Supporting Document
Mandatory Technical Document

Application of Attack Potential to Smartcards

March 2009

Version 2.7
Revision 1

CCDB-2009-03-001

Foreword

This is a supporting document, intended to complement the Common Criteria version 3 and the associated Common Evaluation Methodology for Information Technology Security Evaluation.

Supporting documents may be “Guidance Documents”, that highlight specific approaches and application of the standard to areas where no mutual recognition of its application is required, and as such, are not of normative nature, or “Mandatory Technical Documents”, whose application is mandatory for evaluations whose scope is covered by that of the supporting document. The usage of the latter class is not only mandatory, but certificates issued as a result of their application are recognized under the CCRA.

Technical Editor: BSI

Document History:
V2.7 March 2009 (technical update of rating categories and update on usage of open samples based upon corresponding JIL document version 2.7)
V2.5 December 2007 (explicit statements added that the points for identification and exploitation have to be added at the end to achieve the final attack potential value, references updated)
V2.3 April 2007 (evaluation time guideline and rules regarding the use of open samples added and updated for use with both CC version 2 and 3)
V2.1 April 2006 (classification as mandatory technical document, several updates to the tables)
V1.1, July 2002 (draft indicator deleted, references updated, same content as V1.0)

General purpose:

The security properties of both hardware and software products can be certified in accordance with CC. To have a common understanding and to ensure that CC is used for hardware integrated circuits in a manner consistent with today’s state of the art hardware evaluations, the following chapters provide guidance on the individual aspects of the CC assurance work packages in addition to the Common Evaluation Methodology [CEM].

Field of special use: Smart cards and similar devices

Acknowledgments:

The governmental organisations listed below and organised within the Joint Interpretation Working Group contributed to the development of this version of this Common Criteria Supporting document:

France: *Direction Centrale de la Sécurité des Systèmes d'Information*
Germany: *Bundesamt für Sicherheit in der Informationstechnik*
Netherlands: *Netherlands National Communications Security Agency*
Spain: *Ministerio de Administraciones Públicas and Centro Criptológico Nacional*
United Kingdom: *Communications-Electronics Security Group(CESG)*

They also acknowledge the contribution of the work done by several smart card vendors, evaluation labs, and other companies organised within:

- *International Security Certification Initiative (ISCI)*
- *JIL Hardware Attacks Subgroup (JHAS)*

Table of contents

1 Introduction

1 This document interprets the current version of Common Criteria Methodology [CEM] (annex A.8 for CC v2, annex B.4 for CC v3). This work has been based on smartcard CC evaluation experience and input from smartcard industry through the International Security Certification Initiative (ISCI) and the JIL Hardware Attacks Subgroup (JHAS).

2 This chapter provides guidance metrics to calculate the attack potential required by an attacker to effect an attack. The underlying objective is to aid in expressing the total effort required to mount a successful attack. This should be applied to the operational behaviour of a smartcard and not to applications specific only to hardware or software.

3 This document is compatible with CC v2 and CC v3 [CC].

2 Scope

4 This document introduces the notion of an attack path comprised of one to many attack steps. Analysis and tests need to be carried out for each attack step on an attack path for a vulnerability to be realised. Where cryptography is involved, the Certification Body should be consulted.

3 Identification of Factors

5 Note about CC v3.1 :

6 With Common Criteria version 3.1, there is no more distinction between the identification phase and the exploitation phase but within the smartcard community, the risk management performed by the user of CC certificates clearly required to have a distinction between the cost of “identification” (definition of the attack) and the cost of “exploitation” (e.g. once a script is published on the World Wide Web). Therefore, this distinction is kept when calculating the attack potential for smartcard evaluations. Although the distinction between identification and exploitation is essential for the smartcard evaluation to understand and document the attack path, the final sum of attack potential is calculated by adding the points of these two phases, as both phases together constitute the complete attack.

3.1 How to compute an attack

7 Attack path identification as well as exploitation analysis and tests are mapped to relevant factors: elapsed time, expertise, knowledge of the TOE, access to the TOE, equipment needed to carry out an attack, as well as whether or not open samples or samples with known secrets had been used. Even if the attack consists of several steps, identification and exploitation need only be computed for the entire attack path.

8 The identification part of an attack corresponds to the effort required to create the attack, and to demonstrate that it can be successfully applied to the TOE (including setting up or building any necessary test equipment). The demonstration that the attack can be successfully applied needs to consider any difficulties in expanding a result shown in the laboratory to create a useful attack. For example, where an experiment

reveals some bits or bytes of a confidential data item (such as a key or PIN), it is necessary to consider how the remainder of the data item would be obtained (in this example some bits might be measured directly by further experiments, while others might be found by a different technique such as an exhaustive search). It may not be necessary to carry out all of the experiments to identify the full attack, provided it is clear that the attack actually proves that access has been gained to a TOE asset, and that the complete attack could realistically be carried out. One of the outputs from Identification is assumed to be a script that gives a step-by-step description of how to carry out the attack – this script is assumed to be used in the exploitation part.

9 Sometimes the identification phase will involve the development of a new type of attack (possibly involving the creation of new equipment) which can subsequently be applied to other TOEs. In such a case the question arises as to how to treat the elapsed time and other parameters when the attack is reapplied. The interpretation taken in this document is that the development time (and, if relevant, expertise) for identification will include the development time for the initial creation of the attack until a point determined by the relevant Certification Body. Once a Certification Body has determined this point, no points for the development of the attack (in terms of time or expertise) will be used in the attack potential calculation any more.

10 The exploitation part of an attack corresponds to achieving the attack on another instance of the TOE using the analysis and techniques defined in the identification part of an attack. It is assumed that a different attacker carries out the exploitation, but that the technique (and relevant background information) is available for the exploitation in the form of a script or a set of instructions defined during the identification of the attack. The script is assumed to identify the necessary equipment and, for example, mathematical techniques used in the analysis. This means that the elapsed time, expertise and TOE knowledge ratings for exploitation will sometimes be lower for exploitation than for identification. For example, it is assumed that the script identifies such things as the timing required for a perturbation attack, and hence in the exploitation phase the attacker does not have to spend significant time to find the correct point at which to apply the perturbation. Furthermore, this same information may also reduce the exploitation requirement to one of mere time measurement, whereas the identification phase may have required reverse engineering of hardware or software information from power data – hence the expertise requirement may be reduced. Similarly, knowledge about the application that was used to achieve the timing of an attack may also be included either directly in the script or indirectly (through data on the timing required). As a general rule, no points can be awarded for the exploitation phase at all when, e.g., a secret master key common to all TOEs under investigation has been compromised in the identification phase. This is so as the script defining details to be passed on between the identification and exploitation phase will already contain the information on this master key. An example would be storing a master key in ROM.

11 In many cases, the evaluators will estimate the parameters for the exploitation phase, rather than carry out the full exploitation. The estimates and their rationale will be documented in the ETR.

12 To complete an attack potential calculation the points for identification and exploitation have to be added as both phases together constitute the complete attack. When presenting the attack potential calculation in the ETR, the evaluators will make an argument for the appropriateness of the parameter values used, and will therefore give the developer a chance to challenge the calculation before certification. The final attack potential result will therefore be based on discussions between the developer, the ITSEF and the CB, with the CB making the final decision if agreement cannot be reached.

13 No rigid rules can be given on how much time should be spent on a typical smartcard VLA.4 / VAN.5 evaluation by a competent lab, but the following guidance shall none-the-less be provided in an effort to harmonise evaluations and the various national schemes alike: Assuming the CC vulnerability analysis has already been performed the evaluation testing from scratch for a new IC should take about 3 man months, depending on the complexity of the IC such as the number of cryptographic services, interfaces, etc. The total evaluation time for composite evaluations using a certified IC for VLA.4 / VAN.5 testing activities is of the order of 1-3 man months, depending on the complexity of the platform, such as open platform, native platform, number of APIs, etc.. It is possible to deviate from this guidance, but some reasoning will have to be provided to the CB.

14 It is an assumption of this interpretation that the Certification Bodies will ensure that there is harmonisation not only nationally, but also between national schemes. This is required, for example, where new types of attack are applied and a decision has to be taken as to when the attack is considered 'mature', at which point it will no longer gain points for the time or expertise to develop the attack (as discussed above).

3.2 Elapsed Time

15 Compared to the "Elapsed Time" factor as given in CEM, further granularity is introduced for smartcards. In particular, a distinction is drawn between one week and several weeks. The Elapsed Time is now divided into the following intervals:

	Identification	Exploitation
< one hour	0	0
< one day	1	3
< one week	2	4
< one month	3	6
> one month	5	8
Not practical	*	*

Table 1: Rating for Elapsed Time

16 The CEM defines the term *Not Practical* as "the attack path is not exploitable within a timescale that would be useful to an attacker".

17 In practice an evaluator is unlikely to spend more than 3 months attacking the TOE. At the end of the evaluation the evaluator has to assess the time it would take to carry out the minimum attack path. This computes the estimated time to mount the attack, and not necessarily the time spent by the evaluator to conduct the attack.

18 Where the attack builds on the findings of a previous evaluation, Elapsed Time as well as Expertise have to be taken into account, e.g., a particular attack may have been developed on a smartcard product similar to the TOE. It is not possible to give general guidance here.

19 The question of "Not Practical" may depend on the specific attack scenario as the following two examples show:

(a) Consider a smartcard used for an online system, where the card contains only individual keys and assume further that these keys are deactivated in the system within days after loss of a card was reported. In this case an attack is not even practical for an attacker if he can extract the keys in one week.

(b) Consider a smartcard, which contains system-wide keys, which might be used for fraud even if use of the individual card is blocked after loss. In this case an attack may be successful for the attacker even if it takes a year.

20 So if a general assumption on a time for "Not Practical" is needed, something about 3-5 years is a better worst-case oriented time frame. (This is the time after which a card generation is normally exchanged and system wide keys may be changed in a comparable time frame). However, the best rule seems to be to decide on the meaning of "Not practical" only in a specific attack scenario.

3.3 Expertise

21 For the purpose of smartcards two types of experts are defined:

- an expert with the ability to define new attacks for smartcards (hardware, software, cryptography) and the necessary tools, and
- an expert with a level of knowledge of the TOE commensurate to that of the developer (e.g. knowledge of product standards and specifications).

22 The expertise necessary to carry out an attack may cover several disciplines: chemical, ability to drive sophisticated tools, cryptographic.

	Definition according to CEM	**Detailed definition to be used in smartcard evaluations**
a) Experts	Familiar with implemented • Algorithms • Protocols • Hardware structures • Principles and concepts of security	Familiar with • Developers knowledge namely algorithms, protocols, hardware structures, principles and concepts of security and • Techniques and tools for the definition of new attacks

	Definition according to CEM	Detailed definition to be used in smartcard evaluations
b) Proficient	Familiar with • security behaviour	Familiar with • security behaviour, classical attacks
c) Laymen	No particular expertise	No particular expertise

Table 2: Definition of Expertise

Extent of expertise **(in order of spread of equipment or smartcard related knowledge)**	
Equipment: The level of expertise depends on the degree to which tools require experience to drive them • Oscilloscope • Optical Microscope • Chemistry (etching, grinding), Microprober • Laser Cutter, Radiation • Plasma (etching, grinding), Focused Ion Beam (FIB) • Scanning Electron Microscope (SEM) • Atomic Force Microscope (AFM)	**Knowledge:** The level of expertise depends on knowledge of • Common Product information • Common Algorithms, Protocols • Common Cryptography • Differential Power Analysis (DPA), Differential Fault Analysis (DFA), Electromagnetic Analysis (D/EMA) • Reverse Engineering • Smartcard specific hardware structures • Principles and concepts of security • Developers knowledge

Table 3: Extent of expertise

23 It may occur that for sophisticated attacks, several types of expertise are required. In such cases, the highest of the different expertise factors is chosen.

24 A new level “Multiple Expert” was introduced to allow for a situation, where different fields of expertise are required at an Expert level for distinct steps of an attack. It should be noted that the expertise must concern fields that are strictly different like for example HW manipulation and cryptography.

	Identification	Exploitation
Layman	0	0
Proficient	2	2
Expert	5	4
Multiple Expert	7	6

Table 4: Rating for Expertise

3.4 Knowledge of TOE

25 The CEM v.2.3 states that “to require sensitive information for exploitation would be unusual”. However, it shall be clearly understood that any information required for

identification shall not be considered as an additional factor for the exploitation. In general it is expected that all knowledge required in the Exploitation phase will be passed on from the Identification phase by way of suitable scripts describing the attack.

26 Since all sensitive and critical design information must be well controlled and protected by the developer, it may not be obvious how it assists in determining a dedicated attack path. Therefore, it shall be clearly stated in the attack potential calculation why the required critical information cannot be substituted by a related combination of time and expertise, e.g a planning ingredient for a dedicated attack.

27 The following classification is to be used:

- Public: this is information in the public domain,
- Restricted: this corresponds to assets which are passed about during the various phases of smartcard development. Suitable examples might be the functional specification (ADV_FSP), guidance documentation (AGD) or administrative documents usually prepared for smartcard issuers/customers. (See [CC-IC])
- Sensitive: HLD and LLD information.
- Critical: Implementation representation (Design and Source Code).
- Very critical hardware design: The designs of modern ICs involves not only huge data bases but also sophisticated bespoke tools. Therefore, the access to useful data requires an enormous and time consuming effort which would make detection likely even with the support from an insider. If an attack is based on such knowledge the new level of "Very critical design" is introduced. It has to be decided in a case by case decision, if the knowledge cannot be gained in another way.

28 In this way knowledge shall distinguish between access to high level design, low-level design on the one hand and source code/ schematics of the product on the other hand by taking into account two types of information (HLD/LLD and Implementation Level). (See [CC-IC])

29 It may occur that for sophisticated attacks, several types of knowledge are required. In such cases, the highest of the different knowledge factors is chosen.

	Identification	**Exploitation**
Public	0	0
Restricted	2	2
Sensitive	4	3
Critical	6	5
Very critical hardware design	9	NA

Table 5: Rating for Knowledge of TOE

3.5 Access to TOE

30 Availability of samples (in terms of time and cost) needs to be taken into account as well as the number of samples needed to carry out an attack path (this shall replace the CEM factor "Access to TOE").

31 The attack scenario might require access to more than one sample of the TOE because:

- the attack succeeds only with some probability on a given device such that a number of devices need to be tried out,
- the attack succeeds only after having destroyed a number of devices (on average),
- the attacker needs to collect information from several copies of the TOE.

32 In this case, TOE access is taken into account using the following rating:

	Identification	**Exploitation**
< 10 samples	0	0
< 100 samples	2	4
> 100 samples	3	6
Not practical	*	*

Table 6: Rating for Access to TOE

33 "Not Practical" is explained as follows:

- For identification: not practical starts with 2000 samples or the largest integer less than or equal to n/(1+(log n)^2), n being the estimated number of products to be built.
- For exploitation: not practical starts with 500 samples or the largest integer less or equal to n/(1+(log n)^3), n being the estimated number of products to be built.

34 The Security Policy as expressed in the Security Target should also be taken into account.

3.6 Equipment

35 In order to clarify the equipment category, price and availability has to be taken into account.

- None
- Standard
- Specialized (this type of equipment shall be considered as the type of expensive equipment which universities have in their possession.)
- Bespoke
  - Expensive [CEM]
  - Difficult to keep confidential [CEM] such as PC's linked across Internet.

36 In an ideal world definitions need to be given in order to know what are the rules and characteristics for attributing a category to an equipment or a set of equipments. In particular, the price, the age of the equipment, the availability (publicly available, sales controlled by manufacturer with potentially several levels of control, may be hired) shall be taken into account. The tables below have been put together by a group of industry experts and will need to be revised from time to time.

37 The range of equipment at the disposal of a potential attacker is constantly improving, typically:

- Computation power increase
- Cost of tools decrease
- Availability of tools can increase
- New tools can appear, due to new technology or due to new forms of attacks

38 It may happen that for sophisticated attacks several types of equipment are required. In such cases by default the highest of the different equipment factors is chosen.

3.7 Tools

39 The border between standard, specialized and bespoke cannot be clearly defined here. The rating of the tools is just a typical example. It is a case by case decision depending on state of the art and costs involved. The following tables are just a general guideline.

Tool	Equipment
UV-light emitter	Standard
Flash light	Standard
Low-end visible-light microscope	Standard
Climate chamber	Standard
Voltage supply	Standard
Analogue oscilloscope	Standard
Chip card reader	Standard
PC or work station	Standard
Signal analysis software	Standard
Signal generation software	Standard
High-end visible-light microscope and camera	Specialized
UV light microscope and camera	Specialized
Micro-probe Workstation	Specialized
Laser equipment	Specialized
Signal and function processor	Specialized
High-end digital oscilloscope	Specialized
Signal analyzer	Specialized
Tools for chemical etching (wet)	Specialized
Tools for chemical etching (plasma)	Specialized
Tools for grinding	Specialized

Table 7: Categorisation of Tools (1)

3.7.1 Design verification and failure analysis tools

40 Manufacturers know the purchasers of these tools and their location. The majority of the second hand tools market is also controlled by the manufacturers.

41 Efficient use of these tools requires a very long experience and can only be done by a small number of people. Nevertheless, one cannot exclude the fact that a certain type of equipment may be accessible through university laboratories or equivalent but still, expertise in using the equipment is quite difficult to obtain.

Tool	Equipment
Scanning electron microscope (SEM)	Bespoke
E-beam tester	Bespoke
Atomic Force Microscope (AFM)	Bespoke
Focused Ion Beam (FIB)	Bespoke
New Tech Design Verification and Failure Analysis Tools	Bespoke

Table 8: Categorisation of Tools (2)

42 Note, that using bespoke equipment should lead to a moderate potential as a minimum.

43 The level “Multiple Bespoke” is introduced to allow for a situation, where different types of bespoke equipment are required for distinct steps of an attack.

	Identification	Exploitation
None	0	0
Standard	1	2
Specialized (1)	3	4
Bespoke	5	6
Multiple Bespoke	7	8

Table 9: Rating for Equipment

(1) If clearly different testbenches consisting of specialised equipment are required for distinct steps of an attack this shall be rated as bespoke.

44 Equipment can always be rented but the same quotation applies.

3.8 Open Samples/Samples with known Secrets

3.8.1 Purpose of this section

45 In a composite evaluation as a rule, the properties of the hardware are taken from the information supplied with the documentation from the certification of the underlying platform IC. For this purpose, the CC supporting document [COMPO] specifies the process, called “composite smartcard evaluation”.

46 In general, the ETR-FOR-COMPOSITION should be written so as to contain enough information to evaluate and certify a composite product. In certain cases, it might be

opportune to use “open samples” to speed up the evaluation process. The use of the “open samples” or “samples with known secrets”, its scope, and the implications on the evaluation and the attack rating is described in this section.

3.8.2 Definition of “open samples / Samples with known Secrets”

47 Within the context of a composite evaluation, the term “open samples” stands for samples were the evaluator can put SW on the HW platform at his own discretion that bypasses countermeasures prescribed in the IC guidance. The intention is to use test SW without SW countermeasures but not deactivate any IC inherent countermeasures. In addition, another possibility is to enable the evaluator to define one or more pieces of secret data, such as a PIN or key, where this ability would not be available under the normal operation of the TOE. The SW should serve to highlight IC properties described in the IC ETR-FOR-COMPOSITION considering the special use of the HW in the TOE but not be used to repeat the IC evaluation. If the IC allows different configurations, the configuration implemented in the TOE shall be used. With these samples, it is thus possible to characterise the HW without SW.

48 “Samples with known secrets” refers to a TOE for which the evaluator knows or can define one or more pieces of secrets data, such as a PIN or key for performing either passive (monitoring) or fault attacks.

3.8.3 Use of “open samples / Samples with known Secrets”

49 For a composite evaluation, the TOE is the combination of HW and SW and the attacks during the evaluation have to be directed against this combination. For the definition of the attacks, the evaluator has to have full knowledge of the TOE. For the HW part in a composite evaluation this knowledge is provided by the evaluation results as described in the CC supporting document [COMPO].

50 The documents passed on from the HW evaluation to the composite evaluator describe the protection against threats and states requirements on the environment (especially the SW) necessary to obtain this protection. In addition, these documents will be a guidance on how the HW has to be used to achieve the security objectives.

51 For the vulnerability analysis and definition of attacks he wants to perform, the evaluator of the composite TOE can build on this information.

52 In some special cases the vulnerability analysis and definition of attacks might be difficult, need considerable time and require extensive pre-testing, if only this information is available. For example, samples with known secrets will allow faster characterization and allow a clear demonstration of successful attacks as well as the effectiveness of SW countermeasures.

53 Also, the platform may be used in a way that was not foreseen by the HW developer and the composite evaluator, or the SW provider may not have followed the recommendations provided with the HW.

54 Finally, the composite evaluator has to consider parts of the HW functionality that may not have been covered by the security target of the HW and therefore the HW evaluation.

55 Different possibilities exist to shorten the evaluation time in such cases:

- The composite evaluator can consult the evaluator of the HW and draw on his experience gained during the evaluation
- Separation of vulnerabilities of SW and HW with the use of "open samples" and/or the use of "samples with known secrets".

56 As a rule, a composite evaluation should not require the use of "open samples". However, if an efficient and meaningful evaluation in a maintainable time is only possible with the use of "open samples", then certain rules should be followed:

- The purpose of open samples is to set up tests for the composite evaluation and not to repeat the hardware-evaluation.
- The use of open samples and the information flow between parties is discussed and agreed upon between the certification body, the evaluator, the developer of the composite TOE and the developer of the open samples.
- The time spent on the dedicated "open sample" tests is restricted to one or two weeks.
- The goal and type of the tests is discussed and made known to all parties as defined in the information flow agreement.
- Failures and observations resulting from the tests are communicated and made known at least to the certification body of the composite TOE. The certification body of the composite TOE shall take appropriate steps together with the certification body of the HW evaluation.
- The rating should make provision for the judgement whether or not the attack would have been possible without the use of "open samples".

3.8.4 Implications on the composite evaluation

57 With the use of "open samples", it is possible to factorise attack paths and by that reduce the complexity of an attack. That saves time in the evaluation because it makes it possible to obtain the targeted result much faster.

58 A good example for this is the retrieving of secret information (e.g. keys) by light attacks. In a well-designed product, the HW as well as the SW will have protective mechanisms to avert this attack. In combination, they will make attacks quite difficult. The evaluator will have to try a very high number of combinations and variations of parameters like beam diameter, light frequency, light strength, location for applying the light, position in time for the light flash. This gets especially difficult if the SW contains means to render the TOE inoperable if an attack is detected. An attack could not only prove very time consuming but also require a great number of samples.

59 With "open samples", the situation is quite different. The evaluator can use his own optimised test program and scan the IC for "weak spots" much faster and without risking the destruction of the device (the fact that such "weak spots" exist might even have been stated in the HW evaluation documentation). With the knowledge gained in these tests the attacker can then launch much more directed attacks on the TOE.

60 This example also shows the danger of this approach. Without open samples, the attack on the TOE (combination of HW + SW) might not be realistic and unfeasible. Therefore, this would lead to unjustified rating and in the extreme to a fail of the product.

3.8.5 Implications on the composite rating

61 For the rating two possibilities have to be considered:

- Freely programmable samples of the HW or similar variants are freely available. In that case, the samples are not to be considered as "open samples". They have to be considered just a tool (like e.g. a microscope) for the evaluator. The results can be used without any special treatment in the rating.
- The access to the samples is restricted and controlled and has been evaluated during the IC evaluation. In that case, the rating has to include an additional factor for the use of "open samples" as described in the table below.

3.8.6 Background of the use of "samples with known secret" to accelerate the evaluation

62 An additional possibility to accelerate the evaluation especially where cryptographic operations are involved is the use of "samples with known secret". With these samples the evaluator knows the "secret" (key). This allows either comparing of retrieved data (e.g. as deduced from passive analysis) against the known "secret", or may be useful in a profiling step required for some attacks. The evaluator therefore has a simplified way to determine if his attack has revealed the correct secret. He can stop after retrieving parts of the "secret" and estimate the remaining time to find the complete "secret".

63 However, a rating based on such samples has to be carefully considered because the attack might only be made possible by the availability of "samples with known secret".

64 For instance:

- To extract the complete key might prove to be very time consuming. With some error in the retrieved key and no possibility to decide which part of the secret is not correct an attack might not be possible.
- A profiling stage is sometimes required to perform some attacks, such as template attacks. Knowing the key, and then the intermediate values of the algorithms, may then make an attack possible.

65 In general the rating of "samples with known secret" is comparable to the rating of "open samples". Therefore, both tools are combined here.

3.8.7 Calculating the attack potential

66 As with other aspects of an attack, the evaluator has to estimate the value of the factors (time, access to TOE, etc) for an attacker.

67 Where open samples exist, collusion (or direct attack, such as theft) to obtain them is possible in the same way that the evaluation takes into account a possible collusion or

direct attack for an attacker to get design information (down to the implementation level).

68 A factor “open sample” is therefore defined in the attack potential table, with points in the identification phase for “open samples” used during evaluations. The same factor should be applied for the “samples with known secret”.

69 When rating an attack that makes use of open samples / samples with known secrets, the evaluator must first determine (at least theoretically) and describe the way in which an attacker could carry out the attack on the real TOE (instead of on the open sample / sample with known secret). Having determined this, the evaluator will perform two calculations, using open samples / samples with known secret:

- Estimating the value for each factor for an attacker without access to open samples / samples with known secrets.
- Giving the values for each factor corresponding to what he has done:
  - Time spent, destroyed samples, Expertise, Knowledge of the TOE, equipment
  - Adding the points corresponding to the open samples used

70 Should it turn out that the attack is not practical when not using open samples or samples with known secrets, then that rating has be used and the open samples / samples with know secrets rating discarded. In all other cases the final value will be the minimum of the two calculations. It is expected that the two values are quite close. If this is not the case further analysis is required to decide on the rating.

71 The points corresponding to the availability of open samples are defined by taking into account the protection and the control of these open samples during the entire life cycle.

72 For ICs, the protection level will be analysed during the IC evaluation and stated in the ETR-FOR-COMPOSITION.

73 For “samples with known secret”, defining the protection level is part of the evaluation of the full product.

74 Because of the similarity in the threat to the TOE, the rating for open samples (with and without known secrets) should be defined according to the values of the Knowledge of the TOE factor: PUBLIC, RESTRICTED, SENSITIVE and CRITICAL:

- PUBLIC:
  - Open samples: No protection of the samples, delivered without control (no NDA, no checking of the customer); or the IC is used in non-secure applications (e.g. applications without guarantee of implementing the security recommendations or versions which can be freely programmed with native code).
  - Samples with known secrets: This concerns secrets easily deducible from information already rated in “knowledge of the TOE”.

- RESTRICTED:
    - Open samples: Typically protected as the specifications of the card, as the data sheet of an IC, or delivered without extra control of the people having access to this kind of information.
    - Samples with known secrets: Typically applies to secrets where a specific decision and action is required to release the information (so it is not, for example, automatically available for anonymous access via a website), and where the recipient is made aware that the data is potentially useful to an attacker (e.g. via guidance information). In some cases it may be possible for an attacker to find out or deduce the information, but its availability still provides convenience (and perhaps a saving of time).
- SENSITIVE:
    - Open samples: Protected as the HLD/LLD design levels are.
    - Samples with known secrets: Secrets are only shared by a limited number of clearly defined and identified people or devices, with strong access controls. Handling of the Secret data is governed by specific and appropriate written procedures to protect it, and there is a clear method by which the Secret data is identified as requiring these procedures (e.g. by labelling the data).
- CRITICAL:
    - Open samples: Protected as the implementation level (source code, VHDL, layout). This requires to have very few open samples produced, to have very strong control of their delivery and to have the assurance that the receiving organisation is able to setup a control at the same level.
    - Samples with known secrets: Secrets were generated inside the sample and are only owned by it, or in another module which does not make these secrets available outside the module (except to the sample). These secrets are therefore not available outside the card, and possibly the module, under normal conditions. Only under exceptional conditions could these secrets be known, for example by providing the evaluator with either specific commands to access the secrets (not available in any normal configuration of the TOE), or special samples with static secrets instead of dynamic secrets (fixed in personalization phase for instance). As with Open Samples at the Critical level, this requires that there are very few Open Samples produced, that they have very strong control over authorisation for their release and delivery to the recipient, and to have assurance that the receiving organisation will control the samples so as to provide equivalent limits on their availability.

75 The composite evaluation has also to define if the use of “open samples” **and** “samples with known secret” accumulates the efforts in time and add points for each of them.

The analysis will be done during the ALC_DVS.2 task, checking if a single collusion can be enough or if two different collusions are necessary.

76 The IC evaluation will give a rating for the "open samples" in the ETR-FOR-COMPOSITION. Any indication for a different rating has to be considered in the composite evaluation.

(Identification phase only)	PUBLIC or not required	RESTRICTED	SENSITIVE	CRITICAL
Open Samples	0	2	4	6
Samples with known secret	0	2	4	6

3.8.8 Impact on the evaluation of an IC with guidance

77 As such, the concept of "samples with known secrets" does not apply to HW IC evaluations, since the final application is not known at this point in time. For instance, open, unprotected applications / APDUs residing alongside the secured ones may effectively allow to perform an analysis equivalent to using "samples with known secrets" in the first place.

78 The situation is more complex with regards to the concept of "open samples". Here it is useful to distinguish two different classes of countermeasures that may be described in the IC guidance.

- Simple countermeasures are those that are effectively equivalent to a switch. For instance, the IC guidance may require that the TOE shall only be used with internal clock, or only with a clock-skipping mode enabled. In those cases it may be advantageous for the IC evaluator to switch these features off and thereby speed up the evaluation. The assessment will then involve generating two different ratings, one with an estimate for the time that would have been spent on the attack had the countermeasure been enabled, and a second rating along the rules for open samples. As before, in the end the minimum of the two ratings will be chosen.
- Complex countermeasures are those that require more or less complex SW code to be generated in the final application where it can be expected that some variability will exist from one implementation to another. Typical examples here are countermeasures against fault attacks. In such a case the concept of open samples does not apply to HW IC evaluations, since there is no way of knowing whether a final product based on this IC does implement all SW countermeasures recommended in the HW IC guidance in such a way that no loophole could possibly exist.

3.8.9 Impact on the evaluation of a firmware / crypto library of an IC with guidance

79 Crypto libraries and other supporting routines for a HW IC that are evaluated within the composite evaluation scheme (often separately from the HW IC evaluation) are somewhat special in that they are not a final product in their own right, but rather are to

be used in one or more final products, which in turn may themselves be subject to a composite evaluation.

80 The concept of “samples with known secrets” usually does not apply here since the crypto library has in general no control over the handling of those secrets outside its boundaries. This is different, though, for secrets generated and maintained within the crypto library that are not exported.

81 However, the concept of “open samples” may apply more often, depending on the circumstances. A typical example would be countermeasures against light attacks or SPA-DPA attacks that are applied under the control of the crypto library. Provided these countermeasures cannot be switched off in the final product, the crypto library may be considered to be equivalent to a final product in this respect, and consequently the concept of “open samples” of the composite evaluation scheme applies. It does not matter here whether these are countermeasures that have been actually suggested in the HW IC guidance or not.

3.9 Final Table

Factors	Identification	Exploitation
Elapsed time		
< one hour	0	0
< one day	1	3
< one week	2	4
< one month	3	6
> one month	5	8
Not practical	*	*
Expertise		
Layman	0	0
Proficient	2	2
Expert	5	4
Multiple Expert	7	6
Knowledge of the TOE		
Public	0	0
Restricted	2	2
Sensitive	4	3
Critical	6	5
Very critical hardware design	9	NA
Access to TOE		
< 10 samples	0	0
< 100 samples	2	4
> 100 samples	3	6
Not practical	*	*
Equipment		
None	0	0
Standard	1	2
Specialized (1)	3	4
Bespoke	5	6
Multiple Bespoke	7	8
Open samples (rated according to access to open samples)		
Public	0	NA
Restricted	2	NA
Sensitive	4	NA
Critical	6	NA

Table 10: Final table for the rating factors

82 (1) If clearly different testbenches consisting of specialised equipment are required for distinct steps of an attack this shall be rated as bespoke.

83 * Indicates that the attack path is not exploitable within a timescale that would be useful to an attacker. Any value of * indicates a High rating.

3.10 Range for CC v2

84 The following table replaces table A.8 of CEM, para 1835 for smartcards.

Range of values*	Resistance to attacker with attack potential of:	SOF rating
0-15	No rating	No rating
16-24	Low	Basic
25-30	Moderate	Medium
31 and above	High	High

Table 11: Rating of vulnerabilites for CC v2

85 *final attack potential = identification + exploitation.

3.11 Range for CC v3

86 The following table replaces table B.4 of CEM, para 1869 for smartcards.

Range of values*	TOE resistant to attackers with attack potential of:
0-15	No rating
16-20	Basic
21-24	Enhanced-Basic
25-30	Moderate
31 and above	High

Table 11: Rating of vulnerabilites for CC v3

87 *final attack potential = identification + exploitation.

4 Examples of attack methods

88 The following examples have been compiled by a group of security experts representing the different actor groups involved in the development, production, security evaluation and distribution of a smartcard product (Hardware vendors, Card vendors, OS provider, Evaluation labs, Certification bodies, Service providers**).**

89 The collection represents the current state of the art at that time (Q4/05). As state of the art is not static this document is under review of the same expert group and will be updated if necessary.

90 For the evaluation of a TOE at least these examples have to be considered. This does not mean that in any case all attacks have to be carried out. For each TOE the evaluation lab conducting the evaluation has to select the appropriate attacks from this catalogue in agreement with the certification body. This selection will be dependent on the type of the TOE and additional tests may also be required.

91 In this document only a general outline of the attacks is given. For more detailed descriptions and examples, please refer to the certification bodies. They can also provide examples as reference for rating.

4.1 Physical Attacks

92 Microelectronic tools enable to either access or modify an IC by removing or adding material (etching, FIB, etc). Depending on the tool and on its use the interesting effect for the attacker is to extract internal signals or manipulate connections inside the IC by adding or to cutting wires inside the silicon.

93 Memories could also be physically accessed for, depending on the memory technology, reading or setting bit values.

94 The attack is directed against the IC and often independent of the embedded software (i.e. it could be applied to any embedded software and is independent of software counter measures).

95 The main impacts are:

- Access to secret data such as cryptographic keys (by extracting internal signals)
- Disconnecting IC security features to make another attack easier (DPA, perturbation)
- Forcing internal signals
- Even unknown signals could be used to perform some attacks

96 The potential use of these techniques is manifold and has to be carefully considered in the context of each evaluation.

4.2 Overcoming sensors and filters

97 This attack covers ways of deactivating or avoiding the different types of sensor that an IC may use to monitor the environmental conditions and to protect itself from conditions that would threaten correct operation of the TOE. Hardware or software may use the outputs from sensors to take action to protect the TOE.

98 Sensors and filters may be overcome by:

- Disconnection
- Changing the behaviour of the sensor
- Finding gaps in the coverage of the monitored condition (e.g. voltage), or of the timing of monitoring.

99 Sensors may also be misused, in order to exploit activation of a sensor as a step in an attack. This misuse of sensors is a separate attack.

100 The different types of sensors and filters include:

- Voltage (e.g. high voltage or voltage spike)
- Frequency (e.g. high frequency or frequency spike)
- Temperature
- Light (or other radiation)

101 The main impacts are:

102 The correct operation of a chip can no longer be guaranteed outside the safe operating conditions. The impact of operating under these conditions may be of many sorts. For example:

- Contents of memory or registers may be corrupted
- Program flow may be changed
- Failures in operations may occur (e.g. CPU, coprocessors, RNG)
- Change of operating mode and/or parameters (e.g. from user to supervisor mode)
- Change in other operating characteristics (e.g. changed leakage behaviour; enable other attacks like RAM freezing, electron beam scanning).

103 If a chip returns incorrect cryptographic results then this may allow a DFA attack, see section 4.4. Other consequences are described under general perturbation effects in section 4.3

4.3 Perturbation Attacks

104 Perturbation attacks change the normal behaviour of an IC in order to create an exploitable error in the operation of a TOE. The behaviour is typically changed either by applying an external source of energy during the operation of the IC, or by operating the IC outside its intended operating environment (usually characterised in terms of temperature, Vcc and the externally supplied clock frequency).

105 The attack will typically aim to make cryptographic operations weaker by creating faults that can be used to recover keys or plaintext, or to avoid or change the results of checks such as authentication or lifecycle state checks or else change the program flow.

106 Chapter 4.3 concerns itself more with the methods to induce meaningful faults whereas Chapter 4.4 describes how these induced faults may be used to extract keys from cryptographic operations.

107 Perturbations may be applied to either a hardware TOE (an IC) or a software/composite TOE (an OS or application running on an IC).

108 The main impacts are:

109 For attackers, the typical external effects on an IC running a software application are as follows:

- Modifying a value read from memory during the read operation: The value held in memory is not modified, but the value that arrives at the destination (e.g. CPU or coprocessor) is modified. This may concern data or address information.
- Changing the characteristics of random numbers generated (e.g. forcing RNG output to be all 1's) – see Attacks on RNG 4.9 for more discussion of attacks on random number generators.
- Modifying the program flow: the program flow is modified and various effects can be observed:
  - Skipping an instruction
  - Replacing an instruction with another (benign) one
  - Inverting a test
  - Generating a jump
  - Generating calculation errors

110 It is noted that it is relatively easy to cause communication errors, in which the final data returned by the IC is modified. However, these types of errors are not generally useful to an attacker, since they indicate only the same type of errors as may naturally occur in a communication medium: They have not affected the behaviour of the IC while it was carrying out a security-sensitive operation (e.g. a cryptographic calculation or access control decision).

111 The range of possible perturbation techniques is large, and typically subject to a variety of parameters for each technique. This large range and the further complications involved in combining perturbations means that perturbation usually proceeds by investigating what types of perturbation cause any observable effect, and then refining this technique both in terms of the parameters of the perturbation (e.g. small changes in power, location or timing) and in terms of what parts of software are attacked. For example, if perturbations can be found to change the value of single bits in a register, then this may be particularly useful if software in a TOE uses single-bit flags for security decisions. The application context (i.e. how the TOE is used in its intended operating environment) may determine whether the perturbation effect needs to be precise and certain, or whether a less certain modification (e.g. one modification in 10 or 100 attempts) can still be used to attack the TOE.

4.4 Retrieving keys with DFA

112 DFA is the abbreviation of Differential Fault Analysis. With DFA an attacker tries to obtain a secret by comparing a calculation without an error and calculations that do have an error. DFA can be done with non-invasive and invasive techniques.

113 This class of attacks can be divided in the following stages:

- Search for a suitable fault injection method
- Mounting the attack (performing the cryptographic operation once with correct and once with faulty parameters)
- Retrieving the results and composing a suitable set of data and calculating the keys from that data

114 By applying special physical conditions during the cryptographic operation, it is possible to induce single faults (1 bit, 1 byte) in the computation result.

115 This attack can be carried out in a non-invasive or an invasive manner. The non-invasive method (power glitching) avoids physical damages. The invasive method requires the attacker to physically prepare the TOE to facilitate the application of light on parts of the TOE.

116 The main impacts are:

117 DFA can break cryptographic key systems, allowing to retrieve DES, 3DES and RSA keys for example, by running the device under unusual physical circumstances. The attacker needs to inject an error at the right time and location to exploit erroneous cryptographic outputs.

118 As keys and code are usually present in EEPROM it might be difficult to randomly alter bits without crashing the entire system instead of obtaining the desired faulty results, although code alteration can give results as well. Other techniques may be useful to determine best location and time to inject an error; such as analyzing the power consumption to determine when the cryptographic computation occurs.

4.5 SPA/DPA – Non-invasive retrieving of secret data

119 SPA and DPA stand for ‘Simple’ and ‘Differential Power Analysis’, respectively, and aim at exploiting the information leaked through characteristic variations in the power consumption of electronic components – yet without damaging the TOE in any way what-so-ever. Although various levels of sophistication exist, the power consumption of a device can in essence be simply measured using a digital sampling oscilloscope and a resistor placed in series with the device. The outcome of the attack may be as simple as a characteristic trigger point for launching other attacks (such as DFA), or as much as the secret key used in a cryptographic operation itself. Depending on the goal of the attack it may involve a wide range of methods from direct interpretation of the retrieved signal to a complex analysis of the signal with statistical methods.

120 The main impacts are:

121 It lies in the very nature of SPA and DPA attacks that they may in principle be applied to any cryptographic algorithm – either stand alone, or as part of a composite attack.

Additionally, SPA may serve as a stepping stone for launching further attacks. For instance, SPA may be employed to detect a critical write operation to the EEPROM that needs to be intercepted. An SPA analysis may also be performed as part of a timing attack, or for deducing which branch of a conditional jump has been taken by the program flow. Finally, an SPA attack could be used to determine the proper trigger point for a subsequent glitch or light attack, or as an aide for localising a suitable time window for a physical probing attack.

122 A DPA attack does not need to be entirely successful for it to become dangerous. Given a suitable key search strategy that takes into account imperfect DPA results, it may be enough to retrieve only part of the secret key by DPA, and obtain the rest by brute-force methods.

4.6 Higher Order DPA

123 Implementations that include countermeasures like boolean masking that resist first order DPA may be vulnerable to higher order DPA. This requires that the attacker is able to correlate more than once per TOE computation using hypotheses on intermediate states that depend on secret key parts.

124 The combined statistical analysis may be based on aligned measurements of the same side channel at different times or on aligned simultaneous measurements of different channels like power consumption and electromagnetic radiation of the device during the computation.

125 The attack requires at least the same effort as for a standard DPA attack with respect to expertise, knowledge of the TOE and equipment. Depending on the countermeasures and the implementation the effort increases at least for the identification of the attack.

126 The main impacts are:

127 Although higher order DPA is also a generic approach the analysis must be adapted for every TOE since the information that must be combined for the statistical analysis depends on the implementation including the combination of countermeasures. Nevertheless higher order DPA can be applied to different cryptographic systems such as all kind of secret key (symmetric) algorithms that make use of similar Boolean or arithmetic masking countermeasures. (Of course, algorithms that are vulnerable to first order DPA are vulnerable to higher order DPA too.)

128 The extension of higher order DPA to public key (asymmetric) algorithms seems to be very difficult because of the widely applied blinding measures that make use of algebraic transformations during the calculation that are completely different from masking. Therefore higher order DPA is more adapted and efficient when used to retrieve secret information from symmetric than asymmetric algorithms.

129 With higher order DPA, it may be possible to analyse implementations with countermeasures against standard DPA.

130 However, it seems that a non-negligible part of higher order DPA success is in parallel

- with the observer’s experience,

- his ability to recognize and to interpret the different clues to follow,
- his ability to develop the corresponding tools that will be used to track this information,
- his knowledge and skills in cryptography and the analysis of the hardware design as well as
- with the environment conditions of the experiment itself,

131 in order to gain access to the sought-after information..

4.7 EMA Attacks

132 When an IC is operating, each individual element will emit electromagnetic radiation in the same way as any other conductor with a current flowing through it. Due to the change of the data processed, small changes in the current flow will be the result. These current flow changes lead to an electromagnetic emission depending on the processed data.

133 Electromagnetic Analysis (EMA) attacks measure these electromagnetic emissions from an IC during its operation and inferences to the data processed. It uses similar analysis techniques to those used in power analysis, hence it is sometimes referred to as SEMA (Simple Electromagnetic Analysis, analogous to Simple Power Analysis (SPA)) or DEMA (Differential Electromagnetic Analysis, analogous to Differential Power Analysis (DPA)).

134 The attack may use emissions from the whole IC (chip-EMA), or may focus on the emissions from particular areas of the die, where critical components are located (local-EMA).

135 Experimental evidence show that electromagnetic obtained data (particularly from localised areas of a die) can be different from power trace data, and ICs that are protected against power analysis may therefore be vulnerable to EMA.

136 The attack will typically aim to recover keys or plaintext, but may also be applied to recover other secret data such as PINs, or random numbers generated for use as secrets.

137 The main impacts are:

138 An EMA attack may be used in various ways – the following are examples:

- Used for determination of secret data correlated with emissions, such as keys or PINs, in a similar way to that of SPA or DPA (note that the EMA attack may be more efficient, requiring fewer examples, than a power-based attack)
- Used for identification of power analysis countermeasure activity, enabling them to be removed from power traces hence enabling a power analysis attack to succeed
- Used for identification of distinct localised activity that can be used in breaking security functions – for example, it might be possible to detect different register configurations for squares and multiplies in a coprocessor that supports RSA

- Used for identification of activity that may assist in synchronisation of other attacks – for example, it may be possible to detect actions within a cryptographic algorithm or PIN check that enable the precise synchronisation of a perturbation (see chapter 4.3 Perturbation Attacks).

139 As with power analysis, EMA attacks may be carried out for a hardware TOE (an IC), or a software/composite TOE (an OS or application running on an IC). In the same way as for power analysis, the vulnerability of software to EMA attack should not be predicted from the EMA characteristics of the hardware alone. The way in which software uses the IC functions may make a critical difference to its vulnerability to this type of attack.

4.8 Exploitation of Test features

140 The attack path aims to enter the IC test mode to provide a basis for further attacks.

141 These further attacks might lead to disclosure or corruption of memory content but as this depends on the possibilities of the test mode this is not considered here.

142 The main impacts are:

143 As result of a successful attack, the attacker is able to read out the content of the non-volatile memory using test functions. The implementation of the test functions may have an impact on the usability of the retrieved user data.

144 Another result is the re-configuration of life cycle data or error counters using a test function. Thereby an attacker is able to continue his analysis on the same device.

4.9 Attacks on RNG

145 Attacks on RNGs aim in general to get the ability to predict the output of the RNG (e.g. of reducing the output entropy) which can comprise:

- past values of the RNG output (with respect to the given and possibly known current values),
- future values of the RNG output (with respect to the possibly known past and current values),
- forcing the output to a specific behaviour, which leads to:
  - known values (therefore also allowing for the prediction of the output),
  - unknown, but fixed values (reducing the entropy to 0 at the limit),
  - repetition of unknown values either for different runs of one RNG or for runs of two or more RNGs (cloning) .

146 A RNG considered here can be one of the following types[1]:

- true RNGs (TRNG), the output of which is generated by any kind of sampling inherently random physical processes,

[1] In the context of smart cards the RNG based on some measurements of environment are not considered to be relevant.

- pseudo RNG (PRNG) which output is generated by any kind of algorithmic processing (the algorithm is in general state based, with the initial state (seed) may generated by a TRNG),
- hybrid RNG (HRNG), which consists of a TRNG and a PRNG with a variety of state update schemes,

147 The applicable attack methods vary according to the Type of RNG:

148 A true RNG may be attacked by[2]:

- permanent or transient influence of the operating conditions (e.g. voltage, frequency, temperature, light)
- non invasive exploitation of signal leakage (e.g. signal on external electrical interfaces)
- physical manipulation of the circuitry (stop the operation, force the line level, modify and/or clone the behaviour, disconnect entropy source)
- wire taping internal signals (compromise internal states)

149 A pseudo RNG may be attacked by:

- direct (cryptographic) attack on the deterministic state transition and output function (e.g. based on known previous outputs of the RNG)
- indirect attack on the state transition computation process by employing some side channel information (i.e. leakage on external electrical interfaces)
- attack on the execution path of the processing (modification of the results)
- attack on the seed (prevent reseeding, force the seed to fixed known or unknown (but reproducible) value, compromise the seed value)
- overcome the limit of RNG output volume (e.g. forcing the RNG to repeat values or to produce enough output to enable the attacker to solve equations and based on the solution to predict the output)

150 The attacks on hybrid RNG will be in general a combination of attacks on TRNGs and PRNGs.

151 All RNG designs can be expected to demand also for test procedures to counter attacks like those listed above. The analysis above does not take attacks on test procedures into account, as such attacks will by covered sufficiently by the more general attack scenario on software. Observe that test procedures may be an object on attack like SPA/DFA to reveal the RNG output values.

152 The main impacts are:

153 A successful attack on the RNG will result in breaching the security mechanisms of the chip, which rely on the randomness of the RNG. The mechanisms may be DPA/SPA countermeasures, sensor testing, integrity checking of active shield, bus and/or memory encryption and scrambling. The application software is affected by such attacks

[2] It is here assumed that the direct attack on a true RNG (i.e. guessing the value) is not feasible for any attacker.

indirectly, e.g. sensors and related tests being disabled by an attacker, will generate further attack possibilities.

154 The software developer can rely on the capabilities of the hardware platform for testing the RNG and use these or implement and perform additional tests by himself based on such capabilities. The software developer may implement also tests for repetition of RNG output, but the coverage and feasibility of such tests may depend on the implementation and seems to be a problem. The cloning attack for RNG output on different instances of a RNG cannot be countered by tests, so other mechanisms must be designed as appropriate.

155 In case of TRNGs, sufficient tests should be performed (either by the chip platform itself or by the software developer). [AIS31] is an example of a methodology for assessing the effectiveness of the testing mechanisms. In case of PRNG a special effort on protecting the seed and the algorithm in terms of integrity and confidentiality is required. This effort pertains to the general software and data protection aspects and will be not discussed further in this chapter.

4.10 Ill-formed Java Card applications

156 This logical attack consists in executing **ill-formed applications**, i.e. malicious applications that are made of illegal sequences of byte-code instructions or that do not have valid byte-code parameters.

157 This example is only applicable to Java Cards (although there may be equivalent attacks for other operating systems). If not combined with any other attack such as authentication bypass, this attack has to be applied to Java Cards with known loading keys (these could be considered as open mode samples). In addition, if the card includes an embedded byte-code verifier, this verifier must be disabled. No other specific configuration is required.

158 Ill-formed applications execute a sequence of byte-code that violates the Java rules. Ill-formed applications are usually created from standard applications, in which the byte-code is manually modified. It means that such ill-formed applications cannot be the output of a normal CAP file generator. As a consequence, most Java Card platforms don't enforce the rules during the execution of applications.

159 The main impacts are:

160 In the most favourable cases, the attacker can retrieve information (e.g. a dump of memory), execute functions that usually require specific privileges or even switch to a context giving the full control over the card (JCRE context).

4.11 Software Attacks

161 Most of the examples of attacks in this document require hardware attack steps for all or part of the attack. However, it is clear that there are many relevant attacks that can be made on software alone. This section considers some of these attacks. In many cases software attacks start with source code analysis.

162 In general, it is important to note that most software attacks arise from errors (bugs) in the TOE, either in design or implementation. In these cases, the error will generally

result in a failure to meet the requirements of one (or more) of the ADV families (e.g. ADV_IMP.1.2E: The evaluator shall determine that the least abstract TSF representation provided is an accurate and complete instantiation of the TOE security functional requirements). Hence an error of this sort will cause the TOE to fail evaluation (or, more usually, will require a modification to the TOE to correct the error).

163 In some other cases, a design's specification may be insufficient to meet the TOE security objectives: for example, a protocol specification might itself contain critical vulnerabilities. This would also cause a TOE to fail the evaluation.

164 This section therefore lists a number of attack steps that may be used to discover software errors, but no attack potential examples are given, since if any error is discovered then it must be corrected if the TOE is to pass evaluation.

165 In the text below we consider first an information gathering attack step, which may be relevant to a number of different types of attack. We introduce five specific attack techniques that may exploit software vulnerabilities:

- Editing commands
- Direct protocol attacks
- Man-in-the-middle attacks
- Replay attacks
- Buffer overflow

166 The attacks are of a logical nature, the test environment consists of a smart card reader connected to a PC. The PC runs communication software, a protocol analyser and some development tools to modify communication. This tool set is considered to be standard equipment. Tools are available as freeware on the Internet, and they can be modified quite easily to fit the attackers' needs.

167 To perform such attacks, it is necessary to have:

- a means to listen to message sequences (reader, traffic analyser)
- a means to create messages (information on external API, pattern generator)
- a means to interrupt messages without detection (protocol dependent)

168 Setting up a test environment and identifying an attack is quite simple, as the tools are standard and the commands are often ISO standard, and therefore public knowledge. If the command set is proprietary, the expertise needed is slightly higher because the communication must be interpreted. However, in most cases this would be expected to be relatively straightforward, and this type of 'security by obscurity' would not be considered a valid defence against attack.

4.12 Information gathering

4.12.1 Introduction

169 By their nature, communication protocols are susceptible to information leakage. This unwanted effect is a consequence of the fact that they are designed to pass information. This type of attack tries to use the protocols in ways that were not intended by the protocol developer, by first gathering information and then changing that communication to obtain secret data or other resources.

170 The attack step is usually a non-invasive technique, with the aim of getting information on the communication commands that the smartcard supports or using information from message sequences to enable other attacks. It is noted that the information is assumed to be information not contained in design documents (e.g. undocumented responses to commands). This information may then enable the attacker to modify the interaction or to disclose information (e.g. user data or keys) using weaknesses in the software implementation. This attack step is normally not a full attack path leading to the retrieval of secret data, although it might do in specific cases.

171 This attack step results in gathering information on the operation of the TOE, with possible disclosure of secret data (exposure of secret data in this way would generally be considered a sufficient vulnerability to cause the TOE to fail evaluation[3]). The information gathered is analysed to see whether it can be used to mount an attack to retrieve secret data from the TOE with one of the other mechanisms described in this document. The attacker knows the attack has succeeded by analysing the answers the smartcard gives during the communication.

4.12.2 Attack Step Descriptions

Observing Message Sequences

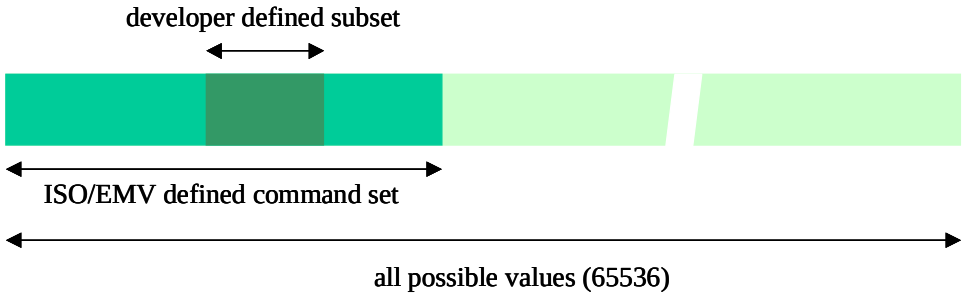
172 Observing message sequences may result in:

- obtaining information on an unknown protocol (e.g. where the interface specification is not public) to prepare an attack
- obtaining information on unknown internal product structures (typically data structures in software) to prepare an attack
- disclosing information, keys, or security attributes during import or export operations
- tracing product activity or user behaviour (e.g. to enable a replay attack).

Command searches

173 The total amount of values that a smartcard can communicate using a typical protocol such as ISO 7816 T=1 is 2^{16}, or 65536 different commands. Of this set, ISO defined a subset as being valid commands. And of this ISO set, a developer defines a subset and documents these commands as being valid commands for this card.

[3] Depending on the scope of the evaluation and the environment, there may be some situations where such information exposure is accepted, e.g. in a protocol for use only in secure personalisation environments.



174 A T=1 test plan should contain the following tests:

- A 'brute force' approach in which all values outside the ISO defined set are tried and it is checked whether the card responds (inopportune behaviour).
- A 'brute force' approach in which all values of the ISO defined set, but outside the developer defined set are tried for a response (undocumented command search).
- Trying all developer documented commands and checking the answers.
- Influencing the communication by sending commands in different sequences.
- Interrupting message from system or from product

175 Attacks that make use of undocumented commands and editing commands are closely related, but distinctive attacks. Finding undocumented or undefined commands is a straightforward brute-force type of attack, where the attacker simply runs the ISO defined set of commands to see if the card replies to one or more commands that it should not answer to.

176 As an undocumented command search can be highly standardized and automated, it should not take much more time than one day. Once all variations of Class, Instruction, Parameter 1 and Parameter 2 are tried and the answers recorded, the attacker analyses if there is any interesting attack mount point. Once an interesting answer has been determined the attacker builds a script to exploit the vulnerability. This could also be done by source code checking.

177 Whether the undocumented command may present attack points depends on the quality of the software (the separation of execution domains) and the type of command that is discovered.

4.13 Editing commands

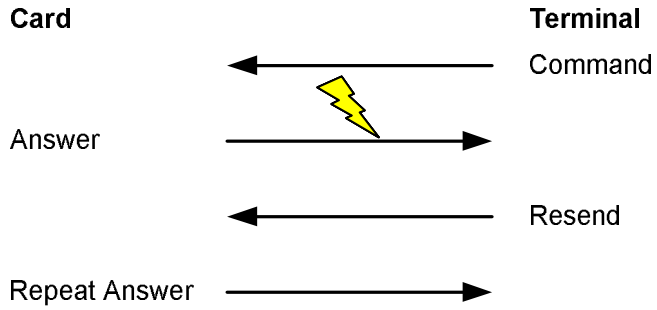
178 Editing commands is an attack step where the attacker tries to modify commands during the communication sequence to see if the card gives an unexpected reply (these commands may be in an interface specification, or they may have been discovered by observing message sequences or a command search as described above). These attack steps may enable vulnerabilities to be discovered and exploited (e.g. editing previously observed messages to supply a parameter that is too long may enable a buffer overflow attack). They may also expose timing differences that assist in reverse engineering of the software.

179 According to the security mechanisms associated to the API and the type of message, it may be easy or complex to forge a message (Mutual authentication, Secure channel, MAC, Ciphering, session key,...). However, as noted earlier, if an attack of this sort can be found then it will generally cause a TOE to fail evaluation.

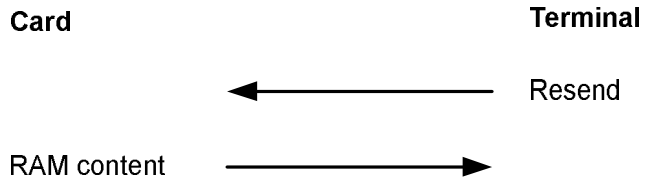
4.14 Direct protocol attacks

180 A typical protocol attack is to try to send commands that the smartcard does not expect in its current state. For example: the ISO 7186-3 and 14443 protocols for smartcards contain a command for handling failure in the communication. Instead of starting a genuine communication, by sending this command an attacker may receive an un-initialized buffer, or the last buffer that was written. This example is shown in the following pictures.

T=1 example valid behavior



T=1 example of security risk (inopportune behavior)



181 Whether the TOE actually dumps the memory contents depends on the proper initialisation of I/O buffer pointer and length. The memory shown in the example might contain residual secret data, for example a DES session key that was just calculated. Therefore this attack may allow an attacker to retrieve secret data from the TOE.

4.15 Man-in-the-middle attacks

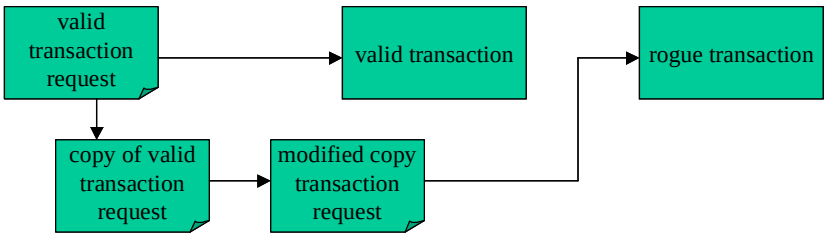
182 In this attack, the attacker hides in the communication path between two entities that are executing a valid communication. The attacker presents himself to either party as the other (valid) party. Some applications of Man in the middle attacks in public literature may be found in the following papers:

- An Example of a Man-in-the-middle Attack Against Server Authenticated SSL-sessions, Mattias Eriksson

- Man-in-the-Middle in Tunnelled Authentication Protocols, N. Asokan, Valtteri Niemi, Kaisa Nyberg, Nokia Research Center, Finland
- Why Cryptosystems Fail, Ross Anderson

4.16 Replay attacks

183 Replay attacks are possible when a mechanism does not check that a command is a genuine part of the current message sequence, or that a complete message sequence has not been used before (in general, a secure protocol should prevent this sort of attack by design[4]). An attacker uses a protocol analyser to monitor and copy packets as they flow between smartcard and reader or host. The packets are captured, filtered and analysed for interesting information like digital signatures and authentication codes. Once these packets have been extracted, the packets are sent again (replayed), thus giving the attacker the possibility to get unauthorized access to resources.



184 The picture shows a situation where the attacker copies a valid transaction request, modifies it and sends a second request using the same (or slightly modified) versions of the messages. In general this type of attack might allow the attacker to get unauthorised access to a user's assets, for example a bank withdrawal or access to protected system resources.

185 The attack may be a full attack path, such as if a bank account withdrawal succeeds. In the case where system resources are accessed, it might be a partial attack path, depending on the nature of the resources that are accessed (e.g. the attacker is now able to communicate as an ordinary user and tries to get elevated privileges).

186 The replay attack might be countered by using sequence numbers with appropriate integrity protection, making the use of recorded valid messages much harder.

4.17 Bypass authentication or access control

187 This type of attack aims at getting unauthorised access to data residing on the smartcard respectively at performing operations which do not match the current life cycle state of processed data objects or of the Operating System. In particular, unauthorised reading or modification of personalisation data stored on the card, or a further (unauthorised) initialisation or personalisation of the product could be the target of such an attack scenario. This type of attack (which may also be whole program sequences) makes use of weaknesses in software implementation and is performed by a

[4] Even where a protocol is designed to be secure, it may be possible to use a replay attack if a further attack step (such as a perturbation) is used to avoid a check that would otherwise detect and reject the replayed commands.

logical or physical attack on the Operating System and its processed data. The tools used are protocol attacks, either e.g. man-in-the-middle, replay, command editing or using commands which are undefined or not allowed in the current life cycle state of the Operating System. Furthermore, logical and physical attacks manipulating the program flow, status information (as the life cycle state of objects and of the Operating System) and access rules for objects processed by the Operating System have to be taken into account.

4.17.1 Description of Attack

188 This type of attack aims to get unauthorised access to data residing on the smartcard respectively to perform operations, which do not match the current life cycle state of processed data objects or of the Operating System. As an example, such an attack aims to read or modify personalised data that reside on the card or targets to perform a further (unauthorised) initialisation or personalisation of the product.

189 Getting unauthorised access to data stored on the smartcard can be obtained by various techniques:

- Impersonating the other side of the communication (known as 'man-in-the-middle'),
- using timing differences (by capturing and replaying commands),
- trying command variations (either editing valid commands or
- finding undefined commands),
- manipulation of access rules themselves,
- circumvention or manipulation of the request and evaluation of access rules during program execution.

190 Executing commands that are not allowed in the current life cycle state of the Operating System or of a data object can be as well obtained by various techniques:

- manipulation of the current life cycle state itself,
- circumvention or manipulation of the request and
- evaluation of the current life cycle state during program execution, and
- trying command variations (either editing valid commands or finding undefined commands).

4.17.2 Effect of Attack

191 The effect of the attack is unauthorised access to data residing on the smartcard respectively the possibility to perform operations, which do not match the current life cycle state of the Operating System or of data objects processed by the Operating System. In particular, such an attack could lead to the disclosure of stored secret data or to a further (unauthorised) initialisation or personalisation of the product. The attacker knows the attack has succeeded by analyzing the answers the smartcard gives during the (following) communication.

192 In general, the described attack scenario aims at the manipulation of the intended security structure integrated in the Operating System, in the applications set up on this Operating System and in the (application) data processed by the Operating System. The integrated access control to data objects and commands is affected.

193 Replay attacks have existed for a long time. Years ago, replay attacks were aimed at stealing passwords. Given the encryption strength of passwords these days, the focus of this type of attack has shifted to stealing digital signatures and keys.

194 The command editing attack aims to find commands that are not documented or using valid commands in a way that breaks the communication mechanisms in the TOE. The attacker may try to find improper bounds checking by sending longer commands than the TOE expects. He may try to send commands with unexpected values, forcing the smartcard to dump memory contents.

195 The manipulation of life cycle state information and access rules themselves, and the manipulation of their request and evaluation can be considered as a direct attack on the access control implemented in the Operating System and the applications running on this platform. In particular, the access control is modified or completely switched off in a way that unauthorised access to secured data or the execution of not allowed commands is possible.

4.17.3 Characteristics of the Attack

196 The manipulations of life cycle state information and access rules require a physical attack on the smartcard and its Operating System and applications. The circumvention and manipulation of the request and evaluation of life cycle state information and access rules bases on a manipulation of the intended program flow what may be achieved by logical or physical means. An active countermeasure for securing life cycle state information and access rules and their request and evaluation during program execution could be to attach an integrity attribute and to check this attribute appropriately during program execution. More details concerning the characteristics of these attacks and effective countermeasures can be found in the sections 4.1 “Physical Attacks” and 4.3 “Perturbation Attacks”.

197 The attacks of logical nature as man-in-the-middle attacks, replay attacks, command editing are considered in detail in the sections Software Attacks 4.11.

4.18 Buffer overflow or stack overflow

198 This attack is applicable to open platforms.

199 Open platforms are defined in this document as smart card operating systems with the capability of running and downloading multiple applications.

200 Open platforms provide to the applications a set of services, in particular services to protect their sensitive data against external applications (unauthorized access and unexpected modification).

201 This attack could be performed through buffer overflow or stack overflow, produced by the execution of a malicious application.

202 Overflow, when not checked by the platform, can have various effects, such as overwriting existing content in the current stack.

203 The expected effect by the attacker here is the malicious application modifies the current execution context and switch to system privileges.

204 Gaining such privileges allow this application to virtually execute every operation and then disclose or modify secret data, e.g. modifying or disclosing the PIN of another application.

5 References

[AIS31] Functionality classes and evaluation methodology for physical random number generator, Version 1, 2001-09-25 and the associated technical document: “A proposal for: Functionality classes and evaluation methodology for true (physical) random number generator”, Version 3.1, 2001-09-25, W. Killmann (T-Systems), W. Schindler (BSI)

[CC] Common Criteria for Information Technology Security Evaluation, Version 2.3, August 2005.

Common Criteria for Information Technology Security Evaluation, Version 3.1, Revision 2, september 2007.

[CEM] Common Methodology for Information Technology Security Evaluation (CEM), Version 2.3, August 2005.

Common Methodology for Information Technology Security Evaluation (CEM), Version 3.1, Revision 2, september 2007.

[CC-IC] The Application of CC to Integrated Circuits, Version 3.0, Rev.1,March 2009, CC Supporting Document CCDB-2009-03-001.

[COMPO] Composite product evaluation for Smartcards and similar devices Version 1.0, Rev.1, September 2007, CC Supporting Document ,CCDB-2007-09-001.