

Detile: Fine-Grained Information Leak Detection in Script Engines

Robert Gawlik, Philipp Koppe, Benjamin Kollenda,
Andre Pawlowski, Behrad Garmany and Thorsten Holz

Horst Görtz Institute for IT-Security (HGI)
Ruhr-Universität Bochum, Germany
{firstname.lastname}@rub.de

Abstract. *Memory disclosure attacks* play an important role in the exploitation of memory corruption vulnerabilities. By analyzing recent research, we observe that bypasses of defensive solutions that enforce control-flow integrity or attempt to detect return-oriented programming require memory disclosure attacks as a fundamental first step. However, research lags behind in detecting such information leaks.

In this paper, we tackle this problem and present a system for fine-grained, automated detection of memory disclosure attacks against scripting engines. The basic insight is as follows: scripting languages, such as JavaScript in web browsers, are strictly sandboxed. They must not provide any insights about the memory layout in their contexts. In fact, *any* such information potentially represents an ongoing memory disclosure attack. Hence, to detect information leaks, our system creates a clone of the scripting engine process with a re-randomized memory layout. The clone is instrumented to be synchronized with the original process. Any inconsistency in the script contexts of both processes appears when a memory disclosure was conducted to leak information about the memory layout. Based on this detection approach, we have designed and implemented DETILE (detection of information leaks), a prototype for the JavaScript engine in Microsoft's Internet Explorer 10/11 on Windows 8.0/8.1. An empirical evaluation shows that our tool can successfully detect memory disclosure attacks even against this proprietary software.

1 Introduction

Over the last years, many different techniques were developed to prevent attacks that exploit spatial and temporal memory corruption vulnerabilities (see for example the survey by Szekeres et al. [52]). As a result, modern operating systems deploy a wide range of defense methods to impede a successful attack. For example, *Data Execution Prevention* (DEP) [38] marks data as non-executable and thus an attacker is prohibited from injecting data into a vulnerable application that is later on interpreted as code. Furthermore, *Address Space Layout Randomization* (ASLR) [43] randomizes the memory layout either once during the boot process or every time a process is started. Since the attacker lacks information about the exact memory layout, it is harder for her to predict where her shellcode or reusable code are located.

Besides these widely deployed techniques, many other defenses were proposed in the literature in the last years [52]. Most notably, the enforcement of *control flow integrity* (CFI) is a promising technique to prevent a whole class of memory corruption vulnerabilities [1]. The basic idea behind CFI is to verify that each control flow transfer leads to a valid target based on a control flow graph that is either statically pre-computed or dynamically generated. Several implementations of CFI with different design constraints, security goals, and performance overheads were published (e.g., [21, 65, 66]).

A general observation is that the first step in modern attacks is based on a *memory disclosure attack* (also referred to as *information leak*): the adversary finds a way to read a (raw) memory pointer to learn some information about the virtual address space of the vulnerable program. Generally speaking, the attacker can then de-randomize the address space based on this leaked pointer (thus bypassing ASLR), use ROP to bypass DEP, and finally execute shellcode of her choice. Modern exploits leverage information leaks as a fundamental primitive. Furthermore, recent CFI and ROP defense bypasses use memory disclosures as well. For example, Snow et al. introduced *Just-In-Time Code Reuse* attacks (JIT-ROP [48]) to bypass fine-grained ASLR implementations by repeatedly utilizing an information leak. *G-Free* [39], a compiler-based approach against any ROP attack, was recently circumvented by Athanasakis et al. [3]. Their technique requires successive information leaks to disclose enough needed information. Gökteş et al. demonstrated several bypasses of proposed ROP defenses and their exploit needs an information leak as a first step [27]. An information leak is also needed by Song et al., who showed that dynamic code generation is vulnerable to code injection attacks [49]. Similarly, *Counterfeit Object-oriented Programming* (COOP [44]) needs to disclose the location of *vtables* to mount a subsequent control-flow hijacking attack by reusing them. Disclosures are also utilized by *memory oracles* to weaken various defenses [24]. *All* of these offensive bypasses utilized an information leak as a first step and implemented the attack against a web browser.

Another general observation is that script engines in web browsers are commonly utilized by adversaries to abuse information leaks in practice. Browser vulnerabilities are prevalent and as the yearly *pwn2own* competition shows, researchers successfully use them to take control of the machine. Notably, most of these attacks are based on vulnerabilities that create an information leak utilizing the script engine.

In this paper, we take these observations into account and propose a technique for fine-grained, automated detection of memory disclosure attacks against script engines at runtime. Our approach is based on the insight that information leaks are leveraged by state-of-the-art exploits to learn the placement of modules—and thereby code sections—in the virtual address space in order to bypass ASLR. Any sandboxed script context is forbidden to contain memory information, i.e., no script variable is allowed to provide a memory pointer. As such, a viable approach to detect information leaks is to create a clone of the to be protected process with a re-randomized address space layout, which is

Table 1: Defenses and offensive approaches utilizing an information leak in browsers to weaken or bypass the specific defense. All mentioned attacks are mitigated by DETILE.

Protection flavor	Defense	Weakened/Bypassed by	Mitigated by DETILE
Address randomization	Fine-grained ASLR [29]	Just-In-Time Code Reuse [48]	✓
Code-reuse protection	RopGuard [23], KBouncer [41], ROPecker [12]	Size Does Matter [27], Anti-ROP Evaluation [45], COOP [44]	✓
Code-reuse protection	G-Free [39]	Browser JIT Defense Bypass [3], COOP [44]	✓
Coarse-grained CFI	CCFIR [65], BinCFI [66]	Stitching the Gadgets [18], Out of Control [26], COOP [44]	✓
Fine-grained CFI	IFCC [53], VTV [53]	Losing Control [13]	✓
Information-hiding	Oxymoron [4]	Vtable disclosure [19], Crash-Resistance [24] COOP [44]	✓
Information-hiding	CPI linear region [33]	Crash-Resistance [24]	✓
Execution randomization	Isomeron [19]	Crash-Resistance [24]	✓
Randomization/Information-hiding	Readactor [15]	Crash-Resistance [24], COOP [44]	✓

instrumented to be synchronized with the original process. An inconsistency in the script contexts of both processes can only occur when a memory disclosure vulnerability was exploited to gain information about the memory layout. In such a case, the two processes can be halted to prevent further execution of the malicious script. An overview of bypassed defenses by specific attacks which are mitigated by our approach is shown in Table 1.

We have implemented a prototype of our technique in a tool called DETILE (detection of information leaks). We extended Internet Explorer 10/11 (IE) on Windows 8.0/8.1 to create a synchronized clone of each tab and enforce the information leak checks. We chose this software mainly due to two reasons. First, IE is an attractive target for attackers as the large number of vulnerabilities indicates. Second, IE and Windows pose several interesting technical challenges since it is a proprietary binary system that we need to instrument and it lacks fine-grained ASLR. Evaluation results show that our prototype is able to re-randomize single processes without significant computational impact. Additionally, running IE with our re-randomization and information leak detection engine imposes a performance hit of $\sim 17\%$ on average. Furthermore, empirical tests with real-world exploits also indicate that our approach is usable to unravel modern and unknown exploits which target browsers and utilize memory disclosures.

In summary, our main contributions in this paper are:

- We present a system to tackle the problem of information leaks, which are frequently used in practice by attackers as an exploit primitive. More specifically, we propose a concept for fine-grained, automated detection of information leaks with per process re-randomization, dual process execution, and process synchronization. An extended version of this paper with more technical details is available as a technical report [25].

- We show that dual execution of highly complex, binary-only software such as Microsoft’s Internet Explorer is possible without access to the source code, whereby two executing instances operate deterministic to each other.
- We implemented a prototype for IE 10/11 on Windows 8.0/8.1. We show that our tool can successfully detect several real-world exploits, while producing no alerts on highly complex, real-world websites.

2 Technical Background

In the following, we briefly introduce several concepts needed to understand the challenges we were confronted with when developing DETILE.

2.1 N-Variant Systems

N-Variant or *Multi-Execution* systems evolved from fault-tolerant environments to mitigation systems against security critical vulnerabilities [9, 14, 30, 54]. Our concept of DETILE incorporates similar ideas like dual process execution and dual process synchronization. However, our approach is constructed specifically for scripting engines, and thus, is more fine-grained: While DETILE operates and synchronizes processes on the scripting interpreter’s bytecode level, n-variant systems intercept only at the system call level. One drawback for these conventional systems is that they are prone to *Just-In-Time Code-Reuse* (JIT-ROP [48]) and *Counterfeit Object-oriented Programming* (COOP [44]) attacks, while DETILE is able to detect these (see Sections 3.1 and 6 and for details).

2.2 Windows ASLR Internals

Address Space Layout Randomization (ASLR) is a well-known security mechanism that involves the randomization of stacks, heaps, and loaded images in the virtual address space. Its purpose is to leave an attacker with no knowledge about the virtual memory space in which code and data lives. Combined with DEP, ASLR makes remote system exploitation through memory corruption techniques a much harder task. While brute-force attacks against services that automatically restart are possible [6], such attacks are typically not viable in practice against web browsers.

In Windows, whenever an image is loaded into the virtual address space, a section object is created, which represents a section of memory. These objects are managed system-wide and can be shared among all processes. Once a DLL is loaded, its section object remains permanent as long as processes are referencing it. This concept has the benefit that relocation takes place once and whenever a process needs to load a DLL, its section object is reused and the view of the section is mapped into the virtual address space of the process, making the memory section visible. This way, physical memory is shared among all processes that load a specific DLL whose section object is already present. In particular, as long as the virtual address is not occupied, each image is loaded at the same virtual address among all running usermode processes.

2.3 WOW64 Subsystem Overview

64-bit operating systems are the systems of choice for today’s users: 64-bit processors are widely used in practice, and hence Microsoft Windows 7 and later versions are usually running in the 64-bit version on typical desktop systems. However, most third-party applications are distributed in their 32-bit form. This is for example the case for Mozilla Firefox, and also for parts of Microsoft’s Internet Explorer. As our framework should protect against widely attacked targets, it needs to support 32-bit and 64-bit processes. Therefore, the *Windows On Windows 64* (shortened as *WOW64*) emulation layer plays an important role, as it allows legacy 32-bit applications to run on modern 64-bit Windows systems.

Executing a user-mode 32-bit application instructs the kernel to create a WOW64 process. According to our observations, it creates the program’s address space and maps the 64-bit and 32-bit *NT Layer DLL* (`ntd11.dll`) and the main executable into it. Even when a program may have been started in suspended mode, these three modules are already available. Afterwards, WOW64 layer DLLs are mapped, which mediate several necessary transitions between 64-bit and 32-bit at runtime [43]. Subsequent 32-bit DLLs are mapped into the address space via `LdrLoadD11` of the 32-bit `ntd11.dll`. The first of them is `kernel32.dll`. The loader assures that it is mapped to the same address in each WOW64 process system wide, using a unique address per reboot. It therefore compares its name to the hardcoded “KERNEL32.DLL” string in `ntd11.dll` upon loading. If the loader is not able to map it to its preferred base address, process initialization fails with a conflicting address error. As process based re-randomization plays a crucial role in our framework, this issue is handled such that each process contains its `kernel32.dll` at a different base address (see Section 4.1). After mapping `kernel32.dll`, all other needed 32-bit DLLs are mapped into the address space.

2.4 Internet Explorer Architecture

IE is developed as multi-process application [64]. That means, a 64-bit main frame process governs several 32-bit WOW64 tab processes, which are isolated from each other. The frame process runs with a medium integrity level and isolated tab processes run with low integrity levels. Hence, tab processes are restricted and forbidden to access all resources of processes with higher integrity levels [37]. This architecture implies that websites opened in new tabs can lead to the start of new tab processes. These have to incorporate our protection in order to protect IE as complete application against information leaks (see Section 4).

2.5 Scripting Engines

In the context of IE, mainly two scripting engines are relevant and we briefly introduce both.

Internet Explorer Chakra. With the release of Internet Explorer 9, a new JavaScript engine called *Chakra* was introduced. Since Internet Explorer 11, Chakra exports a documented API which enables developers to embed the engine into their own applications. However, IE still uses the undocumented internal COM interface. Nevertheless, some Chakra internals were learned from the official API. The engine supports just-in-time (JIT) compiling of JavaScript bytecode to speed up execution. Typed arrays like integer arrays are stored as native arrays in heap memory along with metadata to accelerate element access. Script code is translated to JS bytecode on demand in a function-wise manner to minimize memory footprint and avoid generating unused bytecode. The bytecode is interpreted within a loop, whereby undocumented *opcodes* govern the execution of native functions within a switch statement. Dependent on the opcode, the desired JavaScript functionality is achieved with native code.

ActionScript Virtual Machine (AVM). The *Adobe Flash* plugin for browsers and especially for IE is a widely attacked target. Scripts written in *ActionScript* are interpreted or JIT-compiled to native code by the AVM. There is much unofficial documentation about its internals [7, 34]. Most importantly, it is possible to intercept *each* ActionScript method with available tools [28]. Thus, no matter whether bytecode is interpreted by the opcode handlers or JIT code is executed, we are able to instrument the AVM.

2.6 Adversarial Capabilities

Memory disclosure attacks are an increasingly used technique for the exploitation of software vulnerabilities [47, 48, 51]. In the presence of full ASLR, DEP, CFI, or ROP defenses, the attacker has no anchor to a memory address to jump to, even if in control of the instruction pointer. This is the moment where information leaks come into play: an attacker needs to read—in any way possible—a raw memory pointer in order to gain a foothold into the native virtual address space of the vulnerable program. As soon as the attacker can read process memory, she can learn the base addresses of loaded modules. Then, any code reuse primitives can be conducted to exploit a vulnerability in order to bypass DEP, ASLR, CFI [18] and ROP defenses [11, 27]. Another possibility is to leak code directly in order to initiate an attack and bypass ASLR [48]. Other mitigations like Microsoft’s Enhanced Mitigation Experience Toolkit (EMET) [36] cannot withstand capabilities of sophisticated attackers.

For applications with scripting capabilities, untrusted contexts are sandboxed (e.g., JavaScript in web browsers) and must not provide memory information. Thus, attackers use different vulnerabilities to leak memory information into that context [26, 47, 58]. We assume that the program we want to protect suffers from such a memory corruption vulnerability that allows the adversary to corrupt memory objects. In fact, a study shows that *any* type of memory error can be transformed into an information leak [52]. Furthermore, we assume that the attacker uses a scripting environment to leverage the obtained memory disclosure information at runtime for her malicious computations. This is consistent with

modern exploits in academic research [11, 18, 26, 27, 45] as well as in-the-wild [46, 55, 58–60]. Our goal is to protect script engines against such powerful, yet realistic adversaries.

3 System Overview

In the following, we explain our approach to tackle the challenge of detecting information leaks in script engines. Hence, we introduce the needed building blocks, namely per process re-randomization and dual process execution.

3.1 Main Concept

As described above, information leaks manifest themselves in the form of memory information inside a context which must not reveal such insights. In our case, this is any script context inside an application: high-level variables and content in a script must not contain memory pointers, which attackers could use to deduce image base addresses of loaded modules.

Unfortunately, a legitimate number and a memory pointer in data bytes received via a scripting function are indistinguishable. This leads us to the following assumption: a memory disclosure attack yields a memory pointer, which may be surrounded by legitimate data. The same targeted memory disclosure, when applied to a differently *randomized*, but otherwise *identical* process, will yield the same legitimate data, but a *different* memory pointer. Due to the varying base addresses of modules, different heap and stack addresses, a memory pointer will have a different address in the second process than in the first process. Thus, a master process and a cloned twin process—with different address space layout randomization—can be executed synchronized side-by-side and perform identical operations, e.g., execute a specific JavaScript function. In benign cases, the same data getting into the script context is equal for both processes. When comparing the received data of one process to the same data received in the second process, the only difference can arise because of a leaked memory pointer pointing to equal memory, but having a *different* address. In order to compare the data of the master and twin process, we have to instrument the interpreter loop of the script engine. We can instrument the `call` and `return` bytecodes to precisely check all outgoing data and therefore to detect an information leak.

Based on this principle, our prototype system launches the same script engine process twice with diverse memory layouts (see also Figure 1). The script engines are coupled to run in sync which enables checking for information leaks. In spirit, this is similar to n-variant systems [9, 14] and multi-execution based approaches [10, 17, 20]. However, our approach is more fine-grained since it checks and synchronizes the processed data on the bytecode level of the script context and is capable of detecting the actual information leak, instead of merely detecting an artifact of a successful compromise (i.e., divergence in the control flow).

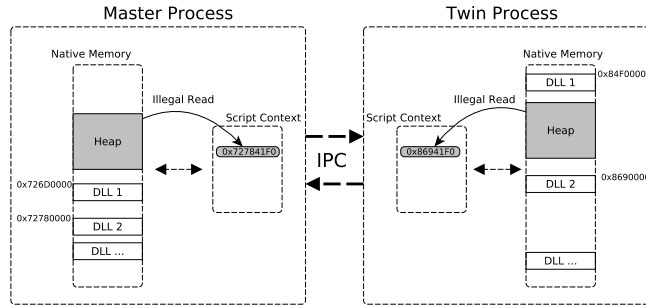


Fig. 1: Overview of our main information leak detection concept: The master process is synchronized with a re-randomized, but otherwise identical twin process. If a memory disclosure attack is conducted in the master, it appears as well in the twin. Due to the different randomization, the disclosure attack manifests itself in different data flowing into the script context and can be detected (0x727841F0 vs. 0x86941F0)

3.2 Per Process Re-randomization

To overcome the dilemma of modules having equal base addresses in different processes, we collect all base addresses of modules a process loads during its runtime. We refer to this first process, which is launched, as *master* process. A second process instance of the application known as the *twin* process is spawned. Upon its initialization, the base addresses gained from the master are occupied in the virtual address space of the twin. This forces the image loader to map the modules to other addresses than in the master process, as they are already allocated. We save us the time and trouble to re-randomize the stack and heap process-wise, as modern operating systems (e.g., Windows 8 on 64-bit) support it natively. Finally, we establish an *inter-process communication* (IPC) bridge between the master and twin process. This enables synchronized execution between them and comparison of data flows into their script contexts.

3.3 Dual Process Synchronization

After the re-randomization phase, both processes are ready to start execution at their identical entrypoints. After exchanging a handshake, both resume execution. In order to achieve comparable data for information leak checking, the executions of script interpreters in both processes have to be synchronized precisely. This is accomplished by intercepting an interpreter's native methods. Additionally, we install hooks inside the bytecode interpreter loop at positions where opcodes are interpreted and corresponding native functions are called. Thus, we perceive any high-level script method call at its binary level. The master drives execution and these hooks are the points where the master and twin process are synchronized via IPC. We check for information leaks by comparing binary data which returns as high-level data into the script context. All input data the master loads are stored in a cache and replayed to the twin process

to ensure they operate on the same source (e.g., web pages a browser loads). Built-in script functions that potentially introduce entropy (e.g., `Math.random`, `Date.now`, and `window.screenX` in JavaScript) interfere with our deployed detection mechanism, since they generate values inside the script context that are different from each other in the master and twin processes, respectively. Additionally, they may induce a divergent script control flow. Both occurrences would be falsely detected as memory disclosure. Thus we also synchronize the entropy of both processes by copying the generated value from the master to the twin process. This way the twin process continues working on the same data as the master process and we are creating a co-deterministic script execution.

4 Implementation Details

Based on the concepts of per process re-randomization and dual process execution, we implemented a tool called DETILE for Windows 8.0 and 8.1 64-bit. The current prototype is able to re-randomize on a per process basis and instrument Internet Explorer 10 and 11 to run in dual process execution mode.

4.1 Duplication and Re-randomization

In order to re-randomize processes and load images at different base addresses, we developed a duplicator which creates a program's master process. It enumerates the master's initial loaded images with the help of the Windows API (`CreateToolHelp32Snapshot`) before the master starts execution. Then, the twin process is created in suspended mode, and a page is allocated in the twin at all addresses of previously gathered image bases. We then need to trick the Windows loader into mapping `kernel32.dll` at a different base in the twin. This is achieved by leveraging the DebugAPI and via manipulating parameters at calls of `RtlEqualUnicodeString` in the 32-bit loader in the `ntdll.dll`. This way, the loader believes that a *different* DLL than `kernel32.dll` is going to be initialized and allows the mapping to a different base. It is the first DLL which is loaded after the WOW64 subsystem. Thus, all subsequent libraries that are loaded and import functions from `kernel32.dll` have no problems to resolve their dependencies using the remapped `kernel32.dll`. The loader maps them to different addresses, as their preferred base addresses are reserved. Although the DebugAPI is used, all steps run in a fully automated way. As a next step, the DebugAPI is detached and the main image is remapped to a different address. As it is already mapped even in suspended processes, this has to be done specifically. Additionally, `LdrLoadDll` in the twin process is detoured to intercept new library loads and map incoming images to different addresses than in the master. Technical details about our remapping can be found in the technical report [25].

We were not able to re-randomize `ntdll.dll` because it is mapped into the virtual address space very early in the process creation procedure. Attempts to remap `ntdll.dll` later on did not succeed due to callbacks invoked by the kernel. The implications of a non re-randomized `ntdll.dll` are discussed in Section 7.

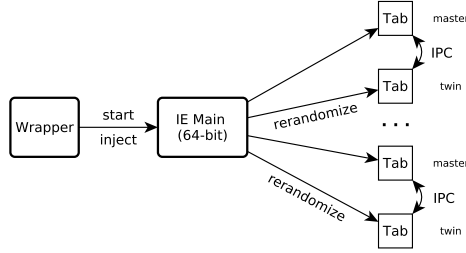


Fig. 2: DETILE running with Internet Explorer. A 64-bit duplicator library is injected into the main IE frame process to enable it creating and rerandomizing twin tab processes for each master tab process, by itself. The main IE frame also injects a 32-bit DLL into each tab process to allow synchronization, communication between master and twin, and information leak detection.

Note that this design works also with pure 64-bit processes. However, frequently attacked applications like tab processes of Internet Explorer are 32-bit and are running in the WOW64 subsystem. Hence, our framework has to protect them as well. The following explains how DETILE achieves this support.

While the above explained logic is sufficient to duplicate and re-randomize a single-process program, additional measures have to be taken in the case of multi-process architecture applications like Internet Explorer. Therefore, we developed a wrapper which starts the 64-bit main IE frame process and injects a 64-bit library, which we named duplicator library (see Figure 2). This way, we modify the frame process, such that each time a tab process is started by the frame process, a second tab process is spawned. The first becomes the master, the second the twin. This is achieved via detouring and modifying the process creation of the IE frame. Additionally, our above explained re-randomization logic is incorporated into the duplicator library to allow the main IE frame process itself to re-randomize its spawned twins at creation time. To protect each new tab which is run by the IE frame, we ensure that each tab is run in a new process and gets a twin. To enable communication, synchronization, and detection of information leaks, the duplicator injects also a 32-bit library into the master and the twin upon their creation by the main IE frame process.

4.2 Synchronization

We designed our prototype to be contained in a DLL which is loaded into both target instances. To reliably intercept all script execution, we hook `LdrLoadDll` to initialize our synchronization as early as possible once the engine has been loaded. After determining the role (master or twin), the processes exchange a short handshake and wait for events from the interpreter instrumentation. While most of our work is focused on the scripting engine, we also instrument parts of `wininet.dll` to provide basic proxy functionality. The twin receives an exact copy of the web data sent to the master to ensure the same code is executed.

Entropy Normalization. The synchronization of script execution relies heavily on the identification of functions and objects introducing entropy into the script context. Values classified as entropy are overwritten in the twin with the value received from the master. This ensures that functions such as `Math.random` and `Date.now` return the exact same value, which is crucial for synchronous execution. While it is obvious for `Date.now`, it is not immediately clear for other methods. Therefore, *entropy inducing* methods are detected and filtered incrementally during runtime. Hence, if a detection has triggered but the cause was not an information leak, it is included into the list of entropy methods.

Rendezvous and Checking Points. Vital program points where master and twin are synchronized are bytecode handler functions. If a handler function returns data into the script context, it is first determined if the handler function is an entropy inducing function. However, the vast majority of function invocations and object accesses do not introduce entropy and are checked for equality between master and twin on the fly. If a difference is encountered that is not classified as entropy, we assume that an information leak occurred and take actions, namely logging the incident and terminating both processes.

4.3 Chakra Instrumentation

The Chakra JavaScript Engine contains a JIT compiler. It runs in a dedicated thread, identifies frequently executed (so called *hot*) functions and compiles them to native code. Our current implementation works on script interpreters, hence we disabled the JIT compiler. This is currently a prototype limitation whose solution we discuss in Section 7.

In order to synchronize execution and check for information leaks, we instrumented the main loop of the Chakra interpreter, which is located in the `Js::InterpreterStackFrame::Process` function. It is invoked recursively for each JavaScript call and iterates over the variable length bytecodes of the JavaScript function. The main loop contains a `switch` statement, which selects the corresponding handler for the currently interpreted bytecode. The handler then operates on the JavaScript context dependent on the operands and the current state. In the examined Chakra versions, we observed up to 648 unique bytecodes. Prior to the invocation of a bytecode handler, our instrumentation transfers the control flow to a small, highly optimized assembly stub, which decides whether the current bytecode is vital for our framework to handle.

We intercept all `call` and `return` as well as necessary `conversion` bytecodes in order to extract metadata such as JavaScript function arguments, return values, and conversion values. `Conversion` bytecodes handle dynamic type casting, native value to JavaScript object and JavaScript object to native value conversions. Additionally, we intercept engine functions that handle implicit type casts at native level, because they are invoked by other bytecode handlers as required and have no bytecode equivalents themselves. Furthermore, all interception sites support the manipulation of the outgoing native value or JavaScript object for the purpose of entropy elimination in the JavaScript context of the twin process.

4.4 AVM Instrumentation

Instrumentation of the AVM is based on prior work of F-Secure [28] and Microsoft [34]. We hook at the end of the native method `verifyOnCall` inside `verifyEnterGPR` to intercept ActionScript method calls and retrieve ActionScript method names. At these points, master and twin can be synchronized. Parameters flowing into an ActionScript method and return data flowing back into the ActionScript context can be dissected, too. They are also processed inside the method `verifyEnterGPR`. Based on their high level ActionScript types, the parameters and return data can be compared in the master and twin. This way, we can keep the master and twin in sync at method calls, check for information leaks and mediate entropy data from the master to the twin.

5 Evaluation

In the following, we present evaluation results for our prototype implementation of DETILE in the form of performance and memory usage benchmarks. The benchmarks were conducted on a system running Windows 8.0/8.1 that was equipped with a 4th generation Intel i7-4710MQ quad-core CPU and 8GB DDR3 RAM. Furthermore, we demonstrate how our prototype can successfully detect several kinds of real-world information leaks.

5.1 Re-randomization of Process Modules

We evaluated our re-randomization engine according to its effectiveness, memory usage, and performance.

Effectiveness. We applied re-randomization to internal Windows applications and third-party applications, to verify that modules in the twin are based at different addresses than in the master. We therefore compared base addresses of all loaded images between the two processes and confirmed that all images in the twin process had a different base address than in the master, except `ntdll.dll`. See the discussion in Section 7 for details on the difficulties of remapping the 64-bit and 32-bit NT Layer DLLs. The extended version of this paper lists important Windows DLLs, re-randomized in different processes running *simultaneously* on a *single* user session [25].

Physical Memory Usage. To inspect the memory overhead of our re-randomization scheme, we measured the working set characteristics for different master and re-randomized twin processes compared to native processes. Figure 3 shows the memory working sets of three applications. *ReASLR* denotes thereby the re-randomization within a single process. DE means that two processes are running, whereby the master’s randomization is kept native while the twin is re-randomized. The applications besides IE are only included to measure the memory overhead and are not synchronized. We calculate the memory overhead of per process re-randomization (*ReASLR*) of a *single* process as follows:

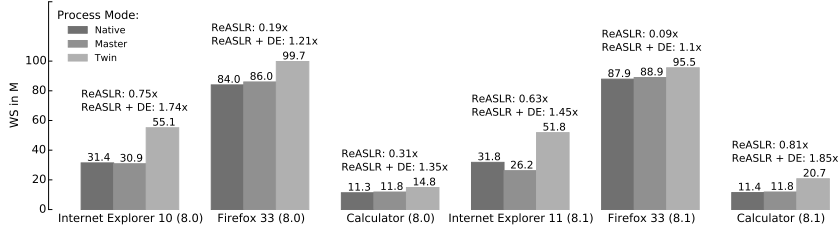


Fig. 3: Memory overhead of re-randomization and dual execution measured via working set (WS) consumption in megabytes (M): Native processes on Windows 8.0 and 8.1 are contrasted to their counterparts running in re-randomized dual execution mode (master and twin).

$$Overhead(ReASLR) = \frac{WS(Twin)}{WS(Native)} - 1$$

Thus, the overall memory overhead based on working sets is 0.46 times. When running a program or process in per process re-randomization *and* dual process execution (*DE*), we have to include both master and twin into the memory overhead calculation. Therefore, the overhead is calculated by

$$Overhead(ReASLR + DE) = \frac{WS(Twin) + WS(Master)}{WS(Native)} - 1$$

Its overall value is 1.45 times. Note that memory working sets can highly vary during an application’s runtime, and thus, are difficult to quantify. The measurements shown in Figure 3 were performed after the application has finished startup, and was waiting for user input (i.e., it was idle and all modules were loaded and initialized). Due to additional twins for master processes, the overall additional memory is about one to two times per protected process. The technical report provides more details on the working set characteristics [25].

Re-randomization and Startup Time. When a program is started the first time after a reboot, the kernel needs to create section objects for image modules. Hence, the first start of a program always takes longer than subsequent starts of the same program. To measure the additional startup and module load times our protection introduces, we first run each program natively once to allow the kernel to create section objects of most natively used DLLs, and close it afterwards. We then start the program natively without protection and measure the time until it is idle and all of its initial modules are loaded. In the same way, we measure the time from process creation until both the master and twin process have their initial modules loaded. The startup comparison can be seen in Table 2. As expected, the startup times of applications protected with our approach are approximately doubled. This is caused by the fact that a twin process needs to be spawned for each master that should be protected.

Table 2: Startup times in seconds and startup slowdowns of native 32-bit applications compared to their counterparts running with per process re-randomization and dual process execution on Windows 8.0 and Windows 8.1 (both 64-bit).

	Native (8.0)	ReASLR+DE (8.0)	Slowdown	Native (8.1)	ReASLR+DE (8.1)	Slowdown
IE tab spawn	0.9163 s	2.0710 s	1.3x	0.5194 s	1.3082 s	1.5x
Firefox	0.9624 s	1.8064 s	0.9x	1.3823 s	1.5441 s	0.1x
Calculator	0.3484 s	0.3610 s	0.0x	0.4391 s	0.6599 s	0.5x

Table 3: Native script execution of IE 11 on Windows 8.1 64-bit compared to the script execution of IE 11 instrumented with DETILE. Execution time is measured in milliseconds using the internal F12 developer tools provided by IE.

Website	google.com	facebook.com	youtube.com	yahoo.com	baidu.com	wikipedia.org	twitter.com	qq.com	taobao.com	linkedin.com	amazon.com	live.com	google.co.in	sina.com.cn	hao123.com
Native	425	774	1196	3674	1108	472	599	2405	645	439	958	254	483	3360	373
DEFILE	482	961	1519	4722	1339	513	623	2724	824	517	1210	275	517	4269	379
Overhead	13.4%	24.1%	27%	28.5%	20.8%	8.6%	4%	13.2%	27.7%	17.7%	26.3%	8.2%	7%	27%	1.6%

5.2 Detection Engine

Next, we evaluate the impact of DETILE on the user experience and its effectiveness in detecting information leaks.

Script Execution Time and Responsiveness. We used the 15 most visited websites worldwide [2] to test how the current prototype interferes with the normal usage of these pages. Besides the subjective impression while using the page, we utilized the F12 developer tools of Internet Explorer 11 to measure scripting execution time provided by the *UI Responsiveness* profiler tab. These tests were performed using Windows 8.1 64-bit and Internet Explorer 11. While we introduce a performance hit of around 17.0% on average, the subjective user experience was not noticeably affected. This is due to IE’s deferred parsing, which results in displaying content to the user before all computations have finished.

Information Leak Detection. We tested our approach on a pure memory disclosure vulnerability (CVE-2014-6355) which allows illegitimately reading data due to a JPEG parsing flaw in Microsoft’s Windows graphics component [61]. It can be used to defeat ASLR by reading leaked stack information back to the attacker via the `toDataURL` method of a `canvas` object. We successfully detected this leak at the point of the call to `toDataURL` in the master and twin process. In the same way, detection was successful for an exploit for a similar bug (CVE-2015-0061 [62]).

To further verify our prototype, we evaluated it against an exploit for CVE-2011-1346, a vulnerability that was used in the pwn2own contest 2011 to bypass ASLR [63]. As this memory disclosure bug is specific for IE 8, we ported the vulnerability into IE 11. An uninitialized `index` attribute of a new `HTML option`

element is used to leak information. Similarly, we successfully detected this exploitation attempt when the `index` attribute was accessed.

Additionally, we tested our prototype on another real-world vulnerability (CVE-2014-0322) that was used in targeted attacks [22]. It is a use-after-free error that can be utilized to increase an arbitrary bit, which is enough to create information leaks. DETILE triggered as a Vtable pointer was returned into the JavaScript context. Therefore, the information leak was detected successfully.

We also constructed a toy example in which our native code creates an information leak by overwriting the length field of an array. Additionally, the image base of `jscrypt9.dll` is written after the array data. In our tests, we reliably detected the out-of-bounds read of the image base and stopped the execution of the process. Exploit details are provided in the technical report [25].

False Positive Analysis. We analyzed the 100 top websites worldwide [2] to evaluate if our prototype can precisely handle real-world, complex websites and their JavaScript contexts without triggering false alarms. None of the tested websites did generate an alert, indicating that the prototype can accurately synchronize the master and twin process.

6 Related Work

In the following, we review work closely related to ours and discuss differences to our approach.

Randomization Techniques. Several approaches have been proposed to either improve address space layout randomization, randomize the data space, or randomize on single instruction level. For example, binary stirring [56] re-randomizes code pages at a high rate for a high performance cost. While it hinders attackers to *use* information leaks in code-reuse attacks, it does not impede their creation by itself. In contrast, our re-randomization scheme reuses the native operating system loader and is the base to allow information leak detection with dual process execution. Other solutions [32, 40, 40] are prone to JIT-ROP code-reuse attacks [48], which are based on information leaks. Address space layout permutation is an approach to scramble all data and functions of a binary [32]. Therefore, a given ELF binary has to be rewritten and randomization can be applied on each run. ORP [40] rewrites instructions of a given binary and reorders basic blocks. As discussed above, it is prone to information leak attacks, which we detect. Instruction set randomization [5, 31] complicates code-reuse attacks as it encrypts code pages and decrypts it on the fly. However, in the presence of information leaks combined with key guessing [48, 50, 57] it can be circumvented. Instruction layout randomization (ILR) [29] randomizes the location of each instruction on each run, but no re-randomization occurs. Thus, the layout can be reconstructed with the help of an information leak. *Readactor* is a defensive system that aims to be resilient against just-in-time code-reuse attacks [15]. It hides code pointers behind execute-only trampolines and code itself is made execute-only, to prevent an attacker building a code-reuse payload just-in-time. However,

it has been shown that it is vulnerable against an attack named *COOP*, which reuses virtual functions [44]. Unlike Readactor, DETILE prevents COOP, as this attack needs an information leak as first step. Crane et al. recently presented an enhanced version of Readactor, dubbed *Readactor++* [16], that also protects against whole function reuse attacks such as COOP. This is achieved through function pointer table randomization and insertion of booby traps. Consequently, an adversary can no longer obtain meaningful code locations that can be leveraged for code-reuse attacks. Readactor++ also does not detect or prevent the exploitation of memory disclosures, which poses a potential attack vector.

Multi-Execution Approaches. Most closely related to our research are *n-variant systems*, which run variants of the same program with diverse memory layout and instructions [14]. Similar work runs *program replica* synchronized at system calls to demonstrate the detection of memory exploits against the lightweight server `thttpd` on the Linux platform [9, 30].

The major drawback of these systems is the detection approach: if a memory error is abused, one of the variants eventually crashes, which indicates an attack. As information leaks *do not* constitute a memory error, they *do not* raise any exception-based signal. Thus, they remain undetected in these systems. One significant implication is that unlike DETILE, *n-variant systems* do not protect against just-in-time code-reuse attacks such as JIT-ROP [48]. Similarly, this is the case with COOP attacks in browsers [44]. *N-variant systems* prevent conventional ROP attacks [42, 54] with multi process execution and disjunct virtual address spaces: An attacker supplied absolute address (e.g., obtained through a remote memory disclosure vulnerability) is guaranteed to be invalid in $n - 1$ replicas. Hence, any system call utilizing this address will trigger a detection. However, JIT-ROP attacks may perform several memory disclosures and malicious computations without executing a system call inbetween, and thus, can evade traditional *n-variant systems*. COOP attacks may as well perform touring-complete computations on disclosed memory without executing a system call and evade these systems.

7 Discussion

In the following, we discuss potential shortcomings of our approach and the prototype, and also sketch how these shortcomings can be addressed in the future.

Further Information Leaks. Serna provided an in-depth overview of techniques that utilize information leaks for exploit development [47]. The techniques he discussed during the presentation utilize JavaScript code. As our prototype leverages the JavaScript engine of the browser itself, each information leak that is based on these techniques is detected. This implies that memory disclosure attacks that leverage other (scripting) contexts (e.g., VBScript) can potentially bypass our implementation. However, in practice exploits are typically triggered

via JavaScript and thus our prototype can detect such attacks. Furthermore, due to the generic nature of our approach, our current prototype can be extended by instrumenting other scripting engines as well.

Prototype Limitations. In the unlikely event one of the functions we classified as entropy source, such as `Math.random` or `Date.now`, contain a memory disclosure bug, our approach can lead to an under-approximation of detected information leaks. In this specific case, the master confuses the leaked pointer with data from the entropy source and transfers it to the twin process. This is an undesirable state, because DETILE does not prevent the memory layout information to leak into the script context. However, the obtained pointer is only valid in the master process. An attempt to leverage the pointer to mount a code-reuse attack crashes the twin. As a consequence, DETILE halts the master process and prevents further damage.

The current prototype disables the JIT Engine as we protect the interpreter only. However, *dynamic binary instrumentation* (DBI [8, 35]) frameworks allow to synchronize processes on the instruction or basic block level, and hence, make it possible to hook emitted JIT code to dispatch our assembly stub in order to synchronize and check within the JIT code.

Asynchronous JavaScript events are currently not synchronized. This is solvable with DBI frameworks as well: If an event triggers in the master process, we let the twin execute to the same point. Then DETILE sets up and triggers the same event in the twin process.

One additional shortcoming of our prototype implementation is the identical mapping of `ntdll.dll` in all processes. As this DLL is initialized already at startup, remapping it is a cumbersome operation. JavaScript, HTML, and other contexts in browsers normally do not interact directly with native `ntdll.dll` Windows structures, and thus internal JavaScript objects, do not contain direct memory references to it. Hence, attackers resort to disclose addresses from libraries other than `ntdll.dll` at first. On the contrary, there might be script engines which directly interact with `ntdll.dll`. Still, the issue is probably solvable with a driver loaded during boot time.

Another technical drawback is the application of re-randomization on every process on the OS, as DLL modules of each process would turn into non-shareable memory and increase physical memory consumption. This can be avoided by protecting only critical processes that represent a valid target for attacks.

Deployment. The current prototype is not meant to be a protection framework for end users of web browsers. It is intended to be deployed as a system for scanning web pages to discover unknown exploits which utilize information leaks. As ASLR needs to be circumvented as a first step of each modern exploit against web browsers, DETILE has the advantage to provide an early detection of the exploit process.

8 Conclusion

Over the last years, script engines were used to exploit vulnerable applications. Especially web browsers became an attractive target for a plethora of attacks. State-of-the-art vulnerability exploits, both in academic research [11,18,26,27,45] and in-the-wild [46,55,58–60], rely on memory disclosure attacks.

In this work, we proposed a fine-grained, automated scheme to reliably detect such information leaks in script engines. It is based on the insight that information leaks result in a noticeable difference in the script context of two synchronized processes with different randomization. We implemented a prototype of this idea for the proprietary browser IE to demonstrate that our approach is viable even on closed-source systems. An empirical evaluation demonstrates that we can reliably detect real-world attack vectors and that the approach induces a moderate performance overhead only (around 17% overhead on average). While most research focused on mitigating specific types of vulnerabilities, we address the root cause behind modern attacks since most of them rely on information leaks as a first step. Our approach thus serves as another defense layer to complement defenses such as DEP and ASLR.

Acknowledgements

We would like to thank the anonymous reviewers for their valuable comments. This work was supported by the European Commission through the ERC Starting Grant No. 640110 (BASTION).

References

1. M. Abadi, M. Budiu, U. Erlingsson, and J. Ligatti. Control-Flow Integrity. In *ACM Conference on Computer and Communications Security (CCS)*, 2005.
2. Alexa. The top 500 sites on the web. <http://www.alexa.com/topsites>, 2014.
3. M. Athanasakis, E. Athanasopoulos, M. Polychronakis, G. Portokalidis, and S. Ioannidis. The devil is in the constants: Bypassing defenses in browser JIT engines. In *Symposium on Network and Distributed System Security (NDSS)*, 2015.
4. M. Backes and S. Nürnbergger. Oxymoron: making fine-grained memory randomization practical by allowing code sharing. In *USENIX Security Symposium*, 2014.
5. E. G. Barrantes, D. H. Ackley, T. S. Palmer, D. Stefanovic, and D. D. Zovi. Randomized instruction set emulation to disrupt binary code injection attacks. In *ACM Conference on Computer and Communications Security (CCS)*, 2003.
6. A. Bittau, A. Belay, A. Mashtizadeh, D. Mazieres, and D. Boneh. Hacking blind. In *IEEE Symposium on Security and Privacy*, 2014.
7. D. Blazakis. Interpreter exploitation: Pointer inference and jit spraying. *BlackHat DC*, 2010.
8. D. Bruening, E. Duesterwald, and S. Amarasinghe. Design and implementation of a dynamic optimization framework for windows. In *4th ACM Workshop on Feedback-Directed and Dynamic Optimization (FDDO-4)*, 2001.
9. D. Bruschi, L. Cavallaro, and A. Lanzi. Diversified process replicæ for defeating memory error exploits. In *Performance, Computing, and Communications Conference, 2007. IPCCC 2007. IEEE International*, 2007.

10. R. Capizzi, A. Longo, V. Venkatakrisnan, and A. P. Sistla. Preventing information leaks through shadow executions. In *Annual Computer Security Applications Conference (ACSAC)*, 2008.
11. N. Carlini and D. Wagner. ROP is still dangerous: Breaking modern defenses. In *USENIX Security Symposium*, 2014.
12. Y. Cheng, Z. Zhou, M. Yu, X. Ding, and R. H. Deng. ROPecker: A Generic and Practical Approach for Defending Against ROP Attacks. In *Symposium on Network and Distributed System Security (NDSS)*, 2014.
13. M. Conti, S. Crane, L. Davi, M. Franz, P. Larsen, M. Negro, C. Liebchen, M. Qunaibit, and A.-R. Sadeghi. Losing control: On the effectiveness of control-flow integrity under stack attacks. In *ACM Conference on Computer and Communications Security (CCS)*, 2015.
14. B. Cox, D. Evans, A. Filipi, J. Rowanhill, W. Hu, J. Davidson, J. Knight, A. Nguyen-Tuong, and J. Hiser. N-variant Systems: A Secretless Framework for Security Through Diversity. In *USENIX Security Symposium*, 2006.
15. S. Crane, C. Liebchen, A. Homescu, L. Davi, P. Larsen, A.-R. Sadeghi, S. Bruntaler, and M. Franz. Readactor: Practical code randomization resilient to memory disclosure. In *IEEE Symposium on Security and Privacy*, 2015.
16. S. Crane, S. Volckaert, F. Schuster, C. Liebchen, P. Larsen, L. Davi, A.-R. Sadeghi, T. Holz, B. D. Sutter, and M. Franz. It's a TRAP: Table randomization and protection against function reuse attacks. In *ACM Conference on Computer and Communications Security (CCS)*, 2015.
17. J. Croft and M. Caesar. Towards practical avoidance of information leakage in enterprise networks. In *HotSec*, 2011.
18. L. Davi, D. Lehmann, A.-R. Sadeghi, and F. Monrose. Stitching the gadgets: On the ineffectiveness of coarse-grained control-flow integrity protection. In *USENIX Security Symposium*, 2014.
19. L. Davi, C. Liebchen, A.-R. Sadeghi, K. Z. Snow, and F. Monrose. Isomeron: Code randomization resilient to (just-in-time) return-oriented programming. In *Symposium on Network and Distributed System Security (NDSS)*, 2015.
20. D. Devriese and F. Piessens. Noninterference through secure multi-execution. In *IEEE Symposium on Security and Privacy*, 2010.
21. U. Erlingsson, M. Abadi, M. Vrable, M. Budiu, and G. C. Necula. XFI: Software Guards for System Address Spaces. In *Symposium on Operating Systems Design and Implementation (OSDI)*, 2006.
22. FireEye. Operation SnowMan. <http://www.fireeye.com/blog/technical/cyber-exploits/2014/02/operation-snowman-deputydog-actor-compromises-us-veterans-of-foreign-wars-website.html>, 2014.
23. I. Fratric. Runtime Prevention of Return-Oriented Programming Attacks. <http://ropguard.googlecode.com/svn-history/r2/trunk/doc/ropguard.pdf>.
24. R. Gawlik, B. Kollenda, P. Koppe, B. Garmany, and T. Holz. Enabling client-side crash-resistance to overcome diversification and information hiding. In *Symposium on Network and Distributed System Security (NDSS)*, 2016.
25. R. Gawlik, P. Koppe, B. Kollenda, A. Pawlowski, B. Garmany, and T. Holz. Technical Report: Detile: Fine-Grained Information Leak Detection in Script Engines. Technical report, Ruhr-University Bochum, 2016.
26. E. Göktaş, E. Athanasopoulos, H. Bos, and G. Portokalidis. Out of control: Overcoming control-flow integrity. In *IEEE Symposium on Security and Privacy*, 2014.
27. E. Göktaş, E. Athanasopoulos, M. Polychronakis, H. Bos, and G. Portokalidis. Size does matter: Why using gadget-chain length to prevent code-reuse attacks is hard. In *USENIX Security Symposium*, 2014.

28. T. Hirvonen. Dynamic flash instrumentation for fun and profit. *Black Hat USA*, 2014.
29. J. Hiser, A. Nguyen-Tuong, M. Co, M. Hall, and J. W. Davidson. ILR: Where'd my gadgets go? In *IEEE Symposium on Security and Privacy*, 2012.
30. P. Hosek and C. Cadar. Varan the unbelievable: An efficient n-version execution framework. In *International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, 2015.
31. G. S. Kc, A. D. Keromytis, and V. Prevelakis. Countering code-injection attacks with instruction-set randomization. In *ACM Conference on Computer and Communications Security (CCS)*, 2003.
32. C. Kil, J. Jim, C. Bookholt, J. Xu, and P. Ning. Address space layout permutation (ASLP): Towards fine-grained randomization of commodity software. In *Annual Computer Security Applications Conference (ACSAC)*, 2006.
33. V. Kuznetsov, L. Szekeres, M. Payer, G. Candea, R. Sekar, and D. Song. Code-pointer integrity. In *Symposium on Operating Systems Design and Implementation (OSDI)*, 2014.
34. H. Li. Inside AVM. In *REcon*, 2012.
35. C.-K. Luk, R. Cohn, R. Muth, H. Patil, A. Klauser, G. Lowney, S. Wallace, V. J. Reddi, and K. Hazelwood. Pin: building customized program analysis tools with dynamic instrumentation. *Acm Sigplan Notices*, 2005.
36. Microsoft. EMET 5.1 is available. <http://blogs.technet.com/b/srd/archive/2014/11/10/emet-5-1-is-available.asp>, 2014.
37. Microsoft. What is the Windows Integrity Mechanism? <http://msdn.microsoft.com/en-us/library/bb625957.aspx>, 2014.
38. I. Molnar. Exec shield, new linux security feature. *News-Forge*, May, 2003.
39. K. Onarlioglu, L. Bilge, A. Lanzi, D. Balzarotti, and E. Kirda. G-free: defeating return-oriented programming through gadget-less binaries. In *Annual Computer Security Applications Conference (ACSAC)*, 2010.
40. V. Pappas, M. Polychronakis, and A. D. Keromytis. Smashing the gadgets: Hindering return-oriented programming using in-place code randomization. In *IEEE Symposium on Security and Privacy*, 2012.
41. V. Pappas, M. Polychronakis, and A. D. Keromytis. Transparent ROP Exploit Mitigation Using Indirect Branch Tracing. In *USENIX Security Symposium*, 2013.
42. M. Prandini and M. Ramilli. Return-Oriented Programming. In *IEEE Symposium on Security and Privacy*, 2012.
43. M. Russinovich, D. Solomon, and A. Ionescu. *Windows Internals, Part 2*. Microsoft Press, 2012.
44. F. Schuster, T. Tendyck, C. Liebchen, L. Davi, A.-R. Sadeghi, and T. Holz. Counterfeit object-oriented programming. In *IEEE Symposium on Security and Privacy*, 2015.
45. F. Schuster, T. Tendyck, J. Pevny, A. Maaß, M. Steegmanns, M. Contag, and T. Holz. Evaluating the effectiveness of current anti-ROP defenses. In *Symposium on Recent Advances in Intrusion Detection (RAID)*, 2014.
46. V. Security. Advanced Exploitation of Mozilla Firefox Use-after-free (MFSA 2012-22). http://www.vupen.com/blog/20120625.Advanced_Exploitation_of_Mozilla_Firefox_UaF_CVE-2012-0469.php, 2012.
47. F. J. Serna. The info leak era on software exploitation. In *Black Hat USA*, 2012.
48. K. Z. Snow, F. Monrose, L. Davi, A. Dmitrienko, C. Liebchen, and A.-R. Sadeghi. Just-in-time code reuse: On the effectiveness of fine-grained address space layout randomization. In *IEEE Symposium on Security and Privacy*, 2013.

49. C. Song, C. Zhang, T. Wang, W. Lee, and D. Melski. Exploiting and protecting dynamic code generation. In *Symposium on Network and Distributed System Security (NDSS)*, 2015.
50. A. N. Sovarel, D. Evans, and N. Paul. Where's the FEEB? the effectiveness of instruction set randomization. In *USENIX Security Symposium*, 2005.
51. R. Strackx, Y. Younan, P. Philippaerts, F. Piessens, S. Lachmund, and T. Walter. Breaking the memory secrecy assumption. In *ACM European Workshop on System Security (EUROSEC)*, 2009.
52. L. Szekeres, M. Payer, T. Wei, and D. Song. SoK: Eternal War in Memory. In *IEEE Symposium on Security and Privacy*, 2013.
53. C. Tice, T. Roeder, P. Collingbourne, S. Checkoway, Ú. Erlingsson, L. Lozano, and G. Pike. Enforcing forward-edge control-flow integrity in gcc & llvm. In *USENIX Security Symposium*, 2014.
54. S. Volckaert, B. Coppens, and B. De Sutter. Cloning your gadgets: Complete rop attack immunity with multi-variant execution. *IEEE Transactions on Dependable and Secure Computing*, 2015.
55. P. Vreugdenhil. A browser is only as strong as its weakest byte - Part 2. <http://blog.exodusintel.com/2013/12/09/a-browser-is-only-as-strong-as-its-weakest-byte-part-2/>, 2012.
56. R. Wartell, V. Mohan, K. W. Hamlen, and Z. Lin. Binary stirring: Self-randomizing instruction addresses of legacy x86 binary code. In *ACM Conference on Computer and Communications Security (CCS)*, 2012.
57. Y. Weiss and E. G. Barrantes. Known/chosen key attacks against software instruction set randomization. In *ACM Conference on Computer and Communications Security (CCS)*, 2006.
58. T. Yan. The art of leaks: The return of heap feng shui. In *CanSecWest*, 2014.
59. Y. Yu. ROPs are for the 99%. In *CanSecWest*, 2014.
60. Y. Yu. Write Once, Pwn Anywhere. In *Black Hat USA*, 2014.
61. M. Zalewski. Two more browser memory disclosure bugs. <http://lcamtuf.blogspot.de/2014/10/two-more-browser-memory-disclosure-bugs.html>, 2014.
62. M. Zalewski. Bi-level TIFFs and the tale of the unexpectedly early patch. <http://lcamtuf.blogspot.de/2015/02/bi-level-tiffs-and-tale-of-unexpectedly.html>, 2015.
63. ZDI. CVE-2011-1346, (Pwn2Own) Microsoft Internet Explorer Uninitialized Variable Information Leak Vulnerability. <http://www.zerodayinitiative.com/advisories/ZDI-11-198/>.
64. A. Zeigler. IE8 and Loosely-Coupled IE (LCIE). <http://blogs.msdn.com/b/ie/archive/2008/03/11/ie8-and-loosely-coupled-ie-lcie.aspx>, 2008.
65. C. Zhang, T. Wei, Z. Chen, L. Duan, L. Szekeres, S. McCamant, D. Song, and W. Zou. Practical Control Flow Integrity & Randomization for Binary Executables. In *IEEE Symposium on Security and Privacy*, 2013.
66. M. Zhang and R. Sekar. BinCFI: Control Flow Integrity for COTS Binaries. In *USENIX Security Symposium*, 2013.