The TSAR procedure rev. 1

Test of Security Assessment Relevance

JOHAN BENGTSSON, KRISTOFFER LUNDHOLM,
JONAS HALLBERG, AMUND HUNSTAD, JACOB LÖFVENBERG

FOI, Swedish Defence Research Agency, is a mainly assignment-funded agency under the Ministry of Defence. The core activities are research, method and technology development, as well as studies conducted in the interests of Swedish defence and the safety and security of society. The organisation employs approximately 1000 personnel of whom about 800 are scientists. This makes FOI Sweden's largest research institute. FOI gives its customers access to leading-edge expertise in a large number of fields such as security policy studies, defence and security related analyses, the assessment of various types of threat, systems for control and management of crises, protection against and management of hazardous substances, IT security and the potential offered by new sensors.

FOI
Swedish Defence Research Agency
Information Systems
Box 1165
SE-581 11 Linköping

Phone: +46 13 37 80 00
Fax: +46 13 37 81 00

www.foi.se

FOI-R-3061
ISSN 1650-1942

Methodology report
December 2010

Information Systems

Johan Bengtsson, Kristoffer Lundholm, Jonas Hallberg, Amund Hunstad, Jacob Löfvenberg

The TSAR procedure rev. 1

Test of Security Assessment Relevance

Titel	TSAR-proceduren revision 1 – Test av relevans avseende säkerhetsvärdering
Title	The TSAR procedure rev. 1 – Test of Security Assessment Relevance
Rapportnr/Report no	FOI-R-3061
Rapporttyp Report Type	Metodrapport / Methodology report
Månad/Month	December
Utgivningsår/Year	2010
Antal sidor/Pages	29 p
ISSN	ISSN 1650-1942
Kund/Customer	Försvarsmakten
Projektnr/Project no	E53077
Godkänd av/Approved by	Anders Törne

FOI, Totalförsvarets Forskningsinstitut
Avdelningen för Informationssystem
Box 1165
581 11 Linköping

FOI, Swedish Defence Research Agency
Information Systems
Box 1165
SE-581 11 Linköping

Sammanfattning

I dagsläget finns ett antal olika säkerhetsvärderingsmetoder. Säkerhetsvärderingsmetoderna skiljer sig bland annat åt genom att ha olika angreppssätt, investeringskostnader med mera. För att underlätta valet av värderingsmetod behövs ett formaliserat sätt att utvärdera säkerhetsvärderingsmetoder.

I denna rapport beskrivs en revision av den tidigare presenterade testproceduren TSAR. Syftet är att testproceduren TSAR ska utgöra ett stöd vid val av metod för att genomföra värderingar av IT-säkerhet. Testproceduren TSAR beskriver olika metoders lämplighet med hjälp av relevansvärden. För att kunna beräkna metoders relevans behövs en uppsättning med metodegenskaper som kan nyttjas för att modellera användares behov avseende värdering av säkerhet. En sådan uppsättning med egenskaper återfinns i denna rapport.

Nyckelord: Säkerhetsvärdering, relevans, testprocedur

Summary

Nowadays there exist a number of different security assessment methods. Different security assessment methods have, for example, different approaches to how to perform security assessments at the same time as the cost of performing an assessment can vary widely. In order to facilitate the choice of security assessment method, a formalized way of evaluating security assessment methods is needed.

This report presents the first revision of the testing procedure TSAR, which is used to evaluate security assessment methods and thereby facilitates the process of choosing a method. The TSAR procedure describes to what degree a security assessment method fulfills the need of security assessment, that is, the relevance of the tested security assessment method. To model the security assessment needs, a set of characteristics is used. The relevance of a security assessment method is decided by comparing the model of the security assessment needs to the characteristics of the method. Such a set of characteristics, to be used for the modeling of security assessment needs, is provided in this report.

Keywords: Security assessment, relevance, testing procedure

Contents

1 Introduction

The widespread use of information systems, and our dependency on them, stresses the need to be able to discuss their security. One way of doing this is to calculate a numerical value which reflects the security of a planned or launched system. A variety of security assessment methods and tools strive to produce such security values for different types of information systems, but the problem is to decide which method is the most appropriate to use in a given situation. To facilitate the choice of security assessment methods a formalized way of evaluating methods would be of considerable value.

The Test of Security Assessment Relevance (TSAR) is a testing procedure which, in a formalized way, evaluates the relevance of different security assessment methods. The test results in numerical values which describes to what degree security assessment methods fulfill important general qualities, which reflects the current needs for security assessment. The TSAR procedure does not provide an in-dept analysis of different assessment methods, but rather makes the initial thinning to identify methods worth analyzing in depth.

1.1 What's new?

This revision of the TSAR procedure is an update of the method presented in (Bengtsson et al, 2008), where the experiences from performed tests using the original version have been regarded (Bengtsson et al, 2009). The following has been changed:

- The validity characteristics have been removed.
- The relevance characteristics have been restructured and updated.
- The simple weighting method for attributes has been updated.

The removal of the validity characteristics is motivated by the intended functionality of the TSAR procedure. It is meant to be a coarse grained filter for rapidly finding a set of methods that are relevant for further study. Finding out if a method is relevant, i.e. assesses the needed areas, can often be decided by a quick study. Finding out if a method is valid, on the other hand, requires a more thorough study of the method along with a multitude of method tests. Since the TSAR procedure was meant to reduce the number of methods that needs to be studied in detail, it is logical to focus on the judgment of relevance.

The restructuring of the relevance characteristics was performed to reduce the importance of a single set of attributes and to present the characteristics in a, for the user, more logical way. The updates of the characteristics were mostly performed to further clarify the aim of each characteristic. The updated characteristics are available in Appendix A.

The update to the simple method involves adopting the hierarchical aspect of the Analytical Hierarchical Process (Saaty, 1994) where the weight of each attribute is dependent on the weight of the parent attribute.

1.2 Report layout

Chapter 2 presents background knowledge and terminology which can come handy while reading this report. Chapter 3 contains a step-by-step guide which describes how to use the TSAR procedure.

2 Background

This chapter presents the use of relevance as a quality for characterizing security assessment methods. Furthermore, the context of security assessment methods and testing procedures is illustrated. Finally, the terminology, related to the area of security assessment, used in this report is presented.

2.1 Relevance

Relevance is a fundamental characteristic regarding information quality (SIS, 2007). The relevance expresses the correspondence between the information needs of the users and the information provided by the presented data.

The test procedure presented in this report focuses on providing quantified values for the relevance of security assessment methods. These values should reflect the correspondence between the relevant needs of prospective users of security assessment methods and the data provided by the tested methods.

2.2 What is a testing procedure?

A testing procedure is used to measure the qualities of security assessment methods and thereby increase the knowledge about the tested methods. Figure 1 illustrates the relations between an information system, a security assessment method and a testing procedure. Two feedback loops are illustrated in order to illuminate how a security assessment method and a testing procedure relate to each other.

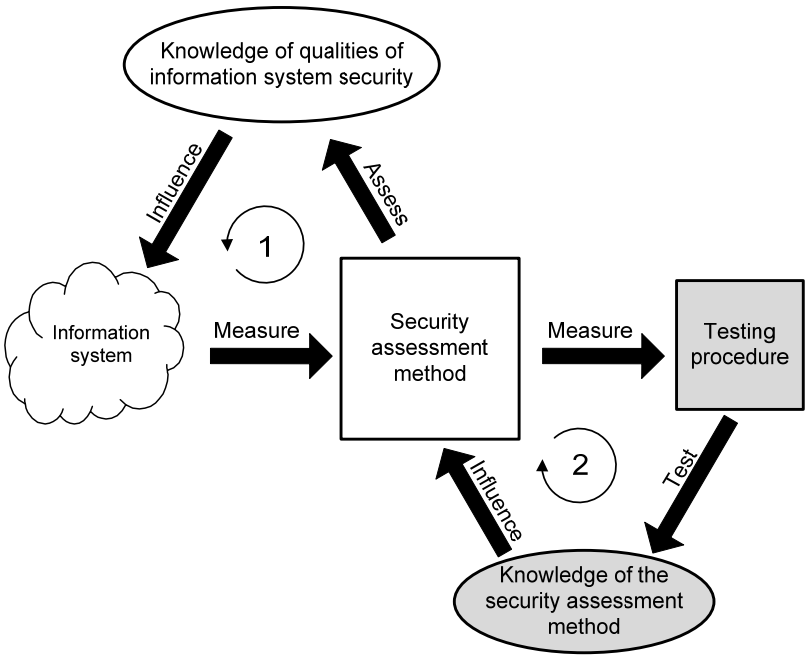

Figure 1: The relation between the testing procedure, security assessment methods, and information systems.

The first loop describes the feedback loop for assessing the security of an information system. The loop starts with a system assessor preparing the input to the assessment method by making measurements on the system. The input is used by the security assessment method in order to perform the assessment. The result of the security assessment is knowledge of qualities of the information system security, which is used by an information system user, administrator, designer etc. in order to influence the system.

The second loop describes the feedback loop for testing a security assessment method. In the first step of the loop, a test supervisor makes measurements on a security assessment method. The measurements are used as input to a testing procedure which is used by the test supervisor to test the security assessment method. The result of the test is knowledge of the security assessment method, which is used by a system assessor, method designer etc. in order to influence the security assessment method.

2.3 Security assessment terminology

Information system
Information systems collect, process, store and distribute information. The term

has a general meaning, but is most often used for computer based information systems. The definition includes the technical equipment of a system as well as its human activities and routines. (Encyclopedia Britannica, 2008)

IT Security assessment
IT security assessments are performed in order to establish how well a system meets specific security criteria. The aim of an IT security assessment is to produce knowledge, which can, for example, be used to improve the security levels of the assessed system. Although perfect security should be the goal, it cannot be achieved. By increasing the knowledge of the assessed system, security assessments improve the validity of the corresponding actors' perception of the information security. Although security assessments cannot guarantee any level of security, they can provide a basis for confidence in the assessed system (Bishop 2003). Thus, the trust in the system may be increased.

Needs
Needs describe activities or resources that are required to be able to perform tasks or reach goals. Needs can be conscious or unconscious, real or imagined, and satisfied or unsatisfied. Outspoken needs are often related to implicit requirement for action or change.

Security assessment
In this report, the term *security assessment* is used in the meaning of IT security assessment.

System
A system consists of cooperating entities working together with a common purpose.

3 The TSAR procedure

This chapter provides a description of how to use the TSAR procedure, starting with the preparation of the input data followed by a step by step description of the procedure.

The aim of the testing procedure is to evaluate the security assessment method in terms of relevance, yielding numerical values for this quality. These numerical values are then interpreted in order to gain more knowledge about the assessment method.

3.1 Preparation of input data

Before using the test procedure it is important to thoroughly prepare the input data. Preparation of the input data is not part of the testing procedure, but since it is essential for getting valid results it is described in short.

As can be seen in Figure 2, the TSAR procedure requires two types of input data; a set of user needs and a set of assessment methods to test. The user needs refers to the needs that should be met by the assessment method. Since the set of assessment methods, which is meaningful to test, will depend on the user needs, the needs should be defined first. The quality of the defined user needs will dictate the quality of the test results, which makes it important to ensure that the user needs have been thoroughly identified.

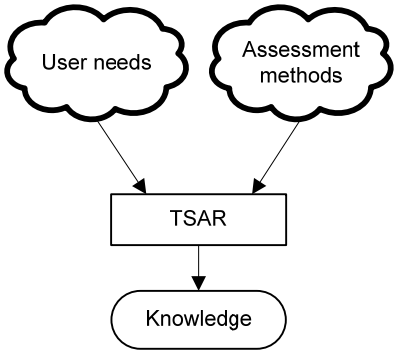

Figure 2: The input to and possible result from the TSAR procedure.

User needs may be stated directly by the user in the form of a set of statements. However, in a more formalized needs analysis, user statements may serve as input to a series of analytical activities that transform the statements into a set of user needs (Hallberg et al, 2005).

The second type of input, the assessment methods, should be selected with the user needs in mind. The most common case, however, will probably be that the set of assessment methods is known in advance. The reason for this assumption is that the problem of having a predefined set of methods from which a choice must be made is what the TSAR procedure was designed to solve.

3.2 Using the TSAR procedure step by step

The TSAR procedure, illustrated in Figure 3, is divided into the five steps (1) Select attributes, (2) Assign weights, (3) Perform measurements, (4) Compute test results, and (5) Interpret the results. The steps are chronologically described in the following five subsections.

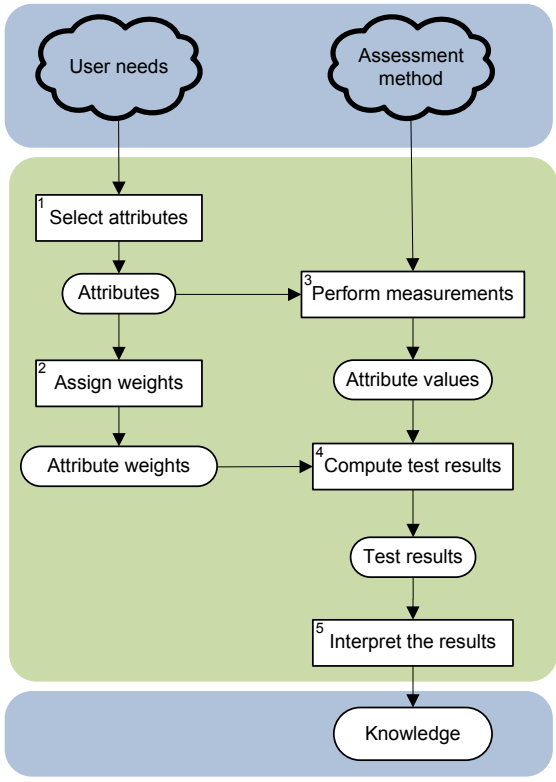

Figure 3: An overview of the TSAR procedure. Clouds represent input, boxes represent activities and rounded boxes represent results.

3.2.1 Select attributes

The TSAR procedure requires a set of characteristics as the basis for the selection of attributes. From this set, the characteristics matching the identified needs are selected and referred to as *attributes*. The selection of attributes should be done by picking the characteristics that best match the identified needs. An example set of characteristics is available in Appendix A.

This activity should be performed by the test supervisor in cooperation with the analysts who identified the needs of security assessment. Thereby the transformation of identified needs into attributes is as accurate as possible.

The result of this step is a set of attributes representing the identified needs of security assessment.

3.2.2 Assign weights

The attributes selected in 3.2.1 will in general be of different importance. To capture this difference of importance, the attributes are assigned weights based on their relative importance. In the next two subsections, two different weighting methods are defined. The two weighting methods differ in how the weight vector is chosen. Note that it is possible to use other methods than the two proposed in 3.2.2.1 and 3.2.2.2 for assigning a weight vector.

The result of this step is a weight vector for the attributes.

3.2.2.1 Type I – AHP weighting of attributes

The first method for defining the weight vectors is to use the mechanisms for criteria weighting in the Analytical Hierarchy Process[1], AHP (Saaty, 1994). AHP is suggested due to its ability to support decision making in scenarios where decision criteria are related in a complex way. AHP takes advantage of the human capability to perform pair-wise comparisons of alternatives and state how much more or less important a certain criterion is compared to another criterion.

The weighting of the attributes starts with the weighting of the leaves of each category. The weight of each leaf is calculated by performing pair-wise comparisons of all leaves within the same category. When the leaves have been weighted, all the categories are pair-wise compared in order to get the weight of each category. The overall weight of each attribute can then be calculated by multiplying the weight of the attribute with the weight of its categories. The weight of all attributes in a category always sums up to 1. Hence the sum of the overall weights for all attributes equals 1. Thereby the importance of a specific

[1] A brief summary of AHP and its advantages using selection of a new car as an illustrative example: http://www.boku.ac.at/mi/ahp/ahptutorial.pdf

attribute can be compared to the importance of any other attribute, regardless of what category it belongs to.

Example
The following example uses AHP for weighting the attributes. Category A consists of the attributes A1, A2 and A3. Category B consists of the attributes B1, B2 and B3. The attributes in category A are pair-wise compared in order to get their relative importance. Then the same thing is done for category B. All leaves now have a weight which shows the relative importance compared to the other attributes in the same category.

To enable comparison of the weights of attributes from different categories, the categories themselves have to be weighted. Therefore the importance of category A is compared to category B, which results in their relative importance.

Examples of weighting results are shown in Table 1 in order to visualize how the calculations are made. By multiplying the weight of a specific attribute with the weight of its category, the overall weight (OW) of that specific attribute is achieved. Based on the weights in Table 1, the OW of attribute A1 would be $3/5 \cdot 2/5 = 12/50$, while the OW of attribute B1 would be $2/5 \cdot 1/2 = 10/50$. Thereby it is possible to find that in this case attribute A1 is of greater importance than attribute B1.

Table 1: Example of weights. W is the weight and OW is the overall weight.

	W	OW
Category A	**3/5**	
Attribute A1	2/5	12/50
Attribute A2	3/10	9/50
Attribute A3	3/10	9/50
Category B	**2/5**	
Attribute B1	1/2	10/50
Attribute B2	1/5	4/50
Attribute B3	3/10	6/50

3.2.2.2 Type II – Simplified weighting of attributes

The second proposed method for finding the weight vectors can be seen as a simplification of the AHP weighting method. Since the testing procedure contains elements of subjective approximation, it may be unnecessarily scrupulous to perform the AHP weighting method, yielding such extremely exact weights. Instead a simpler method is proposed.

As with AHP, the weighting of attributes using the simple method starts by putting weights on the leaves by comparing the attributes in each category. Unlike AHP, this comparison is done for all attributes in a category at once instead of pair wise.

When assigning weight with the simple method, the attributes in each category is divide into two groups: less important (type α) and more important (type β). The groups should contain the same number of attributes, with one more in α or β if there is an odd number of attributes. Define the weight vector such that the weights for each attribute of type β has twice the weight of each type α attribute, while adhering to the rule that the sum of the weight vector elements should be one. Next do the same for each category above the currently weighted, comparing all categories at the same level against each other until the top category is reached.

Example
The following example uses the simple method for weighting the attributes. Category A consists of the attributes A1, A2 and A3. Category B consists of the attributes B1, B2, B3 and B4. The attributes are divided into two groups: less important (type α) and more important (type β). We assume that A1 is of type β and that A2 and A3 is of type α; B2 and B4 are of type β, and B1 and B3 are of type α. We also assume that category A is more important than category B. Thus, the two groups for each category contain roughly the same number of attributes.

For A, with one β and two α, there are $2+2\cdot1=4$ "weight shares" resulting in $\alpha_A=1/4$ and $\beta_A=2/4$. For B with two β and two α there are $2\cdot2+2\cdot1=6$ "weight shares" resulting in $\alpha_B=1/6$ and $\beta_B=2/6$.

Next the category A and B are weighted, resulting in A having 2/3 and B having 1/3 of the weights. From this the overall weights (OW) for each attribute can be calculated by multiplying the weight of the attributes with the weight of their category. The result from this can be seen in Table 2.

Table 2: Example of weights.

	W	OW
Category A	**2/3**	
Attribute A1	2/4	6/18
Attribute A2	1/4	3/18
Attribute A3	1/4	3/18
Category B	**1/3**	
Attribute B1	1/6	1/18
Attribute B2	2/6	2/18
Attribute B3	1/6	1/18
Attribute B4	2/6	2/18

3.2.3 Perform measurements

Performing the measurements consists of determining the *fulfillment* and *coverage* values for the assessment methods. The fulfillment value shows whether an assessment method fulfills an attribute or not, while the coverage value shows whether the description of the assessment method contains enough information to determine the fulfillment value.

Measurements should be performed as follows:

- For each attribute, determine if the tested assessment method can be considered to fulfill the attribute. If so, set the fulfillment value to 1 and the test coverage value to 1. If it is clear that the assessment method does not fulfill the attribute, the fulfillment value is set to 0 and the test coverage value to 1.
- If the description of an assessment method implies that the measured attribute could be fulfilled but does not further describe how, the fulfillment value is set to 0 and the test coverage value is set to 0. This effectively means that the assessment method does not fulfill the attribute based on the current description, but a more detailed description could possibly change this.

The coverage value was incorporated in the test procedure to give a value of a method's potential to increase its current fulfillment value. The coverage value was introduced based on experience from using the TSAR procedure on methods

described in conference papers, since it is not always possible to include all the details in the limited space of such publications.

Example

In Table 3, the different possible values for fulfillment and coverage can be seen. C1 shows that the attribute has been found to be fulfilled. C2 is not fulfilled but the description of the assessment method tested is such that it implies that the attribute could be fulfilled if the description was more thorough. C3 shows that the attribute has been found to not be fulfilled. The weights in the figure do not affect the fulfillment or coverage, but are included for the example presented in the next section.

Table 3: Examples of fulfillment and coverage values.

Attribute ID	Weight	Fulfillment	Coverage
C1	2/5	1	1
C2	2/5	0	0
C3	1/5	0	1

3.2.4 Compute test results

The test results consist of a relevance value and a test coverage. The relevance value is in the range 0 to 1, while the test coverage is expressed as a percentage of the tested attributes.

3.2.4.1 Relevance value

The computation of the relevance value is quite straightforward. The fulfillment value vector is multiplied (inner product) with the weight vector, yielding the relevance value.

Example

Based on the data in Table 3, the relevance value would be calculated in the following way.

Relevance value = $(2/5)\cdot 1+(2/5)\cdot 0+(1/5)\cdot 0 = 2/5 = 0.4$

3.2.4.2 Test coverage

The test coverage is calculated as the sum of the coverage of each attribute divided by the number of attributes.

Example

The example values in Table 3 would result in the following computation.

Test coverage = (1+0+1)/3 = 2/3 ≈ 67%

3.2.5 Interpret the results

Interpreting the results from a test is quite straight-forward since the method with the highest relevance value is the most relevant. However, there are some things to consider.

Firstly, there is no magic threshold value deciding if a method is relevant. A method having a higher relevance value is more relevant than a method having a lower relevance value, based on the performed prioritizations and measurements. However, there is no threshold value deciding if a method with a specific relevance value is relevant or not.

Secondly, the coverage value should be considered when the results are interpreted. A coverage value below 1 means that the description of the method used for testing was not complete and that this method has potential to get a higher score. This has to be taken into account when methods are chosen for further studies.

If the test has been performed by someone else than the user of the test results, it is important to discuss the results so the user understands the reason for the achieved test results.

After the test results have been interpreted and discussed, a decision should be made whether the most relevant method should be used or if a set of the most relevant methods should be further studied.

References

Bengtsson, J., Hallberg, J., Hunstad, A., Löfvenberg, J. (2008). *The TSAR procedure – Test of Security Assessment Relevance and validity.* Scientific report. FOI-R--2624--SE. FOI, Linköping, Sweden.

Bengtsson, J., Hallberg, J., Hunstad, A., Lundholm, K. (2009). *Tester av metoder för värdering av informationssäkerhet (in Swedish).* Scientific report. FOI-R--2901--SE. FOI, Linköping, Sweden.

Bishop, M. (2003). *Computer Security – Art and Science*. Addison-Wesley. ISBN 0-201-44099-7

Encyclopedia Britannica. (2010). *"Information System"*. Encyclopedia Britannica Online. 25 Oct 2010. http://www.britannica.com/EBchecked/topic/287895/information-system

Hallberg, N., Andersson, R., Westerdahl, L. (2005). *Quality-Driven Process for Requirements Elicitation: The Case of Architecture Driving Requirements*. User report. FOI-R--1576--SE. FOI, Linköping, Sweden.

Saaty, T. (1994) *Fundamentals of Decision Making and Priority Theory – with the Analytic Hierarchy Process*, Vol. VI. RWS Publications. Pittsburgh, USA.

SIS (2007). *SIS HB 550: Terminologi för informationssäkerhet, utgåva 3 (in Swedish)*. Technical report, SIS Förlag.

Appendix A Characteristics

This appendix presents a set of characteristics which can be used to perform tests using the TSAR procedure. An overview of the characteristics is presented as a tree in Figure 4.

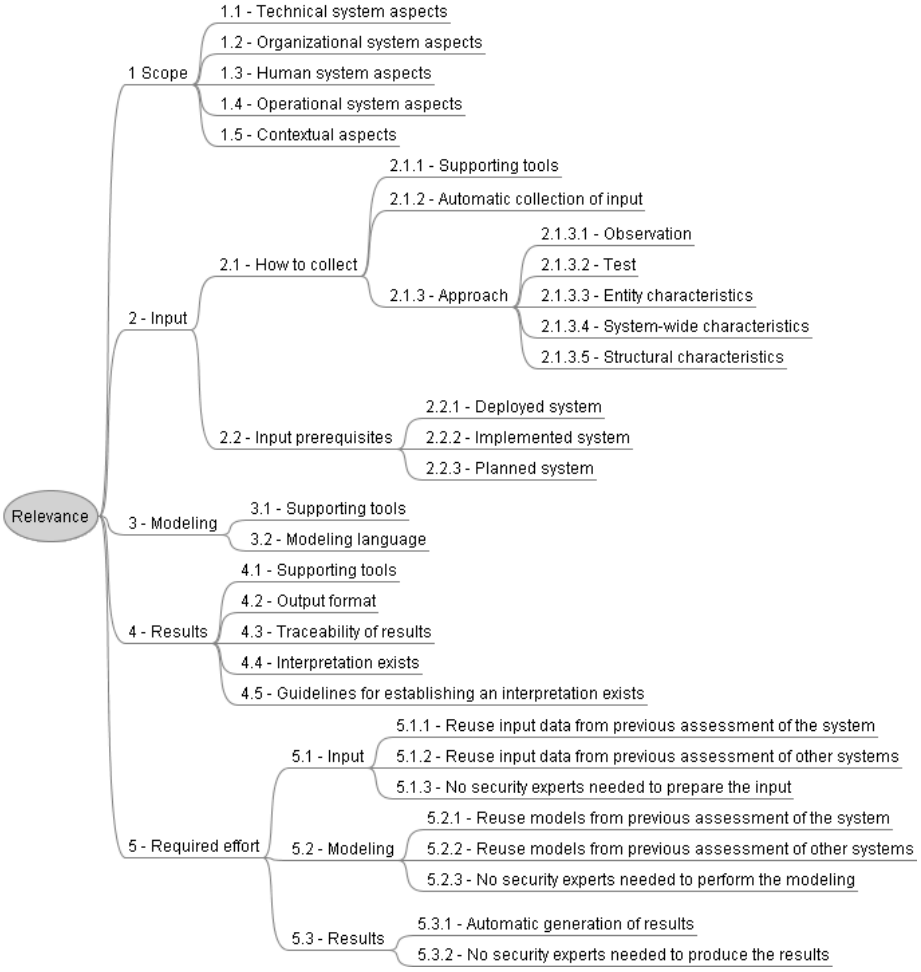

Figure 4: Overview of the proposed set of characteristics.

1 Scope

This is a compound characteristic concerning what the scope for the security assessment is.

The scope of the system security assessment can be limited to one or more system aspects. For example, focusing on technical system aspects will result in a radically different assessment than an assessment based on purely organizational aspects.

1.1 Technical system aspects

This characteristic specifies whether the scope of the assessment includes technical system aspects. Technical system aspects regard if the impact of technical artifacts on the IT security is considered when the assessment is performed.

Technical artifacts include, for example, the choice and configuration of:

- firewalls
- computers
- routers
- web services
- communication protocols

1.2 Organizational system aspects

This characteristic specifies whether the scope of the assessment includes organizational system aspects. Organizational system aspects regard if the impact of organizational artifacts on the IT security is considered when the assessment is performed.

Organizational artifacts include policies, processes, procedures, and routines for information security work.

1.3 Human system aspects

This characteristic specifies whether the scope of the assessment includes human factors. Human system aspects regard whether the impact on the IT security resulting from the interaction between humans and the technical system is considered.

Human factors include the usability of security functions, level of security awareness and training, and employee satisfaction.

1.4 Operational system aspects

This characteristic specifies whether the scope of the assessment includes operational system aspects. Operational system aspects regard how the IT security of the system is affected by the fact that it is in operation.

Operational factors include how information security work is actually carried out, like for example the time it takes until security updates are installed.

The difference to organizational factors is that the operational factors consider the actual work, while the organizational factors consider the routines for said work.

1.5 Contextual aspects

This characteristic specifies whether the scope of the assessment includes contextual aspects. This regards if the impact of contextual artifacts on the possibility to achieve certain levels of IT security is considered when the assessment is performed.

Contextual artifacts refer to factors, which are not part of the system, but still indirectly affect the IT security of the actual system. This can for example include physical protection, laws and regulations.

2 Input

This is a compound characteristic for specifying demands on the input to the security assessment method.

2.1 How to collect

This is a compound characteristic concerning how the input data for the assessment method is collected.

2.1.1 Supporting tools

This characteristic specifies whether the assessment method has explicitly mentioned tools for collecting the input. Supporting tools for the collection of input includes:

- Software for semi-automatic data collection
- Documented procedures for data collection in a structured way

An example of a semi-automatic data collection is when the collected data needs to be processed in Excel before being used for the actual security assessment. Vulnerability scanners, Log analysis tools, Debuggers and AHP are other examples of semi-automatic data collection.

2.1.2 Automatic collection of input

This characteristic specifies whether there exist tools for automatic collection of input.

This characteristic regards reoccurring assessments where the input is automatically collected after an initial configuration. This requires software with fully automated collection of the input data. The software can collect input directly from the system or via other tools such as vulnerability scanners.

2.1.3 Approach

This is a compound characteristic for how the input data to the assessment method is obtained.

2.1.3.1 Observation

This characteristic specifies whether input data is collected through observations. Examples of input data collected through observations are:

- Number of virus infections per time unit
- Number of intrusions per time unit
- Number of confidentiality violations per time unit
- System logs

2.1.3.2 Test

This characteristic specifies whether input data is collected through tests. Examples of tests that can be used are:

- Vulnerability scanners
- Red teams
- Code inspection

2.1.3.3 Entity characteristics

This characteristic specifies whether input data is collected based on entity characteristics.

Entities are subjects, objects or subsystems that perform tasks in a system and the tasks themselves. These can be described by performance characteristics, interfaces etc.

Examples of entities are:

- Computers
- Implemented security functions

2.1.3.4 System-wide characteristics

This characteristic specifies whether input data is collected based on system-wide characteristics.

Examples of system-wide characteristics are:

- Systems ability to withstand attacks
- Update policies and their implications
- Number of computers
- Number of users

2.1.3.5 Structural characteristics

This characteristic specifies whether input data includes structural characteristics.

Examples of structural characteristics are:

- Interactions between entities
- Network routing

2.2 Input prerequisites

This is a compound characteristic concerning what stage of development the assessed system must be in for the assessment method to be applicable.

2.2.1 Deployed system

This characteristic concerns whether it is possible to collect the input data needed by the assessment method from a deployed system. That is, the assessment considers aspects of deployed systems affecting the security.

Deployed systems are those that produce services in a live environment. That is, systems that are relied upon to perform a service.

2.2.2 Implemented system

This characteristic concerns whether it is possible to collect the input needed by the assessment method from an implemented system. That is, the assessment considers aspects of the system, which can be partially or fully implemented.

A fully implemented system is a system which is ready for deployment.

Partially implemented systems can be prototypes of systems or functional sub-components of a system. An example of this is systems under development.

2.2.3 Planned systems

This characteristic concerns whether it is possible to collect the input needed by the assessment method from descriptions of systems that do not yet exist but are planned to be implemented.

Planned systems can be described as, for example:

- System requirements documentation
- Network topology
- System specification

3 Modeling

This is a compound characteristic about how to perform the modeling required for the assessment method.

3.1 Supporting tools

This characteristic concerns the availability of tools that facilitate the modeling required by the assessment method. The modeling is part of the actual

assessment. Thus, the tools used for modeling should be specified in the description of the assessment method.

Examples of modeling tools are Microsoft Visio, MooD and Sparx.

3.2 Modeling language

This characteristic specifies whether created models are described using a standardized or de-facto standard modeling language.

Established modeling languages can for example be:

- Unified Modeling Language (UML)
- Finite state machines

4 Results

This compound characteristic concerns how the result from the assessment method is produced as well as how this result is presented. Further, the interpretation of the result is considered.

4.1 Supporting tools

This characteristic specifies whether tools exist for producing the results of the assessment method. The tools used for producing the assessment results should be specified in the description of the assessment method.

4.2 Output format

This characteristic concerns whether the results from supporting tools are available in an export-friendly format.

Examples of export-friendly formats are:

- XML
- Text file
- Excel

4.3 Traceability of results

This characteristic concerns whether the results produced by the assessment method can be traced back to the factors that caused the results. In order to fulfill this characteristic it should be possible to identify the reasons for the outcome of the assessment.

In a method without traceability of results the assessment states the security level, but not why the security is at this level. A method with traceability is able to motivate the results.

4.4 Interpretation exists

This characteristic concerns whether there is a predefined way to interpret the result from the assessment.

An assessment method can be considered to have a predefined interpretation of the results if there is a mapping from the possible results to qualitative statements about the security.

For example, a method producing a security value X in the range $0 \leq X \leq 1$ is supplemented with the following predefined interpretations of the result:

$0 \leq X < 0.6$: Needs immediate attention

$0.6 \leq X < 0.9$: Should be further checked up on

$0.9 \leq X \leq 1$: No action required

Another type of predefined interpretation is when an assessment tool has the interpretation as output. The result from the assessment method can for example be a list of actions that should be performed to reach an adequate level of security.

4.5 Guidelines for establishing interpretation exists

This characteristic concerns whether there are guidelines for establishing an interpretation of the assessment results.

Guidelines for establishing an interpretation can for example be an algorithm for producing the interpretation based on the assessment results and additional data. Examples of additional data are:

- Statistics on assessment results from other systems
- Previous assessment results
- Security needs of system stakeholders

The difference to a predefined interpretation is that the guidelines do not tell the user how to interpret the assessment results, but rather helps the user to establish a basis for the interpretation.

5 Required effort

This compound characteristic concerns how much effort is required to perform the security assessment. The term effort includes both work hours needed as well as direct financial costs.

5.1 Input

This compound characteristic concerns the effort required to collect the input data.

5.1.1 Reuse input data from previous assessments of the system

This characteristic specifies whether it is possible to reuse input data from previous security assessments of the system. That is, only changes to the system needs to be considered.

Reused data is static properties in the system that has not changed since the last assessment. This data can for example describe:

- Characteristics of system hardware
- Classification of information

5.1.2 Reuse input data from previous assessments of other systems

This characteristic specifies whether it is possible to reuse input data from previous assessments of other systems.

Reused data could for example describe:

- Components that the systems have in common
- Rules or regulations that apply to the systems

5.1.3 No security experts needed to prepare the input

This characteristic specifies whether preparation of the input to the security assessment can be performed *without* consulting experts.

The characteristic should be considered fulfilled if the assessment method does not require experts to prepare the input to the method.

5.2 Modeling

This compound characteristic concerns the effort required to create the models required by the assessment method.

5.2.1 Reuse models from previous assessments of the system

This characteristic specifies whether it is possible to reuse models from previous security assessments of the system.

Reused data could for example be:

- System models from previous assessments which can be modified to reflect the current system
- Prioritizations of security relevant characteristics influencing the assessment results

5.2.2 Reuse models from previous assessments of other systems

This characteristic specifies whether it is possible to reuse models from previous security assessments of other systems.

Reused data could for example be:

- System models from previous assessments which can be modified to reflect the current system
- Prioritizations of security relevant characteristics influencing the assessment results

5.2.3 No security experts needed to perform the modeling

This characteristic specifies whether the modeling required during the assessment can be performed *without* consulting experts. Examples of required modeling include:

- Modeling of systems to be assessed
- Specification of the computations of assessment results

The characteristic is not fulfilled if the assessor, that is, the person responsible for the assessment, has to resort to consulting experts in order to be able to perform the required modeling. In some cases, it may be difficult to draw the line between the processes of acquiring input and consulting security experts. For example, acquiring the number of virus infections during the last month would be considered input data. On the other hand, acquiring data regarding the prioritization of security-relevant system characteristics would be considered to be support from security experts.

5.3 Results

This compound characteristic concerns the effort required to produce and interpret the assessment results.

5.3.1 Automatic generation of results

This characteristic concerns whether the results from the security assessment are produced automatically, without any involvement of humans. The method may require initial, manual setup.

One example is a network scanner that automatically detects and reports on vulnerabilities that needs to be fixed.

5.3.2 No security experts needed to produce the results

This characteristic specifies whether the assessment results can be generated *without* consulting security experts. Examples of tasks that need to be performed without consulting security experts are:

- Computation of aggregated security values
- Interpretation of the computed results
- Presentation of the computed results