KLEESPECTRE: Detecting Information Leakage through Speculative Cache Attacks via Symbolic Execution

GUANHUA WANG*, National University of Singapore
SUDIPTA CHATTOPADHYAY*, Singapore University of Technology and Design
ARNAB KUMAR BISWAS, National University of Singapore
TULIKA MITRA, National University of Singapore
ABHIK ROYCHOWDHURY†, National University of Singapore

Spectre-style attacks disclosed in the early 2018 expose data leakage scenarios via cache side channels. Specifically, speculatively executed paths due to branch mis-prediction may bring secret data into the cache which are then exposed via cache side channels even after the speculative execution is squashed. Symbolic execution is a well known test generation method to cover program paths at the level of the application software. In this paper, we extend symbolic execution with modelling of cache and speculative execution. Our tool KLEESPECTRE, built on top of the KLEE symbolic execution engine, can thus provide a testing engine to check for the data leakage through cache side channel as shown via Spectre attacks. Our symbolic cache model can verify whether the sensitive data leakage due to speculative execution can be observed by an attacker at a given program point. Our experiments show that KLEESPECTRE can effectively detect data leakage along speculatively executed paths and our cache model can make the leakage detection more precise.

CCS Concepts:
• Software and its engineering → Software verification and validation;
• Security and privacy → Software and application security;
• Computer systems organization → Architectures.

Additional Key Words and Phrases: Spectre attacks; Symbolic execution; Cache side-channel; Software security

ACM Reference Format:

1 INTRODUCTION

Speculative execution in modern superscalar microprocessors improves the program performance (by reducing execution time and by increasing throughput) compared to a non-speculative processor by predicting both the outcome and the target of branching instructions. The processor continues executing instructions after the branch where the number of speculatively executed instructions depends on how soon the actual branch condition is evaluated and also on the size of the buffer that holds the resulting states during speculative execution.

* Joint first author.
† Corresponding Author

Authors’ addresses: Guanhua Wang, National University of Singapore; Sudipta Chattopadhyay, Singapore University of Technology and Design; Arnab Kumar Biswas, National University of Singapore; Tulika Mitra, National University of Singapore; Abhik Roychoudhury, National University of Singapore, abhik@comp.nus.edu.sg.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2020 Association for Computing Machinery.
1049-331X/2020/2-ART1 $15.00
https://doi.org/10.1145/nnnnnnn.nnmmnnnn

If the prediction of a branching instruction is incorrect, all effects due to the speculatively executed instructions after the branch instruction are rolled back. To this end, the buffer and pipeline stages are flushed which hold these instructions or their results. However, if the cache content is also modified due to speculatively executed load instructions, the cache state is not fully rolled back. This opens up the possibility of a cache side-channel through which an attacker can obtain sensitive information from a user who shares the same platform with the attacker. The family of Spectre attacks [29] shows that this vulnerability is present in all modern general purpose processors. Such a vulnerability thus poses major concerns from the stand-point of software security.

Symbolic execution [26] is a well-known path exploration method that can be used for program testing and verification. Given a program with un-instantiated or symbolic inputs, it constructs a symbolic execution tree by expanding both directions of every branch whose outcome depends on symbolic variable(s). The leaf nodes of the tree correspond to program paths, and by solving the constraint accumulated along a program path (also called a path condition), a test input to explore the path can be generated.

Symbolic execution can be used to cover program paths (modulo a time budget). However, it does not consider behaviors induced by performance enhancing features of the underlying processor, specifically cache and branch prediction. Due to branch mis-prediction, certain paths may be speculatively executed and then squashed. Such speculatively executed paths are not covered in symbolic execution. As shown via Spectre attacks, such speculative paths may read sensitive data, which can then be potentially exfiltrated by attackers via cache side-channels. Current generation symbolic execution engines, as embodied by tools like KLEE [8] do not demonstrate the presence or absence of such side-channel scenarios. This is because the reasoning in current day symbolic execution engines is solely at the program level. Although a few recent works aim to enhance symbolic execution with the support of speculative semantics [23] or cache model [7], none of these works comprehensively model both the speculative execution and caches to accurately detect information leakage via Spectre-style attacks.

In this paper, we extend symbolic execution with the modelling of speculative execution as well as cache accesses. For an unresolved branch involving a symbolic variable, classical symbolic execution considers two possibilities - the branch is either taken or not taken. In the presence of speculative execution, note that for every unresolved branch we need to consider four possibilities, namely: taken and correctly predicted, taken and mis-predicted, not taken and correctly predicted, not taken and mis-predicted. As explained earlier, since the mis-predicted execution paths are squashed, they only need to be considered in symbolic execution in the presence of cache modelling. We model the behavior of the cache by capturing memory accesses to concrete or symbolic memory addresses; the symbolic memory accesses occur when the accessed memory address depends on a symbolic input such as accessing array element a[i] when i is a symbolic input variable. Given such symbolic memory accesses, the possible cache conflicts (two memory accesses to the same cache set) can be captured as a symbolic formula. By solving such symbolic formula, we can enunciate whether a secret brought into cache in a speculative path continues to linger in the cache (this is when it has not been evicted from the cache due to cache conflicts). Hence we can detect and infer the cache side-channel leakage in Spectre attacks.

The remainder of the paper is organized as follows. After providing a brief background (Section 2) and overview (Section 3) of KLEESpectre, we make the following contributions:

(1) We present KLEESpectre, our methodology to extend state-of-the-art symbolic execution engines with micro-architectural features, specifically speculative execution and caches (Section 4).
(2) We present a symbolic cache model embodied in KLEESpectre to precisely detect and highlight cache side-channel leakage through speculative execution paths, resulting in potential Spectre style attacks (Section 4).

(3) We implement our KLEESpectre approach on top of state-of-the-art and widely used symbolic virtual machine KLEE (Section 5). Our implementation and all experimental data are publicly available: https://github.com/winter2020/kleespectre.

(4) We evaluate KLEESpectre on litmus tests provided by Kocher [28] as well as on real-world cryptographic programs from libTomCrypt, Linux-tegra, openssl and hpn-ssh. Our evaluation reveals that KLEESpectre can effectively and efficiently detect Spectre vulnerable code. Moreover, the cache modelling embodied in KLEESpectre results in a precise leakage detection by ruling out false positives.

After discussing the related work (Section 8), we conclude in Section 9.

2 BACKGROUND AND THREAT MODEL

In this section, we introduce the necessary background regarding the speculative execution and our targeted threat model.

Speculative execution. Speculative execution [22] is an indispensable micro-architectural optimization for performance enhancement in modern superscalar processors. Speculative execution allows the processor pipeline to continue execution even in the presence of some data or control dependency between the current instruction and the unfinished instructions instead of stalling the pipeline. The branch predictor is one of the prediction units in a processor that supports speculative execution. The branch predictor predicts the execution path based on the history of the executed branch instructions. The processor stores a record of the speculatively executed instructions in a so-called Reorder Buffer (ROB). This buffer mainly helps the processor to commit all instructions in-order though they are executed out-of-order. If the outcome of a branch prediction is correct, then the instructions in ROB are committed to the architectural state, otherwise, the results of these instructions are squashed. However, the effect of the load execution unit i.e. the bytes that are read from memory during speculative execution may reside in the cache. The state of the cache is usually not squashed due to performance reasons. Thus, for a mis-predicted branch, even though the functional effects of all speculatively executed instructions are rolled back, the cache state may hold unexpected memory addresses. This phenomenon opens the potential vulnerability of cache side-channel attack.

Bounds Check Bypass (BCB) attack. Spectre-style attacks have proven that the computer can leak secret data through the cache side-channel when it performs the speculative execution. Bound Check Bypass (BCB, also called Spectre variant 1) attack is one such Spectre attack. In this attack, an attacker can induce the victim program to transfer the sensitive data to side-channel by training the branch predictor. Then the data can be recovered by the attacker through performing cache side-channel attack.

Listing 1. Example code of Spectre variant 1.

```c
if (x < array1_size) {
    y = array1[x]; \ VB
    temp |= array2[y * 64]; \ LS
}
```

Listing 1 shows an example code vulnerable to BCB attacks. In this example code, if the condition `x < array1_size` holds, then the statement at line 2 loads `array1[x]` to variable `y`. Finally, the
statement at line 3 reads data from array2[] where the accessed address depends on array1[x]. Normally, the boundary check at line 1 guarantees the absence of out-of-bound memory access. However, in the presence of the speculative execution, such guarantees do not hold. For example, the mis-prediction of the branch instruction at line 1 allows a memory access array1[x] where x ≥ array1 size. Such a memory access may point to a sensitive value. Thus, y may hold a sensitive value when the branch is mis-predicted. Finally, the statement at line 3 changes the cache state using the potential sensitive value y. By observing this cache state, the attacker can reconstruct the potentially sensitive value y. For simplicity, we name the branch potentially causing the BCB attack as Vulnerable Branch (VB), the instruction loading the potential sensitive data as Read Secret (RS) (e.g. statement at line 2) and the instruction leaking the sensitive data to cache state as Leak Secret (LS) (e.g. statement at line 3).

Threat model. Similar to the existing literature on cache side-channel attacks [31], in this work, we assume the victim and the attacker coexist on a machine, and they share the cache. The attacker can execute any code in its security domain (e.g., a process or a virtual machine) and it can learn information from the shared cache side-channel. Besides, in our threat model, we do not consider the data leakage in the normal execution path. Instead, we focus on data leakage only due to the speculative execution.

We assume that all conditional branches in a program are potentially vulnerable. If a branch mis-prediction is intentionally arranged by an attacker, then we call this branch to be mis-trained. This is in line with the existing works on Spectre-style attacks [9] that show the possibility of a branch to be mis-trained either by the victim process or outside the victim process (e.g., by an attacker-controlled process). As a result, any branch in the victim process is potentially vulnerable to mis-training by the attacker. To consider the implication of our threat model, consider the code in listing 2. Since the conditional branch at line 6 is unsatisfiable, the code at lines 7–8 will never be executed without speculation. However, in our considered threat model, the code at lines 7–8 can leak data if the branch at line 6 is mis-trained and the branch is subsequently mis-predicted (thus pointing outside the array bound of array1). We also note that neither the branch nor the memory access at line 7 is controlled by any external input.

Finally, we assume that the attacker can either perform the access-based cache side-channel attack or the trace-based cache side-channel attack [38]. The ability of such attackers depends on the execution points at which (s)he observes cache states. In particular, the access-based attack assumes that an attacker can probe the cache only upon the termination of a program. On the contrary, the trace-based attack assumes that an attacker can snoop the cache after any executed instruction from the victim process. It is worthwhile to note that in general, the attacker capability might fall...
anywhere between observing the cache state after each and every instruction to observing the state only when the execution terminates. We explicitly evaluate the case where the attacker observes at the end of execution. However extending our analysis, to consider attackers who observe cache states at some intermediate execution point(s), is straightforward. More specifically, as the cache model of KLEESpectre captures the memory operation addresses along the execution path, the leakage verification can be performed on the currently stored memory addresses at any customized program point.

3 OVERVIEW

Intuitively, KLEESpectre is an effort to consider and expose the micro-architectural execution semantics at the software layer. Specifically, KLEESpectre enhances the machinery of symbolic execution with branch speculation and cache modelling. In the following, we will use a running example to show the motivation behind the design of KLEESpectre and briefly outline the KLEE-Spectre work-flow. We use the term normal execution to capture the execution semantics embodied in classic symbolic execution tools.

The example: We consider the example code shown in Figure 1(a). The variable x is a user-controlled input. The code performs several memory related operations on two arrays array1 and array2. Although x is user controlled, we note that the access to array1[x] is protected by the bound check (i.e. x < SIZE). Thus, considering the normal execution, the example does not exhibit any out-of-bound access. Figure 1(b) captures the execution tree generated by any classic symbolic execution tool.

Enhancing symbolic execution: Consider the code fragments labelled A in Figure 1(a). Such a code has the following problems that only appear in the presence of branch speculation. Assume the value of the user controlled input is such that x ≥ SIZE. If the branch b1 is mis-predicted, then the memory access array1[x] exhibits an out-of-bound reference. Moreover, if array1[x] captures a sensitive value (e.g., a secret), then the subsequent memory access array2[array1[x]] (cf. Figure 1(a)) refers to a memory address dependent on secret value. Memory addresses that

![Fig. 1. The example code, and its normal execution paths along with the execution paths with branch speculation. (a) example code where b1 and b2 capture branch instructions. A, B and C indicate the corresponding basic blocks. (b) Execution paths explored by classic symbolic execution. pT1, pF1 represent normal paths that go along the true or false leg of a branch. (c) Symbolic execution tree explored by KLEESpectre. spT1, spF1 denote speculative paths that go along the true or false leg of a branch. The node in red color indicates the basic block with potential data leakage.](image-url)
depend on secret values are potentially exposed to cache side-channel attacks. For example, consider the access-based attacker who probes the state of the cache after the end of execution. For such an attack, the attacker might be successful to exfiltrate the value of \texttt{array1[x]} (potentially holding a sensitive value) only if \texttt{array2[array1[x]]} remains in the cache after the execution.

It is evident from the preceding example that detecting the potential leakage of \texttt{array1[x]} is beyond the capability of classic symbolic execution. Specifically, to detect this side-channel scenario, it is crucial to capture both the branch speculation and the cache behaviour while exploring symbolic execution states. In KLEESpectre, we enhance the power of symbolic execution along these two dimensions.

**Speculative symbolic execution in KLEESpectre:** In KLEESpectre, the purpose of speculative symbolic execution is to explore any potential secret that might be accessed due to branch speculation. To investigate the mechanism, consider again the example in Figure 1(a). To incorporate the branch along the normal execution path. Moreover, along the speculative path, we only record memory that are potentially vulnerable to a cache side-channel attack. Each such memory access sequence may involve at least one memory access along the speculative path and multiple memory accesses along the normal execution path. Moreover, along the speculative path, we only record memory.

\[ sp_{T1}: x < SIZE \] is satisfiable and the branch \( b_1 \) is correctly predicted. In this case, the symbolic execution will fork a new state with constraint \( x < SIZE \) and proceeds by executing the code fragment \( A \).

\[ sp_{F1}: x \geq SIZE \] is satisfiable and the branch \( b_1 \) is correctly predicted. In this case, the symbolic execution will fork a new state with constraint \( x \geq SIZE \) and proceeds by executing the code fragment \( C \).

\[ sp_{T1}: x \geq SIZE \] is satisfiable and the branch \( b_1 \) is mis-predicted. In this case, KLEESpectre forks a new state with constraint \( x \geq SIZE \), but proceeds by executing the code fragment \( A \).

\[ sp_{F1}: x < SIZE \] is satisfiable and the branch \( b_1 \) is mis-predicted. KLEESpectre forks a new state with constraint \( x < SIZE \), but proceeds by executing the code fragment \( C \).

\( sp_{T1} \) and \( sp_{F1} \) are the additional symbolic states explored by KLEESpectre at branch \( b_1 \). Figure 1(b) and (c) capture the symbolic execution trees explored by normal symbolic execution and KLEESpectre, respectively, for the code in Figure 1(a).

The symbolic execution along a speculative path spans across only a limited number of instructions. This is because the maximum number of speculatively executed instructions is bounded by the size of the re-order buffer (ROB). In KLEESpectre, we use Speculative Execution Window (SEW) to limit the number of speculatively executed instructions at any branch. It is worthwhile to note that a speculatively executed path may still span over multiple branch instructions (cf. Figure 1(c)) despite the limited size of SEW.

KLEESpectre prunes speculative symbolic states if they do not pose any risk of data leakage. For example, in Figure 1(c), only the execution of code fragment \( A \) under the branch \( sp_{T1} \) exhibits such risk. This is due to the access of array elements \( array1[x] \) and \( array2[array1[x]] \). Also note that, the symbolic states \( sp_{T3}, sp_{F3}, sp_{T4} \) and \( sp_{F4} \) are all discarded once KLEESpectre reaches the limit of speculation window SEW. In this fashion, KLEESpectre can control the explosion in the number of symbolic states due to speculation. Specifically, for the example in Figure 1(c), KLEESpectre only keeps the record of executing code \( A \) under the speculative state \( sp_{T1} \).

The next stage of KLEESpectre computes whether the secret accessed in \( sp_{T1} \) can potentially be exfiltrated by a cache side-channel attacker.

**Cache modelling in KLEESpectre:** KLEESpectre computes the set of memory access sequences that are potentially vulnerable to a cache side-channel attack. Each such memory access sequence may involve at least one memory access along the speculative path and multiple memory accesses along the normal execution path. Moreover, along the speculative path, we only record memory.
accesses that are dependent on secret. This is because KLEESpectre focuses to discover data leakage due to branch speculation.

For example, in Figure 1(c), KLEESpectre computes the following sequence of memory accesses for inspecting the leakage of data:

\[
\langle \langle A, x > SIZE, &array2[array1[x]] \rangle, \langle C, x > SIZE, &array3[8] \rangle \rangle
\]

The triplet \(\langle A, x > SIZE, &array2[array1[x]] \rangle\) captures that the address \&array2[array1[x]] was accessed with the symbolic constraint \(x > SIZE\) in code fragment \(A\). The sequence of memory accesses capture the accesses in the speculative state \(sp_{T1}\) followed by a memory access in the normal state \(p_{F1}\) (cf. Figure 1(c)). Even though the functional states in \(sp_{T1}\) do not affect the computation in \(p_{F1}\), the cache state influenced in \(sp_{T1}\) remains unchanged when the branch is resolved and the execution continues through code fragment \(C\) (cf. Figure 2).

Through our cache modelling, we check the presence of the address \&array2[array1[x]] in the cache when the code segment \(C\) finishes execution. To this end, we check whether memory access \&array3[8] can replace \&array2[array1[x]] from the cache. For the sake of simplicity, let us assume a 1-way associative (i.e. direct-mapped) cache. For direct-mapped caches, a memory address maps to exactly one cache line. In particular, the following symbolic condition is satisfiable if and only if the terminating cache state holds the memory address \&array2[array1[x]]:

\[
(x > SIZE) \land (set(\&array2[array1[x]]) \neq set(\&array3[8])) \\
\lor (tag(\&array2[array1[x]]) = tag(\&array3[8]))
\]

where \(set(x)\) and \(tag(x)\) capture the cache line and cache tag, respectively, for a memory address \(x\). Intuitively, the constraints in Formula 1 can be presented to a satisfiability modulo theory (SMT) solver. KLEESpectre formulates such constraints for each memory access sequence that may access secrets along a speculative path. These constraints are then discharged by an SMT solver to check the presence of data leakage due to speculation.

In the subsequent section, we will elaborate the individual sub-systems within KLEESpectre in detail.
4 CACHE AWARE SPECULATIVE SYMBOLIC EXECUTION

In this section, we describe the design of KLEESpectre. First, we describe the overall speculative symbolic execution process augmented with a symbolic cache model. Subsequently, we discuss in detail the features of the cache model to accurately detect the cache side-channel leakage along a speculative execution path.

4.1 Speculative Symbolic Execution

Algorithm 1 outlines the overall process involved in KLEESpectre. Our methodology takes a program $P$ and symbolically executes $P$ by taking into account the speculation at branches. Moreover, KLEESpectre records memory accesses along the speculatively executed paths to check whether any such memory access may refer to a secret. Finally, the sequence of memory accesses are used to formulate a symbolic cache model $\Gamma_{\text{spectre}}$. The model $\Gamma_{\text{spectre}}$ is satisfiable if a possible secret $s$, accessed along a speculative path, remains in the cache after the program execution. This is because the presence of a speculatively accessed secret $s$ in the cache might result in exfiltrating $s$ via a cache side-channel attack. The construction of the speculative execution revolves around the concept of speculative execution window (SEW). Such a bounded window captures the number of instructions that a processor might speculatively execute beyond a branch before the outcome of the branch is resolved. We note that SEW may span across multiple unresolved branch instructions.

At a broader perspective, Algorithm 1 intercepts each conditional branch instruction $r$ and explores all possible speculatively executed instructions from this branch. To this end, we compute $\Omega_r$. After handling a conditional branch $r$, each element in $\Omega_r$ is a possible sequence of memory accesses that might have occurred during a speculative execution from $r$. Moreover, $\Omega_r$ only records memory accesses that may refer to a secret. If memory accesses do not refer to secrets along speculatively executed paths, then they do not impose any risk related to the leakage of information. Algorithm 1 terminates when the time budget exceeds or KLEESpectre explores all (speculatively) executed paths and $\Omega_r$ is constructed for every conditional branch $r$. In the following, we will discuss some critical features of KLEESpectre.

Identifying Secret Access: In this work, we identify secrets via two different strategies. The first strategy is relevant for applications where secrets may not be clearly identified. In such cases, for each memory-related instruction $r$, we consider that $r$ might access a secret if and only if $r$ points to an out-of-bound memory location along a speculative path. In KLEESpectre, we leverage the explicit checks embodied within the KLEE symbolic execution tool to identify an out-of-bound memory access. Although all such memory accesses may not refer to secrets, these memory accesses capture illegal accesses, a typical target for attacks exploiting speculative execution. We define the leakage of such secrets as $N(\text{LS})$. The definition of $N(\text{LS})$ were implicitly adopted by SpecFuzz [34] and oo7 [40]. In our second strategy, we can selectively mark secrets in the application before invoking KLEESpectre. In our evaluation, we empirically evaluate both strategies for encryption routines. The leakage of explicitly marked secrets is defined as $\Delta(\text{LS})$. More specifically, in the presence of $\Delta(\text{LS})$, attackers can (partially) retrieve the user-customized secrets. We use the function $\text{DEP}(sec, m)$ to capture whether some memory address $m$ is data-dependent on secret $sec$. Concretely, $\text{DEP}(sec, m)$ is true if and only if $m$ is data-dependent on the secret $sec$.

Procedure KLEESpectre. Algorithm 1 outlines the symbolic execution process embodied in KLEESpectre. Intuitively, KLEESpectre modifies the handling of branch instructions within a classic symbolic execution process. For each conditional branch $r$, KLEESpectre maintains a structure $\Omega[r]$. Upon encountering a conditional branch instruction $r$, KLEESpectre explores all possible execution paths that might occur due to branch speculation. This is accomplished.
via the procedure ExpandSpecTree. Consider the symbolic state before the branch instruction
is \( \mu \) and the branch condition is \( \phi_r \). If \( \pi[\mu] \) captures the partial path condition before \( r \), then a
speculative execution may proceed in the following two scenarios. Firstly, the true leg of the branch
might be explored with the constraint \( \pi[\mu] \land \neg \phi_r \). Secondly, the false leg of the branch might be
explored with the constraint \( \pi[\mu] \land \phi_r \). These explorations are accomplished via the two calls to
procedure ExpandSpecTree in Algorithm 1. Upon termination of ExpandSpecTree for a branch
instruction \( r \), the structure \( \Omega[r] \) contains the set of memory access sequences that depend on some
secret. Therefore, these memory accesses are candidates that may leak secret information via cache
side-channel. Each memory access captures a triplet of the form \( \langle r, \pi, \sigma \rangle \) where \( r \) points to the
instruction in the execution trace, \( \pi \) captures the symbolic constraint with which \( r \) was executed and \( \sigma \) captures the symbolic expression of the accessed memory address. Finally, KLEESPECTRE
records all memory accesses that influence the cache state for memory blocks in \( \Omega[r] \). Thus, after
termination of a symbolically executed path, each list \( \Gamma \in \Omega[r] \) contains all memory accesses that
may replace a memory block accessed during the speculation at \( r \).

**Procedure** ExpandSpecTree: Algorithm 2 outlines the overall process of exploring the set of
speculative execution paths. In summary, ExpandSpecTree performs the following operations.
First, it explores all speculative paths until the speculation depth \( SEW \). We note that such an
exploration may involve nested speculation. Secondly, while exploring the speculative paths, we
record memory addresses for checking information leakage through the cache. These are the set of
memory accesses that may depend on some secret \( sec \in SEC \). In our framework, we consider that
any out-of-bound memory access along a speculative path points to a secret. Thus, the procedure
ExpandSpecTree also records the potential secrets during exploration.

**Termination of a speculative state.** The execution of speculative state can be terminated in the
following ways:

1. The speculation window \( SEW \) expires. Since \( SEW \) captures the maximum number of in-
structions that can be executed speculatively, we terminate the exploration of a speculative
execution state after exploring \( SEW \) instructions.

2. A memory fence instruction is executed. The memory fence can stop the speculative execution
triggered due to branch mis-prediction.

3. An exception is raised. When an exception (e.g., division by zero) is raised, the speculative
execution terminates. This is analogous to the termination of normal execution.

Algorithm 1 satisfies the following crucial properties:

**Property 1.** Consider an instance of the procedure call ExpandSpecTree\((\pi, \mu, r, r_s, \Gamma)\). Upon
termination of this call, let us assume \( \Omega[r] = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_n\} \). During an arbitrary execution, further
assume that the conditional branch \( r \) was mispredicted and memory address \( m_s \) was accessed specu-
latively. If \( m_s \) is data-dependent on some secret, then \( \langle \ast, \ast, m_s \rangle \in \Gamma_i \) for some \( \Gamma_i \in \Omega[r] \). In short, \( \Omega[r] \) is guaranteed to be an over-approximation of speculatively accessed memory addresses that are
dependent on secret.

**Property 2.** Consider \( \Gamma \in \Omega[r] \) after the termination of a symbolically executed path with the sym-
bo\$ic state \( \mu \). Let \( \langle \ast, \ast, m_s \rangle \in \Gamma \) where \( m_s \) is data-dependent on some secret. Assume tail \( (\Gamma, \langle \ast, \ast, m_s \rangle) \)
captures the set of elements in the sequence \( \Gamma \) post the element \( \langle \ast, \ast, m_s \rangle \). If \( \langle \ast, \ast, m \rangle \in \text{tail}(\Gamma, \langle \ast, \ast, m_s \rangle) \) and \( m \) is not accessed during speculation, then the memory block \( m \) must be accessed following the
access to \( m_s \) for any concrete execution realizing the symbolic state \( \mu \).
Algorithm 1 Symbolic execution process embodied in KLEESpectre

procedure KLEESpectre($P$, $SEW$)

Let $r$ be the first instruction in $P$
/* $\mu_0$ is the initial state before running $P$ */
/* $\pi[\mu_0]$ is the constraint associated with the state $\mu_0$ */
$\chi := \{\mu_0\}; \pi[\mu_0] := true; Spec := \emptyset$

while $\chi \neq \emptyset$
do

Choose a symbolic state $\mu \in \chi$
/* get the next instruction to symbolically execute */
$r := \text{GetNextInstruction} (P, \mu)$
$\mu := \text{ExecuteSymbolic} (\mu, r)$
if $r$ is a conditional branch then

Let $\phi_r$ be the branch condition
remove $\mu$ from $\chi$
$\Gamma_t := \Gamma_f := \emptyset; \Omega[r] := \emptyset$
$\mu_t := \mu_f := \mu$
$\pi[\mu_t] := \pi[\mu] \land \phi_r; \ \pi[\mu_f] := \pi[\mu] \land \neg\phi_r$
if $\pi[\mu_t]$ is satisfiable then

$r^{t}_{n} := \text{GetNextInstruction} (P, \mu_t)$
/* $r^{t}_{n}$ is executed when $r$ is mis-predicted */
$\text{ExpandSpecTree} (\pi[\mu_t], \mu_t, r, r^{t}_{n}, \Gamma_t)$
$\chi := \chi \cup \{\mu_t\}$
end if
if $\pi[\mu_f]$ is satisfiable then

$r^{f}_{n} := \text{GetNextInstruction} (P, \mu_f)$
/* $r^{f}_{n}$ is executed when $r$ is mis-predicted */
$\text{ExpandSpecTree} (\pi[\mu_f], \mu_f, r, r^{f}_{n}, \Gamma_f)$
$\chi := \chi \cup \{\mu_f\}$
end if
if $\Omega[r] \neq \emptyset$ then

$Spec := Spec \cup \{r\}$
end if
end if
if $r$ is a memory-related instruction then

Let $\sigma$ be the accessed memory address
for each $i \in Spec$ s.t. $\pi[\mu] \implies \phi_i \land \Gamma \in \Omega[i]$ do

Append ($\Gamma, \langle r, \pi[\mu], \sigma\rangle$)
end for
end if
end while
end procedure

4.2 Symbolic Model of Cache

In this section, we model the cache behaviour of an execution path to check whether a secret remains in the cache after program execution. Note that our modified symbolic execution already takes into account the speculative execution semantics. Thus, the obtained execution path already accounts for memory references encountered speculatively. Concretely, the input to our cache model is any memory access sequence $\Gamma \in \Omega[r]$ (see Algorithm 1) where $\Omega[r]$ is constructed for
every conditional branch instruction $r$. In the following, we show the cache modelling for a memory access sequence $\Gamma$. Since $\Gamma$ is arbitrary, the same modelling principle is employed for all the memory access sequences. Concretely, any memory access sequence $\Gamma$ is captured by a sequence of triplets as follows:

$$\Gamma \equiv \langle (r_1, \pi_1, \sigma_1), (r_2, \pi_2, \sigma_2), \ldots, (r_N, \pi_N, \sigma_N) \rangle$$

where $r_i$ is a memory-related instruction, $\pi_i$ is the symbolic constraint with which $r_i$ was executed and $\sigma_i$ is the memory address accessed by $r_i$. We note that $r_i$ can be accessed along a speculative path or a normal path (cf. Algorithm 1). Before discussing the cache model, we first explain the basic design principle behind caches.
Basics of Cache Design: Caches are fast memory employed between the CPU and the main memory. While accessing a memory location, the CPU first checks whether the memory location is cached. If the location is cached, then the CPU fetches the respective value in the location from the cache. Otherwise, it accesses main memory and updates the cache with the accessed memory location and its value. The design parameters of a cache can be captured via a triplet: \( (2^{S}, 2^{B}, \mathcal{A}) \). \( 2^{S} \) captures the number of cache sets and \( 2^{B} \) captures the size of a cache line (in bytes). Each cache set can hold \( \mathcal{A} \) cache lines while \( \mathcal{A} \) is called the associativity of the cache. For any memory-related instruction \( r \), let us assume it accesses the memory address \( m \). The address \( m \) is mapped to the cache set \( \left\lfloor \frac{m}{2^{B}} \right\rfloor \mod 2^{S} \). Since multiple memory addresses can be mapped to the same cache set, each cache line in a cache set stores a tag. This tag is identified via the most-significant \( B \) bits of the memory address \( m \). Once a cache set is full (i.e. holds \( \mathcal{A} \) cache lines) and a new memory location is mapped to the same cache set, then a replacement policy is employed to evict a cache line and make room for fresh memory locations. In this work, we instantiate KLEESpectre for the least recently used (LRU) replacement policy. In LRU, the least recently accessed memory location in a cache set is chosen for eviction to accommodate fresh memory blocks. We define a cache set state as an ordered \( \mathcal{A} \)-tuple where the rightmost element captures the least recently used cache line. For example, in a two-way associative cache, the state \( (L_{1}, L_{2}) \) captures that \( L_{2} \) (respectively, \( L_{1} \)) is the least (respectively, most) recently used cache line.

In line with the preceding description of cache design, we assume the following notations throughout the section:

- \( 2^{S} \): The number of cache sets in the cache.
- \( 2^{B} \): Size of the cache line.
- \( N_{s} \): The set of memory-related instructions accessing symbolic memory addresses (i.e. potential secrets accessed along speculative paths) in memory access sequence \( \Gamma \) (cf. Equation 2).
- \( N_{t} \): The set of memory-related instructions exhibited along normal path in memory access sequence \( \Gamma \) (cf. Equation 2). We note that \(|N_{s} \cup N_{t}| = N\) holds.
- \( \mathcal{A} \): Associativity of the cache.
- \( \sigma_{i} \): Memory address accessed by instruction \( r_{i} \).
- \( \text{set}(r_{i}) \): Cache set accessed by memory-related instruction \( r_{i} \).
- \( \text{tag}(r_{i}) \): Cache tag related to the memory-related instruction \( r_{i} \).

It is worthwhile to note that the symbolic address is defined to be a memory address that is dependent on a secret value. Moreover, as mentioned in the preceding section, we only consider secrets that might be accessed along speculative paths. Thus, if any address dependent on secrets remains in the cache after program execution, the respective program is vulnerable to Spectre attacks. The set of instructions accessing such symbolic addresses, i.e., \( N_{s} \) were identified during our novel symbolic execution stage.

Cache Conflict: The symbolic model of the cache revolves around the notion of cache conflict. Intuitively, the phenomenon of cache conflict influences the states of each cache set. This, in turn, decides whether a value is cached during or after the execution. In the following, we first formally define the notion of cache conflict.

**Definition 3. (Cache Conflict):** Consider memory-related instructions \( r_{i} \) and \( r_{j} \). Let \( \zeta_{i} \) (respectively, \( \zeta_{j} \)) be the cache state immediately after \( r_{i} \) (respectively, \( r_{j} \)) is executed. \( r_{j} \) generates a cache conflict to \( r_{i} \) only if \( r_{j} \) is executed after \( r_{i} \) and executing \( r_{j} \) can influence the relative position of memory block \( \left\lfloor \frac{m_{j}}{2^{B}} \right\rfloor \) within the cache state \( \zeta_{i} \).

Figure 3 captures different scenarios in accounting cache conflict within our cache model. In Figure 3, the notation \( r_{x} : m_{x} \) captures that the memory-related instruction \( r_{x} \) accesses the memory
block $m_x$. As an example of cache conflict, consider Figure 3(a). In this case, $r_3$ generates a cache conflict to $r_1$, as $r_3$ changes the relative position of $m_1$ in $\zeta_3$. In Figure 3(b), both $r_2$ and $r_3$ generate cache conflict to $r_1$, as both can change the relative position of $m_1$ in the cache state after execution. However, special handling is required for the scenario depicted in Figure 3(b). This is because $r_2$ and $r_3$ access the same memory block $m_2$ and a cache conflict from a given memory block needs to be accounted at most once. We discuss this shortly in our cache modelling.

The preceding definition of cache conflict works for arbitrary memory-related instructions $r_i$ and $r_j$. In KLEESpectre, however, our objective is to check whether any symbolic address remains in the cache. To this end, we only need to capture the cache conflict when $r_j \in N_t$ and $r_i \in N_s$. The cache conflicts within normal paths and within the speculative paths are ignored. Similarly, we do not need to check whether a memory block accessed in normal path can be replaced by a memory block accessed along speculative paths. Thus, we can ignore the cache conflict when $r_j \in N_s$ and $r_i \in N_t$. We formalize the aforementioned notion of cache conflict in KLEESpectre via the following definition:

**Definition 4. (Cache Conflict in KLEESpectre):** Consider memory-related instructions $r_i$ and $r_j$. For KLEESpectre, we consider a cache conflict from $r_j$ to $r_i$ if and only if $r_j$ generates a cache conflict to $r_i$ according to Definition 3 and $r_j \in N_t$ and $r_i \in N_s$.

By considering the notion of cache conflict, as defined in Definition 4, we greatly simplify the size of the symbolic cache model and keep the overall complexity of KLEESpectre under check. In the next sections, we shall elaborate on the crucial conditions required for the generation of cache conflicts and usage of such conditions to check the residency of a memory block in the cache. Subsequently, we build upon such conditions to formulate the symbolic model for identifying Spectre vulnerabilities.

**Cache Set and Cache Tag.** We note that due to the symbolic memory addresses, $set(r_i)$ and $tag(r_i)$ can be symbolic expressions. Specifically, $set(r_i)$ and $tag(r_i)$ are computed as follows:

$$set(r_i) = (\sigma_i \gg B) \& (2^S - 1) \text{ subject to } \pi_i$$  \hfill (3)

$$tag(r_i) = \sigma_i \gg (B + S) \text{ subject to } \pi_i$$  \hfill (4)
Cache Conflict and Conflict Propagation. : Our objective is to discover whether any symbolic memory address can be evicted from the cache after being accessed. As stated in Definition 4, KLEESPECTRE only considers cache conflict from memory accesses along normal path (i.e. set \( N_f \)) to the memory accesses along speculative paths (i.e. set \( N_s \)). However, it is not sufficient to check the cache conflict from \( r_j \in N_f \) to \( r_i \in N_s \) to precisely identify Spectre vulnerabilities. To check whether the conflict actually influences the relative position of the memory block till the end of the execution, we need to check whether the memory block accessed by \( r_i \) can be reloaded after \( r_j \) and before the end of the execution. If \( r_i \) is reloaded after \( r_j \), then the cache conflict generated by \( r_j \) is not propagated until the end of the execution. Finally, we need to check whether the memory block accessed by \( r_i \) is replaced from the cache before the execution terminates. This is accomplished by checking whether the number of unique cache conflicts to \( r_i \) that propagate till the end of execution exceeds the cache associativity (\( \mathcal{A} \)). In the following, we will model these phenomena symbolically.

If \( r_j \) generates a cache conflict to \( r_i \), then the following condition must hold: \( r_j \) and \( r_i \) access the same cache set, but have different memory-block tags. This is formalized as follows:

\[
\psi_{\text{cnf}} (r_i, r_j) \equiv (\text{set}(r_i) = \text{set}(r_j)) \land (\text{tag}(r_i) \neq \text{tag}(r_j))
\]

(5)

In Figure 3(a), for instance, \( \psi_{\text{cnf}} (r_1, r_3) \) holds, as \( m_2 \) is mapped to the same cache set as \( m_1 \).

Additionally, we need to check whether \( r_j \) is a unique cache conflict. To this end, we check that none of the memory accesses after \( r_j \) accesses the same memory block as \( r_j \). Thus, we only account for the last memory-related instruction accessing the memory block \( \left[ \frac{\sigma F}{2^p} \right] \). This is formalized as follows:

\[
\psi_{\text{unq}} (r_j) \equiv \bigwedge_{k \in (j, N) \land k \in N_f} (\text{set}(r_j) \neq \text{set}(r_k)) \lor (\text{tag}(r_j) \neq \text{tag}(r_k))
\]

(6)

As an example, consider the scenario in Figure 3(b). In this scenario, we only account for the cache conflict from \( r_3 \) to \( r_1 \). In other words, we observe that \( \psi_{\text{unq}} (r_2) \) is evaluated false, as \( r_3 \) also accesses the memory block \( m_2 \). In contrast, \( \psi_{\text{unq}} (r_3) \) trivially holds and takes into account the cache conflict from \( r_3 \) to \( r_1 \).

Finally, we need to check that \( r_i \) is not reloaded after \( r_j \). Otherwise, the memory block accessed by \( r_i \) will be reloaded to the cache and the conflict due to \( r_j \) would be nullified. This is formalized as follows:

\[
\psi_{\text{rel}} (r_i, r_j) \equiv \bigwedge_{k \in (j, N]} (\text{set}(r_j) \neq \text{set}(r_k)) \lor (\text{tag}(r_j) \neq \text{tag}(r_k))
\]

(7)

As an example, consider the scenario in Figure 3(c). Even though \( r_2 \) generates a cache conflict to \( r_1 \), we observe that \( m_1 \) is reloaded after \( r_2 \). Therefore, the cache conflict generated by \( r_2 \) is nullified via the reloading of \( m_1 \) at \( r_3 \). In other words, \( r_2 \) does not influence the residency of memory block \( m_1 \) in the cache after execution (i.e. after \( r_3 \) in the example).

Combining Equations 5-7, we obtain the symbolic condition where \( r_j \) changes the relative position of the memory block accessed by \( r_i \) and such a change in the relative position of the memory block is also propagated until the end of the execution. Thus, when all the conditions \( \psi_{\text{cnf}} (r_i, r_j) \), \( \psi_{\text{unq}} (r_j) \) and \( \psi_{\text{rel}} (r_i, r_j) \) hold, we can say that the conflict generated by \( r_j \) to \( r_i \) is propagated until the end of the execution. This is symbolically captured as follows:

\[
\Theta_i^+ j \equiv \psi_{\text{cnf}} (r_i, r_j) \land \psi_{\text{unq}} (r_j) \land \psi_{\text{rel}} (r_i, r_j) \Rightarrow (\text{cnf}_{i,j} = 1)
\]

(8)

\[
\Theta_i^{-} j \equiv \neg\psi_{\text{cnf}} (r_i, r_j) \lor \neg\psi_{\text{unq}} (r_j) \lor \neg\psi_{\text{rel}} (r_i, r_j) \Rightarrow (\text{cnf}_{i,j} = 0)
\]

(9)
**Attack Identification.** : We note that $r_j$ is arbitrary in the preceding discussion. To check whether the memory block accessed by $r_j$ can be replaced, we need to repeat the computation of $\Theta^+_{j,i}$ and $\Theta^-_{j,i}$ for any $j \in [i + 1, N]$ where $N$ is the total number of memory accesses in the trace. Finally, we need to check whether the collective sum of $cnf_{i,j}$ for $j \in [i + 1, N]$ exceeds the cache associativity. Let us assume that $spec_i$ is true if and only if the memory block accessed by $r_i$ may remain in the cache after program execution, thus exhibiting a potential Spectre attack. The truth value of $spec_i$ can be symbolically computed as follows:

$$\lambda_i \equiv \left( \sum_{j \in [i+1,N] \land r_j \in N_i} cnf_{i,j} < \mathcal{A} \right) \Rightarrow spec_i$$

In Figure 3, consider that $m_1$ holds a secret. For scenarios captured in Figures 3(a)-(c), $r_3$ generates a cache conflict to $r_1$. However, it is not sufficient, in any of the scenarios, to replace $m_1$ from the cache. Thus, all scenarios in Figures 3(a)-(c) are identified as attacks via our cache modelling. In Figure 3(d), the attack is nullified after memory reference $r_3$. This is because $r_1$ faces two unique cache conflicts from $r_2$ and $r_3$. Such conflicts, in turn, are sufficient to replace $m_1$ from a 2-way set-associative cache.

**Putting it altogether.** : Finally, Spectre attacks can be targeted for any memory-related instruction accessing a symbolic address. Therefore, Equations 5-10 need to account for all such symbolic memory accesses. Recall that $N_s$ captures the set of all memory-related instructions in the trace that access symbolic memory address. Thus, to check the possibility of Spectre attacks for an arbitrary (combination) of memory addresses, the following symbolic model is used:

$$\Gamma_{spectre} \equiv \bigwedge_{r_j \in N_s} \left( \lambda_j \land \left( \bigwedge_{j \in [i+1,N] \land r_j \in N_i} \Theta^+_{j,i} \land \Theta^-_{j,i} \right) \right) \land \left( \bigvee_{r_j \in N_i} spec_j \right)$$

We note that $\Gamma_{spectre}$ is true if and only if any of the symbolic memory address remains in the cache after program execution, thus leading to a potential Spectre attack.

### 5 IMPLEMENTATION

KLEESPECTRE is primarily implemented on top of the state-of-the-art symbolic execution engine KLEE v2.0 [8]. KLEESPECTRE is built from CLang v6.0 and it takes the LLVM bitcode generated with LLVM 6.0 as input. If a subject program contains external function calls, then the program is linked with KLEE- uClibc [3] first, before being passed to KLEESPECTRE. We used the SMT solver STP [20] to check the satisfiability of the path constrains and the symbolic cache model. Broadly KLEESPECTRE makes three major changes in KLEE: generating speculative symbolic states, propagating potentially sensitive data and symbolically modelling the cache behaviour.

**Generating speculative symbolic states.** A symbolic execution engine interprets a single instruction symbolically subject to the constraints imposed on the respective symbolic state. The initial symbolic state is constrained via the logical formula $true$. If the constraint imposed on the current symbolic state is $C$ and the engine encounters a branch instruction with condition $\phi_b$, then traditional symbolic execution engines check the satisfiability of constraints $C \land \phi_b$ and $C \land \neg \phi_b$. If such a constraint is satisfiable, then the engine creates a new symbolic state with the constraint. The new state inherits the state before encountering the branch instruction, but proceeds interpreting the subsequent instructions independently. Our KLEESPECTRE approach generates two extra symbolic states to model the speculative execution. These states are generated to model the speculative paths and they also model nested speculative execution. We also modify
the path selection heuristic in KLEE to take into account the newly generated speculative symbolic states. Specifically, when the scheduler selects a normal state $S_m$ to execute, we check whether the state may be immediately preceded by any speculative state. If such is the case, then KLEESPECTRE selects a speculative state $SS_i$ to process. The normal state $S_m$ is not processed until all preceding speculative states of $S_m$ are handled. KLEESPECTRE can use all existing state selection strategies in KLEE, such as Depth-First Search (DFS), Breadth-First Search (BFS), random path selection (random-path) for both the normal state selection and speculative state selection.

**Propagating potentially sensitive data.** KLEESPECTRE propagates the sensitive data along the execution path to identify the addresses that may leak the sensitive data to the cache state. When a memory load instruction reads a variable $v_s$ from an out-of-bound memory location, we mark $v_s$ as sensitive. All new expressions dependent on $v_s$ are subsequently marked sensitive as well. By tracking these sensitive expressions, we can detect if a memory access leads to the leakage of sensitive data. This is accomplished by checking whether the accessed memory address is constructed from any sensitive expressions.

**Symbolically modelling the cache behaviour.** KLEESPECTRE models the cache to further check whether a cache state impacted by a sensitive address can be observed in an execution point, in particular, at the termination of a program for the access-based cache side-channel attack. Each execution state contains a cache state that symbolically records the cache content along the execution path. The cache modelling of KLEESPECTRE collects all memory load and store addresses except the memory store addresses in a speculative execution. There exists multiple reasons for such a design choice. Firstly, the memory store is not visible to the cache until the speculatively executed instructions are committed in the real execution of a processor. Secondly, our assumption is that all speculative executions in KLEESPECTRE are caused by the branch mis-prediction and all speculatively executed instructions are rolled back. Upon the termination of an execution, the symbolic cache model is constructed in line with the explanation in Section 4.2 and we call the STP solver to check whether the sensitive address may still stay in the cache.

6 EVALUATION

In this section, we perform the effectiveness evaluation of KLEESPECTRE in detecting the Bounds Check Bypass (BCB or Spectre variant 1) vulnerabilities. We aim to answer the following research questions:

1. **RQ1:** Can KLEESPECTRE effectively detect various kinds of BCB vulnerabilities?
2. **RQ2:** How efficient is KLEESPECTRE in detecting the BCB vulnerabilities?
3. **RQ3:** How effective is our cache model in detecting cache side-channel leakage through speculative paths?

Note that BCB vulnerability has not been reported in the wild yet. Therefore, we first run KLEE-Spectre on the litmus tests created by Kocher [28]. These litmus tests are different types of Spectre vulnerable code patterns. Secondly, we run KLEESPECTRE on a set of security-critical benchmarks to check whether KLEESPECTRE can find the potential BCB vulnerabilities. Finally, we evaluate the effectiveness of our cache model in KLEESPECTRE by modifying the litmus tests and the security-critical benchmarks appropriately.

6.1 Evaluation of KLEESPECTRE on litmus tests

No real BCB vulnerability has been reported in the wild. So we first rely on fifteen litmus test programs with Spectre vulnerability created by Kocher [28]. We aim to check whether KLEESPECTRE can successfully detect these different variations of BCB vulnerabilities. These litmus tests were
originally developed to evaluate the effectiveness of the Spectre mitigation in Microsoft C/C++ compiler. The Microsoft compiler uses static analysis to identify the vulnerable code fragments and inserts \texttt{lfence} to repair the vulnerable code. Kocher reports \cite{28} that the Microsoft compiler can only identify two out of 15 vulnerable programs. This is because instead of using precise static analysis, Microsoft C/C++ compiler only performs a simple code pattern matching to identify code fragments related to the BCB vulnerabilities. In contrast, KLEESpectre can correctly detect all the BCB variants in 15 litmus tests produced by Kocher.

The programs used in litmus tests contain no memory access after the sensitive data is leaked and brought into the cache along the speculative path. As a result, our cache modelling has no impact on the detection results and all the litmus tests are correctly confirmed to contain Spectre vulnerability by KLEESpectre. Thus, we design additional experiments to showcase the power of cache modelling in KLEESpectre by introducing memory access code in the litmus test programs after the spectre vulnerability.

For evaluating our cache model, we use a 32 KB set-associative cache with the LRU replacement policy and each cache line has 64 bytes data. We configure the cache to be 2-way, 4-way or 8-way in our experiments. We mainly consider the \texttt{PRIME} + \texttt{PROBE} attack on L1 cache. This attack is used to target both data \cite{35,36} and instruction cache \cite{4}.

A modified litmus test code is outlined in Listing 3. The code contains a vulnerable function \texttt{victim\_fun()} that receives an integer \texttt{idx} as an argument. The if statement at line 8 checks whether \texttt{idx} is less than the \texttt{array1[]} size \texttt{array1\_size}. If the condition holds, then \texttt{idx} is used to access \texttt{array1[]):}. The code between line 8 and 10 exposes a typical BCB vulnerability. Specifically, if the branch at line 8 is mis-predicted, then the access of \texttt{array1[]} with \texttt{id}x value greater than \texttt{array1\_size} can bring in potentially sensitive data. This is because \texttt{array1[]} can point outside of \texttt{array1[]} when the branch at line 8 is mis-predicted. The sensitive data can subsequently be

---

Listing 3. The code for testing KLEESpectre with cache modelling.

```c
1  int array1\_size = 16;
2  char array1[16];
3  char array2[256*64];
4  char array3[512*64];
5  char temp = 0;
6  char victim\_fun(int idx) {
7     register int i;
8     if (idx < array1\_size) {
9        temp &= array2[array1[idx]];
10    }
11    \#define ITER N // N = \{1, ..., 512\}
12    for (i = 0; i < ITER; i++) {
13        temp &= array3[i*64];
14    }
15    return temp;
16 }
17  void main() {
18     int x;
19     klee\_make\_symbolic(&x, sizeof(x), "x");
20     victim\_fun(x);
21 }
```

---
leaked to the cache state by accessing array2[\ldots] at line 9. The question remains whether the leaked data will still remain in the cache after completion of the program execution. This question can be answered by our cache modelling in KLEESpectre.

Thus, to test the effectiveness of the cache model in KLEESpectre, we add a loop at lines 11-15. The loop continuously brings in data to the different cache sets after the leakage of sensitive data at line 9. The memory accesses in the loop may evict the sensitive data introduced into the cache by BCB vulnerability (line 9) after \(N\) iterations. Each iteration brings a memory block to a different cache line, for example, the loop introduces a memory block for each cache set by the first 256 iterations and the entire cache is filled up after performing total 512 iterations for a 2-way (256 sets) cache. We run this litmus test program with different values of \(N\) from 1 to 512 to evaluate the effectiveness of KLEESpectre. Specifically, we aim to detect the eviction of the sensitive data from the cache for different values of \(N\) and different cache associativities (i.e. 2, 4 and 8).

![Figure 4](image)

Fig. 4. The detection result of KLEESpectre with cache model enabled. #\text{-}way represents a cache setting with #\text{-}way set associative cache.

The outcome of our findings is shown in Figure 4. The red solid line denotes that KLEESpectre can detect the sensitive cache state, which means the sensitive data is still in the cache after \(N\) memory accesses in the test code. In contrast, the green dash line indicates that the sensitive data has been evicted from the cache by the additional code (Leakage-free). We can see from Figure 4 that the sensitive data is no longer present in the cache after 260, 388 and 452 memory accesses for cache associativity 2, 4 and 8, respectively.

The result in Figure 4 proves the effectiveness of KLEESpectre cache modelling. As an example, consider the code at line 9 in Listing 3. The data read by array1[idx] is one byte represented as \(B_i\). Thus, \(B_i\) has a value between 0 and 255. The address of the memory access performed by array2[\(B_i\)] is captured via array2 + \(B_i\). As the least significant six bits of \(B_i\) are used for the byte offset in the cache block (64-byte cache block), only two bits of \(B_i\) are used for the two least significant bits of the cache set index. Thus, the address array2 + \(B_i\) can map to one of four selected contiguous cache sets depending on the value of \(B_i\) for any cache associativity. Thus to completely evict array2 + \(B_i\) from the cache for arbitrary values of \(B_i\), we need access to 8, 16 and 32 corresponding cache lines for 2, 4 and 8-way associate caches, respectively.

As shown in Figure 4, for 2-way set associative caches, the leakage is undetectable after 260 memory accesses from the loop at lines 11-13. In a 2-way set-associative cache, array2 + \(B_i\) can
Table 1. Subject benchmarks.

<table>
<thead>
<tr>
<th>Program</th>
<th>Source</th>
<th>Description</th>
<th>LoC</th>
<th>#Branch</th>
</tr>
</thead>
<tbody>
<tr>
<td>chacha20</td>
<td>LibTomCrypt</td>
<td>chach20poly1305 cipher</td>
<td>776</td>
<td>71</td>
</tr>
<tr>
<td>aes</td>
<td>LibTomCrypt</td>
<td>AES implementation</td>
<td>1,838</td>
<td>27</td>
</tr>
<tr>
<td>encoder</td>
<td>LibTomCrypt</td>
<td>encode binary data to ASCII string</td>
<td>134</td>
<td>14</td>
</tr>
<tr>
<td>ocb3</td>
<td>LibTomCrypt</td>
<td>OCB implementation</td>
<td>377</td>
<td>40</td>
</tr>
<tr>
<td>salsa</td>
<td>Linux-tegra</td>
<td>Salsa20 stream cipher</td>
<td>279</td>
<td>20</td>
</tr>
<tr>
<td>camellia</td>
<td>Linux-tegra</td>
<td>camellia cipher</td>
<td>1,324</td>
<td>12</td>
</tr>
<tr>
<td>seed</td>
<td>Linux-tegra</td>
<td>Seed cipher</td>
<td>487</td>
<td>9</td>
</tr>
<tr>
<td>str2key</td>
<td>openssl</td>
<td>Key preparation for DES</td>
<td>385</td>
<td>12</td>
</tr>
<tr>
<td>des</td>
<td>openssl</td>
<td>DES implementation</td>
<td>1,051</td>
<td>11</td>
</tr>
<tr>
<td>hash</td>
<td>hpn-ssh</td>
<td>hash function</td>
<td>304</td>
<td>24</td>
</tr>
</tbody>
</table>

potentially map to four contiguous cache sets depending on the value of $B_i$. Thus, if we want to guarantee the eviction of $array2 + B_i$ from the cache, then we need to fill up these contiguous four cache sets that $array2 + B_i$ may map to. In our experiment, $array2$ was mapped to the first cache set. As a result, 260 memory accesses can completely fill up the first four sets of a 2-way cache. Specifically, the first 256 iterations of the loop (lines 11-13) access memory blocks mapping to all cache sets (256 cache sets for 2-ways cache) and the rest four iterations introduce the second memory blocks for first four cache sets. This guarantees the removal of $array2 + B_i$ from the cache for any value of $B_i$. To the best of our knowledge, none of the existing tools such as oo7 [40] and SPECTECTOR[23] can accurately verify the cache side-channel freedom against BCB attack like KLEESpectre.

6.2 Effectiveness and Efficiency: Detection of BCB Gadgets in Real Programs.

Benchmark selection. To evaluate KLEESpectre on real programs, we select ten cryptography related programs from well known projects: libTomCrypt, Linux-tegra, openssl and hpn-ssh. Table 1 outlines some salient features of the subject benchmarks. All the benchmarks potentially process or contain sensitive data. All of these benchmarks were also used in a recent work [43] to perform the analysis of speculative execution via abstract interpretation. In Table 1, column LoC denotes the lines of code; the collected programs have 134 (encoder) to 1,838 (AES) lines of code. The column #Branch denotes the number of branches in each program ranging from 9 (seed) to 71 (chacha20). For all the benchmarks, we use the internal function klee_make_symbolic() of KLEE to set the input of the programs (e.g., the plaintext and the key in cryptography programs) as symbolic variables. All benchmarks are compiled by Clang-6.0 with “-O1” optimization.

Experimental results. In this experiment, we do not assume the target platform, thus We run KLEE, KLEESpectre with SEW=50, and KLEESpectre with SEW=100 (SEW is the size of the speculative window in terms of the number of instructions) on the benchmarks listed in Table 1 to compare the performance and the effectiveness of KLEESpectre to detect BCB gadgets. The results are shown in Table 2. The column Explored paths denotes the number of explored normal execution paths and the column Explored speculative path indicates the explored speculative execution paths by KLEESpectre.

In each category of KLEE, KLEESpectre 50 and KLEESpectre 100, column Analysis time provides the analysis time of the tool. We conduct our experimental evaluation on Intel Xeon Gold 6126 [2] running at 2.6GHz with 192GB memory. Intel Xeon Gold 6126 is equipped with 12 cores (24 threads)
Table 2. The analysis performance comparison of KLEE, KLEEspectre 50 and KLEEspectre 100 along with the detection results of BCB gadgets. Avg. #inst= The average number of instructions executed on the speculative path. VB= vulnerable branch. UC_VB= user controlled vulnerable branch. RS= Read Secret. N(LS)= Leak Secret and treat all illegal memory accesses as secret.

<table>
<thead>
<tr>
<th>Program</th>
<th>KLEE</th>
<th>KLEEspectre 50</th>
<th>KLEEspectre 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>chacha20</td>
<td>0.58s</td>
<td>1.23s</td>
<td>12.59s</td>
</tr>
<tr>
<td>zts</td>
<td>0.66s</td>
<td>1.46s</td>
<td>2.45s</td>
</tr>
<tr>
<td>encoder</td>
<td>0.88s</td>
<td>2.46s</td>
<td>4.00s</td>
</tr>
<tr>
<td>nuf</td>
<td>0.31s</td>
<td>0.22s</td>
<td>0.26s</td>
</tr>
<tr>
<td>ralsec</td>
<td>0.76s</td>
<td>0.40s</td>
<td>1.41s</td>
</tr>
<tr>
<td>camellia</td>
<td>0.34s</td>
<td>0.35s</td>
<td>1.45s</td>
</tr>
<tr>
<td>vtea</td>
<td>0.14s</td>
<td>0.22s</td>
<td>0.14s</td>
</tr>
<tr>
<td>str2key</td>
<td>0.41s</td>
<td>1.14s</td>
<td>2.10s</td>
</tr>
<tr>
<td>des</td>
<td>0.01s</td>
<td>0.01s</td>
<td>0.01s</td>
</tr>
<tr>
<td>bash</td>
<td>0.32s</td>
<td>0.56s</td>
<td>1.53s</td>
</tr>
</tbody>
</table>

and 19.25MB shared last-level cache (LLC). The machine is running a Ubuntu 16.04 server with Linux kernel 4.4.

Both KLEE, KLEEspectre 50, KLEEspectre 100 complete the analysis within 69 seconds. More specifically, for most benchmarks, KLEEspectre 50 and KLEEspectre 100 have longer analysis time than KLEE; but the analysis time of KLEEspectre is still acceptable. For example, KLEE explores three paths of chacha20 in 0.50s, KLEEspectre 50 explores all three normal paths along with 12,392 speculative paths in 2s. Besides, KLEEspectre 100 always explores more speculative paths than KLEEspectre 50 because KLEEspectre 100 executes more instructions along any speculative path. Moreover, if KLEEspectre encounters branch instructions along the speculative path, then it generates new speculative states (nested speculative execution), resulting in managing a larger number of symbolic states as compared to KLEE. Finally, the speculative execution might be terminated upon receiving an exception or the program exit event. The column Avg. #inst in Table 2 shows the average number of the instructions executed along the speculative path, which is close to the SEW value in most benchmarks (e.g., 47.49 and 95.25 for KLEEspectre 50 and KLEEspectre 100, respectively, while analyzing chacha20).

As for the detection result of BCB Gadgets, the detected number of vulnerable instructions are listed in columns VB, UC_VB, RS and N(LS). VB represents the number of vulnerable branches. The mis-prediction of such branches may result in the secret data to be loaded in the cache. The term UC_VB means that the vulnerable branch can directly be trained via the user controlled input. RS is the abbreviation of Read Secret. Specifically, RS means that the secret can be loaded after executing the respective instruction. LS is an abbreviation of Leak secret wherein an instruction can leak the secret loaded by RS instruction to the cache state. The columns VB, UC_VB, RS and N(LS) in Table 2 are reported as the unique code locations and if one vulnerable code location appears in several speculative execution paths, the code location is only reported once.

We detect VB and RS in all the benchmarks. For example, KLEEspectre 50 found eight vulnerable branches in chacha20 but none of of them is user-controlled. Only the benchmark str2key contains a Leak secret (LS), which means that the secret can potentially be loaded to the cache and observed by the attacker.

Listing 4. Potential Spectre variant 1 vulnerability in str2key; (VB, RS, LS) are highlighted.

```c
1  void DES_set_odd_parity(DES_cblock-key) {
2     int i;
3     for (i=0; i<DES_KEY_SZ; i++) /* VB */
4         (*key)[i]=odd_parity(*key[i]); /* RS, LS */
5 }
```
Table 3. The detection result with cache modelling enabled for different cache configurations.

<table>
<thead>
<tr>
<th>Program</th>
<th>Symbolic address (%)</th>
<th>Analysis time</th>
<th>Detected leakage (NS/LS/AL/LS)</th>
<th>Solver time(%)</th>
<th>Analysis time</th>
<th>Detected leakage (NS/LS/AL/LS)</th>
<th>Solver time(%)</th>
<th>Analysis time</th>
<th>Detected leakage (NS/LS/AL/LS)</th>
<th>Solver time(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>aes</td>
<td>0.5%</td>
<td>9.2s</td>
<td>42.05%</td>
<td>1s</td>
<td>1/1</td>
<td>10.20%</td>
<td>1s</td>
<td>1/1</td>
<td>55.56%</td>
<td>1s</td>
</tr>
<tr>
<td>encoder</td>
<td>28%</td>
<td>1s</td>
<td>20.56%</td>
<td>1s</td>
<td>1/0</td>
<td>65.69%</td>
<td>2/0</td>
<td>2/0</td>
<td>89.19%</td>
<td>156s</td>
</tr>
<tr>
<td>ocb</td>
<td>21.5%</td>
<td>1s</td>
<td>8.02%</td>
<td>1s</td>
<td>1/0</td>
<td>9.56%</td>
<td>1/0</td>
<td>2/0</td>
<td>90.20%</td>
<td>156s</td>
</tr>
<tr>
<td>sal20</td>
<td>6.7%</td>
<td>1s</td>
<td>12.49%</td>
<td>1s</td>
<td>1/0</td>
<td>13.93%</td>
<td>1s</td>
<td>1/0</td>
<td>83.21%</td>
<td>1s</td>
</tr>
<tr>
<td>camellia</td>
<td>7.2%</td>
<td>1s</td>
<td>87.47%</td>
<td>2s</td>
<td>1/1</td>
<td>81.94%</td>
<td>4s</td>
<td>2/1</td>
<td>94.46%</td>
<td>35s</td>
</tr>
<tr>
<td>seed</td>
<td>52.7%</td>
<td>1s</td>
<td>8.26%</td>
<td>1s</td>
<td>1/0</td>
<td>81.94%</td>
<td>2s</td>
<td>2/1</td>
<td>94.46%</td>
<td>35s</td>
</tr>
<tr>
<td>str2key</td>
<td>7.8%</td>
<td>1s</td>
<td>95.10%</td>
<td>2s</td>
<td>1/0</td>
<td>94.48%</td>
<td>3s</td>
<td>1/0</td>
<td>94.46%</td>
<td>35s</td>
</tr>
<tr>
<td>des</td>
<td>26.9%</td>
<td>1s</td>
<td>22.45%</td>
<td>2s</td>
<td>1/1</td>
<td>98.47%</td>
<td>17s</td>
<td>2/1</td>
<td>97.96%</td>
<td>3s</td>
</tr>
<tr>
<td>hash</td>
<td>5.4%</td>
<td>16s</td>
<td>92.66%</td>
<td>6s</td>
<td>1/0</td>
<td>92.58%</td>
<td>6s</td>
<td>1/1</td>
<td>80.32%</td>
<td>15s</td>
</tr>
</tbody>
</table>

Listing 4 shows a potential Spectre variant 1 vulnerability in the str2key benchmark function `DES_set_odd_parity()`. The loop iteratively reads the data pointed by *key and uses the data to index array odd_parity. A mis-prediction of the for loop condition may cause a speculative execution of a few more loop iterations than normal execution. This may lead sensitive data beyond the end of *key (i.e. beyond the size DES_KEY_SZ) to be loaded into the cache. The sensitive data can impact the cache state when it is used to access array odd_parity. Thus, the cache state can potentially be observed by the attacker through probing array odd_parity. The exact amount of the leakage depends on the number of iterations that can be speculatively executed. However, in its current state, KLEESPECTRE does not compute an exact quantification of the leakage.

We also compare the KLEESPECTRE result with oo7 [40] and show that oo7 can only detect data leakage in encoder and ocb3. This is because oo7 only identifies the user-controlled branches as vulnerable branches. However, KLEESPECTRE assumes all branches can be mis-trained by the attacker; for example, the victim process and the attacker process may be scheduled to the same core and the attacker can directly mis-train the branch prediction [9, 19].

### 6.3 Leakage detection with cache modelling

The cache modelling of KLEESPECTRE accurately checks whether the leaked sensitive data can be observed by the attacker through the cache side-channel. Our cache model is not invoked until some sensitive data is identified along the speculative path. As we do not find any data leakage in our benchmarks, in this experiment, we insert several vulnerable functions to the benchmarks and check whether KLEESPECTRE can detect them. More specifically, we randomly choose three Spectre v1 variant functions suggested by Kocher [28], then insert them to the start, middle and the end of each benchmark listed in Table 1. Finally, we run KLEESPECTRE with cache modelling enabled. Each experiment is conducted over three runs for three different cache associativities: 2, 4 and 8. Moreover, we marked the cipher key (e.g., for aes) or program input (if the program does not contain a secret key, e.g., hash) as the user marked secrets.

Table 3 shows the test results of comparing KLEESPECTRE 100 with cache enabled. KLEESPECTRE 100 detects all vulnerable code that leaks the sensitive data in the speculative execution path (str2key contains one original leakage). The column `N(\(LS\))/\(\Delta(\(LS\))\)` denotes the sensitive data leakage for two strategies used in identifying secrets. `N(\(LS\))` captures the leakage where all illegal memory accesses are treated as secrets. In contrast, `\(\Delta(\(LS\))\)` captures the leakage only for user-marked secrets (i.e. a secret key or program input). The results show that KLEESPECTRE can successfully identify the leakage of user marked secrets. All tested programs except encoder, sal1sa20 and seed show that the program can access and leak the user-defined secrets in their speculative execution path.

More importantly, we observe that the number of vulnerable code fragments reduces when we enable the cache modelling. For example, three data leakage scenarios were identified in ocb3...
without cache modelling. However, only two of these leakage scenarios appear in the presence of cache modelling. The remaining one data leakage was identified as a false positive. This means that the sensitive data loaded into the cache was subsequently evicted by other memory accesses. Moreover, program encoder and des have different results when the cache configuration is changed. Specifically, the analysis with 2-ways cache only found one $N(\text{LS})$ style leakage, whereas two such leakage scenarios were discovered in 4-ways and 8-ways cache configurations. This happens as higher cache associativity may prevent some secrets to be evicted, leading to increased leakage.

The observations in the preceding two paragraphs highlight the need for cache modelling alongside symbolic execution. In particular, such a cache modelling facilitates the accurate detection of sensitive data leakage through cache side channel. Specifically, the modelling helps us to eliminate cases where a secret can be loaded along the speculative path. However, it cannot be leaked via the cache side channel (as the secret might be evicted before the attacker observes the cache state). Moreover, with a tunable cache configuration, the designer can gauge the potential spectre attacks on a variety of platforms employing caches.

The precision of KLEESpectre comes with the cost of solving the symbolic cache model. Thus both the analysis time and the solver time increase (as compared to the cache modelling being disabled). The time to solve the symbolic cache model depends on the number of memory accesses and the percentage of symbolic addresses. As observed from Table 3, the percentage of symbolic addresses is relatively low (the maximum is 42.39% for seed). Thus, the solver can finish within an acceptable time. Finally, except for hash and des, we did not observe a noticeable difference in analysis time with increased cache associativity. This means that our symbolic cache model scales well with respect to varying cache configurations.

### 6.4 Case study

```c
static int command_lcp(..., char **argv, ...) { 
    keylen2 = atoi(argv[3]);
    if ((keylen2 < 0) || /* VB */
        (keylen2 > (int)BITSOF(strlen(argv[2])))) { ... return -1; }
    ...
    lcp = fr_trie_path_lcp(..., keylen2, ...);
}
static int fr_trie_path_lcp(..., int keylen2, ...) { ...
    if (end_bit > keylen2) end_bit = keylen2;
    ...
    e2 = end_bit;
    ...
    xor <<= s2;
    xor |= lcp_en_bit[e2 - s2]; /* RS */
    lcp = xor2lcp[xor]; /* LS */
    ...
}
```

Fig. 5. Potential Spectre vulnerability in trie.c within project freeradius reported by oo7 [40]. The triplet ⟨VB, RS, LS⟩ is highlighted.

In this section, we evaluate the effectiveness of KLEESPECTRE in detecting Spectre vulnerabilities on real-world programs. For this purpose, we have chosen two case studies. The programs used in these case studies were reported to be vulnerable to Spectre attacks by oo7 [40]. The first program is trie from freeradius project and the other one is touch from Coreutils.

Freeradius is an open-source Remote Authentication Dial-In User Service (RADIUS) that usually processes sensitive data in its service. The reported leakage was discovered in function fr_trie_path_lcp() within file trie.c as shown in Figure 5. The file trie.c is in a library provided by freeradius and implements a path-compressed, level-compressed patricia tries [1]. We run KLEESPECTRE on tries.c while making all inputs as symbolic. This is to test whether KLEESPECTRE can detect the previously reported leakage and verify whether the leakage can be observed through an access-driven cache attack. Since there is no typical sensitive data in this program, we aim to detect all potential leakage (i.e. N(\(LS\))). This choice is justifiable, as trie belongs to library code and it might be invoked from a program that contains various secrets.

The analysis results are shown in Table 4. KLEESPECTRE analysis terminates in 1.1 hours, 1.2 hours and 1.23 hours for 2-ways, 4-ways and 8-ways cache configurations, respectively. KLEESPECTRE explored 13,751 normal paths and 412,406 speculative paths. Moreover, KLEESPECTRE detects exactly one sensitive data leakage through the cache side-channel. This leakage is observed over several execution paths (6438, 6440 and 6440 execution paths for 2-ways, 4-ways and 8-ways caches, respectively).

Our second case study is chosen from Coreutils project. Coreutils includes basic file, shell commands and text manipulation utilities and is widely adopted by the GNU operating system. We run touch on KLEESPECTRE for 12 hours with 2-ways, 4-ways and 8-ways cache configurations. We aim to test whether KLEESPECTRE can detect the Spectre vulnerability reported in touch. Similar to trie, the program touch does not contain a typical secret. Therefore, our goal is to detect leakage N(\(LS\)).

Our results are shown in Table 5. KLEESPECTRE explored 14,654 normal paths along with 1,331,152 speculative paths. Moreover, KLEESPECTRE detected 388 vulnerable branches, 40 RS and eight N(\(LS\)) style leakage without cache modelling. In the presence of cache modelling, KLEESPECTRE verified only four out of these eight leakage scenarios. Finally, these leakage scenarios were discovered over 9,725, 9,830 and 9,843 different paths for 2-ways, 4-ways and 8-ways caches, respectively.
One vulnerable code fragment in program `touch`, as identified by KLEESPECTRE, is shown in Figure 6. In this code fragment, variable `yyn` is used to access array `yytable[]` only if `yyn` is between 0 and `YYLAST`. However, if the branch that marked as VB is mis-predicted, then the code marked as RS can load the data outside of array `yytable[]` to the memory. This, then leaks the data to the cache side-channel through code `yylen = yyr2[yyn];` as marked LS in the figure.

7 THREATS TO VALIDITY

Path explosion. Path explosion is a major challenge in the symbolic execution. KLEESPECTRE is based on symbolic execution, which does not scale to large programs while exploring all feasible program paths. In particular, KLEESPECTRE forks more paths than the classical symbolic execution for performing the speculative execution. However, only limited number of instructions are executed on the speculative paths in KLEESPECTRE, which is bounded by the Speculative Execution Windows (SEW). Thus, as observed in our evaluation, KLEESPECTRE has similar complexity with the classical symbolic execution. Moreover, the existing methods to alleviate the path explosion can also be used by KLEESPECTRE, for example, state merging [25, 30]. Specifically, the speculative states can be merged similarly as the classical states when the control flows of a program merges. Besides, the symbolic execution can be guided by the low-cost static analysis in such a fashion that a static analysis can be performed to roughly locate the vulnerable code and prune the redundant paths during the construction of symbolic execution tree.

Precise modeling of program behavior. The program behavior running on the hardware may not be the same as it is in the symbolic execution. This is because KLEESPECTRE uses bitcode that may not replicate exactly the same behavior as the final binary code due to the compiler optimization. For example, the program may have more memory accesses when running on the hardware than it is during the symbolic execution due to the register spilling. However, KLEESPECTRE is designed as an over-approximation method that it captures all necessary memory accesses and detects all potentially secret leakage. This results in the absence of false negatives. In other words, KLEESPECTRE guarantees that all leakage in the real execution can be detected. However, KLEESPECTRE can generate false positives. For instance, the leakage detected by KLEESPECTRE may not be exploitable in the real hardware.
Explicit modeling of cache. Our KLEESpectre approach is applicable to execution platforms employing caches. Given the prevalent presence of caches in both mainstream and embedded platforms, KLEESpectre takes an approach to explicitly model the cache within symbolic execution. Such a cache model may facilitate further development of analysis tools (e.g. timing analysis) that requires precise cache modeling. Nevertheless, some cache architectural parameters need to be known (e.g. cache sets, associativity) for such analysis and improperly setting such parameters may lead to both false positives and false negatives.

The setting of Speculative Execution Window (SEW). An LLVM IR instruction may correspond to many micro-instructions in the binary. Thus, improperly chosen SEW affects the analysis result and precision. For example, if the SEW is too small, then KLEESpectre may introduce false negatives. In contrast, when the SEW is set too large, it makes the analysis much slower. Moreover, this may also introduce false positives, as some secrets may never be accessed speculatively, yet such secret accesses might be accounted in KLEESpectre when SEW is configured an unlikely large value. Nonetheless, one LLVM IR instruction is at least transformed to one micro-instruction. As a result, when using the LLVM IR code, the set of memory instructions that conflict in the cache has at least as many corresponding memory instructions in the compiled binary. This means, if a secret remains in the cache while executing the binary code, then the secret will also remain in the cache while executing the corresponding LLVM IR, using the same execution environment. Therefore, if we set the SEW to the size of the Re-order Buffer (RoB) for a given platform, then KLEESpectre should catch all potential data leakage through the cache side-channel.

8 RELATED WORK

Spectre-style attack mitigation in hardware. New hardware-related solutions have been proposed [21, 27, 37, 39, 42, 44] to verify or defend against cache side-channel attacks. IODINE [21] presented a clock-precise, constant-time approach to eliminate the timing side channels in the hardware. InvisiSpec [44] uses a Speculative Buffer(SB) to temporarily hold the data during speculative execution instead of directly loading the data to Cache. CleanupSpec [39] improves InvisiSpec that it performs an “Undo” operation in the cache while encountering a pipeline squash on speculative execution. DAWG [27] partitions the cache ways to limit the data leakage across different security domains. CEASER [37] obfuscates the cache set mapping to prevent the sensitive cache state being leaked to an attacker, and an improved version of CEASER for defending against stronger attacker model is presented in work ScatterCache [42]. However, all proposed hardware solutions are unlikely to be adopted by the legacy systems. Moreover, our approach is orthogonal to hardware-level mitigation to defend against Spectre attacks. In particular, we detect potential information leakage through Spectre-style attacks in arbitrary applications running on any processor employing branch speculation and caches.

Spectre-style attack mitigation in software. As for the software-based approach, Speculative Load Hardening [10] (SLH) is a mitigation technique for Spectre variant 1, adopted by the LLVM compiler. SLH identifies the potentially vulnerable code fragments where memory accesses depend on the conditional branches and then inserts hardening instruction sequence to nullify the pointers that may leak the data. SLH hardens the RS stage of the vulnerable code. Thus, the secret data cannot be loaded into the cache during speculative execution after nullifying the crucial pointers. As SLH repairs the program at every conditional branch and inserts the hardening instructions, SLH significantly increases execution time. Specifically, it introduces 36% performance overhead. Oleksenko et al. [33] present mitigation of Spectre variant 1 attack by delaying the execution of the vulnerable instructions via introducing artificial data dependencies instead of serialization instructions to stop speculative execution altogether. These methods lack accurate analysis and
overestimate vulnerable code fragments leading to a significant performance overhead in the repaired programs. In contrast, KLEESpectre detects the vulnerable code fragments, which, in turn, can aid software-level mitigation. In particular, by focusing on these vulnerable code fragments, existing software-level mitigation can be applied with a reduced performance overhead.

Microsoft Visual C/C++ compiler [15] enables mitigation of Spectre Variant 1 through a compiler option that inserts "lfence" serializing instruction at potentially vulnerable code. However, this technique successfully mitigates only two out of 15 Spectre litmus tests [29]. oo7 [40] is the first work proposed to mitigate Spectre-style attacks via modelling speculative execution in static analysis. oo7 works on binary and leverages taint analysis to track the vulnerable branches and memory operations that lead to Spectre-style vulnerabilities. oo7 can effectively detect and fix Spectre-style attacks, but may still produce false positives due to conservative static analysis. We also empirically show the effectiveness of KLEESpectre over oo7 in our evaluation.

Speculative semantics and formal models. Cheang et. al[13] proposed a formal approach to identify the speculative leakage based on the formulation of a new security property called trace property-dependent observational determinism (TPOD). A program containing transient execution vulnerabilities would violate the TPOD. However, this work is a concept formulation and its implementation aspects are unclear. Disselkoen et al. [16] is another micro-architectural model based on pomsets to model speculative evaluation. This work focuses on the modelling of speculation while considering the relaxed consistency model. Cauligi et al. [11] proposed a model for constant-time programming in the presence of micro-architectural features such as out-of-order and speculative execution. InSpectre [6] presents an infrastructure to model the our-of-order execution by providing the semantics of modern micro-architectural features. InSpectre can model existing and new hardware vulnerabilities. SpecFuzz [34] introduced the speculative execution to fuzz testing for detecting the Spectre vulnerability. However, SpecFuzz does not model cache. SPECTECTOR [23] presents a principled approach using speculative non-interference in symbolic execution to discover data leakage. However, all the aforementioned works focus on modelling the speculative semantics to define the vulnerability in program analysis. These works do not model and check the possibility of cache side-channel attacks. In KLEESpectre, this is achieved via our cache modelling.

Side-channel analysis without cache modelling. Chen et al. [14] presented Quantitative Caresian Hoare Logic-based tool THEMIS to detect side-channel vulnerabilities in Java application. They perform a lightweight static analysis and develop a precise relational verification technique for reasoning about k-safety. This results in checking the non-interference property of a program. Antonopoulos et al. [5] achieve a similar goal by using a novel self-decomposition technique to split a program into partitions. DiffFuzz [32] introduced the fuzzing technique to side-channel analysis. DiffFuzz automatically detects side-channel vulnerabilities by analyzing two versions of the program and using resource-guided heuristics to find inputs that maximize the difference in resource consumption between secret dependent paths. Doychev et al. [17] implemented abstract interpretation based side-channel analysis on binary.

All the techniques, as mentioned in the preceding paragraph, focus on timing-driven attacks and they only consider the timing for each instruction without explicitly modelling the cache behaviour. The lack of an accurate timing model results in less precision and overly conservative analysis in terms of detecting cache-side channel attacks. Moreover, the absence of modelling speculative execution makes these works unable to detect the Spectre-style attacks.

Side-channel attack identification via cache modelling. Casym [7] presents a cache-aware symbolic execution to identify and fix cache side-channel vulnerabilities. Casym provides two cache models: the infinite cache model of caches with infinite size and associativity, and the age model
that tracks the distance of a memory access from its most recent access. However, the description of the cache models in Casym are sketchy and hence prevents reproducibility. Besides, Casym does not consider speculative execution within its cache models. CACHEFIX [12] is another cache side-channel verification tool that can detect and fix timing-driven and trace-driven cache attacks via symbolic model checking. However, CACHEFIX does not consider speculative execution paths during model checking.

Abstract interpretation is a static analysis approach that has been effectively adapted for cache hit/miss modelling in Worst-Case Execution Time (WCET) estimation. Wu et al. [43] introduce abstract interpretation to cache side-channel attack detection by extending it to cover speculative execution. This approach targets timing-based side-channel attacks, but it does not handle Spectre attacks. A similar approach is embodied by CacheAudit [18], however, the CacheAudit approach does not consider speculative execution semantics. Wang et.al [41] proposed a secret-augmented side-channel identification tool CacheS to detect timing-based cache side-channel attacks based on abstract interpretation. CacheS maintains a balance between static analysis and symbolic execution for precision and scalability since it only tracks secrets and dependencies on them to achieve precision, and it collects only coarse-grained information to achieve scalability. Compared with KLEESPECTRE, CacheS does not model speculative execution. Specusym [24] captures speculative symbolic execution for timing attacks and cannot be used for detecting Spectre attacks.

In summary, existing works cannot be used for detecting and confirming Spectre attacks. KLEESpectre can efficiently detect the leakage scenario in the speculative path, and the cache model can further confirm whether the leakage can be observed.

9 CONCLUSION

We have presented a new software testing tool named as KLEESPECTRE to expose the micro-architectural features to the software testing. Micro-architectural features such as speculative execution and caches are generally ignored by traditional software testing. This hides the vulnerabilities caused by invisible micro-architectural behaviours when a program runs on the hardware. KLEESPECTRE makes these behaviours visible in the software testing via modeling the speculative execution and caches within the traditional symbolic execution. Our experiments show that KLEESPECTRE can effectively detect the BCB vulnerabilities and the cache model can make such detection more accurate.

KLEESPECTRE takes a step forward to extend software testing methodologies to systematically discover vulnerabilities dependent on micro-architectural features. Our tool also provides an open platform to extend our methodologies as more Spectre style attacks are being discovered. To enable further research in the area, we have made our tool and subjects publicly available at

https://github.com/winter2020/kleespectre

ACKNOWLEDGMENTS

This research is supported by the National Research Foundation, Prime Minister’s Office, Singapore under its National Cybersecurity R&D Program (Award No. NRF2014NCR-NCR001-21) and administered by the National Cybersecurity R&D Directorate.

REFERENCES


KLEE Spectre: Detecting Information Leakage through Speculative Cache Attacks via Symbolic Execution


