STELLA: Sparse Taint Analysis for Enclave Leakage Detection

Yang Chen[†], Jianfeng Jiang[‡], Shoumeng Yan[‡], Hui Xu[†*]
[†]School of Computer Science, Fudan University
China
[‡]Ant Group
China

ABSTRACT

Intel SGX (Software Guard Extension) is a promising TEE (trusted execution environment) technique that can protect programs running in user space from being maliciously accessed by the host operating system. Although it provides hardware access control and memory encryption, the actual effectiveness also depends on the quality of the software. In particular, improper implementation of a code snippet running inside the enclave may still leak private data due to the invalid use of pointers. This paper serves as a first attempt to study the privacy leakage issues of enclave code and proposes a novel static sparse taint analysis approach to detect them. We first summarize five common patterns of leakage code. Based on these patterns, our approach performs forward analysis to recognize all taint sinks and then employs a backward approach to detect leakages. Finally, we have conducted experiments with several open-source enclave programs and found 78 vulnerabilities previously unknown in 13 projects.

CCS CONCEPTS

• **Computer systems organization** → **Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability.

KEYWORDS

taint analysis, enclave, sparse analysis, privacy leak

ACM Reference Format:
Yang Chen[†], Jianfeng Jiang[‡], Shoumeng Yan[‡], Hui Xu[†]. 2022. STELLA: Sparse Taint Analysis for Enclave Leakage Detection. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX)*. ACM, New York, NY, USA, 12 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Due to the advantages of cloud computing, migrating on-premise software to public cloud services has received more and more attention[10]. However, privacy is still a major concern for many potential cloud users that cannot afford data leakage risks[9]. TEE (trusted execution environment) is a promising solution to this problem provided by chip manufacturers, such as Intel Software Guard Extensions (SGX) [13, 26] and ARM TrustZone [14, 28]. With TEE, cloud users can deploy their security-critical applications in isolated enclaves and prevent them from being accessed by unauthorized parties. In recent years, we have witnessed the application of Intel SGX in many complex systems, including database [30, 40, 47], middleware [6, 21], blockchain [5, 27], and network [4, 20].

Although Intel SGX provides hardware-level protection for the software, its practical effectiveness largely depends on how users code their programs. Any improper use of the SGX interface or bugs in the enclave code may lead to privacy leakage [31]. To our best knowledge, there is still neither systematic work on this problem nor available solutions to detect such bugs. Studying this issue is particularly important because verifying the effectiveness of TEE code is a critical step to the advancement of TEE solutions.

In this study, we first list five bug patterns that lead to privacy leaking in enclave programs (writing sensitive data to ECALL `out`, ECALL `user_check`, OCALL `in`, OCALL return and NULL pointers). To the best of our knowledge, no research has summarized these patterns. A well-known technique for finding privacy leaks is taint analysis [2, 7, 42], and it is straightforward to apply taint analysis to find the bug patterns we define. However, enclave programs' privacy leak exhibits a variety of traits. For instance, taint sinks are data writes to particular pointers that are unknown and require solving. This is distinct from the typical one in that taint sinks are fixed and well-known APIs. Therefore, it is challenging to use taint analysis to tackle this issue.

To find these patterns in enclave programs, we provide an innovative custom taint analysis. We initially investigate the def-use relationship of the variables in enclave programs in order to construct a value flow graph (VFG). In order to identify the data writes to these pointers (i.e., identify taint sinks), we first perform forward analysis starting from the definition nodes of these pointers on the VFG. The written data is examined using backward analysis to determine its sensitivity. We employ VFG rather than control flow graph (CFG) as in conventional static analysis because VFG-based analysis (also known sparse analysis) has been proven in prior studies [33, 34, 36–38] to improve performance without sacrificing accuracy.

To elaborate, the Enclave Definition Language (EDL) [15] files, which serve as interface descriptions, and the LLVM IR files generated by the enclave program are the inputs used by our approach. We parse EDL files to extract pointer parameters that could result in the leakage of private information and to create the VFG, we read the LLVM IR files. Then, starting with the definition nodes associated with the aforementioned pointer parameters, we traverse

*Corresponding author.

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

the VFG and taint the variables in accordance with the rules we established for taint propagation. We designate them as taint sinks if we discover writes to these taint variables in the enclave program. The VFG also provides the corresponding nodes for the data written in a taint sink. To obtain all the leaked variables, we begin from these nodes and move backward via the VFG. If any of them are private, it suggests a possible privacy breach. Our technique provides manual insensitive annotation at the source code level and insensitive variables will be automatically omitted in the backward analysis to reduce false positives because not all variables in the enclave program are sensitive. Furthermore, the decrypted data, the data to be encrypted, and the key for encryption or decryption will be automatically tagged as high-risk sensitive data. In the leak report, the high-risk data will be shown with a greater priority if it has been compromised.

We have implemented a prototype for the **S**parse **T**aint analysis for **Enc**l**L**ave privacy **L**e**A**kage detection, namely STELLA. In addition to employing sparse analysis to increase speed, we have adopted a preference for retrieving the call graph (CG) rather than the call flow graph (CFG) for solving variables on VFG that could reveal secrets because CFG has a significantly higher number of nodes than CG. We used STELLA to analyze 13 open-source SGX-based applications on GitHub to test the viability of our approach, and we discovered 78 privacy disclosure bugs there that had never before been reported.

In short, this article contains several major research contributions as follows.

- Privacy leaks caused by coding errors in enclaves undermine enclave confidentiality guarantees, making it easier for attackers to obtain secrets in enclaves. To the best of our knowledge, our work is the first comprehensive one on this issue.
- We first define patterns of privacy-leaking coding mistakes that enclave developers are prone to make and then propose a novel sparse taint approach to efficiently analyze the vulnerabilities in enclave code based on these patterns. As far as we know, this is the first available approach on the issue.
- We have implemented a prototype tool, STELLA, and released it as open-source on GitHub. Our experimental results show that STELLA can effectively discover privacy-leaking vulnerabilities in enclave programs.

The rest of our paper is organized as follows. We introduce the Intel SGX background in section 2. Section 3 defines the problem of privacy leakage on enclave programs and discusses the challenges of detecting such bugs. Section 4 then elaborates on our sparse taint analysis approach. Section 5 evaluates the performance of our approach. We review related work in Section 6. Section 7 concludes our paper.

2 PRELIMINARY

In this section, we first introduce background about SGX and Intel SGX SDK, then illustrate the privacy leakage issue when developing SGX programs.

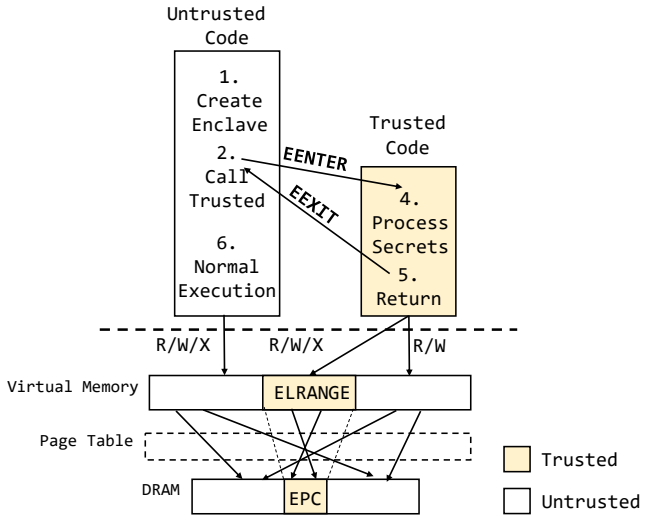


Figure 1: The underlying mechanism of SGX.

2.1 SGX Background

Intel SGX is a set of CPU instructions that enable applications to create hardware-based protected areas, namely enclaves. Enclaves are used to protect private data from modification and closure [43]. Private data within an enclave can only be accessed by the code within the same enclave, and cannot be read or written by programs outside the enclave even such programs run at privilege levels.

A big change in introducing enclaves in a normal program is that we need to separate the application into the untrusted part (outside enclave) and trusted part (inside enclave) and all access to secrets needs to be done in the enclave. As shown in Figure 1, untrusted code can create one or more enclaves. When the secrets need to be processed, the untrusted code needs to call the trusted function like a normal function, and then the CPU is switched to the enclave mode via the EENTER instruction, and the control is transferred to the corresponding trust function. The trusted function in the enclave can directly access these secrets. After processing these secrets, the CPU switches out of enclave mode via EEXIT instruction, and the enclave transfers control to untrusted code and continues normal execution.

2.2 Intel SGX SDK

Intel SGX SDK [16] is a set of toolkits to develop enclave programs. In practice, developers tend to write programs with high-level languages, instead of directly using low-level instructions like EENTER and EEXIT. To this end, SGX SDK provides high-level abstractions on low-level instructions and enables developers to develop enclave programs with C/C++. Currently, enclave programs are mostly developed with Intel SGX SDK.

Intel SGX SDK introduces ECALLs (enclave calls, the calls to enclave functions) and OCALLs (outside calls, the calls to outside-enclave functions) on top of the EENTER and EEXIT instructions. An ECALL calls the EENTER instruction first, then executes the trusted function, and finally calls the EEXIT instruction. OCALL is vice versa. ECALLs and OCALLs are the interfaces between host

```
enclave{
    //ECALLs
    trusted {
        public int foo([in,size = 10] void* in_ptr, [out,size=len]
        void* out_ptr, [user_check] void* uc_ptr, size_t len);
    };
    //OCALLs
    untrusted{
        public void* bar();
    };
};
```

Figure 2: An example EDL file.

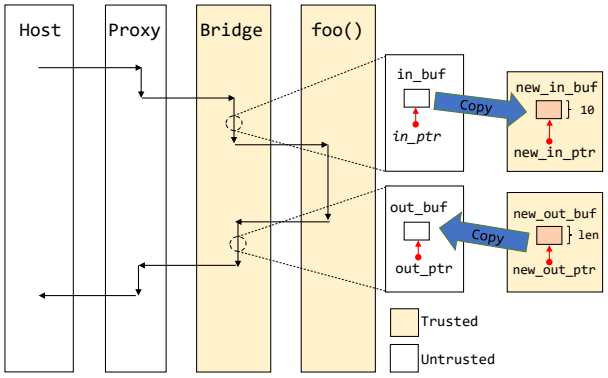


Figure 3: Proxy and bridge functions for the ECALL foo.

```
/***EDL***/
void OCALL_SaveFile([in, size=filename_len]uint8_t
*filename, size_t filename_len, [in,
size=data_len]uint8_t *data, size_t data_len);

/*** enclave code***/
void ECALL_VCFEncryption(const char* dirpath, const
char* filepath, const char* AES_filename, int
FILE_SIZE){
    ...
    // Generate an AES Secret key
    cout = "Create AES Secretkey...";
    OCALL_print(cout.c_str());
    CreateAESSecretKey(AES_SK);
    //Stored AES key outside the enclave without sealing
    ==> leak
    status = OCALL_SaveFile(SK_filename.c_str(),
    SK_filename.length() + 1, AES_SK, 16);
    ...
}

/***application code**/
void OCALL_SaveFile(uint8_t *filename, size_t
filename_len, uint8_t *data, size_t data_len)
{
    ...
    ofstream ofs(filename_str, ios::binary);
    ofs.write((const char*)data, data_len);
    ...
}
```

Figure 4: Motivating example: AES key leakage in BiORAM-SGX.

and enclave, and their prototypes are defined in EDL files. As shown in Figure 2, we declare an ECALL `foo` and an OCALL `bar` in the EDL. The trusted ECALL `foo` is defined in the enclave. The untrusted OCALL is defined in the host.

Intel SGX SDX automatically generates proxy and bridge functions for ECALLs and OCALLs when compiling enclave programs. Take the ECALL `foo()` in Figure 2 as an example, we call its definition in enclave real `foo()`. Intel SGX SDK will generate a pair of functions, an untrusted proxy and a trusted bridge function. When executing this ECALL as Figure 3 shows, The host code first calls the untrusted proxy. Then the untrusted proxy will call the trusted bridge. Finally, the trusted bridge calls the real foo(). The state and return value of the ECALL are propagated back from the opposite direction.

There are two categories of pointer parameters for functions defined in EDL files, namely directed pointers, and raw pointers.

Pointer parameters in ECALLs and OCALLs can be declared with direction (`in`, `out`) and size (`size_t`). The proxy and bridge functions will copy the content of pointers based on the direction and size. Take `foo` in Figure 2 as an example, `in_ptr` has the `in` attribute, and its size attribute is 10. The trusted bridge function will allocate a 10-bytes buffer inside the enclave and copies the first 10 bytes of `in_ptr` to the buffer. Similarly, if the direction is `out`, the trusted proxy will copy data from the enclave to the the untrusted world.

Pointers can also be annotated as `user_check` (*e.g.* `uc_ptr` in `foo` in Figure 2), which means it is a raw pointer. The bridge or proxy will not copy the buffer but directly pass the address. Note that the pointer returned by an OCALL, like `bar` in Figure 2, is also a raw pointer, and developers need to do proper checks before using it. In addition, if an `in` pointer is a pointer to struct, since the buffer copy is a shallow copy, the pointer fields in the struct are all `user_check` pointers.

In general, the pointer direction attributes in EDL determine whether buffer copying or raw pointer transfer occurs between host and enclave. If developers are not careful with these pointers, sensitive data can be accidentally leaked into the untrusted world. In the following subsection, we will demonstrate this problem with a real-world example.

2.3 Issues of Developing with SGX SDK

Although SGX prevents host code from directly accessing data inside the enclave, it is still possible to leak data outside the enclave through the pointer parameter of ECALL or OCALL. For example, writing sensitive data to the `in` pointers in OCALLs may cause privacy leakage.

Such privacy leakage problems are detected in existing enclave programs. BiORAM-SGX [17] is a personal genetic data statistical

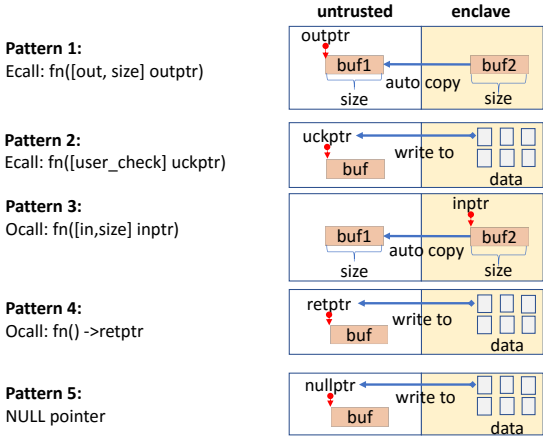


Figure 5: Writing sensitive data to these five pointers in the enclave program will cause privacy leakage. The (ECALL `out` / OCALL `in`) pointer points to the variable in the enclave, but the written data will be copied outside the enclave by the SGX SDK. The (ECALL `user_check` / OCALL return / NULL) pointer points directly outside the enclave.

analysis system using SGX. The system claims that sensitive personal genetic data will not be leaked to the server. But we found the problem of AES key leakage in this system. Figure 4 demonstrates the buggy code snippet. `data` of `OCALL_SaveFile` is an OCALL `in` pointer (line 2), which can pass information from the enclave to the outside of the enclave. The developers must be careful with this pointer and cannot pass confidential information to `data`. However, `ECALL_VCFEncryption` passes the AES key to `data` in plaintext without sealing, calling `OCALL_SaveFile` to save the AES key out of the enclave (line 12). An attacker can easily open the unencrypted file and obtain the AES key, which poses a great threat to user data security.

Note that this vulnerability does not contradict SGX's guarantee against data leakage. SGX SDK requires developers to perform checks when writing sensitive data to pointers. If the developers carelessly fail to check, SGX cannot guarantee that the data inside the enclave will not be leaked.

3 PROBLEM AND CHALLENGES

For enclave programs, developers may mistakenly copy sensitive data out of the enclave, which will cause privacy leakage. Note that the privacy leakage problem is serious for enclave programs because this problem breaks SGX's guarantee of data security. SGX aims to protect private data from access from the host and not allow the leaking of private data to the untrusted world.

Next, we first summarize common code patterns that may cause the problem, then illustrate the challenges to adapt taint analysis to enclave programs.

3.1 Typical Patterns

For enclave programs, most privacy leakages are caused by writing sensitive data to specific pointers. As shown in Figure 5 , by comprehensively studying the coding specification of Intel SGX SDK, we summarize 5 pointer usage patterns that can leak privacy in enclave codes. To the best of our knowledge, no one has yet summarized these patterns.

Write Private Data to ECALL `out` Pointers (P1). As shown in Figure 3, When ECALL *foo* returns, the data written to the ECALL out pointer is copied into untrusted memory by the bridge function. Therefore, when a developer mistakenly writes private data to an ECALL `out` pointer, the private data is leaked.

Write Private Data to ECALL `user_check` Pointers (P2). To achieve better performance, developers may use `user_check` pointers in ECALL to avoid checking pointers and copying buffer by SGX SDK. Instead, developers need to check the `user_check` pointer by themselves. But they may carelessly not check `user_check` pointers and write sensitive data to them. If the pointers point to untrusted memory, private data may leak. Note that we cannot trust the ECALL input according to the Intel SGX threat model, so even if enclave developers assume that these pointers point inside the enclave, they may also be tampered with by an attacker to point to untrusted memory. Note that the pointer field in the `in` structure pointer is `user_check`, so this pattern also includes this case.

Write Private Data to OCALL `in` Pointers (P3). Figure 3 shows how Intel SGX SDK processes the `in` pointer in ECALL. The bridge function copies the untrusted memory data to the enclave. For the `in` pointer of OCALL, the processing is similar but in the opposite direction. The Intel SGX SDK will copy the enclave memory data to the untrusted memory, so developers must be careful, if the OCALL `in` pointer points to sensitive data, it will lead to privacy leakage.

Write Private Data to OCALL Return Pointers (P4). If the return pointer of OCALL points to untrusted memory, and the developer mistakenly writes sensitive data to this pointer, it will lead to privacy leakage.

Write Private Data to NULL Pointers (P5). For general applications, writing data to address 0 (NULL) will be aborted by the OS. However, for SGX applications, the OS is untrusted according to the Intel SGX threat model. Writing sensitive data to NULL pointers in the enclave can lead to privacy leakage.

Taint analysis is a traditional method to detect such patterns. However, since enclave programs differ from traditional programs in many ways, we have to adapt the taint analysis to enclave programs.

3.2 Challenges

The traditional method to detect privacy leakage is via taint analysis, *e.g.*, for android or IoT programs. However, traditional taint analysis cannot be directly ported to detecting privacy leakage on enclave programs. A traditional taint analysis framework includes **taint sources**, **propagation rules**, and **taint sinks**. Enclave programs differ from android or IoT programs in the following three aspects:

Taint sources are variables storing sensitive data. In general, android programs only have a few taint sources such as those variables storing passwords and phone numbers so manually annotating such variables is easy. For enclave programs, considering that SGX protects all data inside the enclave, we should treat all of them except a little public data as taint sources. Hence, there are two main challenges to our task. First, the taint sources will be

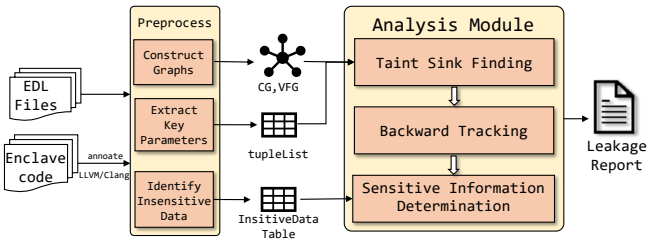


Figure 6: Overall framework of STELLA.

relatively large. Manually marking taint sources is labor-intensive. Second, if all data inside the enclave is marked as sensitive data, we will have many false positives. we should distinguish between sensitive data and public data to avoid false positives.

Propagation rules define how tainted variables are propagated to other untainted variables through CFG. For example, if we convert a tainted variable into its hexadecimal format and store it in another variable, then the latter variable should also be tainted. For enclave programs, encryption functions should be considered during designing propagation rules, and only the leakage of unencrypted data should be viewed as private data leakage. For example, If the parameters of OCALLs are encrypted by some functions *e.g.*, `sgx_rijndael128GCM_encrypt`, it should not be considered a privacy leakage.

Taint sinks are usually APIs that may leak information. For android programs, taint sinks are fixed, like APIs sending SMS messages or remote requests. However, for enclave programs, taint sinks are related to some specific pointers, rather than fixed functions. We have to find pointers that can leak data out of the enclave. Codes that dereference or write to these pointers are the taint sinks we need.

Taint analysis tries to find paths from taint sources to taint sinks. If such a path exists, private data may be leaked.

4 SPARSE TAINT ANALYSIS APPROACH

In general, we use static analysis on SGX programs to detect privacy leakage vulnerabilities. The whole STELLA framework is shown in Figure 6. STELLA receives EDL files and annotated enclave code as input and generates a leakage report. Preprocessing and analysis module comprise the STELLA pipeline.

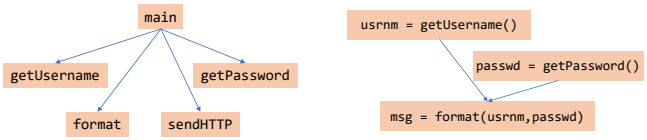
4.1 Preprocess

As shown in Figure 6, STELLA first preprocesses input files, including constructing graphs, extracting key function parameters, and identifying insensitive data. The module accepts compiled bitcode and EDL files as input and outputs graphs and tables required for further analysis.

4.1.1 Construct Graphs. We construct several graphs for analysis, including CG and VFG. STELLA bases graph construction on SVF [37], a static value flow analysis framework. VFG plays a key role in our analysis. STELLA performs an inter-procedural pointer

```
1 string format(string usrnm, string pw){
2     return "username:" + usrnm + ", " + "password:" + pw;
3 }
4 int main(){
5     string usrnm = getUsername(); //source
6     string passwd = getPassword(); //source
7     string msg = format(usrnm,passwd);
8     sendHTTP("http://www...",msg);//sink
9 }
```

(a) Program Code



(b) Call Graph (CG) **(c) Value Flow Graph (VFG)**

Figure 7: Simple taint and source-sink example code and its graphs

analysis to obtain the points-to information after reading the enclave bitcode files. With the points-to information, STELLA constructs a memory SSA form and obtains the def-use chains and value flows to construct the VFG.

Figure 7 demonstrates CG and VFG of an example program. Figure 7a is a simple taint source-sink example program. The program fetches the user name and password, assembles a message, and finally sends it to the outside via HTTP, causing sensitive information leakage. Call graph (CG) describes the calling relationship between functions. As shown in Figure 7b, the `main` function calls `getUsername`, `getPassword`, `format` and `sendHTTP`. Value Flow Graph (VFG) demonstrates the data flow of variables. As shown in Figure 7c, The values of variables `usrnm` and `passwd` both flow to variable `msg`. Our approach is based on CG and VFG. In general, we begin by matching CG nodes with the function name. The taint sink nodes on the VFG are then discovered using these nodes. Finally, leakage of personal information is discovered by resolving the graph reachability problem between sinks and sources.

4.1.2 Extract Key Parameters. We extract the function parameters in ECALLs and OCALLs in EDL files that may potentially leak privacy. For each parameter, we combine its function name, the position index, and the leak pattern into a tuple. For example, for an ECALL, `ecall_func([user_check] void* uc_ptr, [out,size = size] void* out_ptr, int size)`, `uc_ptr` and `out_ptr` are two parameters that may leak privacy, so we will get two tuples, `(ecall_func,0,P2)` and `(ecall_func,1,P1)`. After scanning all EDL files, we will get a list of tuples, namely `tupleList`. It is used for subsequent taint analysis to find sinks.

4.1.3 Identify insensitive Data. We manually annotate some insensitive variables in enclave programs. Since enclaves are used to protect private data, we assume that most data processed inside enclaves is sensitive. However, some variables storing insensitive information are used to interact with the untrusted world, like `user_id` in Figure 8. We will annotate this variable with the

```
#define INSENSITIVE __attribute__((annotate("INSENSITIVE")))
struct type
{
    INSENSITIVE int user_id; //insensitive
    string username;//sensitive
    string password;//sensitive
};
```

Figure 8: Annotate the insensitive variable user_id with INSENSITIVE. Other variables are sensitive by default.

INSENSITIVE prefix. We assume all variables without the annotation store sensitive data. `llvm.var.annotation` [24] is a LLVM intrinsic function. We can use this function to annotate local variables with arbitrary strings. For each variable annotated with INSENSITIVE, after its memory allocation statement (such as `alloca`), the `llvm.var.annotation` function is called once with the variable as the first argument. So we traverse the call graph to find all the code that calls `llvm.var.annotation`, and check whether the function has the argument "INSENSITIVE". Finally, we can collect all annotated variables into the `InsensitiveDataTable`.

4.2 Analysis Module

The analysis module conducts analysis to identify potential privacy leakage paths. It comprises of taint sinks finding, backward tracking, and sensitive information determination.

4.2.1 Taint Sink Finding. We divide the taint sinks in enclave code into three categories, *Pointers from Outside*, *Pointers Declared Inside* and *Explicit Sinks*. We design different search strategies for each category.

- **Pointers from Outside**. From an enclave developer's perspective, ECALL `out` (P1) , ECALL `user_check` (P2), and OCALL return pointers (P4) all come directly from outside the enclave. The strategies for detecting their associated taint sinks are also similar. Note that the enclave code may derive new pointers from these pointers. For example, if an ECALL passes in a `user_check` secondary pointer (e.g. `[user_check]void** p1`), then this pointer is dereferenced to get a new pointer `void* p2 = *p1`. `p2` may also point outside the enclave. Writing sensitive data to `p2` may also lead to privacy leakage. Therefore, to make our analysis as sound as possible, we search not only these pointers but also the pointers derived from them via a taint-style algorithm. The instruction to write data to the pointer in the search result is the taint sink.
- **Pointers Declared Inside.** Pointers defined inside the enclave can also be corrupted, such as null pointers or wild pointers. These pointers may point outside the enclave. We define writing instructions to these pointers as taint sinks. We focus on the privacy disclosure vulnerability caused by writing sensitive data to a NULL pointer in an enclave (*i.e.*, P5).
- **Explicit Sinks.** Calling OCALL functions that accept primitive types or `in` pointers as parameters can leak private data (*i.e.*, P3). We call this type of taint sinks *Explicit Sinks* because they can be located directly by matching function names in the call graph.

Algorithm 1: Taint sinks finding

Input: EDLs: EDL files, prog: the enclave program to be analyzed
Result: sinks: taint sinks

```
1  Function FindSinks(EDLs, prog)
2      tupleList ← ExtractKeyParameters(EDLs)
3      cg, vfg ← ConstructGraphs(prog)
4      sinks ← {}
5      for tuple ∈ tupleList do
6          /* match the cg node by funcname. */
7          node ← GetNode(tuple.funcname, cg)
8          /* Pointers from Outside (ECALL out / user_check / OCALL return)*/
9          if tuple.leakType ∈ {P1, P2, P4} then
10             ptr ← GetPtr(node, tuple.index)
11             /* ptrTaint is shown in Algorithm 2 */
12             sinks ← sinks ∪ PtrTaint(ptr)
13         /* Explicit Sinks (OCALL in pointers)*/
14         else if tuple.leakType ∈ {P3} then
15             ptr ← GetPtr(node, tuple.index)
16             defNode ← GetDefNode(ptr, vfg)
17             sinks ← sinks ∪ {defNode}
18     /* for NULL pointers (P5). */
19     /* malloc without check */
20     node ← GetNode("malloc", cg)
21     ptr ← GetPtr(node)
22     cmpInsts ← GetCmpInsts(ptr)
23     /*if no cmp instruction : no NULL check*/
24     if cmpInsts == ∅ then
25         sinks ← sinks ∪ PtrTaint(ptr)
26     return sinks
```

Algorithm 1 demonstrates our taint sinks finding algorithm. We iterate over the entries of the `tupleList` and use the function name in the entry to match the node on the call graph. For *Pointers from Outside*, we use the node and the parameter index to get the pointer to analyze and then call `PtrTaint` (Algorithm 2) to search for the taint sinks associated with this pointer. For *Explicit Sinks*, it is relatively simple. After we locate the pointer parameter, its definition node on the VFG is the taint sink. For *Pointers Declared Inside*, we focus on NULL pointers due to the enclave not validating the result of `malloc`. After we get the `malloc` pointers in the program, check whether they have corresponding comparison instructions. If not, it may be NULL, and then we call `PtrTaint` to search for taint sinks.

4.2.2 Pointer Tainting. We use a taint-style algorithm to search leak-causing pointers and derivative pointers and their related taint sinks. We design the following 6 propagation rules, which are represented by 6 formulas. The upper part of each formula represents the current situation or conditions, and the lower part of the formula represents the corresponding operation that needs to be performed. `T(v)` represents marking variable `v` as tainted, and `U(v)` represents removing the taint flag of the variable `v`.

$$\frac{T(addr), v = load(addr)}{T(v)} \tag{1}$$

$$\frac{T(v), store(v, addr)}{T(addr)} \tag{2}$$

$$\frac{T(x), y = op(x)}{T(y)} \tag{3}$$

$$\frac{T(x), z = op(x, y)}{T(z)} \tag{4}$$

$$\frac{T(base), addr = gep(base)}{T(addr)} \tag{5}$$

$$\frac{T(x), y = bitcast(x)}{T(y)} \tag{6}$$

In the formula (1), the LLVM load instruction refers to reading a value from memory. When addr is a tainted variable, then the value v read from addr is also marked as a tainted variable.

In the formula (2), the store instruction is used to write memory. The store instruction has two arguments: a value v to store and an address addr which indicates the write location. If v is a tainted variable, then after executing the store instruction, addr should also be marked as a tainted variable.

In the formula (3), for an unary operators op, if the operand x is a tainted variable, the result y need be marked as tainted.

In the formula (4), op is a binary operator. If one of the two operands x is a tainted variable, the result of their operation z need be marked as a tainted variable.

The Get Element Pointer (GEP) instruction provides a way to calculate pointer offsets. base is the base address to start from and addr is the calculated offset pointer. In the formula (5), if base is a tainted pointer, then the derived addr needs to be tainted.

LLVM bitcast instruction converts a value to another type without changing any bits. In the formula (6), if the original value x is a tainted variable,the bitcast result y should be marked as a tainted variable.

The pseudocode for PtrTaint is Algorithm 2. The value-flow graph (vfg) and the pointer (ptr) that we want to track serve as the input. The output is a collection of tainted sinks. The primary purpose of the algorithm is a hierarchical traversal on the def-use chains corresponding to ptr using a queue. New pointers are continually added to tainted during the traversal process using the aforementioned rules. In line 19, it should be noted that if STELLA discovers a store node and the address is tainted but the address is not, the store node is deemed to be a taint sink and is merged into sinks. memcpy and the LLVM store instruction are both represented as store nodes in VFG.

4.2.3 Backward tracking and sensitive information determination. After taint sink finding, we can obtain all code snippets in the enclave that transmit information to the untrusted world. Next, we backward track leaked variables and determine whether they are sensitive. Algorithm 3 shows our backward tracking algorithm. We process each taint sink, traverse from the VFG node of the sink in the reverse direction along the def-use chain, and get all the paths from the sink to the leaf nodes. Then, we start to find out whether the nodes in the paths are variable allocation instructions (*e.g.* alloc). If so, and they are not yet in the InsensitiveTable, then we add the source code location of the sink and leaked variable to the report.

Algorithm 2: Tainting pointers by the propagation rules.

```
   Input: ptr: a pointer, vfg: value flow graph
   Result: sinks: taint sinks result
1  Function PtrTaint(ptr,vfg)
2      defNode ← GetDefNode(ptr, vfg)
3      tainted ← {ptr}
4      visited ← {defNode}
5      sinks ← {}
6      queue.Push(defNode)
7      while queue ≠ ∅ do
8          curNode ← queue.pop()
9          for node ∈ GetChildNodes(curNode, vfg) do
10             srcPtr ← GetSrcPtr(node)
11             dstPtr ← GetDstPtr(node)
12             if type(node) ∈ {Load, Bitcast, Gep, UnaryOp} then
13                 if srcPtr ∈ tainted then
14                     tainted ← tainted ∪ {dstPtr}
15             else if type(node) ∈ {Store} then
16                 if srcPtr ∈ tainted then
17                     tainted ← tainted ∪ {dstPtr}
18                 else if srcPtr ∉ tainted and dstPtr ∈ tainted then
19                     sinks ← sinks ∪ {node}
20             else if type(node) ∈ {BinaryOp, PHI} then
21                 if node.GetSrcPtrs() ∩ tainted ≠ ∅ then
22                     tainted ← tainted ∪ {dstPtr}
23             if node ∉ visited then
24                 visited ← visited ∪ {node}
25                 queue.Push(node)
26     return sinks
```

Algorithm 3: Backward tracking

```
   Input: sinks: taint sinks, vfg: value flow graph
   Result: report: privacy leakage report
1  Function BackTrack(sinks, vfg)
2      for sink ∈ sinks do
3          /*a queue that stores nodes from the sink to leaked variable.*/
4          path ← queue()
5          /*a list that stores visited nodes*/
6          visited ← {}
7          /*backtracking for each sink.*/
8          BackTrackEachSink(sink, visited, path, vfg)
   Input: src: source node, visited: visited nodes, path: trace path , vfg: value flow graph
9  Function BackTrackEachSink(src, visited, path, vfg)
10     /*find a leaked sensitive variable allocation node*/
11     if type(src) ∈ allocStmts and src ∉ InsensitiveDataTable then
12         /*report a leak*/
13         PrintLeakPath(src, path)
14         return
15     /* avoid encryption or seal functions in order to reduce false positives*/
16     if GetFunc(src) ∈ {encrypt, seal, ...} then
17         return
18     /* push the current node to path and set it visited */
19     path.Push(src)
20     visited ← visited ∪ {src}
21     /*continue visit its parents*/
22     for parentNode ∈ GetParentNodes(node, vfg) do
23         /* If parentNode is not visited, continue to trace up recursively.*/
24         if parentNode ∉ visited then
25             BackTrackEachSink(parentNode, visited, path)
26     /*reset the node to unvisited for find all leak paths*/
27     path.Pop()
28     visited.Delete(src)
```

5 EVALUATION

We have implemented a prototype tool STELLA. STELLA can detect the patterns that cause privacy leakage for enclave programs. We perform evaluation experiments to evaluate the effectiveness and

Table 1: The details of selected projects and vulnerabilities we found.

Project Name	Description	ECALL user_check	ECALL out	OCALL in	OCALL return	Null ptr deref	Total
SGX-Tor [19]	secure anonymity network			1	1	7	9
sgx_wechat_app [46]	trusted wechat app			1			1
Fidelius [12]	protect browser users' secrets			1		5	6
sgx-based-mix-networks [11]	hidden anonymization		1				1
sgx-dnet [44]	machine learning inside enclave			1		25	26
SGX_SQLite [25]	secure SQLite database			1			1
TaLoS [3, 29]	secure TLS library	2			1		3
sgx-aes-gcm [32]	SGX AES-GCM usage example		1	1			2
password-manager [22]	password manager using SGX			1		1	2
TACIoT [39]	IoT data protection			1		1	2
BiORAM-SGX [17]	genome analysis system			2			2
PrivacyGuard [41]	data analytic inside enclave			2		7	9
Town-Crier [45]	smart contract using SGX			2		12	14
Total		2	2	14	2	58	78

performance of STELLA. We release our tool as open source, and it is available online [1].

5.1 Experimental Setup

We perform experiments to study two research questions:

RQ1: Can STELLA effectively find privacy leakage bugs in enclave programs?

RQ2: How efficient is STELLA when analyzing real-world projects?

We use STELLA to test the enclave programs developed using the SGX SDK on GitHub. We found bugs in 13 popular projects (the average stars of these projects is 30). Table 1 shows the details of these projects. The code size of these projects varies largely, ranging from a few hundred lines to 527K lines.

Our experiments are done on a ubuntu-20.04 server with an Intel i7-9700T 4.30GHz 8-core CPU. The version of Intel SGX SDK is 2.15. The reported run time is the average of three measurements

5.2 Effectiveness

5.2.1 Overview. STELLA found 78 vulnerabilities in these projects that could leak sensitive data. Table 1 shows the number of vulnerabilities with different patterns. For example, we found that SGX-Tor has two privacy leakage bugs, one of which is to write sensitive data to the OCALL `in` pointer, and the other is to write sensitive data to the pointer returned by OCALL.

5.2.2 Patterns. We find all five patterns in real-world enclave projects. Among all patterns, writing private data to OCALL `in` pointers and to NULL pointers are the most common patterns. Next, we will present a case in real-world projects for each pattern.

Write Private Data to ECALL `out` Pointers (P1). Figure 9 shows a secret disclosure bug in sgx-based-mix-networks. The enclave code incorrectly writes the secret to the ECALL `out` pointer. From the EDL file, we can know that the parameter `result` is an `out` pointer. In `dispatch` function, `message` is sensitive data. But at line 12, `message` is directly copied to `result` without encryption. When the `dispatch` function returns, the `message` plain text in `result` will be copied to the untrusted `content` (line 2), causing sensitive information leaked.

```
1  /***application code***/
2  unsigned char content[KEY_SIZE];
3  ...
4  dispatch(this->eid, &dispatch_response, content, &fan_out,
   &buffer_size, this->fan_all_out);
5
6  /***enclave code***/
7  int dispatch(unsigned char *result, int *fan_out, size_t
   *buffer_size, int fan_all_out)
8  {
9      ...
10     if (*fan_out) {
11             // unencrypted message leaked !!!
12             std::copy(message, message + strlen((char *) message),
               result);
13             return 1;
14     }
15     ...
16     sgx_status_t ret_encrypt =
       sgx_rsa_pub_encrypt_sha256(previous_public_key, result,
       &out_len, message, strlen((char *) message));
17     ...
18 }
```

Figure 9: A P1 privacy leakage vulnerability in sgx-based-mix-networks. The enclave code mistakenly writes unencrypted sensitive information to the ECALL `out` pointer (line 12)

Write Private Data to ECALL `user_check` pointers (P2). TaLoS [29] is a library that enables applications to terminate TLS connection securely. TaLoS protects sensitive data from disclosure by placing sensitive data within a SGX enclave. But for this library, there is still a hidden danger of SSL private key leakage. Figure 10 demonstrates the code snippet that may cause data leakage. Function `ecall_SSL_get_privatekey` is an ECALL and the parameter

[1] https://anonymous.4open.science/r/STELLA-7724

```
1 void ecall_SSL_get_privatekey(EVP_PKEY* pkey, SSL *s) {
2     ...
3     EVP_PKEY* enclave_pkey = SSL_get_privatekey(in_s);
4     // Leak !!!
5     memcpy(pkey, enclave_pkey, sizeof(*pkey));
6     ...
```

Figure 10: A P2 privacy disclosure vulnerability in TaLoS that leaks an SSL private key. pkey is an ECALL user_check pointer, the enclave code erroneously writes the private key enclave_pkey to pkey. TaLoS developers have confirmed the vulnerability.

```
1  /***EDL***/
2  untrusted {
3      ...
4      void ocall_eprint_string([in, string] const char *str);
5      ...
6  };
7
8  /***enclave code**/
9  sgx_status_t ecall_decrypt_secret(sgx_ra_context_t context,
10         const uint8_t *p_src, uint32_t src_len,
11         const sgx_aes_gcm_128bit_tag_t *p_in_mac){
12     ...
13     sgx_ra_key_128_t ra_key;
14     sgx_status = sgx_ra_get_keys(context, SGX_RA_KEY_SK, &ra_key);
15     //ra_key is leaked !!!
16     feprintf("ra key:%s\n", hexstring(&ra_key, sizeof(ra_key)));
17     ...
18     sgx_status = sgx_rijndael128GCM_decrypt(&ra_key, p_src,
       src_len, p_dst, p_iv, SGX_AESGCM_IV_SIZE, NULL, 0, p_in_mac);
19     phone_num = string(hexstring(p_dst, src_len));
20     accid_phone_map[contextid_accid_map[context]] = phone_num;
21     //the decrypted phone_num is leaked !!!
22     feprintf("Phone number:%s\n",
       accid_phone_map[contextid_accid_map[context]]);
23     ...
24 }
25 int feprintf(const char *fmt, ...){
26     ...
27     vsnprintf(buf, BUFSIZE, fmt, ap); // fmt flows into buf
28     ocall_eprint_string(buf); // buf is passed to OCALL in ptr
29     ...
30 }
31
32 /***app code***/
33 void ocall_eprint_string(const char *str){
34     printf_info(felog, "%s", str);// print in untrusted world.
35 }
```

Figure 11: The P3 privacy disclosure vulnerability in sgx_wechat_app will reveal the sgx_ra_key and the decrypted mobile phone number. The enclave code mistakenly passes sensitive information to the OCALL in pointer.

pkey is a user_check pointer that points outside the enclave. In line 5, memcpy copies the private key from enclave_pkey to pkey, *i.e.*,the sensitive data is copied from enclave to untrusted world, resulting in privacy leakage. We report this vulnerability and it is included in the CVE-2022-27102.

Write Private Data to OCALL in pointers (P3). As shown in Figure 11, the function ecall_decrypt_secret in the enclave converts the variable ra_key to hexadecimal format and passes it to the function feprintf at line 16. This ra_key is sensitive data. At line 22, the just decrypted phone_num is also passed to feprintf. In feprintf function, the ra_key and phone_num will be copied into the buf at line 27. Then at line 28, OCALL ocall_eprint_string is called by passing buf to its in pointer parameter. During the execution of ocall_eprint_string, Intel SGX SDK will copy the sensitive data in buf outside enclave (line 34). ra_key and phone_num are finally printed directly on the standard output of the untrust OS, so attackers can easily get them with little effort.

Write Private Data to OCALL Return Pointers (P4). Figure 12 shows a privacy leakage vulnerability in TaLoS. ocall_malloc is an OCALL that returns a pointer (line 3). In the enclave code, ssl_update_cache updates the cache and calls ocall_new_session_callback_wrapper (line 10). Then, ocall_new_session_callback_wrapper will call the OCALL ocall_malloc when ssl_session_outside is NULL to apply for a block of untrusted memory (line 17), and write the SSL session to this memory (line 20), resulting in the disclosure of SSL security information.

Write Private Data to NULL Pointers (P5). As shown in Figure 13, PrivacyGuard's enclave code first uses malloc to apply for a block of trusted memory (line 4) but forgets to check whether the result is successful, so DO_data_key may be NULL, and then generates a 16-byte random number as the key. The key is stored in DO_data_key (line 11), if the memory allocation fails, the key will be leaked. The developers of PrivacyGuard have confirmed the vulnerability.

```
1  sgx_status_t ECALL_enclave_DO_config(int num_DOs){
2      ...
3      // malloc DO_data_key but not check the result.
4      DO_data_key = (sgx_aes_gcm_128bit_key_t *) malloc(num_DOs *
       sizeof(sgx_aes_gcm_128bit_key_t));
5      ...
6  }
7  sgx_status_t ECALL_put_secret_data(sgx_ra_context_t context, uint8_t
   *p_secret, uint32_t secret_size, uint8_t *p_gcm_mac, uint32_t DO_ID){
8      ...
9      //Generate a 16-Byte data encryption key for DO's data
10     //leak!!! DO_data_key may be NULL
11     sgx_read_rand(DO_data_key[DO_ID-1],
       sizeof(sgx_aes_gcm_128bit_key_t));
12     ...
13 }
```

Figure 13: A P5 privacy leakage vulnerability in PrivacyGuard. The enclave code uses malloc to allocate a block of memory to store the encryption key, but forgets to check if the return pointer is NULL. when the host runs out of memory or is attacked, the encryption key can be compromised.

```
/***EDL***/
untrusted {
    void* ocall_malloc(size_t size);
}

/*** enclave code **/
void ssl_update_cache(SSL *s, int mode){
    ...
    int retval = ocall_new_session_callback_wrapper(s);
    ...
}
static SSL_SESSION* ssl_session_outside = NULL;
static int ocall_new_session_callback_wrapper(struct
ssl_st *ssl) {
    if (!ssl_session_outside) {
        //ssl_session_outside points to outside the
        enclave
        ocall_malloc((void**)&ssl_session_outside,
        sizeof(*ssl_session_outside));
    }
        //SSL session leaked!!!
        memcpy(ssl_session_outside, ssl->session,
        sizeof(*ssl_session_outside));
}
/*** application code **/
void* ocall_malloc(size_t size) {
    void* ret = malloc(size); // return a pointer to
    untrusted memory
    return ret;
}
```

Figure 12: A P4 privacy leak in TaLoS. TaLoS applies for a piece of untrusted memory outside the enclave and writes the SSL session to this memory in the enclave. SSL session is important for SSL security, for example, the master key involved in security is included in the SSL session.

5.3 Performance

To answer RQ2, we measured the time it took STELLA to analyze these open-source enclaves, and the results are shown in Table 2. Analysis time is positively related to the number of lines of codes. When the number of lines of code is about 10K, the analysis speed of STELLA is fast, and the analysis can be completed within 10 seconds. When the number of lines of code is around 500K (TaLoS), the analysis time goes up, but it does not exceed 10 minutes. Overall, STELLA is very efficient even when analyzing large programs.

5.4 Threats to Validity

False positives are not statistically analyzed in the experiment. Many leaks in the report are difficult to distinguish whether it is sensitive information because this requires prior knowledge of the enclave programs. The bugs in Table 1 have been manually confirmed to be high-risk sensitive information.

In addition, our NULL pointer detection is not sound. we perform pointer analysis [1] (the point-to set of NULL pointer is empty) and `malloc` analysis (whether the enclave code assumes that `malloc` will succeed), so the actual numbers of P5 vulnerabilities may be more and the problem is more serious. Furthermore, we cannot detect leaks caused by wild pointers and mathematically manipulated pointers pointing outside the enclave.

Table 2: The time to analyze each project. The results are average over three measurements.

Project Name	Enclave LoC	Time(s)
SGX-Tor	491,431	388.84
sgx_wechat_app	307	0.23
Fidelius	14,129	9.11
sgx-based-mix-networks	211	0.09
sgx-dnet	14,344	6.86
SGX_SQLite	213,806	123.57
TaLoS	527,837	407.76
sgx-aes-gcm	136	0.01
password-manager	6,383	0.29
TACIoT	472	0.04
BiORAM-SGX	11,251	3.59
PrivacyGuard	85,015	6.33
Town-Crier	12,275	6.34

6 RELATED WORK

Several previous research also studies privacy leakage problems in enclave programs.

COIN attacks [18] summarizes four interface-oriented attacks: Concurrent, Order, Inputs, and Nested, and implements a testing framework to detect bugs with instruction emulation and concolic execution. COIN detects enclave memory information leakage by checking the length of `memcpy` or a loop condition. This approach can only detect privacy leakage caused by the out-of-bounds copy, but cannot detect the patterns we defined.

TeeRex [8] mainly detects memory corruption vulnerabilities in the enclave code introduced by the interface between the host and the enclave. These vulnerabilities could allow attackers to corrupt function pointers and arbitrary memory writes. In terms of enclave information leakage, TeeRex's work is relatively limited, and it only briefly explains that under the vulnerability of null pointer dereference, malicious `user_check` pointer input causes arbitrary memory read. However, our work demonstrates that even in the absence of malicious third parties, the enclave code may be leaking secrets.

Moat [35] employs formal verification to verify whether the enclave code leaks secrets to an adversary. However, Moat is not flexible and scalable enough to apply to large real-world enclave code. STELLA can efficiently analyze large open-source projects.

DEFLECTION [23] verifies the enclave programs by employing compiler instrumentation to insert some privacy security policies into the enclave programs. DEFLECTION introduces runtime overhead for enclave code. In the worst case, the performance overhead is 39.8%.

7 CONCLUSION

This paper investigates possible privacy leakage in enclave code. The main challenge lies in how to effectively and efficiently identify

the privacy leakage code. We at first define five common privacy-leaking patterns, then propose a novel sparse taint analysis method to identify these leaking patterns. We implement a prototype STELLA and analyze several open-source enclave programs on GitHub with STELLA. Our experimental results show that sparse taint analysis can effectively and efficiently detect privacy leaking bugs. We believe that our method will shed light on further research on enclave privacy protection.

ACKNOWLEDGEMENT

This work was sponsored by CCF- AFSG Research Fund (project id: RF20210017).

REFERENCES

[1] Lars Ole Andersen. 1994. *Program analysis and specialization for the C programming language*. Ph.D. Dissertation. Citeseer.

[2] Steven Arzt, Siegfried Rasthofer, Christian Fritz, Eric Bodden, Alexandre Bartel, Jacques Klein, Yves Le Traon, Damien Octeau, and Patrick McDaniel. 2014. Flowdroid: Precise context, flow, field, object-sensitive and lifecycle-aware taint analysis for android apps. *Acm Sigplan Notices* 49, 6 (2014), 259–269.

[3] Pierre-Louis Aublin, Florian Kelbert, D O'Keffe, Divya Muthukumaran, Christian Priebe, Joshua Lind, Robert Krahn, Christof Fetzer, David Eyers, and Peter Pietzuch. 2017. TaLoS: Efficient TLS Termination Inside SGX Enclaves for Existing Applications. https://github.com/lsds/TaLoS.

[4] Pierre-Louis Aublin, Florian Kelbert, D O'Keffe, Divya Muthukumaran, Christian Priebe, Joshua Lind, Robert Krahn, Christof Fetzer, David Eyers, and Peter Pietzuch. 2017. TaLoS: Secure and transparent TLS termination inside SGX enclaves. (2017).

[5] Gbadebo Ayoade, Vishal Karande, Latifur Khan, and Kevin Hamlen. 2018. Decentralized IoT data management using blockchain and trusted execution environment. In *2018 IEEE International Conference on Information Reuse and Integration (IRI)*. IEEE, 15–22.

[6] Stefan Brenner, Colin Wulf, David Goltzsche, Nico Weichbrodt, Matthias Lorenz, Christof Fetzer, Peter Pietzuch, and Rüdiger Kapitza. 2016. Securekeeper: Confidential zookeeper using intel sgx. In *Proceedings of the 17th International Middleware Conference*. 1–13.

[7] Z Berkay Celik, Leonardo Babun, Amit Kumar Sikder, Hidayet Aksu, Gang Tan, Patrick McDaniel, and A Selcuk Uluagac. 2018. Sensitive Information Tracking in Commodity {IoT}. In *27th USENIX Security Symposium (USENIX Security 18)*. 1687–1704.

[8] Tobias Cloosters, Michael Rodler, and Lucas Davi. 2020. {TeeRex}: Discovery and Exploitation of Memory Corruption Vulnerabilities in {SGX} Enclaves. In *29th USENIX Security Symposium (USENIX Security 20)*. 841–858.

[9] Luigi Coppolino, Salvatore D'Antonio, Valerio Formicola, Giovanni Mazzeo, and Luigi Romano. 2021. <italic>VISE</italic>: Combining Intel SGX and Homomorphic Encryption for Cloud Industrial Control Systems. *IEEE Trans. Comput.* 70, 5 (2021), 711–724. https://doi.org/10.1109/TC.2020.2995638

[10] Chuntao Dong, Qingni Shen, Xuhua Ding, Daoqing Yu, Wu Luo, Pengfei Wu, and Zhonghai Wu. 2022. T-Counter: Trustworthy and Efficient CPU Resource Measurement using SGX in the Cloud. *IEEE Transactions on Dependable and Secure Computing* (2022).

[11] Pedro Escaleira. 2021. Hidden anonymization with Intel SGX based mixes. https://github.com/SabaEskandarian/Fidelius.

[12] Saba Eskandarian, Jonathan Cogan, Sawyer Birnbaum, Peh Chang Wei Brandon, Dillon Franke, Forest Fraser, Gaspar Garcia, Eric Gong, Hung T Nguyen, Taresh K Sethi, et al. 2018. Fidelius: Protecting User Secrets from Compromised Browsers. https://github.com/SabaEskandarian/Fidelius.

[13] Matthew Hoekstra, Reshma Lal, Pradeep Pappachan, Vinay Phegade, and Juan Del Cuvillo. 2013. Using innovative instructions to create trustworthy software solutions. *HASP@ ISCA* 11, 10.1145 (2013), 2487726–2488370.

[14] ARM Holdings. 2009. ARM security technology: Building a secure system using trustzone technology. *Retrieved on June* 10 (2009), 2021.

[15] Intel. 2016. Intel(R) Software Guard Extensions SDK Developer Reference. https://01.org/sites/default/files/documentation/intel_sgx_sdk_developer_reference_for_linux_os_pdf.pdf

[16] Intel. 2019. Intel(R) Software Guard Extensions SDK for Linux*. https://01.org/intel-software-guard-extensions

[17] Daiki Iwata. 2020. A Practical Privacy-Preserving Data Analysis for Personal Genome by Intel SGX. https://github.com/diwata11/BiORAM-SGX.

[18] Mustakimur Rahman Khandaker, Yueqiang Cheng, Zhi Wang, and Tao Wei. 2020. COIN attacks: On insecurity of enclave untrusted interfaces in SGX. In *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*. 971–985.

[19] Seongmin Kim, Juhyeng Han, Jaehyeong Ha, Taesoo Kim, and Dongsu Han. 2017. Tor anonymity network for SGX. https://github.com/kaist-ina/SGX-Tor.

[20] Seongmin Kim, Juhyeng Han, Jaehyeong Ha, Taesoo Kim, and Dongsu Han. 2018. Sgx-tor: A secure and practical tor anonymity network with sgx enclaves. *IEEE/ACM Transactions on Networking* 26, 5 (2018), 2174–2187.

[21] Taehoon Kim, Joongun Park, Jaewook Woo, Seungheun Jeon, and Jaehyuk Huh. 2019. ShieldStore: Shielded In-Memory Key-Value Storage with SGX. In *Proceedings of the Fourteenth EuroSys Conference 2019* (Dresden, Germany) *(EuroSys '19)*. Association for Computing Machinery, New York, NY, USA, Article 14, 15 pages. https://doi.org/10.1145/3302424.3303951

[22] Shiv Kushwah. 2018. Password Manager using Intel SGX SDK's enclave technology. https://github.com/ShivKushwah/password-manager.

[23] Weijie Liu, Wenhao Wang, Hongbo Chen, XiaoFeng Wang, Yaosong Lu, Kai Chen, Xinyu Wang, Qintao Shen, Yi Chen, and Haixu Tang. 2021. Practical and Efficient in-Enclave Verification of Privacy Compliance. In *2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 413–425. https://doi.org/10.1109/DSN48987.2021.00052

[24] LLVM. 2022. LLVM Language Reference Manual. https://llvm.org/docs/LangRef.html

[25] Yerzhan Mazhkenov. 2017. SQLite database inside a secure Intel SGX enclave (Linux). https://github.com/yerzhan7/SGX_SQLite.

[26] Frank McKeen, Ilya Alexandrovich, Alex Berenzon, Carlos V Rozas, Hisham Shafi, Vedvyas Shanbhogue, and Uday R Savagaonkar. 2013. Innovative instructions and software model for isolated execution. *Hasp@ isca* 10, 1 (2013).

[27] Mitar Milutinovic, Warren He, Howard Wu, and Maxinder Kanwal. 2016. Proof of luck: An efficient blockchain consensus protocol. In *proceedings of the 1st Workshop on System Software for Trusted Execution*. 1–6.

[28] Bernard Ngabonziza, Daniel Martin, Anna Bailey, Haehyun Cho, and Sarah Martin. 2016. Trustzone explained: Architectural features and use cases. In *2016 IEEE 2nd International Conference on Collaboration and Internet Computing (CIC)*. IEEE, 445–451.

[29] plaublin. [n.d.]. Efficient TLS termination inside Intel SGX enclaves for existing applications. https://github.com/lsds/TaLoS.

[30] Christian Priebe, Kapil Vaswani, and Manuel Costa. 2018. EnclaveDB: A secure database using SGX. In *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 264–278.

[31] Jaak Randmets. 2021. An Overview of Vulnerabilities and Mitigations of Intel SGX Applications. (2021).

[32] rodolfoams. 2016. Example on how to use the Intel SGX implementation of AES-GCM. https://github.com/rodolfoams/sgx-aes-gcm.

[33] Qingkai Shi, Xiao Xiao, Rongxin Wu, Jinguo Zhou, Gang Fan, and Charles Zhang. 2018. Pinpoint: Fast and precise sparse value flow analysis for million lines of code. In *Proceedings of the 39th ACM SIGPLAN Conference on Programming Language Design and Implementation*. 693–706.

[34] Qingkai Shi, Peisen Yao, Rongxin Wu, and Charles Zhang. 2021. Path-Sensitive Sparse Analysis without Path Conditions. In *Proceedings of the 42nd ACM SIGPLAN International Conference on Programming Language Design and Implementation* (Virtual, Canada) *(PLDI 2021)*. Association for Computing Machinery, New York, NY, USA, 930–943. https://doi.org/10.1145/3453483.3454086

[35] Rohit Sinha, Sriram Rajamani, Sanjit Seshia, and Kapil Vaswani. 2015. Moat: Verifying confidentiality of enclave programs. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*. 1169–1184.

[36] Yulei Sui, Peng Di, and Jingling Xue. 2016. Sparse Flow-Sensitive Pointer Analysis for Multithreaded Programs. In *Proceedings of the 2016 International Symposium on Code Generation and Optimization* (Barcelona, Spain) *(CGO '16)*. Association for Computing Machinery, New York, NY, USA, 160–170. https://doi.org/10.1145/2854038.2854043

[37] Yulei Sui and Jingling Xue. 2016. SVF: interprocedural static value-flow analysis in LLVM. In *Proceedings of the 25th international conference on compiler construction*. ACM, 265–266.

[38] Yulei Sui, Ding Ye, and Jingling Xue. 2012. Static Memory Leak Detection Using Full-Sparse Value-Flow Analysis. In *Proceedings of the 2012 International Symposium on Software Testing and Analysis* (Minneapolis, MN, USA) *(ISSTA 2012)*. Association for Computing Machinery, New York, NY, USA, 254–264. https://doi.org/10.1145/2338965.2336784

[39] Guilherme Araujo Thomaz. 2022. Trusted Acess Control for IoT Data in Cloud using Enclaves. https://github.com/GTA-UFRJ-team/TACIoT.

[40] Yongzhi Wang, Lingtong Liu, Cuicui Su, Jiawen Ma, Lei Wang, Yibo Yang, Yulong Shen, Guangxia Li, Tao Zhang, and Xuewen Dong. 2017. CryptSQLite: Protecting data confidentiality of SQLite with intel SGX. In *2017 International Conference on Networking and Network Applications (NaNA)*. IEEE, 303–308.

[41] Yang Xiao, Ning Zhang, Jin Li, Wenjing Lou, and Y Thomas Hou. 2020. PrivacyGuard: Enforcing Private Data Usage Control with Blockchain and Attested Off-chain Contract Execution. https://github.com/yang-sec/PrivacyGuard.

[42] Zhemin Yang and Min Yang. 2012. Leakminer: Detect information leakage on android with static taint analysis. In *2012 Third World Congress on Software*

Engineering. IEEE, 101–104.

[43] Yuan Yu, Andy Zhao, and Lili Zhang. [n.d.]. Intel Software Guard Extensions SDK for Linux. Retrieved Apr 28, 2022 from https://01.org/intel-software-guard-extensions

[44] Peterson Yuhala, Pascal Felber, Valerio Schiavoni, and Alain Tchana. 2021. SGX-Darknet: SGX compatible ML library. https://github.com/SabaEskandarian/Fidelius.

[45] Fan Zhang, Ethan Cecchetti, Kyle Croman, Ari Juels, and Elaine Shi. 2016. Town crier: An authenticated data feed for smart contracts. https://github.com/bl4ck5un/Town-Crier.

[46] Yao Zhao. 2020. Use sgx to resolve wechat app trust problem. https://github.com/TonyCode2012/sgx_wechat_app.

[47] Wenchao Zhou, Yifan Cai, Yanqing Peng, Sheng Wang, Ke Ma, and Feifei Li. 2021. Veridb: an SGX-based verifiable database. In *Proceedings of the 2021 International Conference on Management of Data*. 2182–2194.