coverage model for cybersecurity planning and management. *IISE Transactions*, 51(12):1303–1317, 2019.

[39] Yueran Zhuo and Senay Solak. Measuring and optimizing cybersecurity investments: A quantitative portfolio approach. 2014.