Proceedings of the Federated Conference on Computer Science and Information Systems pp. 443–447 DOI: 10.15439/2019F84 ISSN 2300-5963 ACSIS, Vol. 18

Formalization of Software Risk Assessment Results in Legal Metrology Based on ISO/IEC 18045 Vulnerability Analysis

Marko Esche, Felix Salwiczek
Physikalisch-Technische Bundesanstalt,
Abbestraße 2-12, 10587 Berlin, Germany
Email: {marko.esche, felix.salwiczek}@ptb.de

Federico Grasso Toro
Federal Institute of Metrology METAS,
Lindenweg 50, 3003 Bern-Wabern, Switzerland
Email: federico.grasso@metas.ch

Abstract—**The Measuring Instruments Directive sets down essential requirements for measuring instruments subject to legal control in the EU. It dictates that a risk assessment must be performed before such instruments are put on the market. Because of the increasing importance of software in measuring instruments, a specifically tailored software risk assessment method has been previously developed and published. Related research has been done on graphical representation of threats by attack probability trees. The final stage is to formalize the method to prove its reproducibility and resilience against the complexity of future instruments. To this end, an inter-institutional comparison of the method is currently being conducted across national metrology institutes, while the weighing equipment manufacturers' association CECIP has provided a new measuring instrument concept, as a significant example of complex instruments. Based on the results of the comparison, a template to formalize the software risk assessment method is proposed here.**

I. Introduction

WHEN MEASURING results obtained from a measuring instrument are used to determine the price to pay for a certain good (such as water, heat, petrol, electricity) in the EU, said instrument is subject to the Measuring Instruments Directive (MID) 2014/32/EU [1]. The essential requirements of the MID include software requirements for protection against corruption, see L 96/173 in [1]. In addition, the MID defines conformity assessment procedures which an instrument has to pass before being made available on the common market. In the frame of most of these assessment procedures, manufacturers are required to conduct a risk assessment demonstrating that their product fulfils the essential requirements. To aid manufacturers whith this task, PTB (Germany's national metrology institute, one notified conformity assessment body for the MID) has developed a software risk assessment procedure [2]. This procedure, specifically tailored to the needs of legal metrology, i.e. the economic sector of measurements subject to legal control, is employed by PTB when performing conformity assessments. To harmonize conformity assessment practice across Europe, the European Cooperation in Legal Metrology (WELMEC) Working Group 7 ”Software“ is currently investigating this software risk assessment method in the frame of an inter-institutional comparison. The aim is to demonstrate the objectiveness of the procedure and the reproducibility of its results. If needed, it is intended to amend the procedure to achieve both goals. To ensure impartial results, generic abstract instruments are used for the comparison. Initial findings indicate that producing objective assessment results for today's simple instruments should be feasible. However, future complex systems will pose a bigger challenge. Most importantly, the simple representation of the assessment result in the form of a single risk score simplifies the assessment process too much. Therefore, it is proposed to improve the investigated method by formalizing the recording of its results, by means of a risk assessment template. Since the procedure closely follows the vulnerability analysis of ISO/IEC 18045 [3], the outcome of this paper should be useful to all assessment procedures (such as ETSI TS 102 165-1 [4]) that are based on the same standard. The remainder of the paper is structured as follows. The basic priniciples of the risk assessment procedure are recapitulated in Section II. Section III details the inter-institutional comparison, the examined generic measuring instruments and describes challenges derived from the results of the comparison. The proposed solution by means of a formalized risk assessment template is detailed in Section IV. Section V summarizes the paper.

II. Basic Principles of the Risk Assessment Procedure

As mentioned in the introduction, manufacturers of measuring instruments shall perform and document a risk assessment of their instruments before submitting a prototype to a NB for conformity assessment, in accordance with Module B (type evaluation) of the MID. To aid manufacturers and NBs in this task, a procedure was developed and published in [2]. With the aim of providing an objective procedure to generate reproducible results, the method is based on the international standards ISO/IEC 27005 [5] and ISO/IEC 18045 [3]. ISO/IEC 27005 provides a principle description of the risk assessment process consisting of three phases:

This work was conducted within the frame of work package 4 of the European Metrology Cloud Project.

TABLE I
TOE RESISTANCE OF MEASURING INSTRUMENTS TO ATTACKS AND ASSOCIATED PROBABILITY SCORE [3].

Sum of points	TOE resistance	Probability score
0-9	No rating	5
10-13	Basic	4
14-19	Enhanced Basic	3
20-24	Moderate	2
≥ 24	High	1

A. *Risk Identification*

During risk identification, unwanted events (so-called threats to assets) are defined based on ”legal and regulatory requirements, and contractual obligations“. Such assets can be derived from the essential requirements given in Annex I of the MID. For convenience reasons, only two such assets are examined here. One asset is the measurement result with the associated security property authenticity, since the MID prohibits the use of measurement results that do not originate from a certified measuring instrument. The other asset is the software critical for the measurement purpose, which shall not be modified or replaced. Therefore, such software can be assigned the security properties integrity and authenticity. A list of all assets applicable to legal metrology is given in [2].

B. *Risk Estimation*

During risk estimation, threats are assigned a quantitative or qualitative risk measure. One possibility to calculate such a measure is given by ISO/IEC 27005 itself, where ”risk is a combination of the consequences that would follow from the occurrence of an unwanted event and the likelihood of the occurrence of the event.“ If unwanted events (threats), have been defined properly, they can be assigned an impact score between 0 (no effect) and 1 (all measurement results affected), signifying the severity of the consequences. The method from [2] uses a score of $\frac{1}{3}$ if only one result is affected by the threat. In addition, a measure for the probability of occurrence is needed. This can be estimated by evaluation of different actions (attack vectors) that an attacker needs to implement for the threat to be realized. The vulnerability analysis provided in Part 2 of ISO/IEC 18045 [3] constitutes one possiblity to quantify the probability of occurrence for such attack vectors by means of point scores assigned in the following categories:

- Elapsed Time (0-19 points)
- Expertise (0-8 points)
- Knowledge of the Target of Evaluation (0-11 points)
- Window of Opportunity (0-10 points)
- Equipment (0-9 points)

An example for a fully evaluated attack vector with assigned scores is given in Table III. The calculated sum score can be mapped to a target of evaluation (TOE) resistance, see [3], and an equivalent probability score between 1 and 5, see Table I. The third column is not part of the original table presented in [3]. Afterwards, by multiplying impact and probability score a risk score can be obtained which will be in the range between 1 (very low risk) and 5 (very high risk).

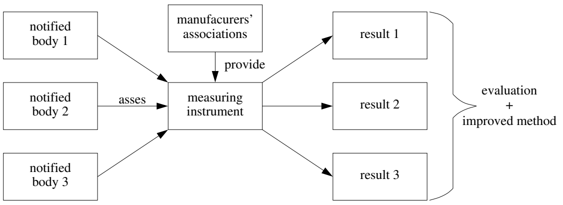

Fig. 1. Anticipated workflow of the risk assessment inter-institutional comparison and its projected outcome.

C. *Risk Evaluation*

During risk evaluation, the calculated risk is put into the context of the field of application for the assessed instrument. Estimated risks are prioritized and a cut off point for the risk assessment is defined. In addition, an initial list of risks to be mitigated by order of importance is produced. As a rule of thumb, PTB will ask manufacturers who have obtained a risk score of 4 or 5 for their instrument to implement additional protective measures. Once these have been implemented, the phases of risk estimation and risk evaluation (including amendments, where necessary) are repeated until the risk score is reduced to 3 or lower. Since attack vectors for real-world measuring instruments might become very complex, they can be decomposed by means of Attack Probability Trees (AtPT), see [6]. These AtPTs can be used by an assessor to subdivide any given attack vector, evaluate the sub-goals and to find the attack probability score for the original complex attack vector.

III. INTER-INSTITUTIONAL COMPARISON AND IDENTIFICATION OF CHALLENGES

Since conformity assessment bodies all across Europe are faced with the challenge of interpreting and evaluating the results of risk assessments, WELMEC Working Group 7 has decided to examine the procedure developed by PTB more closely by means of an inter-institutional comparison with five different NBs. To start the comparison, assessors from these NBs took part in a training exercise. The training covered both the basic procedure [2] as well as AtPTs [6]. Afterwards, see Subsection III-A, two generic measuring instruments were selected for all partners to assess. Subsection III-B describes the examined threats and initial findings. Figure 1 illustrates the workflow of the inter-institutional comparison.

A. *Description of Generic Reference Instruments*

The first instrument assessed is a complex cloud-based measuring system proposed by CECIP, the European weighing instruments manufacturers’ association. WELMEC Working Group 7 anticipates that such systems will be the norm in legal metrology in the near future. The system consists of a number of sensors subject to legal control that send data to a processing software running in the instrument manufacturer’s own cloud, see Figure 2. The cloud offers data storage and a display server (DSP). The DSP sends measurement results to different display devices, e.g. smart phones or general-purpose printers. Communication between the components is realized via Wi-Fi with WPA encryption. Additionally, all transferred data are protected by CRC-16 codes to ensure integrity of transmitted data. Three kinds of users are foreseen for the cloud: administrators with full privileges, maintenance personal with access to log files and backend users. In case data are lost during transmission, all sending devices have sufficiently large buffers for retransmission. Two categories of display devices are established, namely ”full control“ and ”receive only‘, where the prior devices are only accessible to a trustworthy user group. A full system description will be published by CECIP in a future paper. Due to the complexity of the cloud-based instrument and the resulting increased probability for assessment errors, it was agreed to also use a second simpler generic instrument.

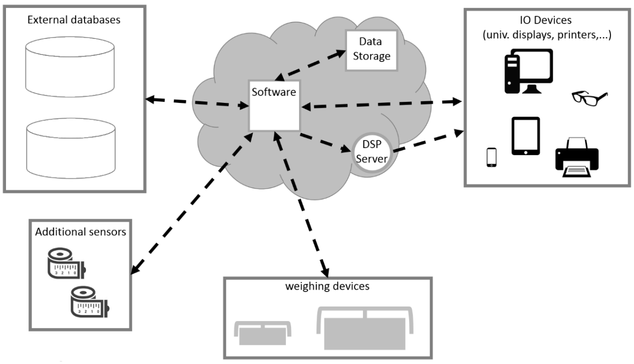

Fig. 2. General structure of the generic complex cloud-based system provided by CECIP, the European weighing instruments manufacturers’ association.

The weighbridge depicted in Figure 3 is an automatic weighing instrument designed for weighing cargo transported on a truck. The measurement is started through the terminal’s GUI consisting of an LCD and eight buttons. The measurement result is directly shown on the LCD. Two load cells measure the weight of front and rear axle of the truck. Two evaluator units interpret the output of the load cells. These units then communicate with the terminal where the final measurement result is computed. Evaluator units and terminal are based on microprocessors, data can be read from the terminal via RS485 or exported to a USB stick. The terminal checks the authenticity of all other units at startup by requesting a CRC-16 of their firmware based on a secret start vector. Legally relevant parameters and software are stored in the terminal unit on a hardware-protected flash memory. All software on the system is subject to legal control. All connections within the system are physically sealed.

B. Experimental Results of the Inter-institutional comparison

To narrow down the scope of the comparison, it was agreed to examine only two threats for both instruments, although in principle, all assets derived from the MID would need to be taken into account:

T1: An attacker introduces false measurement results into the measuring instrument.

T2: An attacker modifies or replaces the software critical for the measuring task.

TABLE II
Assessment results provided by different NBs.

Threat	NB	Impact	Proba-bility score	Risk
complex cloud-based system				
T1	NB1	1	4	4
	NB2	1	1	1
	NB3	1	4	4
	NB4	1	4	4
	NB5	1	3	3
T2	NB1	1	2	2
	NB2	1	1	1
	NB3	1	5	5
	NB5	1	3	3
simple measuring instrument				
T1	NB1	$\frac{1}{3}$	2	1
	NB3	1	3	3
	NB4	1	3	3
	NB5	1	3	3
T2	NB1	1	1	1
	NB3	1	3	3
	NB5	1	3	3

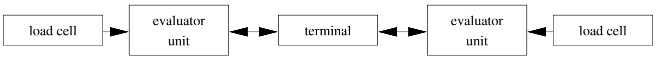

Fig. 3. Components of the generic automatic weighing instrument.

All NBs were asked to identify at least one attack vector per threat per measuring instrument and to evaluate that attack vector as described in Section II. The outcome is a list of point scores for the five mentioned categories. The sum score results in a probability score, see Table I, which produces a risk score when multiplied with the identified impact. Table II summarizes the results provided the NBs for threats T1 and T2. Not all NBs assessed both threats for both instruments. The results from different NBs for threat T1 for the cloud-based system are visualized in Figure 4. For this threat, all NBs selected an attack vector with a permanent effect (impact score of 1). Despite varying sum scores (11 to 17), the probability and risk scores for NB1, NB3, NB4 and NB5 are very close to each other due to the range of sum scores allowed by Table I. The only exception is the attack vector selected by NB2, which appears to be more difficult to implement than the others. When coparing results from NB4 and NB5, another property of the ISO/IEC 18045 vulnerability analysis becomes apparent: A larger score for expertise might be compensated by a smaller score for time, since a layman may take longer to implement a certain attack than an expert. While the results for T1 might suggest that consistent results can be easily obtained by different assessors, the results for threat T2 prove otherwise, see Figure 5. All four NBs concluded that the chosen attack vector would have a permanent effect. For all other scores, the results vary widely. Consequentially, probability and risk scores also differ. One reason for this variability is the imprecisely formulated threat T2, which allows either a partial modification or a complete replacement

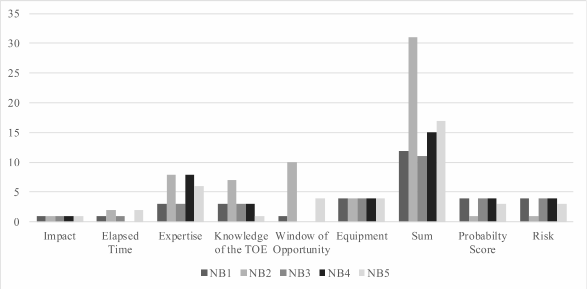

Fig. 4. Results for the complex cloud-based measuring system, evaluation of threat T1 (introduction of false measurement results).

of the software. While a small modification will require less expertise and time (attack examined by NB3), a full redevelopment will require more expertise, time etc. (attack examined by NB1). It is concluded, that properly documented attack vectors and precisely defined threats are key to comparing risk assessment results obtained by different parties. Moreover, assessment results for the cloud-based system do not depend on the perspective of the examiner alone, but also on the chosen attack vector. If an attacker aims to introduce false measurement results by providing them with a valid CRC-16 from a trustworthy source, this will be much less difficult than manipulating data within the WPA-protected Wi-Fi. To solve this disambiguity, all NBs were also asked to perform software risk assessments for the much simpler weighbridge instrument, detailed in Subsection III-A. The results obtained for threat T1 are depicted in Figure 6. NB3, NB4 and NB5 chose to examine attack vectors with a permanent effect, e.g. replacement of a sensor. Again, the point scores vary depending on the selected attacker profile (layman with restricted knowledge vs. expert with publicly available knowledge). Nevertheless, all three NBs arrive at similar sum and probability scores, resulting in identical risk scores. NB1 has examined an alternative attack vector requiring repetition for each measurement (reduced impact of $\frac{1}{3}$). Since this attack also appears to be more complex (writing a specialized software vs. installing a sensor) the risk score obtained is much lower. In this regard, a set of evaluated reference attack vectors could reduce the assessor's required effort and harmonize the outcome of different assessments performed for the same instrument. Concerning threat T2, the same effect can be observed, see Figure 7. While NB3 and NB5 focused on the simple modification of existing software, NB1 expected attackers to implement new software and then replace the original. The differences are due to the different focus of the attack vectors and an imprecisely formulated threat. This could be avoided by formulating individual threats per identified asset and requiring assessors to document chosen attack vectors and their effects for later comparison.

C. Main Challenges

The justification - for rejecting certain assessed attack vectors as unlikely or for quantifying certain scores as wrong - was provided by a review session among the NBs involved. In practice, discussions between NBs about risk analyses provided by different manufacturers are unlikely to happen. Moreover, new examiners may not be familiar with these findings and will be facing the same challenges. As shown in Subsection III-B, an objective comparison of risk assessments is only possible if certain prerequisites are fulfilled:

- Instructions for new evaluators on how to assess risks according to the standard shall be readily available.
- Examples for evaluation of common attack vectors to reduce the workload for evaluators shall be supplied.
- Proper documentation of the complete attack vector and justification for the evaluation shall be required of all assessors for better comparablity of assessment results.

Section IV addresses all three by providing a formalized framework, by means of a software risk assessment template.

IV. Formalization of Risk Assessment Results

Instructions on how to perform a vulnerability analysis are provided by Part 2 clause B.4.2.2 ff of ISO/IEC 18045. Since the method discussed here is based on that standard, the same instructions may be used when performing software risk assessments in legal metrology. However, the standard's guidance is intended for all fields of IT security and is thus kept very general. In the template proposed, a shorter method description is included focused on the needs of legal metrology. Thereby, it is ensured that all assessors are aware of all steps to be performed. The workflow of the template is shown in Figure 8. The template includes a list of all assets

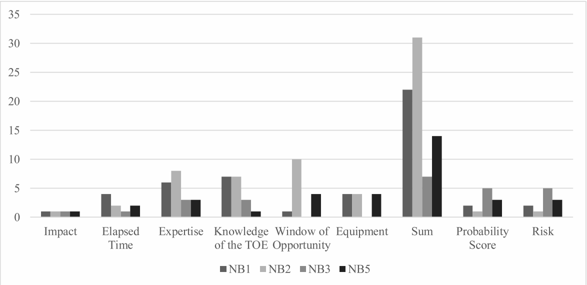

Fig. 5. Results for the complex cloud-based measuring system, evaluation of threat T2 (modification or replacement of software critical for the measurement).

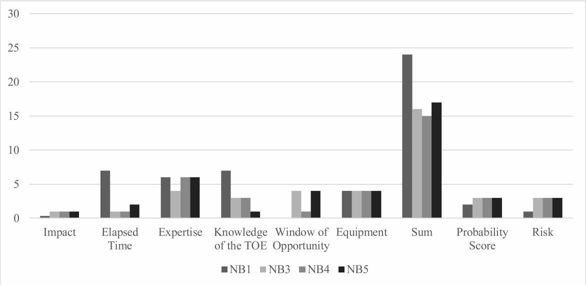

Fig. 6. Results for the simple weighbridge, evaluation of threat T1 (introduction of false measurement results).

TABLE III
EXAMPLE FOR A FULLY EVALUATED ATTACK VECTOR.

Attack vector	Time	Expertise	Knowledge	Window of opport.	Equipment	Justification
Attacker constructs fake results from datasets protected by a CRC32 with a secret start vector.	0	3	3	0	0	Assumed attacker: customer. CRC is a linear operation on binary vectors, an XOR-connection of two datasets automatically produces a third dataset with correct CRC. This can be calculated with standard software by a proficient user. No window of opportunity needed. The CRC is described in the manual.

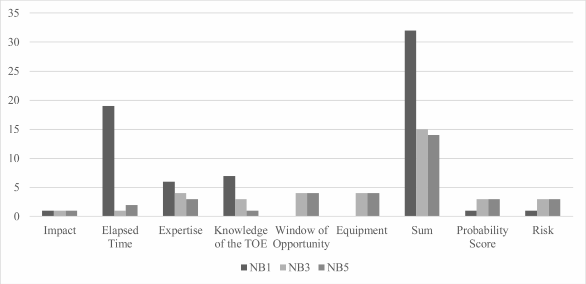

Fig. 7. Results for the simple weighbridge, evaluation of threat T2 (modification or replacement of software critical for the measurement).

and their security properties derived from the MID. These are accompanied by explanations of the assets and examples on how an attacker might invalidate their security properties. The assessor is required to select the applicable assets from the list and to formulate the relevant threats precisely. Even though the identification of attack vectors cannot be standardized, the template provides assessors with some assistance. A number of evaluated reference attack vectors are given, see Table III for an example. Also, assessors are offered the possibility to use AtPTs to decompose attack vectors to facilitate the evaluation. Most importantly, the template requires all assessors to provide justification for each point score alongside the evaluated attack vector. If sufficient details and justification are provided, discussion about assessment results will no longer be necessary, unless a different assessor takes issue with the point score assigned to a specific attribute. In this case, an AtPT can be used to decompose the attack until no room for argument is left. The template will not guarantee uniform results, but if the guidance remarks are observed, there will be sufficient documentation to successfully argue in favor or against a certain assessment result. The template can be found under the following link: https://www.ptb.de/cms/fileadmin/internet/fachabteilungen/abteilung_8/8.5_metrologische_informationstechnik/8.51/Risk_Assesment_Template_v11.docx

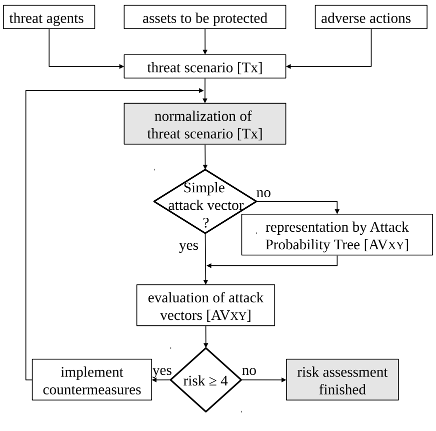
threat agents
assets to be protected
adverse actions
threat scenario [Tx]
normalization of threat scenario [Tx]
Simple attack vector ?
no
yes
representation by Attack Probability Tree [AVXY]
evaluation of attack vectors [AVXY]
implement countermeasures
yes
risk ≥ 4
no
risk assessment finished

Fig. 8. Template workflow mirrors risk assessment procedure [2].

V. SUMMARY

As long as software risk assessment depends on human creativity and judgement, the resulting risk scores will always be biased. Nevertheless, detailed guidance on the assessment steps together with proper documentation of all steps of the assessment may serve as a basis to make software risk assessment results more easily comparable. The vulnerability analysis of ISO/IEC 18045 already provides general remarks on the workflow and on the point scores for specific attributes of assessed attack vectors. These were mapped to the needs of the legal metrology community and augmented by specific detailed examples to help assessors with repetitive tasks. The experimental findings from the inter-institutional comparison and the suggested risk assessment template, should be applicable to any group planning to implement ISO/IEC 18045 vulnerability analysis. To validate the template, WELMEC Working Group 7 is currently performing a second stage of risk assessments using the new template.

REFERENCES

[1] "Directive 2014/32/EU of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of measuring instruments," European Union, Council of the European Union ; European Parliament, Directive, February 2014.

[2] M. Esche and F. Thiel, "Software risk assessment for measuring instruments in legal metrology," in *Proceedings of the Federated Conference on Computer Science and Information Systems*, Lodz, Poland, September 2015. doi: http://dx.doi.org/10.15439/978-83-60810-66-8 pp. 1113–1123.

[3] "ISO/IEC 18045:2008 Common Methodology for Information Technology Security Evaluation," International Organization for Standardization, Geneva, CH, Standard, September 2008, Version 3.1 Revision 4.

[4] "ETSI TS 102 165-1 Telecommunications and Internet converged Services and Protocols for Advanced Networking; Methods and protocols; Part 1: Method and proforma for Threat, Risk, Vulnerability Analysis," European Telecommunications Standards Institute, Sophia Antipolis Cedex, FR, Standard, March 2011, v4.2.3.

[5] "ISO/IEC 27005:2011(e) Information technology - Security techniques - Information security risk management," International Organization for Standardization, Geneva, CH, Standard, June 2011.

[6] M. Esche, F. Grasso Toro, and F. Thiel, "Representation of attacker motivation in software risk assessment using attack probability trees," in *Proceedings of the Federated Conference on Computer Science and Information Systems*, Prague, Czech Republic, September 2017. doi: http://dx.doi.org/10.15439/2017F112 pp. 763–771.