On the Feasibility of Utilizing Security Metrics in Software-Intensive Systems

Reijo Savola,

VTT Technical Research Centre of Finland, Oulu, Finland

Summary
Security measurement of software-intensive systems is an emerging field, rapidly gaining momentum. Well-designed security metrics offer credible and sufficient evidence of security level and performance for security decision-making. In this study, we introduce a novel security metrics feasibility validation approach, consisting of validation criteria and an associated validation process that takes into account the used measurement approaches and the use of security metrics. The approach is based on the identification of needs for and challenges in using security metrics, and the identification of good properties of security metrics from related work.

Key words:
Security metrics, security measurement, feasibility, security engineering, software engineering

1. Introduction

Lord Kelvin said "*If you cannot measure it, you can not improve it*". This fact also applies to information security issues. It is easier to make informed engineering and management decisions concerning security if sufficient and credible security evidence is available. Moreover, as resilience requirements for systems are constantly rising, the needs for adaptive security systems based on adequate security metrics and associated measured data for their operation are increasing.

The main contribution of this study is to introduce feasibility validation criteria and associated validation process for security metrics to be used in the Research and Development (R&D) of software-intensive systems and in the operation of security monitoring systems. The validation approach takes into account the measurement approaches and the final use of metrics.

In this study, the research is organized in four steps: identification of (*i*) needs for security metrics, (*ii*) challenges of them, (*iii*) pre-existing goodness criteria of them, and (*iv*) development of a feasibility validation approach. The remainder of this study is organized as follows: Sections 2 and 3 discuss the needs and challenges of security metrics, respectively; Section 4 presents goodness criteria of security metrics identified from related work; Section 5 proposes our feasibility validation approach; Section 6 discusses the findings of the study; and finally, Section 7 gives conclusions and discusses future work.

2. Needs for Security Metrics

In the following, we discuss the different needs of security metrics in R&D of software-intensive systems. Security metrics offer evidence of the security level and performance of the System under Investigation (SuI). They also support, systematize and organize proactive security engineering and assurance practices. According to Rosenblatt [1], the multifaceted aspects of security issues become clearer with security metrics. In addition to the application of metrics to security engineering, they can be used for online adaptive decision-making and as a means to communicate the state of security management in organization and to diagnose potential problems. In practice, there are no security measurement approaches that would allow the measurement of security as a universal property, due to the inherent complexity of that kind of task. However, security metrics can be developed based on explicit security requirements. Security metrics exist only to provide decision support, whether used online or offline. The information the metrics provide is only useful to the extent it serves that purpose [2].

2.1 Security Engineering and Assurance

In security engineering and software security assurance activities, e.g., testing, monitoring, analysis, the human audience of security metrics consists of most of the personnel associated with software security. The following roles have been identified as being important for software security [3]: security requirements developer, threat analyst, software architect, developer/programmer, tester, verifier, reviewer, auditor, application development manager, configuration manager and tool developer. In addition, project managers, product managers, R&D managers, Chief Information Officers (CIOs) and

Manuscript received January 5, 2010
Manuscript revised January 20, 2010

executive managers belong to the group of stakeholders of software security issues.
According to Landwehr [4], there are three basic ways of providing assurance in a product: (*i*) quality of the people involved in development, (*ii*) quality of the development processes employed, and (*iii*) direct assessment of the product through analysis and testing. Security metrics concentrate mainly on (*ii*) and (*iii*). Security awareness metrics, qualification certificates and reputation information can be used in (*i*).

2.2 Online Security Monitoring and Management

Security metrics can be used for automatic and adaptive online security management in software-intensive systems. For example, a distributed messaging middleware carrying out adaptive security management using different configurations of security-enforcing mechanisms, and especially for authentication, authorization and availability management, is described in [5] and [6]. Adaptive automatic security management approaches can be seen as a step towards security-metrics-based *self-healing software*. Surveys of approaches to adaptive application security and adaptive middleware solutions can be found in [7]. The functionality of Intrusion Prevention and Detection Systems (IPS/IDS) also relies on specific security metrics.

2.3 Information Security Management

In the Information Security Management (ISM) activities of organizations, security metrics can be used for the following purposes [8]: communication of security level and performance to the management of the organization, help in driving performance improvement, measurement of the effectiveness of Information Technology (IT) controls, help in diagnosing security problems, effective decision-making support, increasing accountability, support for resource allocation, demonstration of the state of compliance, and facilitation of benchmark comparisons. The main stakeholders of ISM are CIO, Executive Managers and eventually all employees in the organization. Note that the viewpoint of this study is on security engineering and assurance.

2.4 Security Measurement Objectives

In security engineering, security correctness, security effectiveness and security efficiency can be seen as the main fundamental measurement objectives [9]. They can be defined in the following way [9]:

- **Security correctness** denotes assurance that security-enforcing mechanisms have been correctly implemented in the SuI, and the system, its components, interfaces and the processed data meet the security requirements;
- **Security effectiveness** denotes assurance that the stated security requirements are met in the SuI, and the expectations for resiliency in the use environment are satisfied, while the SuI does not behave in any way other than what is intended; and
- **Security efficiency** denotes assurance that the adequate security quality has been achieved in the SuI, meeting the resource, time and cost constraints.

Security correctness can be seen as an objective for security quality and a *necessary but not sufficient* requirement for both "higher-level" measurement objectives – security effectiveness and security efficiency. If the system meets the technical security specification, we can say that it is correct.

Table 1: Example of Needs for Security Metrics during Design Phases

Target Audience	***Emphasis***	*Specific Needs*
Application development manager	Effectiveness, efficiency	Security and performance requirements
CIO	Effectiveness, efficiency	Security level and costs
Configuration manager	Effectiveness	Effect of configuration changes on security
Developer/ programmer	Correctness, efficiency	Secure coding
Executive manager	Effectiveness, efficiency	Security level, performance and costs
Project/ Product / R&D manager	Correctness, effectiveness, efficiency	Security level, performance and costs
Reviewer/ auditor	Correctness, effectiveness, efficiency	Requirements, associated standards
Security requirements developer	Correctness, effectiveness, efficiency	Prioritization, vulnerability information
Security tester	Correctness, effectiveness, efficiency	Threats, vulnerabilities, requirements
Software architect	Effectiveness, efficiency	Comparison of security-enforcing mechanisms
Tester in general	Efficiency	Performance requirements
Threat analyst	Effectiveness	Threat impacts, system exposure
Tool developer	Correctness, effectiveness, efficiency	Actual security levels, associated standards
Verifier	Correctness	Requirements

Correctness is often discussed together with effectiveness. In some cases, it might be difficult to differentiate them. However, evaluation of effectiveness gauges the strength of the security-enforcing mechanisms to withstand attacks in carrying out their function. Effectiveness requires ascertaining how well the security-enforcing components tie together and work synergistically [10]. Security efficiency dimension can be used to justify the expenditure on security work. Especially effectiveness and efficiency are widely recognized objectives in the security community.

Table 1 shows examples of which security measurement objectives dominate in the needs of different stakeholders in R&D of software-intensive systems. Security effectiveness is the main security measurement objective.

3. Challenges of Security Metrics

It is obvious that security metrics and the methods and tools to develop them are still immature. Therefore, in order to deduce feasibility criteria for security metrics, it is worthwhile to analyze the reasons behind the immaturity, as well as understand the criticism toward utilizing security metrics in general.

3.1 Why Security Metrics are Immature

There are many reasons for the fact that security metrics are still underdeveloped. We divide the reasons into three categories: (*i*) "security in side role syndrome", (*ii*) infancy of security research field, and (*iii*) lack of suitable data.

Security engineering has been carried out in isolation of other system focus areas [11]; consequently, security has been considered as "add-on" property [12]. Part of this problem comes from the history of security engineering: previously, it has focused on high-assurance systems [11], and often exclusively on a single aspect of security [13], especially on confidentiality [11]. Therefore, software-intensive system developers in general have not been involved [12]. Part of this "security in side role syndrome" is also that in some metrics approaches, security has been treated only as just another aspect of software quality [14], although security threats need active countermeasures and focus. In addition, software companies develop applications with only minimum attention paid to security before deploying them. The potential problems are often compensated with perimeter security solutions. Results from penetration testing, carried out on a strict time schedule, are often used as "metrics" [14], although it does not directly indicate the security level or performance, being highly dependent on test case libraries. Obviously, more *holistic and complete* collections of security metrics are needed. Note that in general, software measurement techniques are not yet mature either [12], complicating software security assurance activities as well.

The general infancy of the security research field is also a challenge to security metrics development [12]. There is a dearth of understanding and insight into the composition of security dimensions [12] and required mechanisms [10]. In particular, the reliance on subjective, human and qualitative results is a challenge [10]. Because of this, modeling of system security, not to mention measuring it, is extremely difficult [12]. There is a lack of common and unambiguous notation to describe security, its different components, relationships [12], and a lack of good estimators of system security [10]. Organizational and technical security metrics have not been integrated to provide a comprehensive view [13]. Technical security metrics are the least developed and the most ad hoc [13]. In practice, security is often seen mostly as a reactive field, predominantly focused on responding to incidents. Epstein [15] argues that many current software security metrics either (*i*) have only a distant relationship to vulnerabilities, (*ii*) are retrospective, or (*iii*) have a tendency towards false positives. Obviously, security metrics should show enough *time-dependency* and should be *credible* and *controllable.*

Furthermore, information security, like risk, is a difficult area to measure because there is a lack of suitable incidents – it is practically impossible to measure objectively what might have happened if we had not improved our security mechanisms [16]. Metrics should be able to *show progression*, should *not be overly biased*, and should be as *objective* as possible. Means to obtain measurements are often protracted or delusive [10]. Measurement approaches should be *scalable*, *portable*, *non-intrusive*, *cost-effective* and should allow for *reproducibility of measurements*. There is also a lack of data on variables that can be shown to influence security risk [17]. Chapin and Akridge [18] point out that it has been a tradition to utilize whatever security metrics are available in reporting, demanding a more systematic approach. Rathbun [8] states that a common mistake is to produce a metric that answers a question nobody is asking, or a metric that does not specifically meet the needs of an interested user. Security metrics should be *meaningful*, *contextually specific* and they should *represent* real system characteristics.

Hauser and Katz [19] ask if it is always possible to compute any metrics for a given system. An empirical metric of an operational system is obviously much harder to collect than an analytical one, because the operational system will not always be available for benchmarking or

the costs to conduct the measurement are much higher. However, good security metrics are also *attainable* and their target *measurable*. When developing security metrics, one should also *identify the missing metrics*, which are of substantial value to the users.

3.2 Skepticism towards Security Metrics

The feasibility of using security metrics has been criticized in some contributions. It should be pointed out that this is normal for methods and tools still in the conceptual phase [20]. The skeptics consider the current state of the art of security so low that any attempt to measure it would not be possible [20]. Fortunately, criticism can also be of use, helping in the identification of the areas requiring improvement. Therefore, in order to establish good validation criteria for security, we investigate the skeptic's opinions too.

When developing security metrics, one has to be conscious of the fact that they simplify a complex socio-technical system into models and further to numbers, percentages or partial orders. In other words, information is lost in the projection of the real world into the measurement results. McHugh [21] and McCallam [22] are skeptical of the side effects of such simplification and the lack of scientific results behind it. Today, practical considerations dominate security metrics research, and there has been a lack of scientific interest [23], explaining also the lack of proposal scientific frameworks in general. McHugh opposes the use of security metrics to reflect effectiveness of information security management. McCallam is skeptical about numerical values as a representation of security. In order to avoid these problems, security metrics should be *correct*, *controllable*, and should have *enough granularity* to incorporate the needed information.

Bellovin [24] remarks that developing security metrics is difficult, if not infeasible, because an attacker's effort is often linear, whereas exponential security work is needed. In [25] he stipulates for (*i*) mechanisms to make software less brittle, (*ii*) self-healing software, and (*iii*) science of composition that gives more than linear increase in security strength. He claims that until we have reached at least one of these requirements, we will not be able to establish useful security metrics. Utilizing security metrics in an adaptive security system, we actually make a step toward self-healing software.

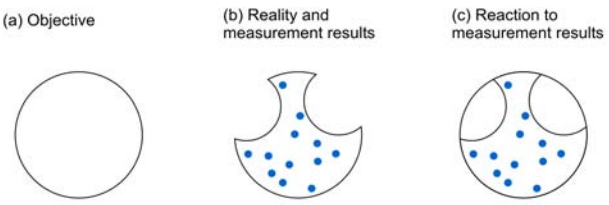

Fig. 1 A visualization of security measurements and a reaction to them.

There is no such thing as perfect security. Another source of security metrics challenges is that luck plays a major role [26], especially in the weakest links of information security solutions. This problem cannot be easily solved, but it can be mitigated by complete and prioritized security metrics collections.

3.3 Limits of Security Metrics

All metrics have limitations, not only security metrics. Metrics are being developed to be able to make justified statements about reality, which is not measurable in its entirety. Consequently, every metric is a simplification: it has a much lower information dimensionality than reality. However, there are means to improve the real-world solutions [27] based on the information offered by the measured data. The used measurement approach is very effective if the appropriate reaction is the result. Fig. 1 illustrates this. The illustration (a) represents the security objectives of the SuI, (b) the security measurement results from the real system, and (c) the desired reaction to the measurement results. Note that the actual security measurement results give only a rough estimate of the reality. However, with to a well-designed decision-making algorithm, the appropriate reaction is enforced, leading to meeting the objectives.

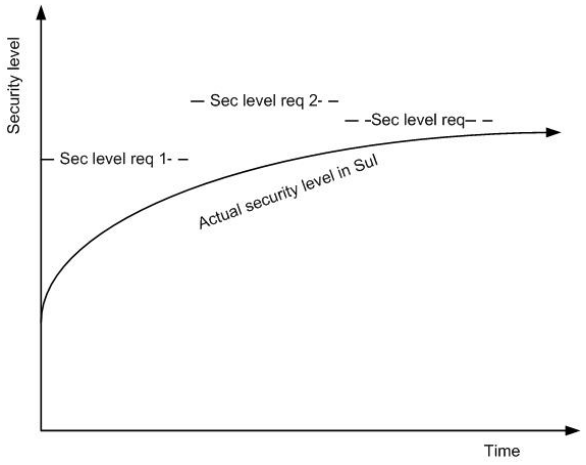

Fig. 2 Security metrics-driven decisions cause the actual security level to approach asymptotically the required security level.

Security metrics should be considered as security indicators, allowing for an increase of the security level, rather than as a means for absolute quantification of security correctness, effectiveness or efficiency. The security requirements that set the reference level to the metrics can change dynamically. See Fig. 2 for a visualization of reactions based on security metrics. The visualization captures the benefits of security measurement, yet showing the impossibility of absolute measurement. In practice, the security requirement level is not absolute, nor it is stable, and the actual security level is asymptotically approaching it, with the help of security metrics.

Table 2: Summary of Desired Characteristics for Security Metrics

FL	***Characteristic***	*Reference*
0	Correctness	Williams and Jelen [28]
0	Granularity	Böhme and Freiling [29]
0	Objectivity and unbiasedness	Schechter [17], Atzeni and Lioy [30]
0	Controllability	Williams and Jelen [28]
0	Time-dependability	Jelen [31], Kanoun *et al.* [32], Henning *et al*. [33]
0	Comparability	*Not found elsewhere*
1	Measurability	Rathbun [8], Williams and Jelen [28], Jelen [31], Jaquith [34]
1	Attainability, availability, easiness	Böhme and Freiling [29], Atzeni and Lioy [30], Jelen [31]
1	Reproducibility, repeatability, scale reliability	Schechter [17], Williams and Jelen [28], Böhme and Freiling [29],Atzeni and Lioy [30], Jelen [31], Henning *et al*.[33]
1	Cost effectiveness	Rathbun [8], Böhme and Freiling [29], Kanoun *et al.* [32], Henning *et al*. [33], Jaquith [34]
1	Scalability, portability	Williams and Jelen [28], Kanoun *et al.* [32]
1	Non-intrusiveness	Kanoun *et al.* [32]
2	Meaningfulness	Rathbun [8], Schechter [17], Henning *et al*.[33]
2	Effectiveness	Williams and Jelen [28]
2	Efficiency	Williams and Jelen [28]
2	Representativeness and contextual specificity	Rathbun [8], Jelen [31], Kanoun *et al.* [32], Henning *et al*. [33], Jaquith [34]
2	Clarity and succinctness	Atzeni and Lioy [30]
2	Ability to show progression	Schechter [17]
	Completeness	Williams and Jelen [28]

Those pursuing the development of security metrics should think of themselves as pioneers and be prepared to adjust strategies as experience dictates [35]. Sanders states that the challenge for security metrics is to "create a scientific foundation, methods and tools for quantitative assessment of security metrics that can be applied to large-scale information-community technology systems throughout their lifecycle" [13]. According to Jansen [10], quick resolution in security metrics is not expected and the likelihood is that not all aspects of the problem are resolvable. As a conclusion, yet having some limitations, security metrics can be used to increase the actual security level and performance in practical systems.

4. Goodness Criteria for Security Metrics from Related Work

Here we discuss the results of a literature survey aiming to identify the most important feasibility properties of security metrics. There are several contributions that propose properties for "good" security metrics; see a summary of them in Table 2. The summary is presented according to the Feasibility Level (FL) classification introduced in Section 5. According to Jelen [31], metrics should be "SMART", i.e. Specific, Measurable, Attainable, Repeatable and Time-dependent. Payne [35] states that truly useful metrics indicate the degree to which security goals are being met, calling out for the feasibility validation assessment results.

According to our knowledge, there are no pre-existing security metrics feasibility validation approaches available. Böhme and Reussner discuss validation of analytical metrics in general in [27]. They define three validation levels for prediction models. This level approach has inspired our three-level feasibility validation approach for security metrics. The Dependability Benchmarking (DBench) project introduces a collection of dependability benchmark validation criteria in [32]. Their criteria focus on the measurement approaches rather than metrics themselves.

5. Proposed Security Metrics Feasibility Validation Approach

For the purposes of this study, we define feasibility of security metrics to mean the technical, economic, and operational credibility, applicability and sufficiency of security metrics, when utilized by the practical measurement approach(es) for the purposes of their final use. We propose 18 generic feasibility validation criteria and an associated feasibility validation process for security metrics. The criteria are classified into three different feasibility levels, each of them incorporating 6 feasibility criteria. This 3×6 matrix of criteria is based on the

security metrics needs, challenges and good properties analysis previously presented. Evaluation of the metrics and the SuI by themselves does not give a sufficient picture of the feasibility of them in different use cases. Therefore, we consider feasibility of security metrics also from the point of view of the measurement approach and the use of them.

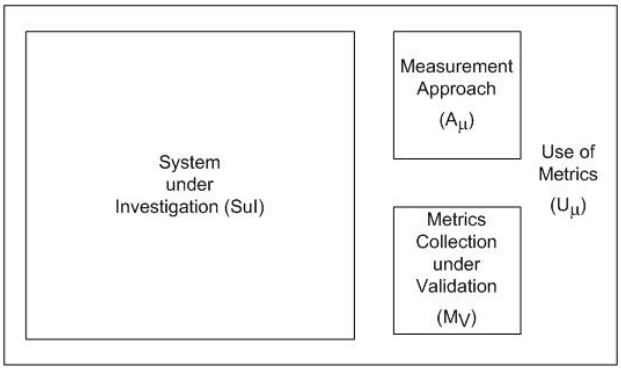

Fig. 3 Entities that are part of the feasibility validation approach.

Fig. 3 represents the main entities needed for feasibility validation: SuI, measurement approach, use of metrics and the actual metrics collection under validation. SuI is the target of the measurement, a system in its use environment. Measurement approach includes the means for gathering and managing measurement data (e.g. monitoring tool), as required by the security metrics. By "use of metrics" we denote all associated decision-making activities carried out using the measurement approaches to obtain security evidence from the system, with the help of the metrics. In the following, we denote security metrics under validation by $\mu_V \in M_V$, the utilized measurement approach by $\alpha_\mu \in A_\mu$, and the use of the metrics and measurement approach by $u_\mu \in U_\mu$, where M_V is the total collection of security metrics under validation, A_μ is the collection of measurement approaches utilizing M_V, and U_μ is the associated collection of all use scenarios specified for the SuI.

Böhme and Reussner distinguish three levels of validation for prediction models in [27]: metric, applicability and benefit validation of security metrics. We apply their approach to the feasibility validation, identifying the following Feasibility Levels (FL), where FL 0 represents the basic level and FL 2 the highest level of feasibility:

- FL 0: **Credibility** of μ_V, used in all associated $\alpha_\mu \in A_\mu$, for the purposes of all $u_\mu \in U_\mu$;
- FL 1: **Applicability** of μ_V to all associated $\alpha_\mu \in A_\mu$; and
- FL 2: **Sufficiency** of μ_V, used in all associated $\alpha_\mu \in A_\mu$, for the purposes of all $u_\mu \in U_\mu$.

FL 0 is a necessary but not sufficient requirement for FL 1, and FL 1 is a necessary but not sufficient requirement for FL 2. The feasibility validation process addresses FL 0 – 2 for each metric, and **completeness** of the security metrics collection.

5.1 Feasibility Level 0 – Credibility

Credibility of security metrics means that they meet the basic validity requirements of metrics in their use environment. The following properties constitute the credibility of security metrics:

1. **Correctness**. The metrics should be correctly implemented and error-free w.r.t. to their *specification*;
2. **Granularity** means that the metrics allow to distinguish at an adequate level the measurement results that differ from each other;
3. **Objectivity and unbiasedness**. Objectivity means that the measurement results should not be influenced by the measurer's will, beliefs or actual feeling [30]. Moreover, unbiasedness means that the results are not influenced by any bias. Absolute objectivity and unbiasedness are impossible to achieve. However, they should be maximized when designing metrics;
4. **Controllability**. The metrics should be controllable, i.e., they should be kept within the defined limits, or the measurement window;
5. **Time-dependability**. The time-dependent behavior of security metrics can be leading, coincident or lagging [10]. Time-dependability of the metrics should be part of their specification. Different timing categories should not be mixed without proper prediction models or heuristics [9]; and
6. **Comparability.** Metrics should support comparison of the targets that they represent. Comparison is often needed in making improvement decisions.

5.2 Feasibility Level 1 – Applicability

Applicability of security metrics and the measurement approaches utilizing them to the final use environment is crucial. Metrics should be designed "hand-in-hand" with the measurement approach: metrics cannot be utilized without measurements, and measurements are useless without interpretation of them using metrics. The following properties of measurement approaches are important for their applicability to the use. The security metrics should support the properties:

1. **Measurability** means that the metrics should be capable of having dimensions, quantity or capacity ascertained [28] in their measurement approaches. Consistent measurability of the parameters included in

the security metrics is a core criteria for the applicability;

2. **Attainability, availability and easiness** of the measurement parameters of the metrics. Attainability means that measurement results can be achieved from the SuI, and availability implies here that they are in general available. Furthermore, easiness here refers to the easiness of achieving the measurement results. At times, the data readily available from the measurement approach can greatly affect the actual security metrics collection. Rathbun [8] reminds that producing an inventory of available data sources can be used to dictate which metrics are produced;
3. **Reproducibility, repeatability and scale reliability**. Reproducibility and repeatability designate that if a measurement is repeated in the same context, with exactly the same conditions, the same measurement result is returned [30]. Scale reliability refers to the reproducibility of the measurement by different persons. Reproducibility, repeatability and scale reliability enable statistical investigation;
4. **Cost effectiveness**. Measurements should be cheap to gather and the measurement approaches utilizing the metrics should be cost effective;
5. **Scalability and portability** of the measurement approach. Scalability means that the measurement approaches and the associated metrics should be applicable to SuIs of different sizes. Portability refers to the applicability of the measurement approaches and the associated metrics to various target systems [32]; and
6. **Non-intrusiveness** of the measurement approach. The measurement approach should not cause harm to the normal operation of the measurement target system, should require only minimum changes to the target system, and should not affect the measurement results.

Note that at the FL 1, the criteria concern mainly the properties of the measurement approaches, and the role of security metrics here is that it should support the criteria.

5.3 Feasibility Level 2 – Sufficiency

Sufficiency of security metrics means that they are able to represent the real-world security issues they are expected to, at an adequate level, as required by the measurement use scenarios. In addition, sufficiency criteria concern the benefits of the metrics to the users. The following properties are important for the sufficiency of the security metrics:

1. **Meaningfulness** denotes that the metrics should be relevant and should respond to the needs. They should address one of the main security measurement objective dimensions: security correctness, effectiveness and efficiency [9]. It is easy to define metrics, but much harder to find meaningful ones [29];
2. **Effectiveness** of the metrics indicates that the expectations for the metrics' sufficiency in the final use environment are satisfied while they do not behave in any way other than what is intended; and
3. **Efficiency** of metrics signals that the adequate requirements of the metrics are achieved with consuming only minimal undesired effort and time for metrics and measurements;
4. **Representativeness and contextual specificity** of the metrics mean how well the metrics correspond to the real system characteristics in SuI in a contextually focused way. Although the contextual specificity and representativeness of security metrics increase clarity, note that according to Epstein [15], it can be better to gather as many security metrics as you can and then decide which of them make sense, rather than argue about which metrics make sense to collect;
5. **Clarity and succinctness** of the security metrics are especially important if they are used for human audience. Succinctness refers to that only important parameters are considered [30]. Clear formulization of metrics to be used in automatic computations is also important, reducing the possibility of errors;
6. **Ability to show progression** is important when using security metrics to demonstrate the security level and performance of the SuI. Illustrating the progression by means effect of corrective security actions is one *raison d'être* of security metrics.

In order to investigate the needs for security metrics for human audience, the users and other stakeholders can be interviewed, learning what evidence is important to them, and then devise a way to produce that information [8].

The feasibility criteria presented above are generic. In addition, there can be security metrics properties that depend heavily on the actual use of the metrics. If *use-dependent* properties are seen important, they should be part of the feasibility validation process. Examples of use-dependent requirements for security metrics are:

- Comparison capability against a specific standard,
- Automation of use, and
- Widely used standardized approach.

Note that standards often direct one toward producing more subjective measurements; not for assessing security effectiveness [8].

5.4 Feasibility Validation Process

We propose the following feasibility validation process for security metrics:

1. **Prioritize every criterion** in FL 0, FL 1 and FL 2 based on security requirements of the SuI. If there are important specific *use-dependent* criteria, make changes to the criteria accordingly. Assign weight values w_{xy}, $0 \le w_{xy} \le 1$, to the criteria according to the prioritization. To obtain the weight values, multiply "raw" weight values by confidence estimates:

$$w_{xy} = \theta_{xy} \cdot w_{xy}', \tag{1}$$

where w_{xy} is the final weight, θ_{xy}, , $0 \le \theta_{xy} \le 1$ the confidence on that weight, and w_{xy}' , $0 \le w_{xy}' \le 1$ the raw weight prior to confidence multiplication of the yth criteria in FL x. In the following matrix W_μ, the weight of the yth criteria of FL x is represented by w_{xy}:

$$W_\mu = \begin{bmatrix} w_{01} & w_{02} & w_{03} & w_{04} & w_{05} & w_{06} \\ w_{11} & w_{12} & w_{13} & w_{14} & w_{15} & w_{16} \\ w_{21} & w_{22} & w_{23} & w_{24} & w_{25} & w_{26} \end{bmatrix}; \tag{2}$$

2. For every $\mu_V \in M_V$ utilized by all $\alpha_\mu \in A_\mu$, for the purposes of all $u_\mu \in U_\mu$ carry out the following:
   a. **Assess the credibility** of μ_V. Aggregate the assessment results using a weighted sum based on the weight assignment of Step 1. Credibility C_μ ($0 \le C_\mu \le 10$) of μ_V is:

$$C_\mu = \frac{1}{6}\sum_{y=1}^{6} w_{0y} \cdot c_{0y}, \tag{3}$$

   where c_{0y} ($0 \le c_{0y} \le 10$) represents the credibility according to yth criteria of FL 0 and w_{0y} its weight. If the aggregated credibility and all its subcomponents are at an adequate level, requirements for FL 0 are met;
   b. **Assess the applicability** of μ_V. Aggregate the assessment results using a weighted sum. Applicability APP_μ ($0 \le APP_\mu \le 10$) of μ_V is the following:

$$APP_\mu = \frac{1}{6}\sum_{y=1}^{6} w_{1y} \cdot c_{1y}, \tag{4}$$

   where c_{1y} ($0 \le c_{1y} \le 10$) represents the applicability according to yth criteria of FL 1 and w_{1y} its weight. If the applicability and all its subcomponents are at an adequate level, requirements for FL 1 are met; and
   c. **Assess the sufficiency** of μ_V. Aggregate the assessment results using a weighted sum based on the weight assignment of Step 1. Sufficiency S_μ ($0 \le S_\mu \le 10$) of μ_V is the following:

$$S_\mu = \frac{1}{6}\sum_{y=1}^{6} w_{2y} \cdot c_{2y}, \tag{5}$$

   where c_{2y} ($0 \le c_{2y} \le 10$) represents the sufficiency according to yth criteria of FL 2 and w_{2y} its weight. If the sufficiency and all its subcomponents are at an adequate level, requirements for FL 2 are met;
3. **Assess the completeness** CO ($0 \le CO \le 10$) of the security metrics collection M_V from the point of view of security requirements. It is desirable to develop a security metrics framework in a way that it is as complete as possible [12], within the time and resource constraints. Identify gaps, i.e. lack of metrics, which are of substantial value to the metrics use; and
4. **Gather and interpret results** from the feasibility validation assessment. The outputs of this assessment are the C_μ, APP_μ and S_μ assessment results of each $\mu_V \in M_V$, the CO results and identification of the missing metrics which are of substantial value to the metrics use.

A feasibility validation plan supports the success of the process. If there are different possibilities to obtain the same security evidence, the most feasible metrics to offer the evidence should be used. If the feasibility of the metrics is the same, the ones most easily within reach should be selected, as suggested in [8].

Feasibility validation of security metric(s) is typically carried out as a stage in the used security metrics development process. In our earlier work [6][9], we have proposed the following process:

1. Carry out a threat and vulnerability analysis,
2. If applicable, utilize the available taxonomical or ontological information,
3. Define the security requirements and carry out modeling (if applicable),
4. Decompose the requirements and/or design,
5. Develop the measurement architecture, the mechanisms to gather the required measurement data from the SuI,
6. Carry out a feasibility analysis, and
7. Develop a balanced and detailed collection of security metrics.

Security metrics feasibility validation (Stage 6) offers input to Stage 7 of the security metrics development process.

5.5 What is Adequate Level?

In the above security metrics feasibility validation process, it is required that the credibility, applicability and sufficiency of them are at an *adequate level.* This term allows a lot of freedom, which in turn, is needed in such a challenging task as security metrics development. As security metrics exist only to provide decision support, the most important consideration in the assessment of "adequate level" is that the correct decisions can be made based on the data that the metrics produce. Obviously, this approach requires feedback from the decision-making, evaluation of right and wrong decisions w.r.t. the measured data. The feedback is often difficult to be arranged in practical systems. Fortunately, decision-making can be simulated before taking the metrics into operational use. First and foremost, we should be able to avoid *unintended consequences* that poorly designed security metrics use can cause. Especially, organizational security metrics can affect decisions even if those decisions inadvertently sacrifice long-term benefits [19]. At worst security metrics can give a false impression of security that leads to inefficient or unsafe implementation of security measures [15].

We would like to especially emphasize the following. Even though the criteria of our approach are all important, they can yet be all very difficult to achieve for such a complex property as security. Therefore, one should take into account these difficulties when defining the adequate levels of different metrics criteria and requirements.

Note that the previous usage of metrics affects what will be required from them in the future. Since rational decision-makers are very inventive to achieve a favorable measurement, metrics can start to lose their role, they can "wear out" [27]. In this case, needs for the "adequate level" of credibility, applicability and sufficiency decrease. At least for a short time, correct decisions can be made even if the metrics are not at an "adequate level". However, in the long run, the "adequate level" should be analyzed and updated as a requirement. "Worn-out" metrics should be replaced by those with more benefits.

6. Discussion

In this study, we have investigated what is the recipe for "good" security metrics. An experimental analysis on this subject is not possible since the state of the art in developing security metrics is still in its early stages, and very little experience in utilizing them in practical systems is available. In other words, the introduced feasibility validation approach cannot be justified experimentally due to the lack of data. Widely-accepted approaches for measuring security are needed, but do not yet exist. Security metrics and measurements should make a move from ad hoc practices to a more systematic process that is capable of responding to constantly changing threats and business demands [9].

The proposed collection of security metrics feasibility criteria relies on the "adequate level" of requirements. Obviously, to better define what is "adequate", more experience is needed from utilizing the metrics in practice. The "adequate level" should be dependent on the ability of the decision-making to carry out the right decisions based on the security metrics.

We have highlighted some challenges of security metrics and investigated criticism towards them. This has offered valuable feasibility improvement information for the development of security metrics. According to our analysis, security metrics are useful, especially if they are seen as a means to offer appropriate evidence for decision-making. It is important to understand that security metrics cannot be used to measure security as a whole, i.e as a universal property. Security solutions are developed based on security requirements that set the desired reference level for security metrics as well. Even though the introduced approach is not complete, lacking thorough experimental validation results, we think that this study is an important step toward developing feasible security metrics, introducing the initial feasibility validation criteria and associated process for security metrics. The basis of the selected criteria is justified by the properties suggested in the literature, the challenge and the criticism analysis.

7. Conclusions and Future Work

We have introduced a novel security metrics feasibility validation approach consisting of validation criteria and an associated validation process. The feasibility criteria are classified into three feasibility levels, which incorporate six criteria each. The feasibility levels emphasize the credibility of security metrics, their applicability to be utilized together with the measurement approach, and the sufficiency of them for their intended use. The used measurement approaches and the final use of the security metrics under validation are taken into account in the criteria. The introduced approach is justified by the "good" security metrics properties suggested in the related work and need, challenge and criticism analysis. The approach is suitable for all kinds of security metrics development, but is probably most valuable to be used in connection with R&D and operational security metrics development for software-intensive systems.

In our future work, we intend to use the introduced feasibility criteria to validate the use of security metrics as a part of an adaptive security monitoring system in practical use cases.

References

[1] J. Rosenblatt, "Security metrics: a solution in search of a problem," EDUCAUSE Quarterly, Vol. 31, No. 3, Jul./Sep. 2008.

[2] W. K. Brotby, "Information security management metrics: a definitive guide to effective security monitoring and measurement," Auerbach Publications, 2009, 200 p.

[3] Practical Software & Systems Measurement Safety and Security Technical Working Group, "Security measurement – white paper," Version 3.0, Jan. 2007, 67 p.

[4] C. E. Landwehr, "Computer security," International Journal on Information Security, 1(1): 3-13, Aug. 2001.

[5] H. Abie, R. Savola and I. Dattani, "Robust, secure, self-adaptive and resilient messaging middleware for business critical systems," ADAPTIVE '09, Athens/Glyfada, Greece, Nov. 15-20, 2009, pp. 153-160.

[6] R. Savola and H. Abie, "Identification of basic measurable security components for a distributed messaging system," SECURWARE '09, Athens/Glyfada, Greece, Jun. 18-23, 2009, pp. 121-128.

[7] A. Elkhodary and J. Whittle, "A survey of approaches to adaptive application security," SEAMS '07, May 20-26, 2007.

[8] D. Rathbun, "Gathering security metrics and reaping the rewards," SANS Institute Information Security Reading Room, Oct. 7, 2009, 21 p.

[9] R. Savola, "A security metrics taxonomization model for software-intensive systems," Journal of Information Processing Systems, Vol. 5, No. 4, 2009, 10 p.

[10] W. Jansen, "Directions in security metrics research," U.S. National Institute of Standards and Technology, NISTIR 7564, April 2009, 21 p.

[11] C. Meadows, "The feasibility of quantitative assessment of security," Conf. on Distributed Computing for Critical Applications (DCCA-4), San Diego, Jan. 1994.

[12] R. Savola, H. Abie, "On-line and off-line security measurement framework for mobile ad hoc networks," Journal of Networks, Vol. 4, No. 7, Sep. 2009, Academy Publisher, pp. 565-679.

[13] W. Sanders, "Security metrics: state of the art and challenges," Slides in Workshop on Evaluation of Dependability and Resiliency, IFIP WG 10.4, Cortina d'Ampezzo, Italy, Jan. 27-31, 2009.

[14] B. Chess, "Metrics that matter: quantifying software security risk," Workshop on Software Security Assurance Tools, Techniques and Metrics, Long Beach, California, Nov. 7-8, 2005.

[15] J. Epstein, "Good enough metrics," Slides in Metricon 1.0, Securitymetrics.org.

[16] G. Hinson, "Seven myths about information security metrics," ISSA Journal, Jul. 2006.

[17] S. E. Schechter, "Computer security strength & risk: a quantitative approach," PhD Thesis, Harvard University, Cambridge, Massachusetts, May 2004.

[18] D. A. Chapin and S. Akridge, "How can security be measured?" Information Systems Control Journal, Vol. 2, 2005.

[19] J. R. Hauser and G. M. Katz, "Metrics: you are what you measure!" European Management Journal, Vol. 16, No. 5, Oct. 1998, pp. 517-528.

[20] N. Bartol, B. Bates, K. M. Goertzel and T. Winograd, "Measuring cyber security and information assurance: a state-of-the-art report," Information Assurance Technology Analysis Center (IATAC), May, 2009.

[21] J. McHugh, "Quantitative measures of assurance: prophecy, process or pipedream?" Workshop on Information Security System Scoring and Ranking (WISSSR), ACSA and MITRE, Williamsburg, Virginia, May, 2001 (2002).

[22] D. McCallam, "The case against numerical measures of information assurance," Workshop on Information Security System Scoring and Ranking (WISSSR), ACSA and MITRE, Williamsburg, Virginia, May, 2001 (2002).

[23] F. C. Freiling, "Introduction to security metrics," in I. Eusgeld, F. C. Freiling and R. Reussner (Eds.): Dependability Metrics, LNCS 4909, Springer-Verlag, 2008, pp. 129-132.

[24] S. M. Bellovin, "On the brittleness of software and the infeasibility of security metrics," IEEE Security & Privacy, Jul./Aug., 2006, p. 96.

[25] S. M. Bellovin, "On the brittleness of software and the infeasibility of security metrics," Slides in Metricon 1.0, Securitymetrics.org.

[26] P. Burris and C. King, "A few good security metrics," METAGroup, 2000.

[27] R. Böhme and R. Reussner, "Validation of predictions with measurements," in I. Eusgeld, F. C. Freiling and R. Reussner (Eds.): Dependability Metrics, LNCS 4909, Springer-Verlag, 2008, pp. 14-18.

[28] J. R. Williams and G. F. Jelen, "A framework for reasoning about assurance," Arca Systems, Inc., 1998, 43 p.

[29] R. Böhme and F. C. Freiling, "On metrics and measurements," in I. Eusgeld, F. C. Freiling and R. Reussner (Eds.): Dependability Metrics, LNCS 4909, Springer-Verlag, 2008, pp. 7-13.

[30] A. Atzeni and A. Lioy, "Why to adopt a security metric? A little survey," 1st Workshop on Quality of Protection (QoP 2005), Milan, Italy, Sep. 2005, pp. 1-12.

[31] G. Jelen, "SSE-CMM security metrics," NIST and CSSPAB Workshop, Washington, D.C., Jun. 13-14, 2000.

[32] K. Kanoun, H. Madeira, Y. Crouzet, M. dal Cin, F. Moreira, J.-C. Ruiz García (Eds.), "DBench dependability benchmarks," Dependability Benchmarking (DBench) Project (EU IST-2000-25425), 2004, 235 p.

[33] R. Henning *et al*., "Proceedings of workshop on information security system, scoring and ranking – information system security attribute quantification or ordering," ACSA and MITRE, Williamsburg, Virginia, May 2001 (2002).

[34] A. Jaquith, "Security metrics – replacing fear, uncertainty and doubt," Addison-Wesley, 2007, 306 p.

[35] S. C. Payne, "A guide to security metrics," SANS Security Essentials GSEC Practical Assignment, Version 1.2e, Jun. 19, 2006.

Mr. Reijo Savola received the M.Sc. degree in Electrical Engineering (with honors) from the University of Oulu, Finland, 1992, and the degree of Licentiate of Technology in Computer Science from the Tampere University of Technology, 1995. He is currently working as a Senior Research Scientist in VTT Technical Research Centre of Finland. He has experience in information and network security, software engineering, telecommunications, multi-technology engineering research topics and in digital signal processing. He has also worked as a digital signal processing consultant for Elektrobit Group Plc. in Oulu, Finland and in Redmond, WA, United States.