# Title: State of the art on methodologies for risk- and model-based security testing

### Summary / Contents:

**Contributors:**

Menz, Rennoch, Großmann, Schieferdecker (Fraunhofer FOKUS)

Maag, Cavalli (Institut Telecom SudParis)

Erdogan, Li, Seehusen, Stølen (SINTEF)

Noponen (VTT)

Weiser (Oulu)
TABLE OF CONTENTS

1. Introduction ................................................................................................................................. 7
2. State-of-the-art on methodology for model-based security testing ........................................... 7
   2.1 Model-driven security ........................................................................................................... 7
   2.2 Security testing methodologies based on formal languages ................................................. 9
       2.2.1 Security testing methodology with TEFSM and Nomad .............................................. 9
       2.2.2 Security testing methodology with ocl4st ................................................................. 11
       2.2.3 From the Or-BAC security rules and the EFSM-based SUT ....................................... 15
   2.3 Network monitoring using machine learning techniques .................................................... 16
       2.3.1 Traffic analysis ............................................................................................................. 17
       2.3.2 The machine learning approach .................................................................................. 17
   2.4 Security testing with GCC exTensions ............................................................................... 18
   2.5 Model Interference Assisted Fuzzing ................................................................................. 19
3. State-of-the-art on methodology for Risk-Based Security Testing ......................................... 21
   3.1 Standards supporting risk-based security analysis ............................................................ 21
   3.2 A literature survey of risk-based testing ............................................................................. 25
       3.2.1 The Prisma method ..................................................................................................... 26
       3.2.2 Redmill's approch ....................................................................................................... 27
       3.2.3 Amland – Risk-based testing ..................................................................................... 27
       3.2.4 Back – Heuristic Risk-Based Testing ......................................................................... 28
       3.2.5 RiteDAP ....................................................................................................................... 28
       3.2.6 Rosenberg et.al. Risk-oriented Object Oriented Testing ............................................ 29
       3.2.7 Murthy et al. – Leveraging Risk Based Testing in Enterprise Systems Security Validation ................................................................. 29
       3.2.8 Zech – Risk-Bases Security Testing in Cloud Computing Environments ................ 29
4. Conclusions ............................................................................................................................... 30
5. References .................................................................................................................................. 30

FIGURES

Figure 1 - Security Testing methodology overview ........................................................................ 11
Figure 2 - UML meta-model of UML4ST ..................................................................................... 11
Figure 3 - Generation methodology from secure diagrams ......................................................... 14
Figure 4 - Attack Potential [11] .................................................................................................. 23
Figure 5 - Calculation of attack potential [11] ............................................................................ 24

TABLES

© Copyright DIAMONDS Consortium
Table 1 Risk-based testing approaches
## HISTORY

<table>
<thead>
<tr>
<th>Vers.</th>
<th>Date</th>
<th>Author</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>22/08/11</td>
<td>Fredrik Seehusen</td>
<td>Template created</td>
</tr>
<tr>
<td>0.1</td>
<td>26/10/11</td>
<td>Stephane Maag</td>
<td>Section 5 added</td>
</tr>
<tr>
<td>0.2</td>
<td>11/11/11</td>
<td>Menz, Rennoch, Großmann</td>
<td>New Secton 6</td>
</tr>
<tr>
<td>0-5</td>
<td>15/11/11</td>
<td>Fredrik Seehusen</td>
<td>Integrated contributions and edited document</td>
</tr>
<tr>
<td>1</td>
<td>12/12/11</td>
<td>Fredrik Seehusen</td>
<td>Feedback from reviewers implemented</td>
</tr>
</tbody>
</table>

## APPLICABLE DOCUMENT LIST

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Title, author, source, date, status</th>
<th>DIAMONDS ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
EXECUTIVE SUMMARY

This document constitutes the first deliverable of work package 4 on risk- and model-based security testing methodologies. While the other work packages of the DIAMONDS project describe techniques/methods and tools, work package 4 describes processes/guidelines for applying these tool and techniques in practice. Thus, the state-of-the-art topics of this deliverable are related to techniques in WP2 that should be supported by methodologies and whose state-of-the-art survey has not already been covered by deliverable D1.WP2.
1. INTRODUCTION

This document constitutes the first deliverable of work package 4 of the DIAMONDS project. Whereas the other work packages focus on techniques (work package 2) and tools (work package 3), work package 4 focuses on methodology, i.e. processes and guidelines for how to apply the techniques and tools in practice.

The deliverable addresses two main research areas: model-based security testing and risk-based security testing. The first topic is related to task 4.2 of work package 4, and is documented in Section 2. The second topic is related to task 4.3 of work package 4 and documented in Section 3.

2. STATE-OF-THE-ART ON METHODOLOGY FOR MODEL-BASED SECURITY TESTING

This section gives a state-of-the-art on methodologies for model-based security testing. The topics are related techniques developed in tasks 2.1 – 2.3 in work package 2 that should be supported by methodologies and whose state-of-the-art has not been already covered in deliverable D1.WP2. In particular, Section 2.1, model-driven security, is related to task 2.2 and 2.3. Section 2.2, security testing methodologies based on formal languages, is related to task 2.1. Sections 2.3 and 2.4, network monitoring using machine learning techniques and security testing with GCC extensions, are related to techniques developed in task 2.2. Finally, Section 2.5, model inference assisted fuzzing is related to task 2.1.

2.1 MODEL-DRIVEN SECURITY

Model-Driven Security (MDS) advocates a methodological approach in which (1) security requirements to be formulated and tested at high-levels of abstraction in the early phases of system development, and (2) security analysis results to be maintained by transformations to lower levels of abstraction. As far as we know, the term MDS was first coined in [16]. Most of the papers addressing MDS consider the specification of access control requirements [16][17][18][20][23][27][36][5]. Perhaps the most notable in this area is the work related to SecureUML [5][16][17][18]. SecureUML is an extension of UML for modeling platform independent access control requirements. It is intended that SecureUML be used together with a system design language (e.g., UML class diagrams or UML statecharts). The SecureUML access control requirements can be transformed into platform specific models. In [17], three platforms are considered: Enterprise JavaBeans (EJB), Microsoft Enterprise Services for .NET, and Java Servlets. An advantage of this approach is that these platforms already have access control enforcement mechanisms. However, it is difficult to precisely characterize what it means that a system satisfies the enforcement mechanisms. For instance, although [17] formalizes what it means that a system adheres to an access control model on the platform independent level, they are unable to do so for the platform specific level. No other approach to MDS, that we are aware of, gives a precise description of what it means that a specification is secure (even for the platform independent models). Consequently, these approaches do not exploit the full potential of MDS.
Fernandez-Medina et. al. [23][24][25][26][35][36] show how platform independent access control requirements for a conceptual data base model can be expressed in an extension of UML and OCL. However, only an informal description of how to interpret the security constructs of the extension is given. A transformation from the conceptual data base model and its associated access control requirements to an XML database schema is also discussed without being precisely defined in the papers. Similarly, [20] proposes a platform independent model for access control and discuss issues that must be considered in order to transform this into a platform specific model, but no precise characterization of transformation or adherence is given. Other approaches to MDS that are not specifically related to access-control requirements are presented in [19][29][30].

Heldal and Hultin [30] presents an approach in which UML diagrams can be annotated by security (in the sense of confidentiality) requirements. The work also discusses the possibility of transforming annotated UML diagrams into Java code that can be validated with respect to confidentiality constraints by the language-based checker Jif (Java Information Flow).

Nakamura et. al. [32] present a tool framework for web service security. Three levels of abstraction are considered: the operation level, the execution level, and the deployment level. At the operation level, UML models can be annotated with security primitives such as “integrity” or “confidentiality”. Security requirements at the execution level are assumed to be written in so-called deployment descriptors of J2EE. At the deployment level, security requirements of the execution level are bound to specific security infrastructures. Transformation rules between the abstraction levels are discussed, but no precise characterization of adherence or transformations is given. Haftner et. al. [19][29] define the abstract syntax of a domain specific language for the design of inter-organizational workflows. The language supports various categories of security patterns. A distinction of platform-independent and platform-specific models is made, and a transformation from the PIM to PSM is discussed. The PSM considered is the abstract syntax of the eXtensible Access Control Markup Language (XACML), a standard supporting the specification of authorization policies to access Web services.

Other works that consider security in a model-driven setting (without addressing transformations) are the access-control related works [14][15][21][22][2][31] and the more general high level security requirements works [13][28][37]. Of these, only [13][14][21][2][31] offer a formal foundation which allows the full potential of MDA to be realized.

Abie et. al. [2] integrate a language for specifying high level security requirements – the Security Requirement Language (SRL) – with UML sequence diagrams. SRL is based on first-order logic extended with a small set of modal operators. Thus a formal foundation for precise security analysis is offered. Integration of SRL and UML sequence diagrams is achieved by annotating the sequence diagrams by so-called tagged values that contain SRL macros defining security requirements.
Alghathbar et. al. [14] present authUML, a framework for analyzing access control requirements of UML use cases. The formal foundation of authUML is based on Prolog logic programming rules. AuthUML has three main phases. In the first phase, a set of access control requirements is transformed into so-called access predicates. In the second phase, all accesses predicates are ensured to be consistent, complete, and conflict-free without considering the operations used to describe their functionality. Finally, in the third phase, authUML analyzes the access control requirements in operations.

Doan et. al. [21] integrate access control requirements into UML use case, class, and sequence diagrams. They propose a number of security assurance rules (SAR’s), which can be used to enforce access control requirements for UML. They also define an algorithm with which to check that UML diagrams adhere to the SAR’s.

Koche et. al. [31] present an approach in which access-control requirements are integrated into the software development process. In their approach, access control requirements are specified in the so-called Abstract Security Model (ASM) which is a graph-based security framework that provides a theoretical basis for verifying security constraints. The ASM model may be integrated with UML use cases, class, and sequence diagrams to obtain a so-called concrete security model that can be checked for consistency.

Jurjens [4][2] presents an approach in which UML diagrams can be labeled with confidentiality, integrity, and secure information flow constraints. He also shows that these kinds of requirements are preserved under refinement. The semantics is based on so-called UML machines which are a kind of state machines.

### 2.2 SECURITY TESTING METHODOLOGIES BASED ON FORMAL LANGUAGES

As mentioned in the deliverable WP2.D1, several formal languages are provided in order to express security models, properties and policies. Most of these languages are dedicated to formally define different kind of security aspects that can be applied through diverse contextual security testing architectures. Based on these formalisms, some methodologies have been defined, implemented and some of them are now tooled. We present in this section the security testing methodologies that take as basis some of the languages defined in the deliverable WP2.D1. We will mainly focus on the Timed Extended Finite State Machine (TEFSM) [51], Nomad [54], Or-BAC (Organizational Based Access Control) [52], OCL4ST (OCL for Smart Testing) [55] and SDL (Specification and Description Language) [53].

#### 2.2.1 Security testing methodology with TEFSM and Nomad

The integration of security rules into a TEFSM model describing the behavioral aspects of a system leads to a TEFSM specification that takes the security policy into account: it is called secure functional specification. The integration process is twofold. At first, the algorithm searches for the rules to be applied on each transition of the TEFSM specification.
Then, it adds some states, transitions or updates the guard of the related transition. These modifications depend on the nature of the rule (prohibition, permission or obligation) and its syntax format. To integrate security rules into a TEFSM specification, the following assumption has to be made: the security rules to be integrated are consistent. We assume that the security rules do not contain any incoherent rules [6]. Besides, the TEFSM may be non deterministic if for instance two outgoing transitions have their guard true at the same time. In that case, one of them, chosen randomly, is fired.

According to the Nomad syntax, there are several possible forms for security rules. It would obviously be tedious to deal separately with each of these forms. Consequently, we classify the Nomad security rules in two main classes described as follows:

1. Basic security rules: in this class we consider security rules of the form \( R(\text{start}(A)|\text{O}^{d}\text{done}(B)) \) where \( A \) and \( B \) are actions, \( R \in \{F;O;P\} \) and \( d > 0 \).
2. General security rules: a general security rule denotes any rule that does not fit into the first class. This means that the rule may contain several contextual and/or timed operators and/or logical connectors.

### 2.2.1.1 Formal testing of security rules

The methodology takes the following three elements as input:
- a TEFSM functional description of the system,
- a set of security rules described using Nomad language,
- the existing implementation of the system.

The objective is to check whether the existing implementation verifies the security rules. It proceeds in four steps as shown in Figure 1.

1. The security rules are integrated into the TEFSM specification according to the different algorithms above mentioned.
2. Abstract test cases are automatically generated from the secured TEFSM specification obtained in the first step. We use TestGenIF tool that implements a formal test case generation approach based on the Hit-or-Jump algorithm [8]. This tool accepts a TEFSM specification encoded in the IF textual formalism [9].
3. The abstract test cases are transformed into an executable script capable of communicating via http (or https) with the implementation under test.
4. The concrete test cases obtained from the instantiation are executed on the implementation under test to check whether it verifies the security rules.
2.2.2 Security testing methodology with ocl4st

2.2.2.1 UML4ST

UML4ST (UML for Security Testing) is a subset of the UML notation as presented in [10]. UML offers a set of diagrams to represent the functional and security aspects of a system under test from static and dynamic points of view. This subset UML4ST is composed of three diagrams: class diagram, object diagram and state-transitions diagrams (Figure 2).
The class diagrams are used to represent the data of the system, their types and links. The state-transitions diagram (or state machine) contains on one hand the formal design of the functional behaviors of the SUT and allows on the other hand to specify the dynamism of the security properties to test. In UML4ST, the execution of a transition of a state machine respects the semantic run-to-completion.

An event emitted on the state machine is selected among a pool of events recognized by the entity carrying the state machine. For each execution step, an event is selected among this pool and triggered. The behaviors depending of this trigger are executed. During the execution of this step (eventually composed of ‘sub-steps’), no other pool event can be triggered till its end (or completion). The execution starts with the event triggering in a stable configuration and also ends in a stable configuration. A stable configuration is characterized by the impossibility to trigger a transition of completion with no events. A run-to-completion step can be seen as a complex transition between two stable configurations of the state machine.

The triggering of a transition \( (evt_t, grd_t, eff_t) \) between a state A and a state B depends on:
- The triggering of the event \( evt_t \)
- The respect of the transition \( grd_t \)

The execution order of the behaviors associated to the triggering of the transition \( t \) is given by the sequence:
1. Execution of the output behaviors from state A
2. Execution of the behaviors of the effect \( eff_t \)
3. Execution of the input behaviors in state B.
2.2.2.2 Methodology with OCL4ST

In order to precise the behaviors of the SUT, UML4ST is accompanied by a subset of the OCL language (Object Constraint Language). OCL is a language of constraints expression de facto dedicated to UML (and used by the OMG to define the meta-model UML). OCL allows to formally express the constraints on the UML entities. It allows to specify constraints on the states of an object or a set of objects.

In UML4ST, OCL is used to formally design the behaviors of an SUT. Concretely, a behavior is an action on the system under test, that can be activated in a particular context (i.e. the activation condition of the action). However, OCL is not an action language (or procedural), but a declarative language. It only allows to describe constraints on events composing a UML model. In order to overcome this concern, UML4ST disposes of its own OCL interpretation. The subset OCL supported by UML4ST is called OCL4ST for OCL for Security Testing.

All the object and class diagrams, completed by the state-transitions diagrams used to design the functional behavior of the system under test is called the test model: it represents the SUT behaviors we want to validate. It has the particularity to be executable or interpretable in order to produce executable test sequences (test cases obtained from the test model). Moreover, it contains the test oracle allowing to establish the verdict of each test when executed on the SUT.

Each of the other state-transitions diagrams used to design the security properties to be tested is called secure diagram. A secure diagram contains the abstract execution traces that are interpretable on the corresponding test model. This type of diagram is strongly dependant of the test model. The security scenarios are thus defined allowing to express the specific test objectives which complete the functional objectives automatically computed from the test model.

Figure 3 illustrates the entire generation approach from secure diagrams (red) and indicates its complementarity with the common test sequences generation approach.
The security scenarios are written by the testing engineer expert from the testing requirements (test plan and/or functional and security requirements) and the operation/events of the test model. The basis of the functional test cases generated from the test model is then enriched by the security test cases concretizing the defined security scenarios. A specific report may then be generated to deliver the testing coverage rate.

The security scenarios described by the dedicated state machines are some abstract traces whose the concretization is executable on the system under test. In this methodology, such a trace is composed of an operation sequence and/or events described in the test model.

We may establish a hierarchy of the scenario expressiveness with regard to the abstraction level of the sequence defined by the testing expert. Four expressiveness levels of a scenario are then proposed:

- **Type 1**: completely valuated trace: an operation sequence whose the input parameters are all valuated.
- **Type 2**: trace on an operation sequence with free valuation: a type 1 trace in which the operation parameters can also be free.
• **Type 3**: trace on partial operation sequence: a type 2 trace whose the operation sequence can be parameterized (e.g. in specifying the optional presence of an operation in the trace).

• **Type 4**: trace constrained by regular expression: a type 3 trace augmented by constraints (guards) on states reached by the generated test(s). The constraints are expressed in OCL and deal with the state variable of the system.

### 2.2.3 From the Or-BAC security rules and the EFSM-based SUT

To achieve its active testing methodology, security experts need to edit the set of security rules that system under test (SUT) has to respect. These rules can be specified in Or-BAC or Nomad models and the SUT is also specified from its functional point of view based on the formal language SDL (or IF2.0).

#### 2.2.3.1 From SDL representing the SUT

To describe a communicating system based on the SDL language, several tools are available. But, whatever the tool chosen by the testing expert in order to edit, verify and test the SUT, the methodologies rather follow the same steps. The communicating system is specified by means of states and transitions (based on the Extended Finite State Machine formalism). One can also verify the specification syntactically and semantically. The syntactic analysis ensures that the specification complies with the syntactic rules of SDL, whereas the semantic verification ensures the consistency of the specification. This step is carried out not only by static analysis but also by an automatic exhaustive exploration of the specification. This is performed by testing all possible ways of system execution, with a certain number of rules and the cases of violations such as deadlocks, loops etc. During verification, the main properties analyzed are:

- Safety (absence of deadlock, unspecified reception, blocking cycles, etc). Deadlock takes place when a state of the system, reachable from the initial state cannot trigger a transition anymore.
- Promptness (livelock). A state is known as alive if it can be reached starting from all the states of the global system.

#### 2.2.3.2 From Or-BAC

The Or-BAC model introduced in the deliverable WP2.D1 has been integrated in a testing security policies methodology. Besides, an associated tool called MotOrBAC [1] has been developed to help designing and implementing security policies using the Or-BAC formalism. First, the security policies are designed, uploaded and stored. They are later simulated to verify their consistencies. While the Nomad-based security rules are currently edited textually (no specific tool exists to perform these tasks yet), the contexts evaluation through the security policies are managed through APIs. The following steps are then followed.

---


© Copyright DIAMONDS Consortium
• Create/edit the Or-BAC security policies (XML RDF files can be performed). The editorial operations must take into account the administration policy associated if this latter is activated.
• Contexts evaluation. The contexts defined in the Or-BAC model are evaluated, which are the temporal contexts declared by the user and the contexts expressing a condition on the attributes of the concrete entities. The composition of the contexts may also be processed if required.
• Checking of the concrete security policy and its administration policy.

2.3 NETWORK MONITORING USING MACHINE LEARNING TECHNIQUES
Detecting network anomalies and intrusions in a network environment is possible with the help of machine learning techniques. However, the network environment should be stable and somewhat isolated for the best efficiency. Networks in industrial control systems meet these requirements. The main method of our study is to exploit the traffic captures from a real network. The method we are using is machine learning combined with passive monitoring and a priori knowledge of protocols used. Passive monitoring is required due to the nature of the environment. It is important that no measuring device or monitoring system interferes with the ICS environment under scrutiny.

![Figure ICS Environment](image-url)
2.3.1 Traffic analysis

The main method of the study is to exploit the traffic captures available to us. In the initial testing and feasibility studies we have so far used tools such as NetAI, NetMate and WEKA 3. NetAI and NetMate are used as provided by the NetMate-flowcalc bundle [56], [57], [58], [59].

2.3.2 The machine learning approach

We argue that the factory networks for ICS that are functioning properly and without serious design flaws can be defined as nearly closed environments. When the factory network would be in a normal status without serious incidents there is typically very little noise. Again, if the network architecture is well defined and implemented, most of the network traffic on the ICS level of the network should be more deterministic than that of open networks such as the office networks.

The most challenging aspect in the initial phase of using machine learning is feature extraction and selection. We have been identifying possibilities which include usual features used by many IDS’s. Properly done feature extraction is one of the most important steps in machine learning, producing a classifier with higher generalization capability by excluding redundant or irrelevant attributes. The feature selection process benefits from extensive testing of the recorded live data.

The usage of payload form and payload data has some significant challenges, subject to future work. However, they could be used to very accurately monitor the sanity of the system and conformance to security policies. Some features studied for possible use include: throughput, IP address and port pairs in a flow, average size of the packets, timing, flow direction, Average duration of flow between endpoints, Payload form, payload data, MAC to IP mapping, networking protocol, protocol settings and connectivity number.

For the model of the ICS network environment all attributes would have to take into account the possible periodic nature of the traffic. Depending on the system being monitored, there might be variations caused by maintenance, periodic processes or environmental fluctuations.

The main argument for using machine learning approaches is the more closed nature of ICS networks. Any benign changes in the traffic are likely caused by actions that can be informed to the system a priori. Malign or anomalous activity, on the other hand, is any new form of traffic that has not been taught or programmed to the monitoring system. We are planning to develop a proof of concept of the traffic analysis and machine learning functionality on top of our existing system. The system already implements support for a vector machine and data pre-processing. We plan to implement new algorithms when needed, or use open source implementations when possible. To be able to implement the planned system, a number of steps must be taken before the implementation can be started. For the next few steps, a study will be completed regarding the specific algorithms available and their strengths and weaknesses in ICS network environment. In addition to
this, a deeper analysis of the different aspects of the network layers present in ICS environments must be accomplished.

### 2.4 SECURITY TESTING WITH GCC EXTENSIONS

GNU Compiler Collection [60] is a very widely used compiler. GCC is open source and it compiles C, C++, Java, Fortran, Ada and Objective-C. GCC versions 4.5 and later support plug-in modules. Plug-ins have full control over the program’s presentation and utilize all the information extracted and generated by the earlier compiler passes. Plug-ins make possible that the program analysis and instrumentation can be added and removed relatively easily. We use plugins to extract information about the program structure and execution. VTT has built a set of libraries that can be used to build external tools that utilize this information.

![CFG and Trace-plugins](image)

**CFG (Control Flow Graph) plugin** is used to instrument the program to output representation of its global control flow graph and some additional link time information when it is executed. The GFC contains more information than what is generally available at compile time. CFG plugin offers much lower level and easier to use representation than the original source code.

**Trace** is a sequence of function calls, returns and visited basic blocks that describes the program’s execution path. This information can be combined with the CFG to reconstruct a specific instance of program’s execution.

The plugin-modules can be used in various ways:
• Automatic test case generation - searching test case inputs based on execution traces.
• Automatic test case selection – based on changes in the program structure and test case’s execution trace.
• Model checking – e.g. correct use of API calls
• Fault detection – Static or dynamic checks to detect errors early
• Fault injection – for generating “hard to detect”-failures for testing purposes
• Program analysis – simplify debugging and gaining information about the program structure.

2.5 MODEL INTERFERENCE ASSISTED FUZZING

One example of model based security testing is fuzzing (or fuzz testing), which exposes tested software with malformed input. The aim is to evaluate if weaknesses in the software can be found, for example denial-of-service conditions, which can be further exploited to a full compromise of the software. Fuzz testing is negative testing – we do not aim to demonstrate presence of functionality, but would like to see about the absence of specific vulnerabilities.

Miller et al. have shown in their work that even very simplistically fuzzing models are able to disclose a large amount of parsing errors [38][39]. To systematize work, the Oulu University Secure Programming group had started in 1999 the PROTOS project [40] and has developed an approach to systematically test implementations of protocols in a black-box fashion. Several successful test suites have been released and this research led to a spin-off company: Codenomicon Ltd.

Our previous work in robustness testing of protocol implementations has shown that manually designed structural mutations and exceptional element values are an efficient way to expose errors in software. Unfortunately, while powerful, manual test design has some bottlenecks: i) it requires some kind of format specification as a basis; and, ii) poorly documented formats must be reverse-engineered before test designers can write a model-based test suite for the format. The human factor also brings in the danger of tunnel vision, as the power of manually designed cases is largely dependent on the expertise and imagination of the designer. On the other hand, blind random fuzzing has a considerably lower entry barrier, but is hindered by the impossibility of efficiently addressing a virtually infinite input space in finite time.

In a subsequent direction, the aim was to automatize the test case generation. Radamsa [41] is an example of this kind of black-box fuzzing tool. As an input it requires valid data stream samples, either in file format or as network data. Radamsa infers a model from this data and then generates – based on predefined heuristics – test cases. As these test cases are not required to trigger predefined situations, Radamsa has leeway in generating such model inference.
The results have been promising: more than 28 findings are reported in the Common Vulnerabilities and Exposures (CVE), a database of computer-security vulnerabilities. Open and closed source software has been affected.
3. STATE-OF-THE-ART ON METHODOLOGY FOR RISK-BASED SECURITY TESTING

The second topic, risk-based security testing, is related to task 4.2 of work package 4, and is in this document presented in Sections 3.1-3.2. These sections are related to task 2.4 of work package 2. Section 3.1 presents standards supporting risk-based security analysis while Section 3.2 presents a literature survey of risk-based testing.

3.1 STANDARDS SUPPORTING RISK-BASED SECURITY ANALYSIS

The evaluation and certification of a product is the advantageous method of attesting its quality through a standardized process. Since the certification is generally carried out by an independent certification body, the customer is now provided with a meaningful testament of the products quality. At the same time the developer benefits from the certification process by increasing their market potential and gaining additional confidence in the product due to the expertise from evaluation and certification labs.

The Common Criteria for Information Technology Security Evaluation is an international standard (ISO/IEC 15408) for the certification of IT-security products. The process conducted by the certification body is internationally agreed upon and the certificates internationally accepted. The evaluation process is highly driven by developer documentation and focuses on product development, security testing and vulnerability assessment. Besides creating trust in the product’s quality, the evaluation results also allow the customer to compare the security functionality of similar products.

The ETSI TVRA method [12] developed by ETSI benefit from the CC work, e.g. by using a generic catalogue of security functional requirements (SFRs). TVRA is characterized by the following concepts and approaches:

- Threat types: Interception, manipulation, Denial of Service, repudiation of sending, repudiation of receiving
- Security objective types: Confidentiality, integrity, availability, authenticity, accountability
- Attack potential is defined by the openness of a system to attack, attackers expertise and resources, systems availability.
- UML is used to model relationships within systems.
- Methods to analyze/evaluate system security including: threats, risks, vulnerabilities
- Calculation of attack potential
- TVRA uses CC taxonomy (family – class – component): Security requirements taxonomy (SFR)
- Lacks on details about generation of security tests.

The following table provides a comparison between CC and TVRA. The most important items have been marked with yellow.

**Table 1: Comparison between CC and TVRA.**

<table>
<thead>
<tr>
<th>Field of application</th>
<th>CC: IT-Security products</th>
<th>TVRA: Telecommunications system</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong></td>
<td>- To answer the question if the TOE fulfills certain security needs.</td>
<td>Determine how open to attack a system is.</td>
</tr>
<tr>
<td></td>
<td>- Making two TOEs comparable.</td>
<td></td>
</tr>
<tr>
<td><strong>Focus Point</strong></td>
<td>Resistance to attack of the system</td>
<td>Impact of an attack on the system</td>
</tr>
<tr>
<td><strong>Target of Evaluation</strong></td>
<td>IT-Product</td>
<td>System under Standardization</td>
</tr>
<tr>
<td><strong>Output</strong></td>
<td>IT-Security certificate</td>
<td>Quantified measure of the risks to the assets and a set of detailed security requirements that will minimize the risks</td>
</tr>
<tr>
<td><strong>EAL</strong></td>
<td>Evaluation is based on a single evaluation level</td>
<td>Evaluation level can be expressed as a range: EAL3 – EAL5</td>
</tr>
<tr>
<td><strong>Terminology</strong></td>
<td>- Objectives must or shall</td>
<td>should</td>
</tr>
<tr>
<td></td>
<td>- Countermeasures IT countermeasures (firewalls, smart cards) and non-IT countermeasures (guards and procedures)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>only technical security countermeasures are considered</td>
<td></td>
</tr>
</tbody>
</table>

CC and TVRA are both lacking on details on how to derive security tests from the TOE description. The following illustration presents the relationship between the SUT/TOE, its requirements, interfaces and security tests.
The security analysis as part of the Common Criteria evaluation is performed during the vulnerability assessment aspect AVA and is closely linked to the aspects development (ADV) and guidance documents (AGD). The analysis is performed by the external evaluation body and supported by testing. Its goal is to determine whether exploitable flaws and weaknesses exist in the system, i.e. if the target of evaluation is resistant to penetration attacks. Publicly available information about known weakness for the specific product and product type serve as the basis for the analysis.

Following the principal of evaluation assurance levels (EAL) with an increasing evaluation depth, the vulnerability analysis during a particular evaluation only considers flaws and weaknesses that can be exploited with the attack potential relevant for the assurance level chosen for this evaluation. These different levels are shown in Figure 4 - Attack Potential [11]

<table>
<thead>
<tr>
<th>Vulnerability Component</th>
<th>TOE resistant to attacker with attack potential of:</th>
<th>Residual vulnerabilities only exploitable by attacker with attack potential of:</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAN.5</td>
<td>High</td>
<td>Beyond High</td>
</tr>
<tr>
<td>VAN.4</td>
<td>Moderate</td>
<td>High</td>
</tr>
<tr>
<td>VAN.3</td>
<td>Enhanced-Basic</td>
<td>Moderate</td>
</tr>
<tr>
<td>VAN.2</td>
<td>Basic</td>
<td>Enhanced-Basic</td>
</tr>
<tr>
<td>VAN.1</td>
<td>Basic</td>
<td>Enhanced-Basic</td>
</tr>
</tbody>
</table>

The attack potential score is calculated based on
- the time it takes to identify and exploit a vulnerability,
- the required expertise of the attacker,
- the needed knowledge of the systems design and operation,
- the needed window of opportunity as well as
- hard- and software required for the exploitation.

Figure 5 is taken from the Common Methodology for Information Technology Security Evaluation and associates those five factors with numeric values to allow for the calculation of a total score for the evaluation of the products resistance to vulnerabilities.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elapsed Time</td>
<td></td>
</tr>
<tr>
<td>&lt;= one day</td>
<td>0</td>
</tr>
<tr>
<td>&lt;= one week</td>
<td>1</td>
</tr>
<tr>
<td>&lt;= two weeks</td>
<td>2</td>
</tr>
<tr>
<td>&lt;= one month</td>
<td>4</td>
</tr>
<tr>
<td>&lt;= two months</td>
<td>7</td>
</tr>
<tr>
<td>&lt;= three months</td>
<td>10</td>
</tr>
<tr>
<td>&lt;= four months</td>
<td>13</td>
</tr>
<tr>
<td>&lt;= five months</td>
<td>15</td>
</tr>
<tr>
<td>&lt;= six months</td>
<td>17</td>
</tr>
<tr>
<td>&gt; six months</td>
<td>19</td>
</tr>
<tr>
<td>Expertise</td>
<td></td>
</tr>
<tr>
<td>Layman</td>
<td>0</td>
</tr>
<tr>
<td>Proficient</td>
<td>3</td>
</tr>
<tr>
<td>Expert</td>
<td>6</td>
</tr>
<tr>
<td>Multiple experts</td>
<td>8</td>
</tr>
<tr>
<td>Knowledge of TOE</td>
<td></td>
</tr>
<tr>
<td>Public</td>
<td>0</td>
</tr>
<tr>
<td>Restricted</td>
<td>3</td>
</tr>
<tr>
<td>Sensitive</td>
<td>7</td>
</tr>
<tr>
<td>Critical</td>
<td>11</td>
</tr>
<tr>
<td>Window of Opportunity</td>
<td></td>
</tr>
<tr>
<td>Unnecessary / unlimited access</td>
<td>0</td>
</tr>
<tr>
<td>Easy</td>
<td>1</td>
</tr>
<tr>
<td>Moderate</td>
<td>4</td>
</tr>
<tr>
<td>Difficult</td>
<td>10</td>
</tr>
<tr>
<td>None</td>
<td>4</td>
</tr>
<tr>
<td>Equipment</td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>0</td>
</tr>
<tr>
<td>Specialised</td>
<td>4</td>
</tr>
<tr>
<td>Bespoke</td>
<td>7</td>
</tr>
<tr>
<td>Multiple bespoke</td>
<td>9</td>
</tr>
</tbody>
</table>

Figure 5 - Calculation of attack potential [11]
3.2 A LITERATURE SURVEY OF RISK-BASED TESTING

In this section, we review state-of-the-art approaches to risk-based testing. Each approach is classified in Table 1. This table has the following categories:

- **Security specific:** Indicates whether or not the approach is specifically aimed at security.
- **Black box/white box:** Indicates whether the approach is based on black-box or white-box testing (white-box testing is typically used under system development).
- **Degree of structure:** Indicates how structured the approach is. Three values are possible: low, medium, and high.
- **System specification language:** Indicates whether the approach is supported by any particular system specification languages.
- **Risk specification language:** Indicates whether the approach is supported by any particular risk analysis languages.
- **Test specification language:** Indicates whether the approach is supported by any particular test specification languages.
- **Tool support:** Indicates whether the approach is supported by a tool.
- **System to risk:** Indicates whether the approach is based on starting with the risk and identifying system parts that needs to be tested, or based on starting with the systems parts and indentifying risks for those.
- **Test case derivation:** Indicates whether or not the approach describes how test cases can be derived from a risk model.

<table>
<thead>
<tr>
<th>Method/attributes</th>
<th>Security specific</th>
<th>Black-box/white box</th>
<th>Degree of structure</th>
<th>System spec. language</th>
<th>Risk spec. language</th>
<th>Test spec. language</th>
<th>Tool support</th>
<th>System to risk</th>
<th>Test case derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prisma</td>
<td>No</td>
<td>Both</td>
<td>Medium</td>
<td>No</td>
<td>Tables</td>
<td>No</td>
<td>Yes (for prioritizing)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Redmill</td>
<td>No</td>
<td>White box/under development</td>
<td>Low</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Amland</td>
<td>No</td>
<td>White box/under development</td>
<td>Medium</td>
<td>No</td>
<td>Tables</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Bach</td>
<td>No</td>
<td>Both</td>
<td>Low</td>
<td>No</td>
<td>Tables</td>
<td>No</td>
<td>No</td>
<td>Both</td>
<td>No</td>
</tr>
<tr>
<td>RiteDAP</td>
<td>No</td>
<td>White box/under development</td>
<td>Low (for risk analysis part)</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

© Copyright DIAMONDS Consortium
Based on the state-of-the-art analysis, we conclude that risk-based testing is still an immature research area with a many opportunities for further development. In particular,

- Most approaches to risk-based testing target functional testing, only 2 out of 8 consider security.
- Few approaches to risk-based testing are sufficiently documented and structured to serve as an easy to use industrial-setting process. For instance, many of the approaches read more like a general discussion on the topic risk-based testing than a detailed description of steps involved in carrying out the process/approach.
- Only 1 out of 8 approaches are supported by tools.
- Few of the approaches are supported by any particular languages for specifying systems, tests, or risk-analysis results.

In the following we give a summary of the approaches.

3.2.1 The Prisma method
The Prisma method is presented in a white paper [48]. The method suggests prioritizing based on the most important areas of a product, and the part of the product which is likely to have the most defects. Areas with most defects are likely to be: complex areas (>200 complexity measures exist), changed areas, areas with new technology or methods, areas developed by inexperienced people involved, areas developed under time pressure and high defect history.

The prisma method has 5 steps:
- **Planning**: Gather input documents, identify risk items (i.e. parts of the systems that may impact risks) through interview and reading documents, determine impact and likelihood factors, define a weight for each factor, select stakeholders.
- **Individual preparation**: Each participant assigns values to factors per risk item. The participants score by selecting (the description of) the value that fits (supported by Excel sheets)
- **Gather individual scores**: Do an initial check to see of the scores are OK, then process individual scores
- **Define a differentiated test approach**: Prioritize risk items than need to be tested based on location in risk matrix.
3.2.2 Redmill’s approach

Redmill has written two papers on risk-based testing. In the first paper [43], Redmill argues that risk-based testing is not consistently defined nor supported by literature on either theory or practice, and proceeds with an informal discussion of risk-based testing. In his second paper [44], Redmill presents three approaches to risk-based testing

- **Single-factor analysis: Consequence.** This approach is based on estimating consequence for risks values only (not likelihood). The steps of the approach are:
  - Consequence identification: For each service and each stakeholder, determine consequence of failure.
  - Consequence analysis 1: Use Hazop guidewords to distinguish between the different ways a service may fail.
  - Consequence analysis 2: Distinguish between services, the functions that provide the services, and the software items of which the functions are composed. Find the relationship between these types of entities to determine the relationship between services and software items. Identify potential causal links between software items and various types of service failure.
  - Assessment on the basis of consequence: Determine how the consequence value is translated into test plans for software items. Assign a category to each software item and then use this category to define an appropriate test programme.

- **Single-factor analysis: Probability.** This approach is based on estimating the likelihood of risks. The steps of the approach are:
  - Relevant factors and their attributes: For each software item and each relevant quality attribute (such as complexity, structure, comments), assign a quality level.
  - Using the information: Each software item’s quality factor is used to inform decisions on what risk-management action to take. One approach is for quality factors to be equated with test programmes.

- **Two-factor analysis.** This approach is based on the combination of the previously mentioned approaches in which software items are placed in a matrix of integrity level and quality factors
  - Combine the two previous mentioned approaches and place software items in a matrix of integrity levels and quality factors.

3.2.3 Amland – Risk-based testing

Åmland describes a six step process which has been applied to a financial application case study [50]. The steps of the method are the following:
• Planning (risk identification/risk strategy): Define the test item tree, i.e. a hierarchical breakdown of the functions and features in the system to be tested. Establish test plan and overall risk strategy.

• Identify risk indicators (risk assessment): Define a set of indicators (e.g. size of function, number of changes since previous release, complexity) that can be used to assess the probability of failure of a function. Then for each function, and each indicator, assign an indicator value. Then use this information to calculate the probability of failure for each function.

• Identify cost of fault (risk assessment): For each function, estimate the consequence of failure.

• Identify risk elements (risk assessment): Calculate a risk expose for each function based on estimated probability and consequence of failure. Use this to prioritize the functions.

• Test execution (risk mitigation): Start testing based on the prioritized list of functions.

3.2.4 Back – Heuristic Risk-Based Testing
Bach defines risk based testing as the following process [42]:

- Make a prioritized list of risks.
- Perform testing that explores each risk.
- As risks evaporate and new one emerge, adjust your test effort to stay focused on the current crop.

Back then proceeds by presenting two different approaches to risk-based testing. One approach is based on starting with the system and then identifying the risks based on an identification of find vulnerabilities, threats, and victims. The other approach is based on starting with the risks to which parts of the system they apply to. Bach suggest that the latter approach could be aided by three kinds of lists: A list of quality criteria categories, a generic risk list, and a risk catalog (containing risks that belong to particular domains)

3.2.5 RiteDAP
RiteDAP [46] is an approach to risk-based testing that allows for the automatic derivation of system test cases from activity diagrams as well as their prioritization based on risk. The RiteDAP process has the following steps:

- Specify activity diagrams which can be used as test models.
- Annotate the activity diagrams with risks.
- Derive a set of unordered test case scenarios form the test model.
- Order the test scenarios based on the risk information in the test model.

The last two steps can be automated.
3.2.6 Rosenberg et.al. Risk-based Object Oriented Testing
Rosenberg et. al. [45] proposes measurement criteria that can be used to estimate the complexity of object-oriented code. It is suggested that a high-level of complexity could imply a high likelihood of failure. Consequence analysis is not considered.

3.2.7 Murthy et al. – Leveraging Risk Based Testing in Enterprise Systems Security Validation
Murphy et. al. [47] proposes an iterative process to risk based testing that consists of the activities:
- Risk analysis: In this activity, risks are identified and evaluated according to the Risk Analysis method proposed by NIST. The output of the activity is a documented list of high level risks.
- Threat modeling: In this activity, risks are detailed by modeling threats to the system under evaluation according to Microsoft’s Threat Modeling process. The system is then decomposed into sub-systems to identify assets, entry points, and trust levels. The output of the activity is a detailed list of categorized risks.
- Test design: In this activity, misuse cases are used to identify security test scenarios. Also, any Security Controls suggested as part of the application design will automatically translate into a test scenario.
- Test execution: In this activity, test scenarios are translated into more detailed test cases which are categorized an prioritized according to the risk of categorized risks.
- Reporting: In this activity, the output of the test execution is captured that details the vulnerabilities found along with their severity level.

3.2.8 Zech – Risk-Bases Security Testing in Cloud Computing Environments
Zach proposes a model-driven methodology for the security testing of cloud environments [49]. The main steps of the method are:
- Step 1: Perform a risk analysis of the Cloud Under Test (CUT), possibly with the help of a vulnerability repository.
- Step 2: Transform the risk model (generated in step I) into a set of negative requirements (more or less textual descriptions) using a model to model (M2M) transformation.
- Step 3: Transform the negative requirements into misuse cases using an M2M transformation.
- Step 4: Automatically transform the misuse cases into test cases using an M2M transformation.

The transformations mentioned in the steps are not presented in detail; it merely suggests that such transformations could be used.
4. CONCLUSIONS

In this document, we have presented state-of-the-art related to methodologies for risk- and model-based security testing. The state-of-the-art is related to techniques developed in WP2 that should be supported by methodologies and whose state-of-the-art survey has not already been covered by deliverable D1.WP2.

5. REFERENCES


[7] Amel Mammar, Wissam Mallouli, and Ana Cavalli. A systematic approach to integrate common timed security rules within a TEFSM-based system specification. Published in Information and Software Technology Journal. ISSN = 0950-5849, Elsevier Editor, August 2011


[12] TVRA ETSI TS 102 165-1 v4.2.3 (2011-03) TISPAN Methods and Protocols; Part 1 Meth-


Review of security testing tools

Deliverable ID: D


[54] F.Cuppens, N.Cuppens, T.Sans, Nomad: A Security Model with Non Atomic Actions and


