

Automated Verification of Dynamic Root of Trust Protocols

Sergiu Bursuc¹, Christian Johansen^{2(✉)}, and Shiwei Xu³

¹ University of Bristol, Bristol, UK

² University of Oslo, Oslo, Norway
cristi@ifi.uio.no

³ Wuhan Digital Engineering Institute, Wuhan, China

Abstract. Automated verification of security protocols based on dynamic root of trust, typically relying on protected hardware such as TPM, involves several challenges that we address in this paper. We model the semantics of trusted computing platforms (including CPU, TPM, OS, and other essential components) and of associated protocols in a classical process calculus accepted by ProVerif. As part of the formalization effort, we introduce new equational theories for representing TPM specific platform states and dynamically loaded programs.

Formal models for such an extensive set of features cannot be readily handled by ProVerif, due especially to the search space generated by unbounded extensions of TPM registers. In this context we introduce a transformation of the TPM process, that simplifies the structure of the search space for automated verification, while preserving the security properties of interest. This allows to run ProVerif on our proposed models, so we can derive automatically security guarantees for protocols running in a dynamic root of trust context.

1 Introduction

A hardware root of trust, including dynamic measurement of programs and their protected execution, is a promising concept for ensuring the integrity of a platform and the privacy of sensitive data, despite powerful software attackers [19]. This relies on the idea that hardware is more difficult to compromise than software, and therefore, it can play a crucial role in protocols for handling sensitive data. When a secure computing platform is needed, a special sequence of instructions allows for a trusted piece of hardware to attest the integrity of the software to be run and to give access to data in a protected environment.

However, turning this idea into a secure design and implementation is not easy, as various attacks have shown [13, 29]. For more assurance, one could use models and tools that allow automated verification of desired properties against trusted computing protocols and implementations. One main challenge for automated verification is the size and number of components involved in running programs protected by a dynamic root of trust. Furthermore, messages of such protocols consist not only of data, but also of programs that are to be executed

on the platform, and that can be supplied by an attacker or by an honest participant. At the same time, modelling the platform configuration registers (PCR) of the trusted platform module (TPM) [20] poses problems, because PCRs can be extended an unbounded number of times. Even the most efficient symbolic methods struggle with the structure of the resulting search space [6, 12].

Our contributions. We propose a formal model in the ProVerif process calculus [7] for the technology and for the security properties of a dynamic root of trust (as instantiated by Intel’s Trusted Execution Technology or AMD’s Secure Virtual Machine). Our model is more realistic than [12] and it covers aspects of trusted computing that [10] does not cover (Sect. 4). We show how a platform state can be naturally represented as a term in ProVerif (or applied pi-calculus [1, 27]) and how operations on the platform state can be expressed as equations in a term algebra (Sects. 4.3 and 4.4). Furthermore, we show how to model the dynamic loading of protected programs. Our model is simple and does not require heavy encodings, being based on the classic idea of processes as data, with a twist to take protection into account (Sect. 4.2).

We propose a new abstraction to model the extension of PCR registers that allows automated verification for a larger class of protocols than in [12]. We show how to over-approximate the model of the TPM such that the structure of the search space is simplified, without losing possible attacks or introducing false attacks. The main idea is that we can let the attacker set the PCR to *any* value, as long as it is “big enough” (Sect. 5).

Putting the formalisation and the abstraction together, we obtain the first automated verification for a realistic model of a dynamic root of trust. As security properties, we prove code integrity (the PCR values correctly record the measurement of the platform) and secrecy of sealed data (only a designated program can access data that has been sealed for its use in a protected environment).

2 Related Work

A programming language and a logic for specifying trusted computing protocols and properties are proposed in [10]. The setting is quite expressive and it allows the analysis of protocols similar to the ones that we study in this paper. [10] does not consider the seal/unseal functions of the TPM, but their language could be extended to capture them. However, the formal analysis of [10] is manual, and considering the complexity of the proofs involved, the lack of automation can be a limitation. We also believe some of their axioms (like those linking the PCR values to a late launch action) could be decomposed into more atomic formulas, in closer relation to the computational platform. Their security properties include correctly reading PCR values and the ability of honest parties to launch roots of trust; our property of code integrity, modeled as a correspondence assertion, can be seen as an additional constraint for these two events.

The analysis of [12] is automated with ProVerif and is based on a Horn clause model. Microsoft’s Bitlocker protocol is shown to preserve the secrecy of data sealed against a static sequence of PCR values. Their model considers a static

root of trust, and cannot handle dynamically loaded programs. Furthermore, there is no way to express a program that has access to data in a protected environment. Without a richer model of platform states, code integrity properties cannot be expressed either. To help with automation, [12] shows that, for a specific class of Horn clauses, it is sound to bound the number of extensions of PCR registers. Since our model is in applied pi-calculus and our security properties are different, we cannot directly rely on their result, and we propose a new way of handling the unbounded PCR extension problem.

Information-flow security and computational models. [14] presents a secure compiler for translating programs and policies into cryptographic implementations, distributed on several machines equipped with TPMs. A computational model capturing functionalities similar to ours, in conjunction with additional features such as authenticated key exchange, was recently proposed in [5]. Our models are more abstract, yet could be related to particular implementations - a closer connections between formal and computational models could be explored in future.

Unbounded search space. Several works tackle the problem of an unbounded search space for automated verification, but technically they are all based on principles that cannot be translated to PCR registers. In [25], it is shown that, for a class of Horn clauses, verification of protocols with unbounded lists can be reduced to verification of protocols with lists containing a single element. In [9], it is shown that to analyse routing protocols it is sufficient to consider topologies with at most four nodes. These are strong results, based on the fact that the elements of a list or the nodes in a route are handled uniformly by the protocol. Similar results, in a different context, are shown in [15,16]. Their reductions are based on the principle of data independence for memory stores. In [22] and respectively [2], it is shown how to handle an unbounded number of Diffie-Hellman exponentiations and respectively reencryptions in ProVerif. Surprisingly, the underlying associative-commutative properties of Diffie-Hellman help in [22], while [2] can rely on the fact that a re-encryption does not change the semantics of a ciphertext. Another case where an unbounded number of operations is problematic is file sharing [8]. In order to obtain an automated proof, [8] assumes a bound on the number of access revocations, without providing justifications for soundness. A sound abstraction for an unbounded number of revocations, in a more general setting, is proposed in [24]. Still, it is specialized to databases and it seems to rely on the same principle as several results mentioned above: it does not matter what the data is, it only matters to what set it belongs.

Tools and models for non-monotonic state. StatVerif [3] is aimed specifically for the verification of protocols relying on non-monotonic states, encoding the semantics of applied pi-calculus enriched with states into a set of Horn clauses for input to ProVerif. Tamarin [28] is based on multiset rewriting and inherently allows specification and automated reasoning for non-monotonic states, where the set of facts can both augment and decrease. SAPIC [21] takes as input

a stateful variant of applied pi-calculus and produces a multiset-based model, which is then analysed using Tamarin.

StatVerif [3], SAPIC [21], and Tamarin directly [23], have been used with success to verify security protocols that rely on non-monotonic states or trusted hardware: *PKCS#11* for key management [26], YubiKey for user authentication [32], and protocols for contract signing [17]. Our models, on the other hand, are tailored for direct input to ProVerif, while extending the scope of formal models for platform state operations and dynamic root of trust protocols based on a TPM [18–20]. It is one of our main interests for future work to see how the models of this paper can be analysed with tools like [3, 21, 28], in order to obtain a closer alignment with the state semantics of real systems.

3 Preliminaries

3.1 Trusted Computing

We first describe the required computing platform (hardware and software) and then describe the considered class of dynamic root of trust protocols.

A. Computing platform. We consider a general purpose computing platform equipped with a CPU and a TPM (both trusted), as well as a generic untrusted operating system.

Trusted hardware. Trusted computing relies on the CPU and the TPM¹ to perform certain operations whose integrity cannot be compromised by any software attacker. Regarding the TPM, two of its trusted features are fundamental for the applications that we consider in this paper: the ability to record a chain of values in its *platform configuration registers* (PCR) and the ability to *seal data* against specified values of the PCR.

The TPM allows the PCR to be *reset* only by the CPU or by a system reset. On the other hand, the PCR can be *extended* with any value by software. If a PCR records a value p and is extended with a value v , the new value of the PCR is $h((p, v))$, i.e. the result of applying a hash function to the concatenation of p and v . Crucially, these are the only two ways in which the values of a PCR can be modified. The role of the PCR for the protocols that we consider in this paper is to store the measurement of programs, recording a chain of loaded programs. When data d is *sealed* against some specified value v of the PCR, the TPM stores d internally and can release it in future only if the value recorded in its PCR matches the value v against which d was sealed.

For the purpose of formal verification, we are flexible about who exactly of the CPU or the TPM is doing a trusted operation, like measuring, sealing, etc. This depends on the implementation, e.g., the Intel SGX can do all the operations of a TPM. Changing the formalization from this paper to fit a particular implementation should be easy.

¹ See recent book [4] detailing the TPM version 2.0 specification and implementations.

Privileged software. When a system interrupt is triggered (e.g. by network communication or user interface action), all physical memory can be accessed by the system management interrupt (SMI) handler. This means that any memory protection mechanism, in particular the protocols that we consider in this paper, must either disable interrupts for their whole duration (not practical in general) or else rely on the fact that the SMI handler cannot be compromised. That is why the SMI handler is stored in a memory area called SMRAM, which enjoys special hardware protection. Still, as shown in [13, 29], the security guarantees of trusted computing can be violated using the CPU caching mechanism to compromise the SMI handler. Roughly, these attacks work because the protection of the SMRAM is not carried on to its cached contents. A countermeasure against such attacks, that we also adopt in this paper at an abstract level, is a software transfer monitor (STM) [18]. It also resides in the SMRAM, but it cannot be cached while a dynamic root of trust is running (special registers of the CPU should ensure that), and its role is to protect some memory regions from the SMI handler.

B. Dynamic root of trust.

We consider the technology of dynamic measurement and protected execution, also called dynamic root of trust (DRT), as instantiated for example in Intel’s Trusted Execution Technology (TXT) or AMD Secure Virtual Machine (SVM), and as illustrated in Fig. 1.

The goal of DRT is to establish a protected execution environment for a program, where private data can be accessed without being leaked to an attacker that controls the operating system. Assume a program, that we will call PP (called `measured launch environment` on Intel and `secure kernel` on AMD), needs to be loaded in a protected environment. The first entry point of the DRT protocol is a trusted instruction of the CPU (called `GETSEC[SENTER]` on Intel and `SKINIT` on AMD), that takes as input the program PP. To help with the establishment of a protected environment, the CPU also receives as input another program, that we will call INIT (called `SINIT authenticated code module` on Intel and `secure loader` on AMD). The DRT launch and execution sequence can then be summarized as follows:

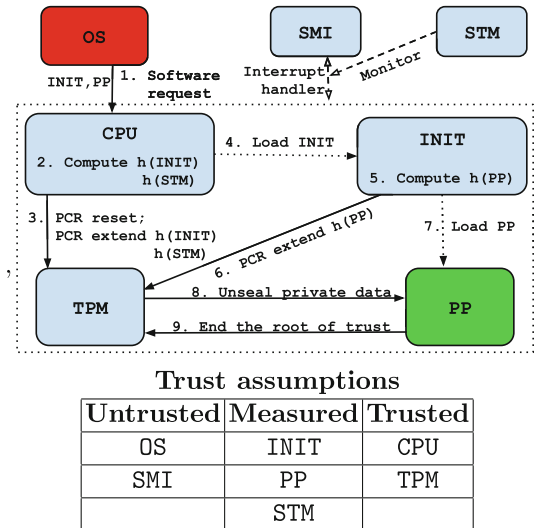


Fig. 1. Execution flow in DRT

The DRT launch and execution sequence can then be summarized as follows:

1. The CPU receives a request from the operating system containing the INIT code and the PP code. The system interrupts are disabled at this step, as an additional protection against untrusted interrupt handlers.
- 2-3. A software attacker that controls the operating system could compromise INIT and the STM, and that is why the CPU computes their measurement and extends the result into the TPM, to keep a trace of programs responsible for the DRT. Measuring a program means applying a hash function to its source code. This computation is performed on the CPU and is trusted, entailing that the resulting value is a correct measurement of INIT and STM. The CPU communicates with the TPM on a trusted channel and requests that the PCR is reset and extended with the resulting value $(h(\text{INIT}), h(\text{STM}))$.
- 4-7. The INIT program is loaded and it computes the measurement of the PP program, extending it into the PCR. The communication between INIT and the TPM is performed on a private channel established by the CPU. INIT also allocates protected memory for the execution of PP and loads it.
8. The PP program can re-enable interrupts once appropriate interrupt handlers are set. Furthermore, it can now request the TPM to unseal data that has been sealed against the current PCR value, and it can have access to that data in a protected environment. The communication between PP and the TPM is performed on a private channel established by the CPU.
9. Before ending its execution, the PP program extends the PCR with a dummy value, to record that the platform state is not to be trusted any more.

Since the OS is untrusted it can supply malicious programs INIT and PP. Therefore, INIT, PP and the STM are not trusted, but they are *measured*. If their measurement does not correspond to some expected *trusted* values, this will be recorded in the TPM and secret data will not be unsealed for this environment.

Security goals. Let us summarize the two main security goals of the DRT.

Code integrity: In any execution of the platform, if the measurements recorded in the PCR value of the TPM correspond to the sequence of programs $\mathcal{P}_{\text{INIT}}$, \mathcal{P}_{STM} , \mathcal{P}_{PP} , then the platform is indeed running a DRT for the protected execution of \mathcal{P}_{PP} in the context of $\mathcal{P}_{\text{INIT}}$ and \mathcal{P}_{STM} . In particular, this means that the programs \mathcal{P}_{PP} , $\mathcal{P}_{\text{INIT}}$ and \mathcal{P}_{STM} cannot be modified while a DRT is running.

Secrecy of sealed data: Any secret data that is sealed only against a PCR value recording the sequence of programs $\mathcal{P}_{\text{INIT}}$, \mathcal{P}_{STM} , \mathcal{P}_{PP} , is only available for the program \mathcal{P}_{PP} , in any execution of the platform.

3.2 ProVerif Process Calculus

We review ProVerif [6,7] and the special way in which we use (a restriction of) its input calculus in our modelling.

A. Terms, equational theories and deducibility. We consider an infinite set of *names*, a, b, c, k, n, \dots , an infinite set of *variables*, x, y, z, \dots and a possibly infinite set of *function symbols* \mathcal{F} . Names and variables are *terms*; new terms

are built by applying function symbols to names, variables and other terms. We split \mathcal{F} into two disjoint sets of *public* functions \mathcal{F}^{pub} and *private* functions $\mathcal{F}^{\text{priv}}$. Public functions can be applied by anyone to construct terms, including the attacker, whereas private functions can be applied only as specified by the protocol. When $\mathcal{F}^{\text{priv}}$ is not explicit, we assume that all functions are public.

A *substitution* σ is a partial function from variables to terms. The replacement of every variable x with $x\sigma$ in a term T is denoted by $T\sigma$. A *context* is a term $\mathcal{C}[_]$ that contains a special symbol $_$ in place of a subterm. For a context $\mathcal{C}[_]$ and a term T , we denote by $\mathcal{C}[T]$ the term obtained by replacing $_$ with T in $\mathcal{C}[_]$. For any formal object \mathcal{D} , we denote by $\text{sig}(\mathcal{D})$ the set of function symbols appearing in \mathcal{D} , and by $\text{top}(T)$ the outer-most function symbol in term T .

An equational theory \mathcal{E} is defined by a set of rewrite rules $U_1 \rightarrow V_1, \dots, U_n \rightarrow V_n$, where $U_1, \dots, U_n, V_1, \dots, V_n$ are terms with variables. A term U rewrites to V in one step, denoted by $U \rightarrow V$, if there is a context $\mathcal{C}[_]$, a substitution σ and an index $i \in \{1, \dots, n\}$ such that $U = \mathcal{C}[U_i\sigma]$ and $V = \mathcal{C}[V_i\sigma]$. Several rewrite steps from U to V are denoted by $U \rightarrow^* V$. We consider only convergent equational theories, i.e., for any term T there exists a unique non-reducible term $T\downarrow$ s.t. $T \rightarrow^* T\downarrow$. We write $U =_{\mathcal{E}} V$ iff $U\downarrow = V\downarrow$. ProVerif also allows operations on sequences: for all n , from any terms T_1, \dots, T_n , one can derive the term (T_1, \dots, T_n) , and conversely.

Deduction. Given an equational theory \mathcal{E} , a set of terms S and a term T , the ability of an attacker to obtain T from S is captured by the deduction relation $S \vdash_{\mathcal{E}} T$ (or simply $S \vdash T$ when \mathcal{E} is understood) defined as being true iff:

- there exists a term $T' \in S$ such that $T' =_{\mathcal{E}} T$, or
- there are terms T_1, \dots, T_n such that $S \vdash_{\mathcal{E}} T_1, \dots, S \vdash_{\mathcal{E}} T_n$ and a function symbol $f \in \mathcal{F}^{\text{pub}}$ such that $f(T_1, \dots, T_n) =_{\mathcal{E}} T$.

B. Processes and operational semantics. *Processes* of the calculus are built according to Fig. 2. Replication spawns instances of a process: $!P$ is formally equivalent with $P \mid !P$. Names introduced by **new** are called *bound* or *private*; they represent the creation of fresh data. Names that are not bound are called *free*, or *public*. The term T in an input $\text{in}(U, T)$ allows to specify filters for messages received on U : a message M will be accepted only if there is a substitution σ such that $M = T\sigma$. A variable x is *free* in a process P if P neither contains x in any of its input patterns nor does it contain any term evaluation of the form $x = T$. Consecutive term evaluations can be written together as **let** $(x_1, \dots, x_n) = (T_1, \dots, T_n)$ **in** P . The notions of substitution, contexts and normal forms translate to processes as expected.

Operational semantics is defined as a transition system on configurations of the form $(\mathcal{N}, \mathcal{M}, \mathcal{P})$, where: \mathcal{N} is a set of fresh names created during the execution of a process; \mathcal{M} is the set of terms made available to the attacker; and \mathcal{P} is the set of processes executing in parallel at a given point in time. We write $(\mathcal{N}, \mathcal{M}, \mathcal{P}) \rightarrow^* (\mathcal{N}', \mathcal{M}', \mathcal{P}')$ if the configuration $(\mathcal{N}', \mathcal{M}', \mathcal{P}')$ can be reached from $(\mathcal{N}, \mathcal{M}, \mathcal{P})$ in zero or more executions steps. Such a sequence of execution steps is called a trace of P .

$P, Q, R ::=$			
0	null process	$\text{in}(U, T); P$	message input on U
$P \mid Q$	parallel composition	$\text{out}(U, T); P$	message output on U
$!P$	replication	$\text{if } U = V \text{ then } P \text{ else } Q$	conditional
$\text{new } n; P$	name restriction	$\text{let } x = T \text{ in } P$	term evaluation

Fig. 2. Process algebra, with n a name, x a variable, and T, U, V terms.

C. Security properties. The ability of an attacker to learn a term T by interacting with a process P is denoted by $P \models \text{Att}(T)$, defined as true iff there exists a process Q , with $\text{sig}(Q) \cap \mathcal{F}^{\text{priv}} = \emptyset$, such that $(\mathcal{N}_{\text{init}}, \emptyset, \{P \mid Q\}) \rightarrow^* (\mathcal{N}', \mathcal{M}', \mathcal{P}')$ and $\mathcal{M} \vdash_{\mathcal{E}} T$, for some configuration $(\mathcal{N}', \mathcal{M}', \mathcal{P}')$. Intuitively, Q represents any computation that can be performed by the attacker.

A (simplified) *correspondence assertion* [7] is a formula of the form

$$\text{Att}(T) \implies \text{false} \quad \text{or} \quad \text{Att}(T) \implies (U = V).$$

For a correspondence assertion $\text{Att}(T) \implies \Phi$ as above, we have

$$P \models \text{Att}(T) \implies \Phi \quad \text{iff} \quad \forall \sigma. [(P \models \text{Att}(T\sigma)) \implies \Phi\sigma]$$

Correspondence assertions of the first type model the *secrecy* of T , while those of second type enforce the constraint $U = V$ for deducible terms matching the pattern T (typically the terms U, V will share variables with T).

4 Formalisation

Our formal specification for the trusted computing platform and protocols described in Sect. 3.1 assumes an attacker that controls the operating system and can execute a DRT any number of times, with any INIT and PP programs. Moreover, using the CPU cache, the attacker can compromise the STM and SMI handler, and use them to access protected memory. The attacker has access to all TPM functions. However, we assume that the attacker cannot compromise the CPU nor the TPM, and that the platform state can only be modified according to the equations that we present in Sect. 4.4.

We model a system state as a term that can be updated by the CPU process, the TPM process and, once it has been output on a public channel, by the attacker. Multiple system states can be explored in parallel by the attacker, whose knowledge monotonically accumulates the set of all reachable states. This is an abstraction with respect to a real platform, where the CPU and the TPM have their own internal state, part of a global, non-monotonic system state. We also have a simplified model of TPM sealing: in reality, it relies on encryption with a TPM private key and refers to a specific system state; in our model, it is represented by the pair of public/private functions `seal/unseal`. For unsealing, the TPM process will require the input of a system state and check that the corresponding unseal request is valid for that state.

4.1 Cryptographic Primitives and Platform Constants

To model cryptographic primitives and various constants on the platform state, we consider the signature $\mathcal{F}_{\text{data}}$, where $\mathcal{F}_{\text{data}}^{\text{priv}} = \{\text{unseal}/2\}$ and

$$\mathcal{F}_{\text{data}}^{\text{pub}} = \{\text{p}_s/0, \text{p}_d/0, \text{true}/0, \text{false}/0, \text{h}/1, \text{senc}/2, \text{sdec}/2, \text{seal}/2\}.$$

We also consider the set of rewrite rules $\mathcal{E}_{\text{data}}$:

$$\begin{aligned} \text{sdec}(\text{senc}(x_{\text{val}}, x_{\text{key}}), x_{\text{key}}) &\rightarrow x_{\text{val}} \\ \text{unseal}(\text{seal}(x_{\text{val}}, x_{\text{pcr}}), x_{\text{pcr}}) &\rightarrow x_{\text{val}} \end{aligned}$$

The constant p_d (resp. p_s) represents the result of a dynamic (resp. static) PCR reset. A dynamic reset marks the start of a dynamic root of trust, and can only be performed by the CPU. The functions **senc** and **sdec**, and the corresponding rewrite rule, model symmetric key encryption. The symbol **h** represents a hash function. Anyone can seal a value, while the corresponding rewrite rule and the fact that **unseal** is private ensure that a value can be unsealed only according to the specification of the TPM.

4.2 Dynamically Loaded Programs

To model the fact that arbitrary programs can be dynamically loaded on the platform state (e.g. for the roles of **INIT** and **PP**), we consider a new public function symbol **prog**/1 and an infinite signature of *private constants* $\mathcal{F}_{\mathcal{P}}$, containing a different constant n_P for every possible process P . Intuitively, the term **prog**(n_P) is a public and unique identifier for the program P . In a computational model, such an identifier can for example be obtained by hashing the source code of P . The first action of a process that models a program will be to output the corresponding program identity **prog**(n_P) on a public channel.

On the other hand, the constant n_P represents a *private entry point* for the program P . Specifically, we consider a private function **get_entry** and the rewrite rule **get_entry**(**prog**(x)) $\rightarrow x$. The idea is that a trusted loader of programs (the CPU in our case) has access to the private function **get_entry** and, using this rewrite rule, it can gain access to the private entry point of any program. Now, n_P can play the role of a private channel between the trusted loader and the loaded program. Furthermore, we can store program identifiers in the platform state, to record what programs are loaded. Then, we can rely on n_P to model the ability of certain loaded programs to affect the platform state (shown in Sect. 4.4). We denote by $\mathcal{E}_{\text{prog}}$ the equational theory defined in this subsection: $\mathcal{F}_{\text{prog}} = \{\text{prog}/1\} \cup \mathcal{F}_{\mathcal{P}}$, $\mathcal{E}_{\text{prog}} = \{\text{get_entry}(\text{prog}(x)) \rightarrow x\}$.

4.3 Platform State

To model a platform state, we consider the signature:

$$\mathcal{F}_{\text{state}} = \{\text{state}/4, \text{tpm}/1, \text{cpu}/2, \text{smram}/2, \text{drt}/3\}$$

where all the symbols of $\mathcal{F}_{\text{state}}$ are private. This ensures that a platform state can be constructed or modified only according to the specification, relying on equations that we present in Subsect. 4.4. Intuitively, a term of the form

$$\text{state}(\text{tpm}(T_{\text{PCR}}), \text{cpu}(T_{\text{INT}}, T_{\text{CACHE}}), \text{smram}(T_{\text{STM}}, T_{\text{SMIH}}), \text{drt}(T_{\text{INIT}}, T_{\text{PP}}, T_{\text{LOCK}}))$$

represents a platform state where:

- T_{PCR} is a term that represents the value of the PCR register of the TPM;
- T_{INT} is the value of a register of the CPU showing if interrupts are enabled;
- T_{CACHE} represents the contents of the CPU cache;
- T_{SMIH} represents the program for the SMI handler and STM represents the STM program, which are located in SMRAM ;
- T_{LOCK} is showing if a dynamic root of trust is running;
- T_{INIT} represents the INIT program;
- T_{PP} represents the protected program PP.

4.4 Read and Write Access

The read access is universal: any agent who has access to a platform state

$$\text{state}(\text{tpm}(T_{\text{PCR}}), \text{cpu}(T_{\text{INT}}, T_{\text{CACHE}}), \text{smram}(T_{\text{STM}}, T_{\text{SMIH}}), \text{drt}(T_{\text{INIT}}, T_{\text{PP}}, T_{\text{LOCK}}))$$

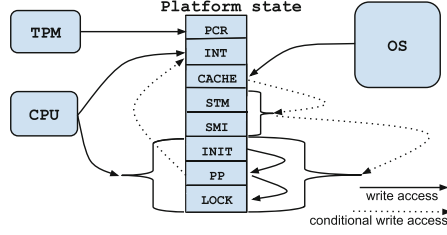
can read any of its components relying on the public unary function symbols $\mathcal{F}_{\text{read}} = \{\text{pcr}, \text{int}, \text{cache}, \text{stm}, \text{smi}, \text{init}, \text{pp}, \text{lock}\}$ and associated rewrite rules:

$$\begin{aligned} \text{pcr}(\text{state}(\text{tpm}(y), x_1, x_2, x_3)) &\rightarrow y \\ \text{int}(\text{state}(x_1, \text{cpu}(y_1, y_2), x_2, x_3)) &\rightarrow y_1 \\ \text{cache}(\text{state}(x_1, \text{cpu}(y_1, y_2), x_2, x_3)) &\rightarrow y_2 \\ \text{init}(\text{state}(x_1, x_2, \text{drt}(y_1, y_2, y_3), x_3)) &\rightarrow y_1 \\ \text{pp}(\text{state}(x_1, x_2, \text{drt}(y_1, y_2, y_3), x_3)) &\rightarrow y_2 \\ \text{lock}(\text{state}(x_1, x_2, \text{drt}(y_1, y_2, y_3), x_3)) &\rightarrow y_3 \\ \text{stm}(\text{state}(x_1, x_2, x_3, \text{smram}(y_1, y_2))) &\rightarrow y_1 \\ \text{smi}(\text{state}(x_1, x_2, x_3, \text{smram}(y_1, y_2))) &\rightarrow y_2 \end{aligned}$$

The write access to the platform state is restricted by the equational theory described and illustrated in Fig. 3, where tpm_acc and cpu_acc are private constants and all other new symbols are public.

PCR. Only the TPM can reset, extend or set the value of the PCR. This capability of the TPM is modeled by the *private constant* tpm_acc , which will be used only in the TPM process, described later in Fig. 4.

INT. The interrupts can be enabled or disabled by the CPU, whose capability is modeled by the *private constant* cpu_acc . Additionally, if a DRT is running, then the corresponding protected program PP also has the ability to enable or disable interrupts. This is modeled in the second set_int equation, by relying on the fact that, if $\text{prog}(x)$ represents the public identity of a program (as explained in Sect. 4.2), then x represents a private entry point for that program. Therefore, we can use x to model the ability of $\text{prog}(x)$ to change certain elements of the platform state when it is loaded.



$$\begin{aligned}
& \text{reset}(\text{state}(\text{tpm}(y), x_1, x_2, x_3), \text{tpm_acc}, p_s) \rightarrow \text{state}(\text{tpm}(p_s), x_1, x_2, x_3) \\
& \text{reset}(\text{state}(\text{tpm}(y), x_1, x_2, x_3), \text{tpm_acc}, p_d) \rightarrow \text{state}(\text{tpm}(p_d), x_1, x_2, x_3) \\
& \text{extend}(\text{state}(\text{tpm}(y), x_1, x_2, x_3), \text{tpm_acc}, v) \rightarrow \text{state}(\text{tpm}(h((y, v))), x_1, x_2, x_3) \\
& \text{set_pcr}(\text{state}(\text{tpm}(y), x_1, x_2, x_3), \text{tpm_acc}, v) \rightarrow \text{state}(\text{tpm}(v), x_1, x_2, x_3) \\
& \text{set_int}(\text{state}(x_1, \text{cpu}(y_1, y_2), x_2, x_3), \text{cpu_acc}, v) \rightarrow \text{state}(x_1, \text{cpu}(v, y_2), x_2, x_3) \\
& \text{set_int}(\text{state}(x_1, \text{cpu}(y, z), x_2, \text{drt}(z_1, \text{prog}(z_2), \text{true})), z_2, v) \\
& \quad \rightarrow \text{state}(x_1, \text{cpu}(v, z), x_2, \text{drt}(z_1, \text{prog}(z_2), \text{true})) \\
& \text{cache}(\text{state}(x_1, \text{cpu}(y_1, y_2), x_2, x_3), v) \rightarrow \text{state}(x_1, \text{cpu}(y_1, v), x_2, x_3) \\
& \text{flush_stm}(\text{state}(x_1, \text{cpu}(y_1, v), \text{smram}(z_1, z_2), \text{drt}(w_1, w_2, \text{false}))) \\
& \quad \rightarrow \text{state}(x_1, \text{cpu}(y_1, v), \text{smram}(v, z_2), \text{drt}(w_1, w_2, \text{false})) \\
& \text{flush_smi}(\text{state}(x_1, \text{cpu}(y_1, v), \text{smram}(z_1, z_2), x_2)) \\
& \quad \rightarrow \text{state}(x_1, \text{cpu}(y_1, v), \text{smram}(z_1, v), x_2) \\
& \text{set_init}(\text{state}(x_1, x_2, x_3, \text{drt}(y_1, y_2, y_3)), \text{cpu_acc}, v) \\
& \quad \rightarrow \text{state}(x_1, x_2, x_3, \text{drt}(v, y_2, y_3)) \\
& \text{set_pp}(\text{state}(x_1, x_2, x_3, \text{drt}(y_1, y_2, y_3)), \text{cpu_acc}, v) \\
& \quad \rightarrow \text{state}(x_1, x_2, x_3, \text{drt}(y_1, v, y_3)) \\
& \text{set_pp}(\text{state}(x_1, x_2, x_3, \text{drt}(\text{prog}(y_1), y_2, y_3)), y_1, v) \\
& \quad \rightarrow \text{state}(x_1, x_2, x_3, \text{drt}(\text{prog}(y_1), v, y_3)) \\
& \text{set_pp}(\text{state}(x, \text{cpu}(\text{true}, z), \text{smram}(\text{prog}(z_1), \text{prog}(z_2))), \text{drt}(y_1, y_2, y_3)), (z_1, z_2), v) \\
& \quad \rightarrow \text{state}(x, \text{cpu}(\text{true}, z), \text{smram}(\text{prog}(z_1), \text{prog}(z_2)), \text{drt}(y_1, v, y_3)) \\
& \text{set_lock}(\text{state}(x_1, x_2, x_3, \text{drt}(y_1, y_2, y_3)), \text{cpu_acc}, v) \\
& \quad \rightarrow \text{state}(x_1, x_2, x_3, \text{drt}(y_1, y_2, v)) \\
& \text{set_lock}(\text{state}(x_1, x_2, x_3, \text{drt}(y_1, \text{prog}(y_2), y_3)), y_2, v) \\
& \quad \rightarrow \text{state}(x_1, x_2, x_3, \text{drt}(y_1, \text{prog}(y_2), v)) \\
& \text{set_lock}(\text{state}(x, \text{cpu}(\text{true}, z), \text{smram}(\text{prog}(z_1), \text{prog}(z_2))), \text{drt}(y_1, y_2, y_3)), (z_1, z_2), v) \\
& \quad \rightarrow \text{state}(x, \text{cpu}(\text{true}, z), \text{smram}(\text{prog}(z_1), \text{prog}(z_2)), \text{drt}(y_1, y_2, v))
\end{aligned}$$

Fig. 3. Write access to the platform state.

CACHE. Any values can be cached. The cache values can then be copied into the contents of the SMI handler and, when a DRT is not running, into the STM component of the state.

INIT. Only the CPU has the ability to load an INIT program on the platform.

PP. The PP program can be loaded by the CPU (the first equation for `set_pp`) or by an INIT program, if the latter is already loaded on the platform (the second equation for `set_pp`). Furthermore, the SMI in conjunction with the STM can also modify the PP program, if the interrupts are enabled (the third equation for `set_pp`).

LOCK. Similarly, the DRT lock can be set/unset by the CPU, by the running PP, or by the SMI in conjunction with the STM, if the interrupts are enabled.

We denote by $\mathcal{E}_{\text{state}}$ the equational theory defined in this subsection.

4.5 Communication Channels

The public constant `os` models a communication channel for platform states and other messages that may be intercepted, modified or provided by the intruder as inputs to the CPU or the TPM. A private constant `cpu_tpm` models the secure channel between the CPU and the TPM. A private function `tpm_ch` models the ability of the CPU to establish a private channel between a loaded program and the TPM. Generally, these channels will be of the form `tpm_ch(prog(t))` and the CPU will send this term both to the program represented by `prog(t)` (on channel `t`) and to the TPM (on channel `cpu_tpm`). We also use message tags that will be clear from the context.

4.6 The Trusted Platform Module

We model the TPM by the process in Fig. 4. A PCR reset request can come either from the CPU, and then the PCR is reset to the value p_a marking a dynamic root of trust, or else from the operating system. A PCR extend request can come from the CPU, from the operating system or from a private channel that the CPU can establish between the TPM and some other process. To unseal a value, the TPM relies on the value of the PCR registers recorded in the platform state that is associated to an unseal

```

TPM      = !TPMRESET | !TPMEXTEND | !TPMUNSEAL
TPMRESET = let (ch, rv) = (cpu_tpm, pa) in !PCRRESET |
  let (ch, rv) = (os, ps) in !PCRRESET
PCRRESET = in(ch, (reset_req, nonce, pf_state));
  let new_st = reset(pf_state, tpm_acc, rv) in
  out(ch, (reset_resp, nonce, new_st))
TPMEXTEND = let ch = cpu_tpm in !PCREXTEND |
  let ch = os in !PCREXTEND |
  !(in(cpu_tpm, (ext_channel, ch)); !PCREXTEND)
PCREXTEND = in(ch, (extend_req, nonce, pf_state, v));
  let new_st = extend(pf_state, tpm_acc, v) in
  out(ch, (extend_resp, nonce, new_st))
TPMUNSEAL = in(os, pf_state);
  if lock(pf_state) = true then
    let ch = tpm_ch(pp(pf_state)) in UNSEAL
  else let ch = os in UNSEAL
UNSEAL    = in(ch, (tag_unseal, blob));
  let v = unseal(blob, pcr(pf_state)) in
  out(ch, (tag_plain, v))

```

Fig. 4. The TPM process

request. The corresponding equation for `unseal` ensures that this operation will succeed only if the PCR values from the state match the PCR values against which plain data was sealed. If a DRT is running, we perform the unseal for the protected program PP, on the private channel `tpm_ch(pp(pf_state))`; otherwise, the unsealed value is made public on channel `os`.

4.7 Dynamic Root of Trust: Launch

The procedure for launching a dynamic root of trust, i.e. steps 1–7 from Fig. 1, is modeled by the processes CPU and INIT, from Fig. 5. The CPU receives a request including the INIT and PP programs and the platform state where the DRT is to be launched. If a DRT is not already running in the corresponding platform state, then the CPU disables the interrupts and sets the DRT lock (step 1). Next, the CPU measures the INIT and STM programs and extends the result into the PCR (steps 2–3). In step 4a, the INIT program is loaded and we use the term $\text{tpm_ch}(\text{init})$ to model an established private channel between the TPM and the running INIT program. We use the program abstraction introduced in Sect. 4.2 to model the loading and the execution of INIT, relying on the private constant Tinit . In turn, the loaded INIT program measures the PP program, records the measurement into the TPM, and loads PP on the platform state (steps 4b–7a). After the INIT program has measured the PP program and loaded it into memory, the CPU gets back

the new platform state and sets up the private channel for communication between the loaded PP and the TPM (step 7b).

```

CPU = ! (** The CPU process **)
(* Step 1: receive a DRT request *)
in(os, (drt_req, init, pp, pf_state))
if lock(pf_state) = false then
let s'_0 = set_int(pf_state, cpu_acc, false) in
let s_0 = set_lock(s'_0, cpu_acc, true) in

(* Step 2: measure INIT and the STM *)
let measure = (h(init), h(stm(pf_state))) in

(* Step 3: reset and extend the PCR *)
new nonce; out(cpu_tpm, (reset_req, nonce, s_0));
in(cpu_tpm, (reset_resp, nonce, s_1));
out(cpu_tpm, (extend_req, nonce, s_1, measure));
in(cpu_tpm, (extend_resp, nonce, s_2));
(*Step 4a: load INIT & grant TPM access*)
let s_3 = set_init(s_2, cpu_acc, init) in
let einit = get_entry(init) in
out(einit, (nonce, s_3, tpm_ch(init), pp));
out(cpu_tpm, (ext_channel, tpm_ch(init)));
(* Step 7b: establish TPM access for PP *)
in(einit, (drt_resp, nonce, new_state));
let epp = get_entry(pp(new_state)) in
out(epp, (new_state, tpm_ch(prog(epp))));
out(cpu_tpm, (ext_channel, tpm_ch(prog(epp))))

INIT = (** A trusted INIT program **)
out(os, prog(Tinit)); out(os, prog(Tstm));
(* Step 4b: receive PP and TPM channel *)
in(Tinit, (nonce, pf_st, tpmc, pp));
(* Steps 5-6: extend h(PP) into PCR *)
let measure = h(pp) in newnonce_1;
out(tpmc, (extend_req, nonce_1, pf_st, measure));
in(tpmc, (extend_resp, nonce_1, ext_st));
(* Step 7a: load PP on platform state *)
let new_st = set_pp(ext_st, Tinit, pp) in
out(exp_init, (drt_resp, nonce, new_st));
out(os, new_st)

```

Fig. 5. DRT process for CPU and INIT

4.8 Dynamic Root of Trust: Execution

We illustrate the execution of a trusted PP program with an example in Fig. 6, where step 8 is an example of some useful execution of PP, i.e., unsealing and decrypting, whereas the rest is behaviour we expect from any protected program. The private constant `Tpp` represents the private entry point of PP according to the model from Sect. 4.2.

In Fig. 7 we consider a fresh symmetric key k_{pp} and assume that this key has been sealed against the measurement of the trusted PP program, with identity `prog(Tinit)`, of the trusted INIT program, with identity `prog(Tinit)`, and of the trusted STM program, with identity `prog(Tstm)`. This is represented by the term `sealed_key` in the process DATA (see the code in the figure below), which we publish on the channel `os`. We also assume that some private message hi_{pp} is encrypted with k_{pp} and `senc(hipp, kpp)` is made publicly available on channel `os`.

In the context of a DRT, the program PP should be able to unseal the key k_{pp} , decrypt and publish hi_{pp} . Before the execution of PP ends, the DRT lock is set to false, and also the PCR is extended with a dummy value in order to leave the PCR in a state which is not to be trusted any more. We verify, in Sect. 4.9, that secret DATA sealed for this program remains secret.

The SETUP process ties everything together, i.e., it loads and publishes an initial state, and runs any DRT request from the operating system. We call EXEC, all the processes put together, whereas the TPM is the one providing the trusted functionalities of reset, extend, and unseal. We use $DRT = (TPM \mid EXEC)$.

4.9 Security Properties in the Formal Model

Reachability. The reachability of a state in the platform can be expressed as a (non-)secrecy property: a state is reachable when a corresponding state term can be obtained by the attacker after interacting with the process DRT modulo the theory $\mathcal{E}_{drt} = \mathcal{E}_{data} \cup \mathcal{E}_{prog} \cup \mathcal{E}_{state}$, expressed as a formula of the form

$$DRT \models_{\mathcal{E}_{drt}} \text{Att}(\text{state}(T_{tpm}, T_{cpu}, T_{sram}, T_{drt})).$$

```

PP = (* Example of protected program *)
(* Step 7c: launch and get TPM access *)
out(os, prog(Tpp));
in(Tpp, (pf_state0, tpmc));
(* Re-enable interrupts *)
let pf_st = set_int(pf_state0, Tpp, true)
in out(os, pf_st);

(* Step 8: unseal and decrypt *)
in(os, x_seal); in(os, x_enc);
out(tpmc, (tag_unseal, x_seal));
in(tpmc, (tag_plain, x_k));
let mess = sdec(x_enc, x_k) in out(os, mess);

(* Step 9: Ending the execution *)
new rand; out(tpmc, (extend_req, rand, pf_st, ⊥));
in(tpmc, (extend_resp, rand, exts));
let ends = set_lock(exts, Tpp, false) in
out(os, ends)

```

Fig. 6. DRT execution

```

DATA = (* Seal and encrypt private data *)
new kpp; new hipp; out(os, senc(hipp, kpp));
let sealed_key = seal(kpp, hchain) in out(os, sealed_key);
(* where hchain = h(h(pd, (h(prog(Tinit)), h(prog(Tstm)))), h(prog(Tpp))) *)

SETUP = (* Launching the system *)
(* Load the initial state *)
in(os, xstm); in(os, xsmi);
out(os, state(tpm(pd, cpu(true, ⊥), smram(xstm, xsmi), drt(⊥, ⊥, false)));
(* Run a DRT with any loaded programs *)
in(os, init); in(os, pp); in(os, pf_state); out(os, (drt_req, init, pp, pf_state));
(* The main processes put together *)
EXEC = ( CPU | ! INIT | SETUP | DATA | ! PP )           DRT = ( TPM | EXEC )

```

Fig. 7. DRT setup and full process.

The property that the $DRT = (TPM \mid EXEC)$ process can reach an expected state where some trusted programs INIT and PP have been correctly measured and loaded on the platform can be expressed as follows:

$$\begin{aligned}
DRT \models_{\mathcal{E}_{art}} \text{Att}(\text{state}(\text{tpm}(\text{h}(\text{h}(\text{p}_d, v_1)), v_2)), \text{cpu}(\text{true}, x), \text{smram}(\text{prog}(\text{Tstm}), \text{prog}(y)), \text{drt}(\text{prog}(\text{Tinit}), \text{prog}(\text{Tpp}), \text{true})) \quad \text{where} \\
v_1 = (\text{h}(\text{prog}(\text{Tinit})), \text{h}(\text{prog}(\text{Tstm}))) \\
v_2 = \text{h}(\text{prog}(\text{Tpp})).
\end{aligned}$$

An additional reachability property of interest is whether the program PP has succeeded to unseal the key k_{pp} , decrypt the private message hi_{pp} and output it on the public channel os . This is captured by the following (non-)secrecy formula:

$$DRT \models_{\mathcal{E}_{art}} \text{Att}(hi_{pp}).$$

Code integrity. We say that the trusted platform ensures code integrity if the measurement contained in the PCR value correctly reflects the state of the platform. Specifically, we require that whenever a dynamic root of trust is active with a PCR value of p_d extended with the expected measurements v_1 and v_2 , then only the corresponding PP, INIT and STM are running on the platform, and they cannot be modified. This can be expressed by the following correspondence assertion, which we will *denote* by Φ_{int} in the rest of the paper:

$$\begin{aligned}
DRT \models_{\mathcal{E}_{art}} \text{Att}(\text{state}(\text{tpm}(\text{h}(\text{h}(\text{p}_d, v_1)), v_2)), \text{cpu}(x, y), \text{smram}(x_{stm}, x_{smi}), \\
\text{drt}(x_{init}, x_{pp}, \text{true})) \implies (x_{init}, x_{pp}, x_{stm}) = (p_1, p_2, p_3)
\end{aligned}$$

where $p_1 = \text{prog}(\text{Tinit})$, $p_2 = \text{prog}(\text{Tpp})$, $p_3 = \text{prog}(\text{Tstm})$.

Note that we ensure the property only for trusted programs. Indeed, if any of PP, INIT or STM are malicious, they could use their privileges to reach a platform state that does not reflect the PCR values. This is fine, because the PCR values will correctly record the identity of running programs in the chain of trust.

In particular, our property shows that untrusted DRT programs cannot make the PCR values record the measurement of trusted programs.

Secrecy of sealed data. We also verify that data sealed for PP, i.e. the key k_{pp} , remains secret (we denote this formula by Φ_{sec}):

$$(\Phi_{sec}) \quad \text{DRT} \models_{\mathcal{E}_{drt}} \text{Att}(k_{pp}) \implies \text{false}.$$

5 Process Transformation for Automated Verification

ProVerif does not terminate for the DRT process and the equational theory \mathcal{E}_{drt} . The main reason is the rewrite rule from \mathcal{E}_{state} that allows an unbounded number of PCR extensions, reflecting a problem first noticed in [12]. In this section, we propose a general transformation of processes that allows a more efficient exploration of the search space by ProVerif. The transformation is based on a general observation formalised in Proposition 1: we can replace a process P with a process Q as input for ProVerif, as long as Q and P are equivalent with respect to the security properties of interest. Concretely, we will replace the process DRT with a process DRT^b that bounds the number of PCR extensions, while allowing a direct way for the attacker to set the PCR to any value that is bigger than the considered bound.

For a process P , let $\text{Att}(P) = \{T \mid P \models \text{Att}(T)\}$ be the set of terms that can be obtained by the attacker when interacting with P . For a set of terms \mathcal{M} , we let $\text{Att}(\mathcal{M}) = \{T \mid \mathcal{M} \vdash T\}$. We notice the following.

Proposition 1. *Let P, Q be processes and $\text{Att}(T) \implies \Phi$ be a correspondence assertion such that, for any substitution σ ,*

$$T\sigma \in \text{Att}(P) \setminus \text{Att}(Q) \implies \Phi\sigma \text{ and } T\sigma \in \text{Att}(Q) \setminus \text{Att}(P) \implies \Phi\sigma.$$

Then we have: $P \models \text{Att}(T) \implies \Phi$ if and only if $Q \models \text{Att}(T) \implies \Phi$.

The proof of Proposition 1 follows immediately from definitions, yet this result is crucial to make our models amenable for ProVerif. We are thus allowed to transform the process DRT into a process DRT^b , that is equivalent to DRT with respect to code integrity and secrecy properties Φ_{int} and Φ_{sec} , and whose search space can be handled by ProVerif. It will be easier to express DRT^b using some additional rewrite rules. In conjunction with Proposition 1, we will then rely on the following result for soundness and completeness:

Proposition 2. *Let \mathcal{P} be a process, \mathcal{E} be an equational theory and $\text{Att}(T) \implies \Phi$ be a correspondence assertion. Assume \mathcal{E}^b is a set of rewrite rules such that $\forall U \rightarrow V \in \mathcal{E}^b : \text{top}(U) \in \mathcal{F}_{priv}$, i.e., is a private symbol. Then we have:*

$$P \models_{\mathcal{E}} \text{Att}(T) \implies \Phi \text{ if and only if } P \models_{\mathcal{E} \cup \mathcal{E}^b} \text{Att}(T) \implies \Phi.$$

Notation. We denoted a term of the form $h(\dots h((T_0, T_1)), \dots, T_n)$ by $\text{chain}(T_0, \dots, T_n)$, using $\text{chain}(T_0)$ for T_0 . We define $\text{length}(\text{chain}(T_0, \dots, T_n)) = n$, representing the number of extensions of a PCR.

Problematic rewrite rule. We recall the rewrite rule that poses non-termination problems for ProVerif:

$$\text{extend}(\text{state}(\text{tpm}(y), x_1, x_2, x_3), \text{tpm_acc}, v) \rightarrow \text{state}(\text{tpm}(h((y, v))), x_1, x_2, x_3)$$

Intuitively, ProVerif does not terminate because it is unable to make an abstract reasoning about the introduction of the term $h((y, v))$ in the right hand side of this rewrite rule. We propose a transformation of the TPM process into a process TPM^b that allows more values to be written into the PCR, overapproximating the effect of the problematic rewrite rule. This transformation will be sound and complete (satisfying the conditions of Proposition 1) based on the observation that, once it exceeds a certain bound, the value of the PCR does not matter for Φ_{sec} and Φ_{int} – thus, we can let the attacker have complete control over it.

Proposed transformation. For a given natural number b , we would like the following behaviour of the TPM^b process: if an extend request is received for a platform state $\text{state}(\text{tpm}(T_1), T_2, T_3, T_4)$ and a value V :

- if the length of the PCR is smaller than b , i.e. $\text{length}(T_1) < b$, then execute this request normally, using the function `extend`. The updated platform state returned by the TPM^b should now be $\text{state}(\text{tpm}(h((T_1, V))), T_2, T_3, T_4)$.
- if the length of the PCR value T_1 is greater or equal to b , i.e. $\text{length}(T_1) \geq b$, then output T_1 and V to the attacker and wait for a new value T'_1 as a response. If the length of T'_1 is big enough, i.e. $\text{length}(T'_1) > b$, the updated platform state returned by the TPM^b should now be $\text{state}(\text{tpm}(T'_1), T_2, T_3, T_4)$. In a normal execution, we would have $T'_1 = h((T_1, V))$. However, the attacker has the choice to set T'_1 to any value.

Formally, the TPM^b process relies on the private function `is_small` to detect if the value of the PCR is lower or higher than the bound, and treat the two cases differently. The following set of rewrite rules, for all $0 \leq i < b$, define `is_small`: $\text{is_small}(\text{chain}(v_0, \dots, v_i)) \rightarrow \text{true}$, where $v_0 \in \{\mathbf{p}_s, \mathbf{p}_d\}$ and v_1, \dots, v_i are mutually distinct variables. We also need to check if some value to be extended into the PCR is big enough. For this, we introduce the private function `is_big`, together with the rewrite rule: $\text{is_big}(\text{chain}(v_0, \dots, v_{b+1})) \rightarrow \text{true}$, where v_0, \dots, v_{b+1} are mutually distinct variables.

The only difference from the normal TPM process is in $\text{PCR}_{\text{EXTEND}}^b$, which first detects if the current value of the PCR is small or big: if it is small, the extension process proceeds normally (the process $\text{TPM}_{\text{EXTEND}}^{\text{SMALL}}$); if it is bigger than the given bound, then the TPM requests that the operating system combines `pcr` and `val` itself (the process $\text{TPM}_{\text{EXTEND}}^{\text{BIG}}$). Upon receiving the response from the `os`, the TPM first checks that the value provided is indeed big (the compromised operating system may be cheating). Only then, it updates the PCR to the requested value.

We denote by $\mathcal{E}_{\text{drt}}^b$ the equational theory \mathcal{E}_{drt} augmented with the rules for `is_small`, `is_big` and `set_pcr` introduced in this section and we assume that these new symbols are private (they are used only by TPM^b).

DRT^b	$= \text{TPM}^b \mid \text{EXEC}$
TPM^b	$= \text{TPM} \{ \text{PCR}_{\text{EXTEND}} \mapsto \text{PCR}_{\text{EXTEND}}^b \}$
$\text{PCR}_{\text{EXTEND}}^b$	$= \text{in}(\text{ch}, (= \text{extend_req}, \text{nonce}, \text{pf_state}, \text{val}));$ $\quad \text{let } \text{pcr} = \text{pcr}(\text{pf_state}) \text{ in}$ $\quad \text{if } \text{is_small}(\text{pcr}) = \text{true} \text{ then } \text{PCR}_{\text{EXTEND}}^{\text{SMALL}} \text{ else } \text{PCR}_{\text{EXTEND}}^{\text{BIG}}$
$\text{PCR}_{\text{EXTEND}}^{\text{SMALL}}$	$= \text{let } \text{new_st} = \text{extend}(\text{pf_state}, \text{tpm_acc}, \text{val}) \text{ in}$ $\quad \text{out}(\text{ch}, (\text{extend_resp}, \text{nonce}, \text{new_st}))$
$\text{PCR}_{\text{EXTEND}}^{\text{BIG}}$	$= \text{out}(\text{os}, (\text{pcr}, \text{val})); \text{in}(\text{os}, \text{new_pcr})$ $\quad \text{if } \text{is_big}(\text{new_pcr}) = \text{true} \text{ then}$ $\quad \text{let } \text{new_st} = \text{set_pcr}(\text{pf_state}, \text{tpm_acc}, \text{new_pcr}) \text{ in}$ $\quad \text{out}(\text{ch}, (\text{extend_resp}, \text{nonce}, \text{new_st}))$

5.1 Sketch of Correctness Proofs

We have to show that, for $\Phi \in \{\Phi_{\text{sec}}, \Phi_{\text{int}}\}$, we have $\text{DRT} \models_{\mathcal{E}_{\text{drt}}} \Phi \Leftrightarrow \text{DRT}^b \models_{\mathcal{E}_{\text{drt}}^b} \Phi$. We note that soundness (direction \Leftarrow) is the property that is necessary to derive the security guarantees for DRT, while completeness is secondary: it explains why we don't get false attacks against DRT^b with ProVerif. Since $\text{Att}(\text{DRT}) \subseteq \text{Att}(\text{DRT}^b)$, soundness is easy to prove, while completeness requires careful analysis of terms in $\text{Att}(\text{DRT}^b) \setminus \text{Att}(\text{DRT})$. We show that such terms are roughly limited to what we explicitly release in DRT^b : state terms with big PCR values; they cannot be used by the attacker to violate Φ_{sec} and Φ_{int} .

First, from Proposition 2 and the definition of $\mathcal{E}_{\text{drt}}^b$, we can easily translate between \mathcal{E}_{drt} and $\mathcal{E}_{\text{drt}}^b$, thus the notions and results that follow are modulo $\mathcal{E}_{\text{drt}}^b$.

Corollary 1. *For any Φ , we have $\text{DRT} \models_{\mathcal{E}_{\text{drt}}} \Phi \Leftrightarrow \text{DRT} \models_{\mathcal{E}_{\text{drt}}^b} \Phi$.*

Terms T with $\text{top}(T) = \text{state}$ are called state terms (or states). For a state term $T = \text{state}(\text{tpm}(T_1), \text{cpu}(T_2, T_3), \text{smram}(T_3, T_4), \text{drt}(T_5, T_6, T_7))$, we let $\text{Comp}(T) = \{T_1, \dots, T_7\}$. For a set of terms \mathcal{M}_1 , we say that a set of state terms \mathcal{M}_2 is \mathcal{M}_1 -saturated if for any $T \in \mathcal{M}_2$ we have $\forall U \in \text{Comp}(T) : \mathcal{M}_1 \vdash U$.

Lemma 1. *Let \mathcal{M}_1 be a set of terms and \mathcal{M}_2 be an \mathcal{M}_1 -saturated set of state terms. Then we have $\text{Att}(\mathcal{M}_1 \cup \mathcal{M}_2) = \text{Att}(\mathcal{M}_1) \cup \mathcal{M}_2$.*

Lemma 1 formalizes the intuition that, without access to TPM or CPU, the only operation that an attacker can perform on a state is to extract its components. The proof follows by a straightforward inspection of rewrite rules. To help in the sequel, we consider several restrictions of attacker's power against DRT^b :

- $\text{Att}_0(\text{DRT}^b)$ is the set of terms that can be obtained by an attacker interacting with DRT^b , while not being allowed to use terms in $\text{Att}(\text{DRT}^b) \setminus \text{Att}(\text{DRT})$ when constructing inputs for DRT^b . That is, $\text{Att}_0(\text{DRT}^b)$ can be seen as a passive attacker with respect to the additional functionality in DRT^b .
- $\text{Att}_1(\text{DRT}^b)$ is the knowledge of the previous attacker whose power is augmented with the ability to unseal terms from $\text{Att}_0(\text{DRT}^b)$, with TPM_UNSEAL , relying on state terms from $\text{Att}(\text{DRT}^b) \setminus \text{Att}(\text{DRT})$. This attacker is not allowed to use terms from $\text{Att}(\text{DRT}^b) \setminus \text{Att}(\text{DRT})$ in any other way.

- $\text{Att}_2(\text{DRT}^b)$ is the knowledge of a *state respecting* attacker against DRT^b : the attacker is given unrestricted access to DRT^b and can use any terms from $\text{Att}(\text{DRT}^b) \setminus \text{Att}(\text{DRT})$ to construct his inputs; however, the attacker can only use state terms according to the specification of an honest behaviour while interacting with the TPM, the CPU, or the equational theory.

Note that $\text{Att}_0(\text{DRT}^b) \subseteq \text{Att}_1(\text{DRT}^b) \subseteq \text{Att}_2(\text{DRT}^b) \subseteq \text{Att}(\text{DRT}^b)$. We denote by \mathcal{M}^b the set of state terms returned to the attacker by the $\text{PCR}_{\text{EXTEND}}^{\text{BIG}}$ process. Note that \mathcal{M}^b is an $\text{Att}(\text{DRT})$ -saturated set of state terms with $\forall T \in \mathcal{M}^b : \text{length}(\text{pcr}(T)) > b$.

Lemma 2. *For any b , we have $\text{Att}(\text{DRT}) \subseteq \text{Att}_0(\text{DRT}^b) \subseteq \text{Att}(\text{DRT}) \cup \mathcal{M}^b$.*

The first inclusion follows easily from the definition of DRT^b , which is able to simulate any normal PCR extension performed by DRT, without access to any terms in $\text{Att}(\text{DRT}^b) \setminus \text{Att}(\text{DRT})$. For the second inclusion, relying on the fact that \mathcal{M}^b is $\text{Att}(\text{DRT})$ -saturated, we use Lemma 1 to deduce $\text{Att}_0(\text{DRT}^b) \subseteq \text{Att}(\text{Att}(\text{DRT}) \cup \mathcal{M}^b) \subseteq \text{Att}(\text{DRT}) \cup \mathcal{M}^b$.

Lemma 3. *For $b \geq 2$, we have $\text{Att}_1(\text{DRT}^b) \subseteq \text{Att}_0(\text{DRT}^b)$.*

By definition, $\text{Att}_1(\text{DRT}^b) \setminus \text{Att}_0(\text{DRT}^b) \subseteq \{U \mid \text{seal}(U, V) \in \text{Att}_0(\text{DRT}^b)\}$. Note that the only sealed term in $\text{Att}_0(\text{DRT}^b)$ that does not originate from the attacker is $\text{seal}(k_{\text{pp}}, \text{hchain})$, with $\text{length}(\text{hchain}) = 2$. For any other term $\text{seal}(U, V) \in \text{Att}_0(\text{DRT}^b)$, we have $U \in \text{Att}_0(\text{DRT}^b)$, and therefore $U \notin \text{Att}_1(\text{DRT}^b) \setminus \text{Att}_0(\text{DRT}^b)$. From Lemma 2, the definition of $\text{TPM}_{\text{UNSEAL}}$, and the fact that $\forall T \in \mathcal{M}^b : \text{length}(\text{pcr}(T)) > b$, we also deduce that $k_{\text{pp}} \notin \text{Att}_1(\text{DRT}^b) \setminus \text{Att}_0(\text{DRT}^b)$, so we can conclude $\text{Att}_1(\text{DRT}^b) \subseteq \text{Att}_0(\text{DRT}^b)$.

Lemma 4. *For $b \geq 2$, we have $\text{Att}_2(\text{DRT}^b) \subseteq \text{Att}_1(\text{DRT}^b) \cup \mathcal{M}^b$.*

New terms $U \in \text{Att}_2(\text{DRT}^b)$ come from using a state term $V \in \text{Att}_1(\text{DRT}^b)$ in $\text{TPM}_{\text{RESET}}, \text{TPM}_{\text{EXTEND}}$ or CPU. From Lemmas 2 and 3, we have either $V \in \text{Att}(\text{DRT})$ or $V \in \mathcal{M}^b$. In both cases, we can show that $U \in \text{Att}_1(\text{DRT}^b) \cup \mathcal{M}^b$.

Corollary 2. *For $b \geq 2$, we have $\text{Att}(\text{DRT}) \subseteq \text{Att}(\text{DRT}^b) \subseteq \text{Att}(\text{DRT}) \cup \mathcal{M}^b \cup \mathcal{M}^f$, where \mathcal{M}^f is a set of terms such that any term $T \in \mathcal{M}^f$ contains a state term T' with $\text{pcr}(T') > b$.*

The set \mathcal{M}^f represents the additional terms that a non state respecting attacker can derive from \mathcal{M}^b . The property of \mathcal{M}^f is due to the fact that $\mathcal{E}_{\text{drt}}^b$ and the DRT^b process do not have effect on state terms that are used outside their intended scope. Such terms will end up as harmless subterms of attacker's knowledge.

Corollary 3. *For $b \geq 2$, DRT and DRT^b satisfy the conditions of Proposition 1 with respect to both Φ_{sec} and Φ_{int} .*

Corollary 2 shows that it is sufficient to check that conditions of Proposition 1 are satisfied for terms T in $\mathcal{M}^b \cup \mathcal{M}^f$. For Φ_{sec} , this follows from the fact that such terms T are either state terms, or contain state terms, and therefore the key k_{pp} cannot be among them. For Φ_{int} , this follows from the fact that those state terms have PCR lengths bigger than 2, while the precondition of Φ_{int} is a state term with PCR length 2. From Corollary 3 and Proposition 1, we deduce:

Corollary 4. *For $\Phi \in \{\Phi_{\text{sec}}, \Phi_{\text{int}}\}$, we have $DRT \models_{\mathcal{E}_{\text{drt}}^b} \Phi \Leftrightarrow DRT^b \models_{\mathcal{E}_{\text{drt}}^b} \Phi$.*

From Corollaries 1 and 4, we conclude:

Theorem 1. *For $\Phi \in \{\Phi_{\text{sec}}, \Phi_{\text{int}}\}$, $DRT \models_{\mathcal{E}_{\text{drt}}} \Phi \Leftrightarrow DRT^b \models_{\mathcal{E}_{\text{drt}}^b} \Phi$.*

6 Verification

The ProVerif code for the DRT^b process and the security properties defined in Sects. 4 and 5 is available online². It uses the equational theory $\mathcal{E}_{\text{data}} \cup \mathcal{E}_{\text{prog}} \cup \mathcal{E}_{\text{state}}^b$, with $b = 2$. The verification of each security property terminates in order of minutes, returning the expected result. From these results (implying there is no attack on DRT^b modulo $\mathcal{E}_{\text{drt}}^b$) and from Theorem 1 (implying there is no attack on DRT modulo \mathcal{E}_{drt}), we derive:

Theorem 2. *The DRT process satisfies, modulo $\mathcal{E}_{\text{data}} \cup \mathcal{E}_{\text{prog}} \cup \mathcal{E}_{\text{state}}$, the properties of code integrity and data secrecy defined in Sect. 4.9.*

In order to check the reachability properties $DRT \models \Phi$ defined in Sect. 4.9, we give $\neg(DRT \models \Phi)$ as input query for ProVerif - an attack with respect to this query would be a witness trace for the desired reachability property. When returning such a trace, ProVerif can either confirm that it is valid (*attack found*) or cannot confirm it. Our models fall in the latter case, and we have to further inspect the output trace to see how its steps can be used to reconstruct a valid trace: we do observe in the output trace the expected intermediary messages on the channels `cpu_tpm` and `os`, and we can follow the source of these messages up to a dynamic root of trust request, of whose validity we have to again make sure. By a similar analysis of attack traces returned by ProVerif, we can observe the attacks of [13, 29] in our models, when we allow the STM to be modified arbitrarily.

7 Further Work

While our model takes into account at an abstract level the attacks and mitigations of [13, 29], further refinements and soundness results are necessary in order to be able to conclude that attacks such as these or as [30, 31] are not possible in practice. We need to develop models that are abstract enough to allow clear specifications and automated reasoning, and realistic enough to capture for

² www.dropbox.com/s/cvq4op3w106868t/drt.pi (using ProVerif version 1.85).

instance implementation flaws. We plan to see how the models of this paper can be expressed in richer frameworks like StatVerif [3] and SAPIC [21], in order to capture more closely the state semantics of real platforms. We think the process transformation that we have presented in Sect. 5 is an instance of a more general result, whose exploration would also be fruitful for future applications.

Acknowledgements. We would like to thank Cas Cremers and several reviewers for helping improve this work.

References

1. Abadi, M., Fournet, C.: Mobile values, new names, and secure communication. In: Proceedings of the 28th ACM Symposium on Principles of Programming Languages (POPL 2001), pp. 104–115, January 2001
2. Arapinis, M., Bursuc, S., Ryan, M.D.: Reduction of equational theories for verification of trace equivalence: re-encryption, associativity and commutativity. In: Degano and Guttman [11], pp. 169–188
3. Arapinis, M., Ritter, E., Ryan, M.D.: StatVerif: verification of stateful processes. In: CSF, pp. 33–47. IEEE Computer Society (2011)
4. Arthur, W., Challener, D., Goldman, K.: A Practical Guide to TPM 2.0. APress, Berkeley (2015)
5. Barbosa, M., Portela, B., Scerri, G., Warinschi, B.: Foundations of hardware-based attested computation and application to SGX. In: IEEE European Symposium on Security and Privacy, EuroS&P, Saarbrücken, Germany, 21–24 March 2016, pp. 245–260. IEEE (2016)
6. Blanchet, B.: An efficient cryptographic protocol verifier based on Prolog rules. In: Computer Security Foundations Workshop (CSFW 2001) (2001)
7. Blanchet, B.: Automatic verification of correspondences for security protocols. *J. Comput. Secur.* **17**(4), 363–434 (2009)
8. Blanchet, B., Chaudhuri, A.: Automated formal analysis of a protocol for secure file sharing on untrusted storage. In: IEEE Symposium on Security and Privacy, pp. 417–431. IEEE Computer Society (2008)
9. Cortier, V., Degrieck, J., Delaune, S.: Analysing routing protocols: four nodes topologies are sufficient. In: Degano and Guttman [11], pp. 30–50
10. Datta, A., Franklin, J., Garg, D., Kaynar, D.: A logic of secure systems and its application to trusted computing. In: 30th IEEE Symposium on Security and Privacy, pp. 221–236. IEEE (2009)
11. Degano, P., Guttman, J.D. (eds.): POST 2012. LNCS, vol. 7215. Springer, Heidelberg (2012)
12. Delaune, S., Kremer, S., Ryan, M.D., Steel, G.: Formal analysis of protocols based on TPM state registers. In: Proceedings of the 24th IEEE Computer Security Foundations Symposium (CSF 2011), Cernay-la-Ville, France, pp. 66–82. IEEE Computer Society Press, June 2011
13. Dufлот, L., Grumelard, O., Levillain, O., Morin, B.: ACPI and SMI handlers: some limits to trusted computing. *J. Comput. Virol.* **6**(4), 353–374 (2010)
14. Fournet, C., Planul, J.: Compiling information-flow security to minimal trusted computing bases. In: Barthe, G. (ed.) ESOP 2011. LNCS, vol. 6602, pp. 216–235. Springer, Heidelberg (2011). doi:[10.1007/978-3-642-19718-5_12](https://doi.org/10.1007/978-3-642-19718-5_12)

15. Franklin, J., Chaki, S., Datta, A., McCune, J.M., Vasudevan, A.: Parametric verification of address space separation. In: Degano and Guttman [11], pp. 51–68
16. Franklin, J., Chaki, S., Datta, A., Seshadri, A.: Scalable parametric verification of secure systems: how to verify reference monitors without worrying about data structure size. In: IEEE Symposium on Security and Privacy, pp. 365–379. IEEE Computer Society (2010)
17. Garay, J.A., Jakobsson, M., MacKenzie, P.: Abuse-free optimistic contract signing. In: Wiener, M. (ed.) CRYPTO 1999. LNCS, vol. 1666, pp. 449–466. Springer, Heidelberg (1999). doi:[10.1007/3-540-48405-1_29](https://doi.org/10.1007/3-540-48405-1_29)
18. Grawrock, D.: Dynamics of a Trusted Platform: A Building Block Approach. Intel Press, Hillsboro (2009)
19. Trusted Computing Group. TCG Architecture Overview, Specification revision 1.4 (2007). www.trustedcomputinggroup.org
20. Trusted Computing Group. TPM main specification (2011). www.trustedcomputinggroup.org
21. Kremer, S., Künnemann, R.: Automated analysis of security protocols with global state. In: IEEE Symposium on Security and Privacy, pp. 163–178. IEEE Computer Society (2014)
22. Küsters, R., Truderung, T.: Using ProVerif to analyze protocols with Diffie-Hellman exponentiation. In: 22nd IEEE Computer Security Foundations Symposium (CSF), pp. 157–171. IEEE Computer Society (2009)
23. Meier, S.: Advancing automated security protocol verification. PhD Thesis, ETH Zürich (2013)
24. Mödersheim, S.: Abstraction by set-membership: verifying security protocols and web services with databases. In: Al-Shaer, E., Keromytis, A.D., Shmatikov, V. (eds.) ACM Conference on Computer and Communications Security, pp. 351–360. ACM (2010)
25. Paiola, M., Blanchet, B.: Verification of security protocols with lists: from length one to unbounded length. In: Degano and Guttman [11], pp. 69–88
26. RSA Security Inc., v2.20. PKCS #11: Cryptographic token interface standard, June 2004
27. Ryan, M.D., Smyth, B.: Applied pi calculus. In: Cortier, V., Kremer, S. (eds.) Formal Models and Techniques for Analyzing Security Protocols, Cryptology and Information Security Series. IOS Press (2011)
28. Schmidt, B., Meier, S., Cremers, C.J.F., Basin, D.A.: Automated analysis of Diffie-Hellman protocols and advanced security properties. In: Chong, S. (ed.) 25th IEEE Computer Security Foundations Symposium (CSF), pp. 78–94. IEEE Computer Society (2012)
29. Wojtczuk, R., Rutkowska, J.: Attacking INTEL trusted execution technology. In: Black Hat DC (2009)
30. Wojtczuk, R., Rutkowska, J.: Attacking INTEL TXT via SINIT code execution hijacking. Invisible Things Lab (2009)
31. Wojtczuk, R., Rutkowska, J., Tereshkin, A.: Another way to circumvent INTEL trusted execution technology. Invisible Things Lab (2009)
32. Yubico, A.B.: Kungsgatan 37, 111 56 Stockholm Sweden. The YubiKey manual - Usage, configuration and introduction of basic concepts (version 2.2) (2010)