Quantitative Assessment of Software Vulnerabilities Based on Economic-Driven Security Metrics

Hamza Ghani, Jesus Luna, Neeraj Suri
Technische Universität Darmstadt, Germany
{ghani, jluna, suri}@deeds.informatik.tu-darmstadt.de

***Abstract*—Vulnerability exploits cost organizations large amounts of resources, mainly due to disruption of ICT services, and thus loss of confidentiality, integrity and availability. As security managers in the industry usually have to operate with limited budgets allocated to information security, they need to prioritize their investment efforts regarding the response mechanisms to the existing vulnerabilities. The utilization of quantitative security vulnerability assessment methods enables efficient prioritization of security efforts and investments to mitigate the discovered vulnerabilities and thus an opportunity to lower expected losses. State of the art approaches for vulnerability assessment such as the Common Vulnerability Scoring System (CVSS), which is the de facto standard quantifying the severity of vulnerabilities, do not consider the economic impact in case of a vulnerability exploit. To this end, our paper targets the quantitative understanding of vulnerability severity taking into account the potential economic damage a successful vulnerability exploit can cause. We propose a novel approach for a systematic consideration of the relevant cost units (associated costs) for the economic damage estimation of vulnerability exploits. Our approach utilizes Multiple Criteria Decision Analysis (MCDA) methods to perform a prioritization of the existing vulnerabilities within the target system. The evaluation results show the potential cost savings w.r.t. the mitigation costs using our approach. Our method supports managers and decision makers in the process of prioritizing security investments to mitigate the discovered vulnerabilities.**

***Index Terms*—CVSS, economic-driven security metrics, MCDA, security quantification, vulnerability assessment.**

I. Introduction

Security vulnerabilities are inherent to software systems. The growing number of vulnerabilities is an ever increasing challenge to both public and private organizations. According to the U.S. National Vulnerability Database (NVD) [2], there have been 5281 disclosed vulnerabilities in 2012 compared to 4151 in 2011. However, not all disclosed vulnerabilities have been mitigated/fixed, i.e., through patches or software updates. According to [3], 38% of all disclosed vulnerabilities were not patched at the end of 2012. Furthermore, security managers operate usually within limited budgets so that there is a need to prioritize the discovered vulnerabilities, and thus the mitigation actions to be considered. A vulnerability should be addressed by a response process that is appropriate to its severity, and more severe vulnerabilities should be prioritized over less severe ones. Vulnerability response processes are not uniform and can differ regarding response timeliness, involved roles, impact on production process and operations, and especially the total response costs. It is not uncommon that technically critical vulnerabilities, do not have the highest economic impact on the organization. Thus, the process of vulnerability severity assessment and prioritization is a real challenge and delicate task. Vulnerability prioritization has been discussed in the state of the art literature and the need for vulnerability prioritization in organizations is widely recognized [15], [16], [19], [20]. However, to define the severity of a vulnerability w.r.t. the underlying context, one needs to determine the relevant criteria to be used to assess that vulnerability. Existing vulnerability scoring approaches, e.g., the Common Vulnerability Scoring System (CVSS), help perform such vulnerability prioritization *from a pure technical perspective* using quantitative scores. The common usage of CVSS, e.g., in publicly available vulnerability databases and scanning tools, is restricted to the so called *base score* and omits all information about a vulnerability's context. Thus, it outputs the same severity scores regardless of the specificities of the affected organization. In practice however, vulnerability impact differs greatly among various organizational contexts. We further explain the missing context using a motivating example in Section II. Previous works such as [20], [21] acknowledge that problem and advise that prioritizing vulnerabilities based on measures such as CVSS scores should be used with caution.

Empirical research also has shown that the actual impact of vulnerability exploits varies significantly among different types of organizations, businesses and users [17], [18]. Since different organizations perceive the severity of a particular vulnerability differently, they also prioritize its mitigation differently [19]. From a technical perspective, CVSS can account for these differences to a certain extent if the user security requirements, which are captured by the optional *environmental metrics*, are applied in the scoring process along with the *base metrics*. Therefore, by considering user's security requirements, the quality/customization of the scores can be improved because they better reflect the actual impact of a vulnerability in particular organization's context [21], and thus improve the prioritization of vulnerabilities from a security management perspective. There is a need for adding the missing context information, especially the economic perspective on the potential costs, in order to perform credible, real-world compliant vulnerability severity assessments. The choice of appropriate response mechanisms implies choosing the ones mitigating the targeted vulnerabilities while causing the lowest total costs, and, *without* ignoring the potential

978-1-4799-3488-1/13/$31.00 ©2013 IEEE

economic damage that a vulnerability exploit could cause. The main driver of this paper is to take into consideration additional contextual information within the vulnerability assessment process, especially the economic aspects reflecting (a) vulnerability response costs, and (b) potential economic damage if the vulnerability is exploited. Furthermore, special interest is given to user security requirements, which vary significantly among different organizational contexts. We outline the research questions addressed in this paper as follows:

- RQ1: Which economic factors and thus economic-driven metrics need to be taken into account within the vulnerability assessment process?
- RQ2: Once the new economic-driven metrics are identified, how to aggregate all the relevant criteria for the severity assessment (e.g., CVSS (sub)metrics); thus enabling a quantitative assessment and ranking of the existing vulnerabilities in the organization's software systems?

Paper contributions: The main contributions of this paper are the following:

- C1: A systematic approach eliciting the economic-driven metrics for vulnerability assessment.
- C2: A quantitative technique to aggregate the technical and economical metrics in a holistic way in order to rank vulnerabilities and reason about their mitigation priorities within an organization.

The main objective is to obtain a context-aware quantitative ranking of existing vulnerabilities affecting a real-world software system. The proposed technique is inspired by the Multiple Criteria Decision Analysis (MCDA) [10], [12], which is an established methodology widely used in fields like operations research and quantitative management methods.

Paper organization: The remainder of this paper is organized as follows: Section II illustrates based on a motivating example the need for our approach. Then we introduce basic concepts and terminology in Section III. Section IV provides an overview of our economic-driven metrics, and Section V our integrative MCDA-based vulnerability ranking method. The results of the theoretical analysis of our approach as well as a case study are presented in Section VI. Section VII presents the existing related approaches and Section VIII provides conclusions for the work.

II. Motivating Example

Let us consider the following business scenario: A small and medium-sized enterprise (SME X) is specialized in providing online games for consumers. Among the main information technology infrastructure components used by SME X there is a Cisco universal broadband router uBR10000 series.

Let us consider the following recently disclosed vulnerability CVE-2013-1189, which has a CVSS score of 5.7 (Medium)[1]. According to the NVD [2], this vulnerability allows remote attackers to cause a denial of service (routing-engine reload) via unspecified changes to IP address assignments. There is a mismatch between the given CVSS score of 5.7 and the high availability requirement for SME X, as an exploit of CVE-2013-1189 could mean a complete productivity stop and thus huge economic losses. Using contextual information, e.g., only adding the high availability requirement for SME X in order to ensure its business continuity, and using the CVSS *optional* environmental metrics, we get a modified CVSS score of 7.9 (High). Taking contextual information into consideration makes the score closer to reality for information security operations staff of SME X and guides them to trigger the appropriate response mechanisms for this category of vulnerabilities.

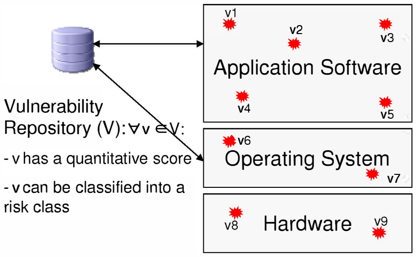


Fig. 1. System Model

From this simple scenario it becomes clear that the pure technical CVSS base score could be misleading and is insufficient w.r.t. the organization's requirements and the overall context. Furthermore, if we consider also the economic loss that could be caused by a potential degradation of availability, then this could eventually lead to classifying the severity of that vulnerability as *critical*. The scenario shows that for each vulnerability assessment, there is a need for integrating contextual information, especially the economic dimensions of a potential exploit. The next Section III introduces some basic concepts, upon which our approach is based.

III. Basic Concepts & Terminology

We present first our system model. Then, we briefly present the basics of CVSS, which constitutes the basis of our approach. Next an overview of our proposed approach is introduced.

A. System Model

We consider software systems that support the business processes inside public and/or private organizations. The organizations' assets need to be protected from attackers seeking to exploit existing software vulnerabilities. ISO 17799 defines a vulnerability as a "weakness of an asset or group of assets that can be exploited by one or more threats" [8].

Figure 1 shows the different elements related to our system model. We assume the availability of experts (in general CTOs and other C-level managers) who can estimate the values of different basic qualitative metrics of a vulnerability v such as the damage costs and Ex-post response costs, with high confidence. The considered vulnerabilities are weaknesses that can be exploited by attackers to compromise the target system. The next section provides a brief overview on CVSS, as it constitutes the starting point of our approach.

[1] http://web.nvd.nist.gov/view/vuln/detail?vulnId=CVE-2013-1189

B. Common Vulnerability Scoring System (CVSS)

CVSS [13] is a set of metrics used to quantitatively describe and compare different vulnerabilities regarding various attributes. The so called *base score* is the primary CVSS metric which is published by the vulnerability databases (e.g., NVD [2], OSVDB [3]) and is used to reflect the criticality of a vulnerability and how difficult it is to exploit. It consists of the subscores *impact* (Equation 2) and *exploitability* (Equation 3). Besides the base score, there are two *optional* CVSS metrics (i) the *temporal score* is a metric used to describe the current threat level, as it may change over time (e.g., depending on the availability of known exploits or remediation); and (ii) the *environmental score* captures the characteristics of a vulnerability that are associated with a specific IT environment; i.e., a vulnerability can be more or less severe for an environment, e.g., depending on the number of affected systems or the specific user security requirements. The base score is calculated using Equation 1 and utilizes a scale from 0-10, where 0-3.9 indicates *Low* severity vulnerability, 4.0-6.9 indicates *Medium* severity vulnerability, and 7.0-10.0 indicates *High* severity vulnerability.

$$\begin{aligned} &BaseScore = (0.6 \times Impact \\ &+0.4 \times Exploitability - 1.5) \times f(Impact) \\ &f(Impact) = \begin{cases} 0 & \text{if } Impact = 0 \\ 1.176 & \text{otherwise} \end{cases} \end{aligned} \tag{1}$$

$$\begin{aligned} &Impact = 10.41(1-(1-ConfImpact) \\ &\times(1-IntegImpact) \times (1-AvailImpact)) \end{aligned} \tag{2}$$

$$\begin{aligned} &Exploitability = 20 \times AccessComplexity \\ &\times Authentication \times AccessVector \end{aligned} \tag{3}$$

CVSS has been widely adopted by the information technology community. CVSS is mandated for use in evaluating the security of payment card systems worldwide [7]. The U.S. National Vulnerability Database [2] uses it for scoring vulnerabilities and mandates its use by products in the Security Content Automation Protocol (SCAP) validation program [6]. CVSS has also been adopted by various software vendors and service providers [5]. According to its authors, CVSS scores are intended to provide a relative comparison of vulnerability severity, not exact measurements [9]. Our proposed approach uses CVSS (sub)scores as input amongst others, and follows its intuitive and widely used scoring principles for eliciting the proposed economic-driven metrics.

C. Overview of Our Approach

Our proposed approach for an economic-driven, context-aware vulnerability assessment is composed of two building blocks (cf. Figure 2):

1) Economic-driven metrics for vulnerability assessment (Section IV).
2) Holistic vulnerability assessment integrating the technical, and economic-driven metrics (Section V).

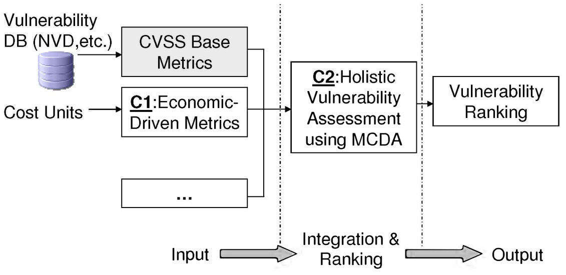

Vulnerability DB (NVD,etc.)
CVSS Base Metrics
Cost Units
C1:Economic-Driven Metrics
...
C2:Holistic Vulnerability Assessment using MCDA
Vulnerability Ranking
Input
Integration & Ranking
Output

Fig. 2. Overview of the Building Blocks of our Approach

In order to perform the holistic technical and economical vulnerability assessment, we have identified different criteria that need to be taken into account:

- CVSS base metrics: they include the following two subscores (i) *exploitability* metric, containing: access vector, access complexity, authentication, and (ii) *impact* metric, containing: confidentiality, integrity, and availability impacts.
- A set of proposed economic-driven metrics containing (i) vulnerability response costs, and (ii) vulnerability damage estimation. These metrics are introduced in Section IV.

As these criteria do not equally influence the vulnerability severity, they need to be weighted following a user-centric approach taking into account his/her security requirements regarding the expected confidentiality, integrity, and availability levels, and the environment specificities. Having a set of decision criteria with different weights, and a set of objects (vulnerabilities) to be ranked according to those criteria; this is a typical problem formalization that can be solved by adopting an MCDA approach to prioritize the vulnerabilities as detailed in Section V). The next Section IV describes our proposed economic-driven metrics.

IV. Economic-Driven Vulnerability Metrics

Analogous to the technical base metrics proposed by CVSS, the metrics that we introduce in this section are meant to capture the economic impact of vulnerability exploits on the business; i.e., economically quantifying their damage potential and occurring costs, as damage estimation is an essential part of risk analysis. We first introduce the requirements that our metrics should fulfil, before detailing the proposed metrics.

A. Requirements

Inspired by the CVSS base metrics [9], we have identified the following requirements for the economic-driven metrics to be used in our holistic vulnerability assessment process:

- **Rq.1:** Score diversity, i.e., avoiding that too many vulnerabilities have the same score; e.g., by making the metrics more granular.
- **Rq.2:** Scoring process should not be more complex than necessary.
- **Rq.3:** Scoring should be intuitive and consistent among different analysts.

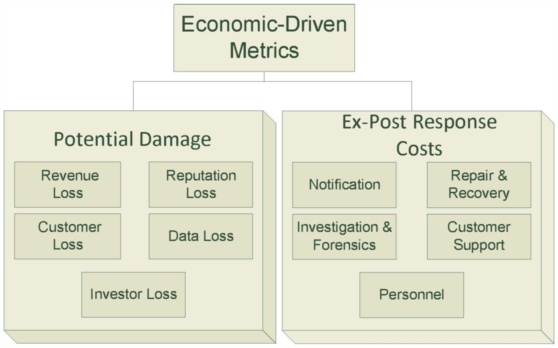


Fig. 3. Proposed Classification of the Economic-Driven Metrics for Vulnerability Assessment

B. Set of Identified Economic-Driven Metrics

In order to define economic-driven metrics reflecting the economic impact of vulnerability exploits, one needs to consider the potential costs a vulnerability exploit could cause. The basis for the definition of these metrics is the empirical work of Innerhofer et al. [23], where the authors define a set of 91 cost units based on an empirical study on publicly known security incidents. We find there different cost types listed, which are actually difficult to distinguish/differntiate, e.g., "Development and release of patch" and "Cost of developing patches". Since different organizations might have various nomenclatures for their own cost types, a unified cost classification helps avoid ambiguity and supports performing cost comparison and analysis within and across organizations. In order to be intuitive and easy to classify for most users, we define a small set of higher level main cost classes aggregating the cost types of [23]. Figure 3 depicts the result of the aggregation. We distinguish between two main cost classes (i) potential damage/losses, and (ii) Ex-post response costs, which could result from a vulnerability exploit. Each one with different subclasses. These cost classes are the result of surveying and scrutinizing the cost types cited in the state of the art literature [23], [24]. The lack of a common terminology and cost categorization regarding damage estimation in the SotA drives the need for a unified cost classification for security incidents. A unified cost classification provides the basis for a common terminology and ensures that there are accurate definitions of the used terms. Each one of the cost types determined in the state of the art (e.g., [23], [24]) can be categorized into one well-defined class/subclass introduced in Figure 3. The interested reader is referred to [25] for a detailed description of all cost units and the corresponding defined cost classes.

C. Metrics Valuation

In order to meet Rq.1-3 and to be inline with the CVSS scoring philosophy, which is widely used [5] and the de facto standard scoring method [2], we propose to use an intuitive scale of 4 possible values (*low, medium, high, critical*) to evaluate the different metrics. Furthermore, as monetized metrics have the advantage of (i) allowing easy numerical comparison between alternative scenarios within the same company, and (ii) are directly understandable by managers and executives with less technical affinity, we propose a mapping (cf. Table I) between our proposed qualitative scale (low, medium, high, critical) and a company-dependent monetized scale (analogous to CVSS scales). The rationale is that absolute monetary terms do not allow an objective comparison across companies of different sizes; e.g., a cost of $100K$ EUR might be *critical* for an SME, but of *low* effect for a large multinational company. Organizations could define their specific interval values c_x for the monetized mapping. For the calculation of our metrics, one needs also quantified factors to be mapped to the proposed scale (cf. Table I). This is important for the calculations performed in our integrative approach described in Section V. The quantitative scale thresholds are defined in such a way that, analogous to the CVSS thresholds, the scoring diversity is taken into consideration [9] and the intuitive and widely accepted CVSS scoring scheme is respected. E.g., for "Customer Support Costs" the user can choose a qualitative value (low, medium, high or critical), and according to the mapping depicted in Table I, a quantitative value to be utilized for the score calculations used will be assigned. We define here calculation formulas to support the monetized calculation of the damage potential metrics listed in Figure 3:

TABLE I
MAPPING PROPOSED SCALE - MONETIZED SCALE

Qualitative Scale	Monetized Scale (EUR)	Quantitative Scale
Low	$[0, C_{medium}[$	3.5
Medium	$[C_{medium}, C_{high}[$	6.1
High	$[C_{high}, C_{critical}[$	7.1
Critical	$[C_{critical}, \infty[$	10

1) *Potential Revenue Loss* (*PRevL*): Systems generating revenue for the business as defined in Section II constitute prototypical scenarios for our analysis in this section. Let c be the number of customers, and r the rate corresponding to the number of customers concluding a revenue generating transaction. The revenue might be lost due to two causes (i) loss due to service unavailability, and (ii) loss due to customer defection caused by high service response times. Let A be the service availability, where A=1 means availability. We assume, as in [30] that whenever the service comes back online after an outage, all affected customer transactions are lost, making no contribution to revenue. The lost revenue due to service unavailability, is thus $PRevL = c \times r \times (1 - A)$.

2) *Potential Reputation Loss* (*PRL*): is one of the most difficult concepts to build measurements around. A good estimator for reputation loss is the historical impact of vulnerability exploits and security incidents on stock price for companies that are listed on the stock exchange [26]. Let *ise* be the average impact of vulnerability exploits on stock prices, which is defined as follows: $ise = \frac{\sum_{t=0}^{n} p_t}{n} - p_{after}$, where p_t is the average stock price

at period t (before the incident), and p_{after} the stock price after the incident. It is worth noticing that if the value of $ise \leq 0$ then it is set to 0, as we do not consider exceptional cases, where there is an increase of the stock price after a security incident.

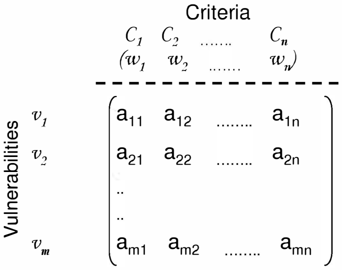
3) *Potential Customer Loss* (*PCL*): expresses the fraction of customers who are security sensitive, i.e., who will stop collaborating with the company if a vulnerability exploit occurs and is made public. $PCL = ssc \times acr_t$, where ssc is the estimated number of security sensitive customers, and acr_t is the average customer revenue per time period t.

4) *Potential Investor Loss* (*PIL*): reflects the number of investors who will stop investing in the company if a vulnerability exploit occurs and is made public by the media. $PIL = ssi \times ai_t$, where ssi is the estimated number of security sensitive investors, and ai_t is the average investment amount per investor in the last time period t.

5) *Potential Data Loss* (*PDL*): $PDL = avr \times nlr$, where avr is the Average Value per data Record, and nlr the Number of Lost data Records. To determine avr, company-internal historical accounting data concerning previous security incidents should be used. An alternative data source might be the publicly available estimations that are periodically published by research institutes, such as Ponemon [2], e.g., 2012 was the U.S. average $avg = 194$ USD, which can serve as orientation for estimating avr.

V. Holistic Vulnerability Prioritization Using MCDA

The basis for our holistic vulnerability assessment method are the technical CVSS base (sub)metrics and the proposed economic-driven metrics (cf. Section IV). As we have no conflicting criteria but only "cost" criteria, i.e., criteria for which the values need to be minimized in an ideal situation, this constitutes a typical problem that can be addressed by the well established prioritization methods called MCDA. They are concerned with the task of ranking a finite number of alternatives (vulnerabilities in our case), each of which is explicitly described in terms of different characteristics (also called decision criteria) which have to be taken into account simultaneously. The Multiplicative Analytic Hierarchy Process (MAHP) [10], [11] is one of the most widely used, and most accurate MCDA methodologies according to the analytical study performed by the authors of [12]. The MAHP is designed to handle the decision environments in which some subjective judgments are inherent in the decision making process. Moreover, it has the ability to handle input from multiple decision makers. It is the essence of MAHP that human judgments, and not just the underlying information, can be used in performing the evaluations. As user requirements w.r.t. weighting security goals mainly depend on human judgements, utilizing MAHP fulfils that need. MAHP converts these evaluations to numerical values that can be processed and compared. Numerical priorities (scores) are calculated for each of the vulnerabilities to be prioritized. These scores represent the vulnerabilities' relative priority, so they allow a straightforward comparison.



$$\begin{array}{c|cccc} & C_1 & C_2 & \ldots\ldots & C_n \\ & (w_1 & w_2 & \ldots\ldots & w_n) \\ \hline v_1 & a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ v_2 & a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ & .. & & & \\ & .. & & & \\ v_m & a_{m1} & a_{m2} & \ldots\ldots & a_{mn} \end{array}$$

Fig. 4. Multiple Criteria Matrix

Our multiple criteria vulnerability assessment problem can be formalized as follows: We have a number, say m, of vulnerabilities to be assessed and prioritized in terms of a number, say n, of decision criteria. The vulnerabilities are denoted as v_i (for $i = 1,2,3,...,m$) and the criteria as C_j (for $j = 1,2,3,...,n$). Each criterion is associated with a weight of importance, denoted as w_j (for $j = 1,2,3,...,n$). The higher the weight is, the more important the criterion is assumed to be. These weights are normalized so they add up to one: $\sum_{j=1}^{n} w_i = 1$. The above data are best summarized in a decision matrix as depicted in Figure 4. The corresponding MAHP formula used to calculate the quantitative score P_{v_i} of each v_i is given by Equation 4:

$$P_{v_i} = \prod_{j=1}^{n} (a_{ij})^{w_j} \tag{4}$$

VI. Evaluation

In this paper we validate our proposed techniques by investigating the validity of the hypotheses formulated below:

H1. *Considering solely the technical criteria of our approach (i.e., CVSS' base (sub)scores), MCDA results are consistent with a pure CVSS-based ranking of vulnerabilities.*

H2. *Vulnerability prioritization that considers economic-driven metrics might result in significant changes in the vulnerability severity classification.*

Furthermore, we show on the basis of an application scenario how our methodology can be applied to prioritize real-world vulnerabilities in an organization. We show also that a vulnerability assessment through our approach might lead to a severity class change (*low, medium, high*) compared to CVSS leading to more score diversity. This is shown by running simulations using synthetic data applied to the scenario, as it is in general widely accepted that it is hard to validate security quantification methods using real-world data [27].

A. Validation of MCDA Rankings Using CVSS Score Data

Regarding hypothesis H1, the first type of experiments is related to assessing the soundness of the proposed MCDA method w.r.t. the vulnerability ranking using all theoretically possible CVSS scoring data combinations. We consider the

[2] http://www.ponemon.org/

technical vulnerability metrics, which constitute the basis of the CVSS base score, and compare the ranking results obtained by applying our methodology with the results obtained using the CVSS scores. For the MAHP calculations, we have used weights for the considered criteria depicted in Table II, where $w_s = \frac{2}{3}$, i.e. the weight given to the overall CVSS score is $\frac{2}{3}$, and the remaining $\frac{1}{3}$ is weighted using the coefficients adopted by CVSS for the *impact* and *exploitability* subscores ($w_{is} = \frac{6}{10}$ and $w_{es} = \frac{4}{10}$ respectively).

TABLE II
WEIGHTS OF THE CONSIDERED CRITERIA

Criteria	CVSS Base Score	Impact Subscore	Exploitability Subscore
Weights	w_s	$(1-w_s) \times w_{is}$	$(1-w_s) \times w_{es}$

We run the experiments using CVSS scoring data. To this end, we generate theoretical scoring distributions for CVSS by considering all the possible sets of metric values and calculating the corresponding scores. First, we count the number of possible combinations of metric values (729 cases). However, vulnerabilities with all impact metrics set to *None* are not possible in practice because each vulnerability must have some impact, so we subtract those and have a final count of 702 possible scoring combinations. We then calculate the score for each combination. The possible resulting score values are 101 scores $s \ \varepsilon \ [0, 10]$. We then use this data set, let us call it D, as a basis for calculating MAHP results for these theoretical combinations. It is worth noticing that this experiment setup is also adopted by the authors of CVSS in [9]. The resulting MAHP matrix is depicted in Table III.

TABLE III
MAHP MATRIX FOR CVSS THEORETICAL SCORE DATA

Vulnerability (v_i)	CVSS Base Score (s_i)	Impact Subscore (is_i)	Exploitability Subscore (es_i)
Weights	w_s	$(1-w_s) \times w_{is}$	$(1-w_s) \times w_{es}$
v_1	s_1	is_1	es_1
v_2	s_2	is_2	es_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$
v_{702}	s_{702}	is_{702}	es_{702}

Figure 5 depicts the results of our experiment using all theoretically possible CVSS score data in ascending order. The obtained results show that using the same data set D, both ranking schemes deliver consistent scores. The Mean Squared Error (MSE) of MAHP w.r.t. CVSS scores quantifying the difference between the results of both schemes is $MSE = 0.032$.

B. Case Study

The objective of this case study is twofold (i) verify the validity of H.2, and (ii) show the potential cost savings w.r.t. the mitigation costs that could be achieved by applying our vulnerability prioritization approach. In this section we revisit the scenario setting defined in Section II. For the sake of simplicity, let us consider that SME X deploys the following simple software configuration: Microsoft Windows 7 64-bit Service Pack 1, MySQL Connector/ODBC 5.1.6, Oracle MySQL Connector J 5.1.15.0, Oracle MySQL Connector Net 6.4.4, Oracle MySQL Documents 5.5.17, Oracle MySQL Examples And Samples 5.5.15, Oracle MySQL Installer 1.0.18.0, MySQL Community Server 5.0.45, Oracle MySQL Workbench Ce 5.2.35, Oracle JDK 1.6.0 Update 25, Oracle JRE 1.6.0 Update 25, Microsoft Internet Explorer 8.0.6001, Adobe Acrobat Reader 9.5, Microsoft .NET Framework 2.0 Service Pack 2, Microsoft .NET Framework Version 3.0 Service Pack 2, Microsoft .NET Framework 3.51 Service Pack 1, Microsoft .NET Framework 4.0, Microsoft Windows Media Player 11.0.6000.6324.

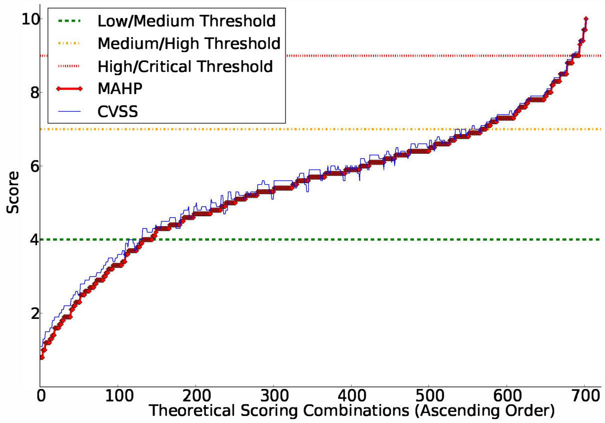


Fig. 5. Theoretical Scoring Combinations: MAHP vs. CVSS

We utilize the widely used vulnerability scanning tool Nexpose [29] to detect possible vulnerabilities existing in our system. Furthermore, we manually investigate the NVD [2] for up-to-date vulnerability reports affecting our real system configuration. The results of this exercise on a real system with the configuration described above show that there is a total of 304 vulnerabilities in our system. The calculations shown next are based on that total number of detected vulnerabilities.

1) Economic-Driven Metrics Data Basis: For the economic-driven metrics, we utilize as a proof of concept synthetic data estimations about SME X. This assumption is in line with the authors of [14] who acknowledge the execution of a simulation with artificial data. Let us assume that SME X adopts the threshold values for the mapping between our proposed vulnerability severity scale (low, medium, high, critical) and the company-dependent monetized scale that are depicted in Table IV.

For the vulnerability response mechanisms' costs, an efficient mechanism is the one resolving the vulnerability in a timely manner while generating the lowest total costs compared to alternative solutions [21]. According to the security policy of SME X, vulnerabilities that are classified as *critical* are addressed immediately by using quick response processes that might require the interruption of production

TABLE IV
MAPPING PROPOSED SCALE-MONETIZED SCALE, AND RESPONSE COSTS, USED FOR THE EVALUATION

Vulnerability Severity Class	Monetized Scale (EUR)	Average Response Costs (EUR)
Low	$[0, 10K[$	700
Medium	$[10K, 50K[$	1.5K
High	$[50K, 70K[$	3K
Critical	$[70K, \infty[$	7.5K

TABLE V
MAHP MATRIX WITH CVSS SCORE DATA, SECURITY REQUIREMENTS, AND ECONOMIC-DRIVEN METRICS

	CVSS Subscores			User Sec. Req.			Econ. Metr.	
v_i (w_j)	s_i $\frac{10}{15}$	is_i $\frac{3}{15}$	es_i $\frac{2}{15}$	$C.Rq.$ $\frac{2}{10}$	$I.Rq.$ $\frac{2}{10}$	$A.Rq.$ $\frac{6}{10}$	DP $\frac{6}{10}$	RC $\frac{4}{10}$
v_1	s_1	is_1	es_1	c_1	i_1	a_1	dp_1	rc_1
v_2	s_2	is_2	es_2	c_2	i_2	a_2	dp_2	rc_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
v_m	s_m	is_m	es_m	c_m	i_m	a_m	dp_m	rc_m

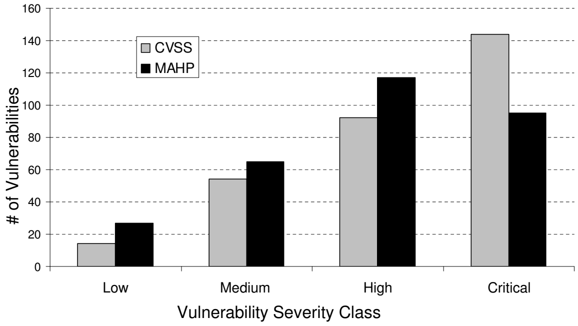


Fig. 6. Vulnerability Severity Classification: CVSS vs. MAHP

TABLE VI
POTENTIAL COST SAVINGS USING MAHP

	CVSS		MAHP	
Severity Class	# of Vuln.	Costs (EUR)	# of Vuln.	Costs (EUR)
Low	14	9,800.00	27	18,900.00
Medium	54	81,000.00	65	97,500.00
High	92	276,000.00	117	351,000.00
Critical	144	1,080,000.00	95	712,500.00
Total	304	1,446,800.00	304	1,179,900.00

processes. Lower vulnerability severity classes are addressed during scheduled maintenance time plans without interrupting production processes. Table IV shows synthetic accounting data for SME X consisting of the average costs associated with mitigating each class of vulnerability.

Based on the proposed mappings shown in Table I and after a careful examination and economic assessment of the potential costs for the detected 304 vulnerabilities based on our proposed economic-driven metrics (cf. Section IV), the MAHP matrix values are determined, then the MAHP scores are calculated applying Equation 4.

2) Vulnerability Prioritization: Table V depicts the MAHP matrix integrating the CVSS score data, SME X's security requirements (confidentiality *C.Rq.*, integrity *I.Rq.*, and availability *A.Rq.*), and our proposed economic-driven metrics (damage potential *DP*, and response costs *RC*), with their respective user defined weights (cf. Section V). Both *DP* and *RC* metrics are calculated as weighted average of their composing submetrics respectiveley. That is $DP = \sum_{i=1}^{5} PL_i \times w_i$, where PL_i represents the submetrics of Potential Damage and w_i the corresponding user defined weights. Similarly, $RC = \sum_{j=1}^{5} RC_j \times w_i$, where RC_j represents the submetrics of Response Costs and w_j the corresponding weights (cf. Section IV). The application of MAHP and the calculation of the score is similar to the CVSS approach. The entries of the MAHP matrix are normalized and weighted using user-defined weights depending on the organization where it is applied.

To verify hypothesis H2. we examine the vulnerability severity assessment results, we observe that there are significant changes in the severity classification of the detected 304 vulnerabilities. The most important changes are observed in the *critical* severity class. We observe that the majority of the scores (209 vulnerabilities 68.75%) have increased or decreased by a magnitude not leading to severity classification change (e.g., from *medium* to *high*, etc.). Nevertheless, a significant decrease of the number of critical vulnerabilities from 144 to 94 (34.02%) is notable. Furthermore, the number of low, *medium* and *high* severity vulnerabilities has also increased leading to a more score diversity, as stated by Rq.1 (cf. Section IV). Figure 6 depicts the changes in the vulnerabilities severity classification comparing the results obtained using our proposed approach with respect to the traditional CVSS scores. We investigate the economic repercussions of this classification change in the next section.

3) Cost Saving Potential - Discussion: To investigate the potential cost savings using our approach, we compare the total response costs of resolving all vulnerabilities in both cases, i.e., using the standard CVSS approach (case A) and using our economic-driven MAHP approach (case B). The data basis for performing this comparison is depicted in Table IV. The cost difference between the two cases constitutes the potential savings that can be achieved through utilizing our proposed economic-driven approach as depicted in Table VI.

The costs are calculated by multiplying the number of vulnerabilities with the corresponding average response costs from Table IV. The calculation results using synthetic data of SME X show that more costly response processes were triggered significantly less using our approach (case B) than in the standard CVSS-based approach (case A). Especially the number of *critical* response processes that might interrupt production processes was reduced by more than 34%. This reduction is relevant for security management, as it reduces the risk that staff members get accustomed to critical alarms and just turn down their severity level [21].

VII. Related Work

Several previous works support the argument that different organizations evaluate vulnerabilities differently, based on their specific contexts [15], [16], [19], [20]. The authors of [17] and [18] empirically showed that the impact of security vulnerability exploits varies with a company's context. Given the challenges w.r.t measuring the costs and severity of vulnerabilities in absolute terms, the usage of relative metrics is a practical alternative. The Common Vulnerability Scoring System (CVSS) [13] provides such relative metrics; nevertheless omitting context information. The authors of [21] propose a method to estimate the effects of adding context information on the quality of vulnerability prioritization. the proposed method enables practitioners to estimate the improvements of adding the missing context information in CVSS. Our approach contributes to the state of the art by proposing a methodology to integrate economic-driven metrics in the vulnerability assessment process, thus supporting decision makers in the process of prioritizing security investments to mitigate the discovered vulnerabilities and saving costs.

VIII. Conclusion

As the ever increasing cyber threats exploiting security vulnerabilities necessitate assessment and quantification, the current vulnerability severity quantification approaches do not provide comprehensive context aware assessment capabilities. Especially they omit the economic repercussions of vulnerability exploits, which could lead an organization to turmoil. Thus, our developed economic-driven vulnerability assessment methodology considers the potential economic damage and response costs an exploit could cause. Furthermore, it takes into account the user security requirements, which are dependent on the organization's context. Our proposed method especially fills the gap between the pure technical and the business views on security vulnerabilities. The main objective of the evaluation experiments was twofold (i) using all 702 theoretically possible CVSS scores our MCDA-based approach delivers results that are congruent with the CVSS scores while using the technical criteria only; and (ii) our MCDA-based approach could achieve potential cost savings by triggering the appropriate and necessary response mechanism for vulnerabilities with different severity classes. In that way we reduce the problem of having too many vulnerabilities with the highest possible CVSS score, thus reducing the risk of getting accustomed to alerts. As future work, empirical studies are needed to investigate the estimated cost savings using our approach and compare them with actually realized savings in firms to provide evidence supporting the use of our economic calculations. Furthermore, a sensitivity analysis of the relative effects of our new introduced metrics on scoring is needed.

Acknowledgments

Research supported in part by the EC FP7 project ABC4Trust (Grant Agreement no. 257782) and Loewe TUD CASED.

References

[1] Acquisti, A. et al.: Is there a cost to privacy breaches? An event study. In: Proc. of WS on Economics of Information Security (WEIS). 2006.

[2] National Vulnerability Database, http://nvd.nist.gov/, 2013.

[3] Open Source Vulnerability Data Base, http://www.osvdb.org/, 2013.

[4] Cheng, P. et al.: Aggregating CVSS base scores for semantics-rich network security metrics. In: Proc. of the IEEE Intl. Symposium on Reliable Distributed Systems (SRDS). 2012

[5] Forum of Incident Response and Security Teams: CVSS Adopters. http://www.first.org/cvss/eadopters.html. 2013.

[6] National Institute of Standards and Technology: Security Content Automation Protocol (SCAP). http://scap.nist.gov/. 2013.

[7] Payment Card Industry Security Standards Council. Payment Card Industry Data Security Standard: Technical and Operational Requirements for Approved Scanning Vendors, Version 1.1. https://pcisecuritystandards.org/tech/supporting_documents.htm. 2013

[8] ISO/IEC: Information technology - Code of practice for information security management. ISO/IEC 17799:2005. 2005.

[9] Scarfone, K. and Mell, P.: An analysis of CVSS version 2 vulnerability scoring. in 3rd International Symposium on Empirical Software Engineering and Measurement (ESEM). 2009.

[10] Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill. 1980.

[11] Saaty, T.L.: Fundamentals of decision making and priority theory with the analytic hierarchy process. RWS publications Pittsburgh. 1994.

[12] Triantaphyllou, E.; Baig, K: The Impact of Aggregating Benefit and Cost Criteria in Four MCDA Methods. In IEEE Transactions on Engineering Management. 2004.

[13] Mell, P. et al.: A Complete guide to the Common Vulnerability Scoring System. http://www.first.org/cvss/cvss-guide.pdf/. 2007.

[14] Hevner, A. et al.: Design science in information systems research. In: Management Information Systems Quarterly. Vol. 28. 2004.

[15] Chen, Y.: Stakeholder value driven threat modeling for off the shelf based systems. In Proc. of Intl. Conf. on Software Engineering. 2007.

[16] Eschelbeck, G.: The Laws of Vulnerabilities: Which security vulnerabilities really matter. Information Security Technical Report, vol. 10, 2005.

[17] Ishiguro, M. et al.: The Effect of information security incidents on corporate values in the japanese stock market. In Proc. of Intl. WS on Economics of Securing the Information Infrastructure (WESII). 2006.

[18] Telang, R. et al.: An empirical analysis of the impact of software vulnerability announcements on firm stock price. In Proc. of IEEE Transactions on Software Engineering, vol. 33. 2007.

[19] Lai, Y. et al.: Using the vulnerability information of computer systems to improve the network security. In Computer Communications, vol. 30. Jun. 2007.

[20] Rieke, R.: Modelling and analysing network security policies in a given vulnerability setting. In Critical Information Infrastructures Security. pp. 67-78. 2006.

[21] Fruehwirth, C. et al.: Improving CVSS-based vulnerability prioritization and response with context. In Proc. of Third International Symposium on Empirical Software Engineering and Measurement, 2009.

[22] Anderson, R.: Why information security is hard - an economic perspective. In Proc. of 17th Annual Computer Security Applications Conference (ACSAC). 2001.

[23] Innerhofer, F. et al.: An empirically derived loss taxonomy based on publicly known security incidents. In: Intl. Conf. on Availability, Reliability and Security (ARES), 2009.

[24] Van Eeten, M. et al.: Damages from internet security incidents. In: OPTA Research reports. http://www.opta.nl/nl/actueel/alle-publicaties/publicatie/?id=3083, 2009.

[25] Ghani, H. et al.: Threat modeling-based economic damage estimation of software security vulnerabilities. Technical Report TR-TUD-DEEDS-07-03-2013. 2013.

[26] Shostack, A. et al.: The New School of Information Security. Addison-Wesley Professional, 1st edition, ISBN-13: 978-0321502780. 2008.

[27] Verendel, V.: Quantified security is a weak hypothesis: A critical survey of results and assumptions. In proc. of new security paradigms workshop. 2009.

[28] Holm, H. et al.: Empirical analysis of system-level vulnerability metrics through actual attacks. In IEEE Transactions on dependable and secure computing, vol. 9, no. 6, 2012.

[29] Rapid7.: Nexpose. http://www.rapid7.com/products/nexpose/, 2013.

[30] Sauvé, J. et al.: Business-Driven Design of Infrastructures for IT Services. In Journal of Netw. Syst. Management, vol. 17, no. 4, 2009.