

Unity is Strength: Enhancing Precision in Reentrancy Vulnerability Detection of Smart Contract Analysis Tools

Zexu Wang, Jiachi Chen, Peilin Zheng, Yu Zhang, Weizhe Zhang, Zibin Zheng, *Fellow, IEEE*

Abstract—Reentrancy is one of the most notorious vulnerabilities in smart contracts, resulting in significant digital asset losses. However, many previous works indicate that current Reentrancy detection tools suffer from high false positive rates. Even worse, recent years have witnessed the emergence of new Reentrancy attack patterns fueled by intricate and diverse vulnerability exploit mechanisms. Unfortunately, current tools face a significant limitation in their capacity to adapt and detect these evolving Reentrancy patterns. Consequently, ensuring precise and highly extensible Reentrancy vulnerability detection remains critical challenges for existing tools. To address this issue, we propose a tool named ReEP, designed to reduce the false positives for Reentrancy vulnerability detection. Additionally, ReEP can integrate multiple tools, expanding its capacity for vulnerability detection. It evaluates results from existing tools to verify vulnerability likelihood and reduce false positives. ReEP also offers excellent extensibility, enabling the integration of different detection tools to enhance precision and cover different vulnerability attack patterns. We perform ReEP to eight existing state-of-the-art Reentrancy detection tools. The average precision of these eight tools increased from the original 0.5% to 73% without sacrificing recall. Furthermore, ReEP exhibits robust extensibility. By integrating multiple tools, the precision further improved to a maximum of 83.6%. These results demonstrate that ReEP effectively unites the strengths of existing works, enhances the precision of Reentrancy vulnerability detection tools.

Index Terms—Reentrancy Detection, Symbolic Execution, Path Pruning, Smart Contracts

1 INTRODUCTION

SMART contracts offer several unique features that distinguish them from traditional software programs, especially in finance and permission management scenarios [1], [2]. Numerous smart contract vulnerabilities have been discovered through real-world attacks or theoretical analysis, including the Reentrancy vulnerability [3], [4]. Since the DAO attack in 2016, which resulted in the theft of approximately 150 million dollars in digital assets, the Reentrancy vulnerability has caused significant asset losses.

Reentrancy vulnerabilities arise when external (malicious) contracts exploit reentrant function characteristics to bypass permission control checks [5]. This allows external contracts to enter the same function multiple times, manipulate

contract logic, and steal assets. A variety of technologies have been developed to detect Reentrancy vulnerabilities in smart contracts, which can be broadly divided into two categories, i.e., static analysis [6]–[8] and dynamic analysis [9]–[12]. Static analysis techniques often collect incomplete program state information, which can lead to false positives due to the loss of the state in contract interactions. Conversely, dynamic vulnerability detection models frequently struggle with deep-state search and comprehensive state analysis in cross-contract vulnerability scenarios. Zheng et al. [5] conducted a large-scale empirical study on existing popular Reentrancy vulnerability detection tools and found that these tools produced false positives as high as 99.8%, with 55% caused by *incorrect permission control verification* and 41% due to *the lack of external contract function analysis*. Accurate permission control checks and cross-contract state analysis continue to challenge existing detection tools. Consequently, reducing the false positives of Reentrancy vulnerability detection remains a major topic in smart contract security research.

Despite significant research efforts directed toward detecting reentrancy vulnerabilities, the constant evolution of exploitation mechanisms has led to the emergence of new Reentrancy attack patterns in recent years. This complexity and variability necessitate a continuous expansion of vulnerability detection patterns within existing tools to ensure reliable detection of Reentrancy vulnerabilities. Unfortunately, many existing tools (such as Oyente [9], Osiris [13], Manticore [10], etc.) have long struggled to effectively detect Reentrancy vulnerabilities in real-world contracts due to their outdated detection patterns.

- Zexu Wang, Jiachi Chen, Zibin Zheng, Peilin Zheng are with School of Software Engineering, Sun Yat-sen University, China.
E-mail: {wangzx97, zhengpl3}@mail2.sysu.edu.cn
E-mail: {chenjch86, zhizibin}@mail.sysu.edu.cn
- Weizhe Zhang, Yu Zhang are affiliated with School of Computer Science and Technology, Harbin Institute of Technology, China.
E-mail: {yuzhang, wzzhang}@hit.edu.cn
- Zexu Wang, Yu Zhang, and Weizhe Zhang are also affiliated with Peng Cheng Laboratory, China.
- Jiachi Chen is also affiliated with The State Key Laboratory of Blockchain and Data Security, Zhejiang University.
- Zibin Zheng is also affiliated with Guangdong Engineering Technology Research Center of Blockchain.
- Jiachi Chen is the corresponding author.

Manuscript received ; revised . The work described in this paper is supported by the Open Research Fund of The State Key Laboratory of Blockchain and Data Security, Zhejiang University, the National Natural Science Foundation of China (62302534, 62332004), and the Major Key Project of Peng Cheng Laboratory (PCL2023A05-2).

To address the above challenges, we introduce ReEP, a tool designed to reduce false positives for Reentrancy vulnerability detection. ReEP evaluates results from existing tools and validates vulnerability likelihood to reduce false positives. When new vulnerability patterns emerge, ReEP integrates the corresponding detection tools to cover different vulnerability patterns. ReEP consists of two phases: *target state search* and *symbolic execution verification*. In the *target state search* phase, ReEP uses program analysis to assist in pruning paths. ReEP performs program dependency analysis on the vulnerability functions provided by *Origin Tools*, generating function sequences related to vulnerability triggering to guide the symbol execution. *Origin Tools* refer to existing tools, such as Mythril [14], and Slither [15]. Meanwhile, ReEP utilizes CFG Pruner to construct SMC-CFG (State Maximal Correlation CFG), which can optimize path traversal. In the *symbolic execution verification* phase, ReEP implements program instrumentation to collect and analyze path constraints, enabling cross-contract symbolic execution analysis. By utilizing symbolic execution to verify the reachability of these paths, we can reduce the false positives and enhance the precision of vulnerability detection for the *Origin Tool*. Moreover, ReEP boosts strong extensibility by integrating new detection tools to expand the coverage of vulnerability detection patterns, thereby enhancing its capability to detect Reentrancy vulnerabilities.

We evaluated the effectiveness of ReEP by examining its ability to improve the detection precision of *Origin Tools*, its capability to integrate multiple tools, and understanding the impact of each stage within the ReEP framework. The experimental results showed that when integrated with ReEP, the average precision of Origin Tools increased from 0.5% to 73%, significantly improving precision without sacrificing recall. Furthermore, ReEP is able to merge multiple Reentrancy detection tools to enhance its capabilities. By integrating six tools, ReEP achieves a peak precision of 83.6%, while the best performance of the current state-of-the-art tools is only 31.8%, demonstrating its effectiveness and extensibility in improving detection precision. In addition, we conducted ablation experiments to understand how each stage within ReEP affects overall effectiveness. In general, ReEP provides a robust solution for improving the ability to detect Reentrancy vulnerabilities in smart contracts.

The main contributions of our work are as follows:

- We designed a tool called ReEP to reduce the false positives for Reentrancy vulnerability detection. At the same time, it has strong extensibility in merging multiple tools to expand its capacity for detection.
- We propose an approach that uses symbolic execution for verifying vulnerability path reachability. It combines program dependency analysis to guide path pruning, achieving efficient Reentrancy vulnerability verification.
- We applied ReEP to eight state-of-the-art Reentrancy vulnerability detection tools, experimental results show that it can significantly reduce the false positive rates and improve the precision of existing tools.
- We publicize the ReEP's source code and the experimental dataset at <https://github.com/ReEP-SC/ReEP>.

This paper is organized as follows. In Section 2, we describe some necessary background and explain the challenges faced in Reentrancy vulnerabilities through motivation examples. In Section 3, we introduce the workflow of our proposed method and delve into the technical details of ReEP. In Section 4, we evaluate the performance of ReEP. We discuss threats to validity in Section 5 and summarize the related work in Section 6. In Section 7, we conclude the paper and outline future works.

2 BACKGROUND AND MOTIVATION

In this section, we provide background knowledge and the motivation behind the design of the ReEP tool. Additionally, we summarize the challenges faced in Reentrancy vulnerability detection.

2.1 Reentrancy

Since the DAO attack in 2016, detecting Reentrancy vulnerabilities has been a critical research topic in smart contract security. Attackers often exploit Reentrancy attacks to illegally acquire substantial amounts of digital assets, especially in DeFi applications where smart contracts manage significant volumes of digital assets. A Reentrancy attack is a type of malicious behavior that exploits a vulnerability in smart contracts, in which permission controls are inadequately checked when called by an external (malicious) contract. In such attacks, the attacker repeatedly enters the function through one function call to obtain considerable profit.

```

1 function withdraw(uint _amount) public {
2     require(balance[msg.sender] >= _amount);
3     (bool success, ) = msg.sender.call.value(
4         _amount)("");
5     require(success);
6     balance[msg.sender] -= _amount;
7 }

```

Fig. 1. Simple example of Reentrancy

Figure 1 shows a function named *withdraw* that contains a Reentrancy vulnerability. In the *withdraw* function, the contract first checks whether the *caller* (represented by *msg.sender*) has a sufficient balance (in L2). It then transfers the requested ether to the *caller* (in L3) and deducts the transferred amount from the *caller's* balance recorded in the user balance variable (in L5). However, Solidity introduces a special mechanism called the *fallback function*, which can be used to execute code when the contract receives ether from other addresses. The *fallback function* provides an opportunity for exploiting the Reentrancy vulnerability. In Figure 1, the *call.value()* (in L3) automatically invokes the *fallback function* of the *caller* contract, allowing the *caller* to take control of the control flow. Attackers can deploy malicious code in the *fallback function* to repeatedly call the *withdraw()* function. Note that in the second invocation of *withdraw()*, the code in L5 has not been executed since the invocation begins at the *call.value()* in L3, and thus the user balance has not been changed at this time. As a result, the condition check (in L2) of the second invocation passes, and the victim contract will repeatedly transfer ether to the *caller* until the contract's balance is drained.

2.2 Motivation

We use the following example to illustrate practical applications and key contributions of ReEP to smart contract development.

Alice, a smart contract developer, rigorously assesses her contracts for potential Reentrancy vulnerabilities before deploying them on Ethereum. To ensure a comprehensive analysis, she employs a variety of detection tools. However, different tools often report different vulnerability locations, compelling Alice to engage in thorough manual verification (as studies indicate that existing tools produce false positives as high as 99.8% [5]). Furthermore, with the emergence of new Reentrancy attack patterns and corresponding tools, Alice needs to add them to her detection toolkit. However, the incorporation of these new tools may also generate new false positives, further increasing her workload.

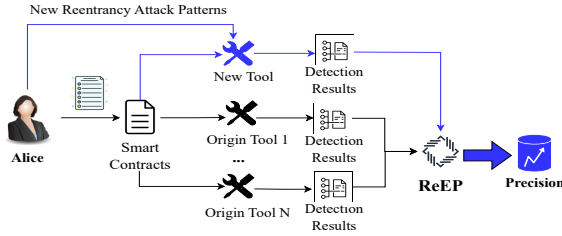


Fig. 2. The use case of ReEP

At this stage, Alice can employ ReEP to efficiently reduce the false positives while consistently addressing the detection of new Reentrancy attack patterns, ultimately lightening her workload. ReEP is an automated verification tool, designed to validate the results of Reentrancy vulnerability detections from existing tools (*Origin Tools*), verifying vulnerability likelihood to reduce false positives, thus enhancing precision. Moreover, ReEP has excellent extensibility. When new vulnerability patterns emerge, by incorporating the corresponding detection tools (*New Tools*), Alice can further enhance Reentrancy vulnerability detection capabilities, ensuring broader coverage.

ReEP alleviates the manual verification of false positives, significantly cutting down Alice's workload. This underscores ReEP's practicality in streamlining audits and enhancing precision, suitable for large-scale contract detection tasks.

2.3 Challenges

In this section, we investigate cases to uncover the main causes of false positives in current Reentrancy detection tools, and outline the related challenges.

2.3.1 Lack of permission control check

Figure 3 shows a code snippet that leads to false positives in many detection tools. The main reason is that lacking of checks on the *msg.sender*'s permission control. Smart contract permission control mainly includes who has the right to call the function, what operations can be performed, and restrictions on contract operations. The form of permission control is diverse and extensive, mistakenly recognizing permission controls can readily lead to inaccurate detection results. Analyzing permission control correctly is the main challenge in improving tool detection accuracy.

```

1 modifier onlyOwner{
2     require(msg.sender == owner);
3 }
4
5 ...
6 function execute( address _to, uint _value, bytes
7     _data) external onlyOwner {
8     ...
9     _to.call.value(_value)(_data);
10 }

```

Fig. 3. Lack of permission control check

Motivation: Relying on matching the single pattern without analyzing permission control can easily result in false positives. For example, in Figure 3, the *execute* function is modified with the *onlyOwner* modifier, which restricts its execution to the contract owner. However, Mythril [14] and Manticore [10] did not capture this permission control logic and assumed that anyone could call this function, resulting in a false positive.

Challenge: Different permission control mechanisms increase complexity and slow down path traversal. To improve detection efficiency with symbolic execution, it's vital to optimize path selection and enable targeted analysis.

2.3.2 Unable to analyze execution logic of external contract functions

Figure 4 shows an example from [5] illustrating false positives caused by the inability to analyze external contract functions. The *getTokenBal* function, querying user balances via an external *balanceOf* function without state changes or transfers, is often incorrectly flagged for Reentrancy vulnerabilities due to many tools' limitations in analyzing external functions.

```

1 contract ForeignToken {
2     function balanceOf(address _owner) constant
3         public returns (uint256);
4 }
5 contract Bitcash {
6     ...
7     function getTokenBal(address tokenAddr, address
8         who) constant public returns (uint){
9         ForeignToken t = ForeignToken(tokenAddr);
10        bal = t.balanceOf(who);
11        return bal;
12    }
13 }

```

Fig. 4. Unable to analyze execution logic of external contract functions

Motivation: Many detection tools struggle to understand how external contract functions work, especially in cross-contract interactions. For instance, in Figure 4, the *balanceOf* function of the *tokenAddr* address is called to check the balance of the *who* address and update the *bal* variable (in L9). However, these tools often can not determine if the external function involves transfers, so they rely on "state changes after external calls" to identify Reentrancy vulnerabilities, leading to false positives.

Challenge: Cross-contract calls are challenging due to the difficulties in analyzing external contract functions. This often leads to incomplete state analysis and false positives.

3 METHODOLOGY

In this section, we will introduce the workflow and delve into the technical details of ReEP.

3.1 Overview

We propose ReEP to verify vulnerability information from existing tools' detection reports, enhancing the precision of Reentrancy vulnerability detection. As shown in Figure 5, ReEP takes smart contracts as input and provides Reentrancy vulnerability detection results as output. The overall process is divided into two stages with four steps. In the first phase, *Origin Tools* like Mythril [14] report Vulnerable Function Names, which ReEP integrates from different tools. Based on these function names, program dependency analysis is performed to generate the sequence of functions related to the vulnerability trigger, guiding symbolic execution. Subsequently, the CFG Pruner constructs the SMC-CFG (State Maximum Correlation Control Flow Graph) to optimize path traversal for symbolic execution. In the second phase, cross-contract interactions are monitored and analyzed to collect global path constraints. Constraint solvers (SMT) are then accessed to verify the reachability of vulnerability paths, determining the existence of the vulnerability.

In general, *Step*①, *Step*②, *Step*③ contribute to improve the efficiency of search, and *Step*④ aims to improve the accuracy of detection.

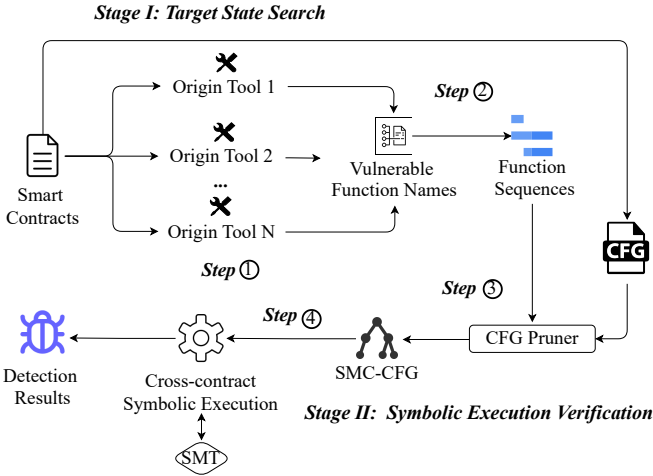


Fig. 5. The workflow of ReEP

3.2 Stage I: Targeted State Search

ReEP takes smart contracts as input and, by running *Origin Tools*, obtains functions related to Reentrancy vulnerabilities to improve the efficiency of state search. Simultaneously, it conducts program dependency analysis on vulnerable functions and generates function sequences to guide symbolic execution. Additionally, ReEP constructs the SMC-CFG to expedite path traversal during symbolic execution.

3.2.1 Step① Vulnerable Functions Collection

Initially, ReEP runs *Origin Tools* to detect vulnerable smart contracts and collect their function names that are susceptible to Reentrancy. It then analyzes the faulty instructions, capturing the variables that are read and written by these

functions, which are crucial for the execution of vulnerabilities. Based on the read-write dependency relationships of these variables among different functions within the contract, ReEP aggregates related functions and generates function sequences to guide symbolic execution.

ReEP collects Vulnerable Function Names reported by *Origin Tools*, supporting incremental extensions through modular design. When integrating new tools, ReEP's main efforts are: 1) Ensuring new tool's execution through environment configuration or Docker images, and 2) Formatting tool's results. ReEP does not need to consider the detection process of each tool. ReEP ensures their normal execution and further formats each tool's results, integrating the reported Vulnerable Function Names. The selection of *Origin Tools* is based on the criteria outlined by Zheng et al. [5], as these tools generate vulnerability detection reports containing location and function name for Reentrancy vulnerabilities.

3.2.2 Step② Function Sequence Generation

To guide symbolic execution, ReEP analyzes functions and variables related to vulnerable functions through program dependency analysis to generate the sequence of functions. Therefore, it is necessary to collect information on the functions and variables related to the vulnerable functions in the contract. This information includes:

- *Target variables*: Variables of the functions from vulnerability detection reports, denoted as V_{Target}^* .
- *Target-related variables*: Variables that have a program dependency relationship (including control dependencies or data dependencies) with V_{Target}^* , denoted as $V_{Target_Related}^*$.
- *Target functions*: Functions that read or assign variables that come from *Target variables* or *Target-related variables*, denoted as F_{Target}^* .

To collect the functions and variables related to the vulnerable function in the contract, ReEP performs a search based on whether they have program dependencies, including data and control dependencies, with the vulnerable functions. The following process is used:

- When $V_{Target_Related}^* = \emptyset$, search for the function that operates on the variables in V_{Target}^* , write the function into the set F_{Target}^* , and write the state variables of the function to the set $V_{Target_Related}^*$.
- When $V_{Target_Related}^* \neq \emptyset$, search for the function that operates on the variables in $V_{Target_Related}^*$, write the function to the set F_{Target}^* , write the state variables of that function to the set $V_{Target_Related}^*$, and continue repeating until F_{Target}^* or $V_{Target_Related}^*$ has no more new elements written to it, as $V_{Target}^* \cap V_{Target_Related}^* = V_{Target}^*$.

To facilitate symbolic execution guidance, ReEP utilizes the FDG (Function Dependency Graph) to analyze the dependencies between vulnerable functions to generate the sequence of functions. The algorithm for generating the FDG is presented in Algorithm 1, which takes $V_{Target_Related}^*$ and F_{Target}^* as inputs and produces the FDG as output. The algorithm examines the dependency relationships between the contract's functions. Specifically, it determines whether there are same variables between two functions to generate

the contract's FDG, as depicted in L3–L9 of Algorithm 1. The edge weight is determined by the number of common variables shared by the two functions. By generating the FDG, ReEP can identify functions that are directly or indirectly related to the vulnerable functions, which enables the construction of a sequence of functions related to the execution of the vulnerable functions.

The function sequence is created by sorting functions in the FDG (Function Dependency Graph), where nodes symbolize functions and edges denote shared state variables. Edge weights reflect the count of common state variables, while a node's weight, summing its edges' weights, shows the operation frequency on these variables. Sorting nodes by their weights yields the function sequence. To avoid uninitialized states, the nodes are sorted in ascending order of weight, creating the corresponding function sequence. ReEP combines symbolic execution and function sequence guidance to efficiently target and analyze functions critical to the execution path of vulnerabilities.

Algorithm 1: Generating Function Dependency Graph

Input: $V_{Target_Related}^*$ as V , F_{Target}^* as F
Output: FDG

```

1 Function Main( $V, F$ ):
2   FDG  $\leftarrow$  empty graph;
3   foreach  $f_i$  in  $F$  do
4     foreach  $f_j$  in  $F$  do
5       if  $modify(f_i, V) \cap modify(f_j, V) \neq \emptyset$ 
6         then
7            $FDG.add\_edge(f_i, f_j, weight)$ 
8         end
9     end
10  return FDG;

```

3.2.3 Step③ Path Pruning

To alleviate the problem of path explosion in symbolic execution, ReEP employs function sequences and the CFG Pruner to prune paths. Function sequences assist in eliminating irrelevant function access, while the CFG Pruner generates the SMC-CFG (State Maximum Correlation CFG) to prune the CFG of the function, thus enhancing the efficiency of symbolic execution. The SMC-CFG retains the CFG branch pointing to the key block where state updates occur and prunes irrelevant paths to minimize the cost of unnecessary path forking. At the CFG branch, it is checked whether the succeeding block is a key block. If all the succeeding blocks are key blocks, then all branches are preserved. Specifically, the CFG Pruner assigns weights to the jump edges of each basic block in the contract, reflecting the block's relevance to state updates. Blocks containing instructions that write or read state variables (*SSTORE*, *SLOAD*, and *CALL*) are considered key blocks, while blocks without state operations or containing *REVERT*/*INVALID* instructions are deemed irrelevant. If the condition for the vulnerability's existence is met, the symbolic execution path traversal is halted. By using the SMC-CFG, the path searching is accelerated by prioritizing the exploration of key blocks.

The algorithm for generating the SMC-CFG is presented in Algorithm 2. Initially, in L3–L9, all basic blocks in the CFG are traversed to calculate the weight of each edge, which is then recorded in the two-dimensional array SMC-CFG. In the *Count_weight* function in L12–L20, the opcode in each basic block is traversed. If the opcode is an instruction in *Key_instrs*, the weight of the corresponding connecting edge (*JUMP_W*) for that basic block is incremented by 1. Simultaneously, the PC value is updated to the position of the first opcode in that block. By utilizing the *Count_weight* function, each basic block establishes weighted connections to its successor blocks, ultimately generating of the SMC-CFG.

Algorithm 2: Generating the SMC-CFG

Input: the CFG of the contract
Output: the SMC-CFG of the contract

```

1 Function Main (CFG):
2   SMC-CFG  $\leftarrow$   $\emptyset$ ;
3   Blocks = getAllBlocks(CFG);
4   foreach block in Blocks do
5     foreach blk in block.successors do
6        $JUMP\_W, Fir\_pc = Count\_weight(blk);$ 
7        $SMC-CFG[blk][Fir\_pc] = JUMP\_W;$ 
8     end
9   end
10  return SMC-CFG
11 Function Count_weight (blk):
12   $JUMP\_W \leftarrow 0, PC \leftarrow 0;$ 
13  Key_instrs = [SSTORE, SLOAD, CALL];
14  foreach opcode in blk.opcodes do
15    if opcode in Key_instrs then
16       $JUMP\_W++;$ 
17       $PC = blk.first\_opcode.pc;$ 
18    end
19  end
20  return JUMP_W, PC

```

Figure 6 illustrates a part of the SMC-CFG for the *withdraw* function. The basic blocks in the SMC-CFG are represented by white rounded rectangles, while pruned blocks are shown as gray rectangles. The red solid lines indicate jump relationships between basic blocks, and the blue solid lines represent pruned jump relationships. The

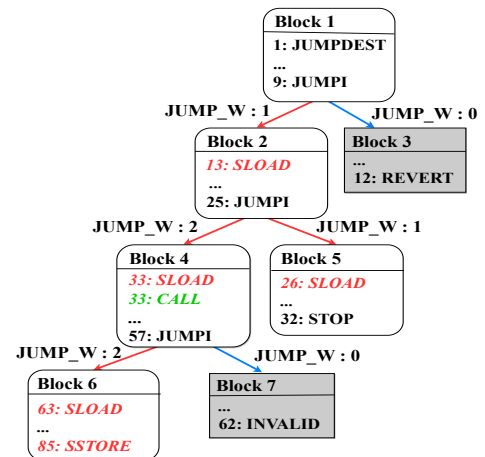


Fig. 6. The SMC-CFG of the function *withdraw*

number on each edge denotes the weight value of the jump edge (*JUMP_W*). In Figure 6, for Block 4, the *JUMPI* opcode branch has jump edges with weight values of 0 and 2. During symbolic execution using the SMC-CFG, the jump edge with a weight value of 2 is explored, and the block with a weight value of 0 is omitted. If both succeeding blocks possess non-zero weight values, they are explored in the order of their weight values. By utilizing the jump edge weight values assigned to each block by the SMC-CFG, the path branching is prioritized by selecting the jump edge with a greater weight value, enabling symbolic execution to efficiently access paths with more state operations.

3.3 Stage II: Symbolic Execution Verification

To verify the existence of vulnerabilities, ReEP employs cross-contract symbolic execution to verify the reachability of the vulnerability path. It combines the SMC-CFG to accelerate path traversal. Additionally, it uses program instrumentation with function sequences to analyze the logic of functions and collect global path constraints. Moreover, the SMT is accessed to validate the reachability of the path and determine the presence of any vulnerabilities.

In the second phase, ReEP’s symbolic execution module is mainly implemented by improving Manticore [10]. During symbolic execution, Manticore generates execution inputs through simulation testing but suffers from low vulnerability detection efficiency and poor performance. We improved its path search module and combined SMC-CFG and function sequence information through program instrumentation to enhance symbolic execution efficiency. Simultaneously, ReEP implements symbolic state storage by constructing the contract’s global storage and designing a Global Offset Table (GOT) to ensure effective symbolic state modeling.

3.3.1 Step④ Cross-contract Symbolic Execution

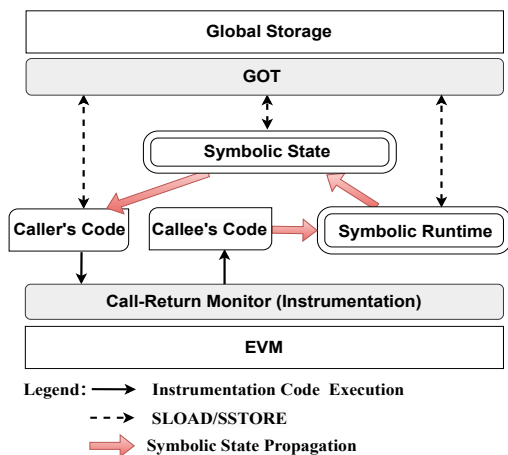


Fig. 7. The overall process of cross-contract symbolic execution

Analyzing the execution logic of functions and collecting global path constraints is of vital importance for cross-contract symbolic execution. Figure 7 illustrates the overall process of cross-contract symbolic execution in ReEP. To identify and ensure the correct switching of different contract contexts (including *msg* and *storage*), ReEP employs the *Call-Return Monitor*, a program instrumentation designed for cross-contract bytecode analysis, with function sequences to

guide the switching of different contexts. The *Global Storage* ensures the accurate writing and reading of distinct contract data, preventing data confusion arising from different contracts sharing a single storage. The *Symbolic State Propagation* addresses the issue of symbolic parameters when calling across contracts. These modules work collectively to ensure the analysis of external functions and accurate collection of global path constraints.

Call-Return Monitor. In order to ensure that the context of different contracts is switched correctly in cross-contract interaction, the instrumentation code needs to receive feedback on the start, and end positions, and return value information of the cross-contract call. As smart contracts use the Call-Return paradigm, which can be summarized into three patterns based on return values:

```
1 function Collect1(uint _am) public payable {
2     if (balances[msg.sender] >= _am) {
3         if (msg.sender.call.value(_am)()) {
4             balances[msg.sender] -= _am;
5             Log1.AddMessage(msg.sender, "Collect1");
6         } } }
```

Fig. 8. Call has no return value or does not use a return value

Call has no return value or does not use the return value: the function called in the external contract (callee) does not return the value, or the return value is not used. This pattern is shown in L5 of the function *Collect1* in Figure 8. The *AddMessage* function in the *Log1* contract does not return a value, and there is no need to pass the value across contracts.

```

1  function Collect2(uint _am) public payable{
2      if (balances[msg.sender] >= _am) {
3          if (Log2.AddMessage(msg.sender, "Collect2")) {
4              if (msg.sender.call.value(_am) ()) {
5                  balances[msg.sender] -= _am;
6              } } } }

```

Fig. 9. Call uses the return value but does not assign it

Call uses the return value but does not assign it: the function called in the external contract (callee) has the return value, which is used in the caller, but not assigned. This pattern is shown in L3 of the function *Collect2* in Figure 9. The *AddMessage* function in the *Log2* contract is checked to see if it returns a value of *True*, but its return value is not assigned.

```

1  function Collect3(uint _am) public payable{
2      if(balances[msg.sender]>=_am){
3          success = Log3.AddMessage(msg.sender,"
4              Collect3");
5          if(success){
6              if(msg.sender.call.value(_am)()){
7                  balances[msg.sender]-=_am;
8              }
9          }
10     }
11 }

```

Fig. 10. Call uses a return value and assigns it

Call uses a return value and assigns it: the function called in the external contract (callee) has the return value, which is used in the caller and assigned. This pattern is shown in L3–L4 of the function *Collect3* in Figure 10. The return value of the *AddMessage* function in the *Log3* contract is assigned to the *success* variable.

The Call-Return Monitor employs static analysis of the bytecode stream to extract essential information about cross-

contract function calls, including start and end positions, as well as the return values of external function calls. The Call-Return Monitor switches different contracts' *storage* and *msg* information based on the start and end positions and types of external calls. By analyzing the return value information of external function calls, it determines whether there is cross-contract parameter passing, ensuring the correct cross-contract interaction and avoiding state loss caused by cross-contract interactions.

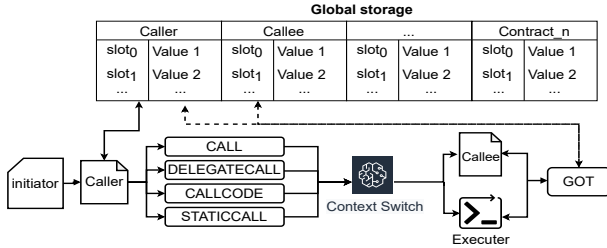


Fig. 11. Global storage for call context switch

Global Storage. To ensure the accurate writing and reading of different contract data and to prevent data confusion, ReEP implements a global storage mechanism. Many dynamic detection tools fail to save data after the function call, leading to data loss issues in subsequent transaction execution and global state analysis, as the triggering of vulnerabilities frequently stems from multiple transactions. Moreover, different types of cross-contract calls, such as *CALL*, *DELEGATECALL*, *CALLCODE*, and *STATICCALL*, necessitate different contexts for the *msg* and the *storage* [16]. Insufficient data or incorrect information can result in missed critical paths and inaccurate detection outcomes. To address this issue, ReEP combines the Global Storage with the GOT (Global Offset Table), which assists in locating and retrieving global data, thereby ensuring the correct storage and reading of various contract data.

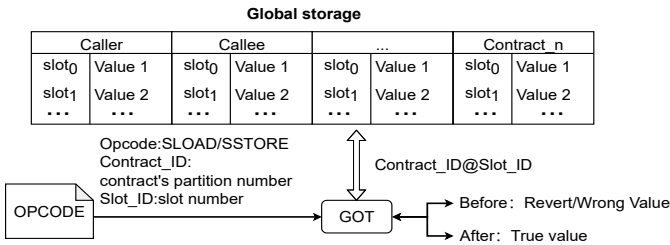


Fig. 12. GOT (Global Offset Table)

Figure 11 illustrates the functioning of global storage in the call context switch. The instrumentation code dynamically switches the contract context according to the various types of function calls between the caller and callee (including *CALL*, *DELEGATECALL*, *CALLCODE*, and *STATICCALL*) to ensure the accuracy of the cross-contract analysis. Additionally, the GOT is used to determine the position of different contract data in the global storage. The design of the GOT is depicted in Figure 12, which aids in identifying the location of various contract data. The storage location of the global variable is calculated using the formula $L = GOT(Contract_ID@Slot_ID)$, where *Contract_ID* represents the partition index of the contract in the global storage, and *Slot_ID* corresponds to the slot

index in the storage of that contract. The instrumentation code monitors the *SLOAD* and *STORE* instructions and the corresponding stack top value in the bytecode, which are employed to compute *Slot_ID*. By using GOT, true values can be obtained from global storage, avoiding issues such as reverts caused by incorrect or lost data.

Symbolic State Propagation. To address the issue of passing symbolic parameters during cross-contract calls, ReEP returns the associated constraints of symbolic parameters as return values and passes them to the caller. During cross-contract transactions, the input parameters of the called function may be symbolic, and the output values or updated states may also be in the form of symbolic constraints, making immediate computation challenging. Symbolic state propagation transforms the contract operations on symbolic parameters into constraint expressions, which can be utilized for subsequent computations.

ReEP converts the execution logic of external functions into symbolic constraint expressions and propagates them between different contracts. Z3's Bit-vector is employed to store symbolic constraint expressions, and the constraints of the input parameters (symbolic values) of the function are saved and propagated, resolving dependency issues of the program on symbolic parameters during symbolic execution. Figure 13 illustrates a simple example of a caller invoking a callee function with symbolic parameters, where the primary logic involves storing the sum of the transferred ether amount (*CALLVALUE*) and 6. The bytecode block displayed in Figure 13 stores the *caller_addr* in *slot0*, *callee_addr* in *slot1*, and adds the transferred ether *CALLVALUE* (a symbolic value) to 6, which is then stored in *slot2*. ReEP converts the calculation operations on *CALLVALUE* into constraint expressions for subsequent computation and storage. Since *CALLVALUE* is a symbolic input parameter, the storage in *slot2* contains a symbolic constraint expression.

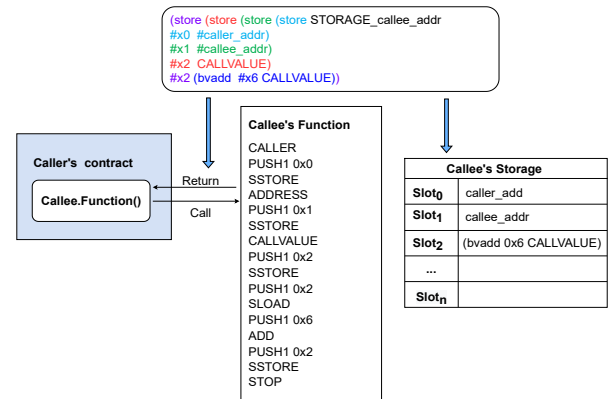


Fig. 13. Storage and propagation of symbolic states

Path constraints are resolved by accessing the constraint solver (SMT). The completeness of the set of constraints during the path constraint search phase is essential for verifying the reachability of path constraints to determine the existence of vulnerabilities.

4 EVALUATION

In this section, we evaluated the efficacy of ReEP by investigating its ability to enhance detection precision of the *Origin*

TABLE 1
Statistics of detection results of the Origin Tool with ReEP

Tool	Oyente		Mythril		Securify1		Securify2		Smartian		Saifish		Slither		EThor	
	Origin	Origin+	Origin	Origin+	Origin	Origin+	Origin	Origin+	Origin	Origin+	Origin	Origin+	Origin	Origin+	Origin	Origin+
# TP	25	25	26	26	15	22	3	3	7	7	19	21	31	31	26	28
# FP	488	8	15474	8	2372	8	2489	5	15	5	2270	8	4587	9	3269	9
# Reported	513	33	15500	34	2387	30	2492	8	22	12	2289	29	4618	40	3295	37
Precision	4.9%	75.8%	0.2%	76.5%	0.6%	73.3%	0.1%	37.5%	31.8%	58.3%	0.8%	72.4%	0.7%	77.5%	0.8%	75.7%

Tools, its ability to integrate multiple tools, and understand the impact of each stage within the ReEP framework. We address these considerations by answering the following research questions.

- RQ1.** How does ReEP improve the Reentrancy detection precision of the *Origin Tools*?
- RQ2.** What is the impact of ReEP on the recall rate?
- RQ3.** What is the extensibility of ReEP when merging multiple tools?
- RQ4.** What is the impact of different stages within ReEP?

4.1 Environment Setup and Dataset

The experiments were conducted on a machine running Ubuntu 18.04.1 LTS and equipped with 16 cores (Intel(R) Xeon(R) Gold 5217). The experimental environment was set up by either downloading Docker images or manually building the tool with the help of construction manuals (for Securify2 and Smartian). To ensure that the experiments were not overly time-consuming, we followed the time budget settings from [5], capping the maximum runtime to 120 seconds and keeping the parameters at their default settings.

We adopted two datasets in our study to comprehensively evaluate ReEP's performance in detecting Reentrancy vulnerabilities. Dataset DB1 was employed to validate RQ1, RQ3, and RQ4, while dataset DB2 was used to validate RQ2.

DB1. *Zheng et al.'s Dataset* [5]. Zheng et al. initially collected 230,548 verified contracts from Etherscan. These contracts were subsequently analyzed by using six state-of-the-art Reentrancy detection tools, and 21,212 contracts were flagged as potentially vulnerable to Reentrancy vulnerabilities by at least one of the tools. After that, two rounds of manual checks were conducted involving 50 participants, resulting in 34 contracts being confirmed as true positives (TP).

DB2. *SmartBugs Dataset* [17]. The widely used SmartBugs dataset contains 143 contracts. Among these contracts, 31 were identified as containing Reentrancy vulnerabilities. The inclusion of these labeled contracts allowed us to evaluate ReEP's recall performance in detecting Reentrancy vulnerabilities.

For tool selection, we adopted the criteria proposed by Zheng et al. [5], selecting the same set of tools: Oyente [9], Mythril [14], Securify [18] (both V1 and V2 versions), Smartian [11], and Saifish [19]. Additionally, we included Slither [15] and EThor [20] as part of the *Origin Tools*, making a total of eight tools used to generate Reentrancy vulnerability detection reports. These selected tools are all presented at top software engineering or security conferences and cover various techniques, such as symbolic execution, formal verification, and fuzz testing. Some tools do not

explicitly define Reentrancy vulnerabilities as "Reentrancy". For instance, Mythril reports two vulnerabilities related to Reentrancy, namely, *external call to user-supplied address* and *state access after external call*. To ensure consistency, we adopted the same classification criteria provided by Zheng et al. [5], which offers a standardized classification scheme for these Reentrancy vulnerability detection tools.

4.2 RQ1: Improvement on Precision

We use DB1 to evaluate the impact of ReEP on improving the Reentrancy detection precision, we compared the results of eight *Origin Tools* with and without ReEP. The results are provided in Table 1, where TP and FP stand for True Positive and False Positive, respectively. *Origin+* and *Origin* represent the *Origin Tool* with or without ReEP, and *Precision* refers to the detection precision. The precision calculation formula is: $Precision(PRE) = TP / (TP + FP)$.

As shown in Table 1, a total of 21,212 Reentrancy cases were reported from 230,548 contract detections by the *Origin Tools*. At the same time, ReEP can significantly reduce false positives for all of them, especially for Mythril, where false positives dropped from 15,474 to 8. It is worth noting that DB1(Zheng et al.'s Dataset) reported only 34 TP, while ReEP identified 7 additional contracts (a 20% increase) with Reentrancy vulnerabilities that were missed in their manual checks. We notified the authors of the dataset, and they confirmed that these 7 contracts do contain Reentrancy vulnerabilities and subsequently updated their dataset.

In conclusion, the results show that ReEP can significantly improve the precision of the *Origin Tools*. Mythril exhibited the maximum increase in precision from 0.2% to 76.5%, while Smartian showed the minimum increase, also increasing from 31.8% to 58.3%. After using the ReEP, the average precision of all *Origin Tools* rose from 0.5% to 73%, indicating that ReEP significantly improves the precision of the detection results for the *Origin Tools*.

Answer to RQ1: Experimental results indicate a significant enhancement in *Origin Tools'* precision with ReEP integration, rising from 0.5% to 73% on average, with Mythril showing the most notable improvement from 0.2% to 76.5%. These results highlight ReEP's exceptional capability in enhancing the precision of the *Origin Tools*.

4.3 RQ2: Impact on Recall

DB1 comprises 21,212 suspicious contracts selected from 230,548 verified contracts. It may not provide a fair evaluation of ReEP's recall impact, as these 21,212 contracts were marked as potentially susceptible to Reentrancy vulnerabilities by at least one of the Original tools. Therefore, we utilized DB2 (SmartBugs dataset), the most commonly used dataset, which consists of 31 contracts with Reentrancy

TABLE 2
Statistics of Detection Results on SmartBugs Dataset

Tool	Oyente		Mythril		Securify1		Securify2		Smartian		Saifish		Slither		EThor	
	Origin	Origin+	Origin	Origin+	Origin	Origin+	Origin	Origin+	Origin	Origin+	Origin	Origin+	Origin	Origin+	Origin	Origin+
# TP	21	21	10	10	17	17	6	6	15	15	19	19	29	29	25	25
# FP	43	8	48	7	31	8	47	9	19	7	24	8	38	9	55	9
# FN	10	10	21	21	14	14	25	25	16	16	12	12	2	2	6	6
# TN	69	104	64	105	81	104	65	103	93	105	88	104	74	103	57	103
Precision	32.8%	72.4%	17.2%	58.8%	35.4%	68.0%	11.3%	40.0%	44.1%	68.2%	44.2%	70.4%	43.3%	76.3%	31.3%	73.5%
Recall	67.7%	67.7%	32.3%	32.3%	54.8%	54.8%	19.4%	19.4%	48.4%	48.4%	61.3%	61.3%	93.5%	93.5%	80.6%	80.6%
F1	44.2%	70.0%	22.5%	41.7%	43.0%	60.7%	14.3%	26.1%	46.2%	56.6%	51.4%	65.5%	59.2%	84.1%	45.0%	76.9%

vulnerabilities and 112 contracts without such vulnerabilities, to assess the recall impact. The recall calculation formula is: $Recall = TP / (TP + FN)$.

Table 2 presents the result statistics of ReEP on DB2. As shown in the table, ReEP significantly enhances the precision and F1 scores of the *Origin Tools*, while having no impact on Recall. Among the eight tools, Slither demonstrated the highest recall performance on SmartBugs, remaining unchanged at 93.5%. This indicates that ReEP maintains the Recall of *Origin Tools* while improving precision.

Warning information is utilized for path pruning, combined with symbolic execution to verify path reachability, leading to fewer false positives and an enhanced F1 score, without affecting Recall of *Origin Tools*.

Answer to RQ2. ReEP utilizes warning information from *Origin Tools* to guide path pruning and targeted analysis, including path symbol execution verification, resulting in a significant reduction of false positives in the detection results without compromising Recall rate.

4.4 RQ3: Extensibility of ReEP

To analyze ReEP's extensibility, we evaluated its capacity for enhanced detection by combining multiple tools. Previous two RQs showed that while ReEP can improve precision by reducing false positives (FP), it cannot improve recall. The main reason is that ReEP relies on vulnerability information from *Origin Tools*, and a single tool's report on Reentrancy vulnerabilities may limit ReEP's potential. However, ReEP can overcome this limitation by combining multiple tools, providing more comprehensive vulnerability information and ensuring effective detection.

We conducted experiments on ReEP with DB2, thoroughly evaluating it by combining different *Origin Tools*. The extensibility of ReEP was thoroughly assessed by fusing different sets of *Origin Tools*, including combinations of two, four, six, and eight tools, resulting in a total of 127 unique combinations. The experimental data is available in our GitHub repository. When combining multiple tools, ReEP analyzes the detection results from the different tools and uses a logical OR operation to determine the final result.

Figure 14 displays ReEP's precision and recall boxplots for combinations of N *Origin Tools* (N = 2, 4, 6, 8), showing improvements with more tools integrated. Of the 28 combinations in Merger2, Mythril and Slither's merge yielded the highest precision at 67.1%. The six-tool combination (Securify1, Securify2, Mythril, Oyente, Saifish, and Slither) reached the peak precision of 83.6%, the highest among all. This suggests that adding more tools doesn't always enhance precision due to potential increases in false positives.

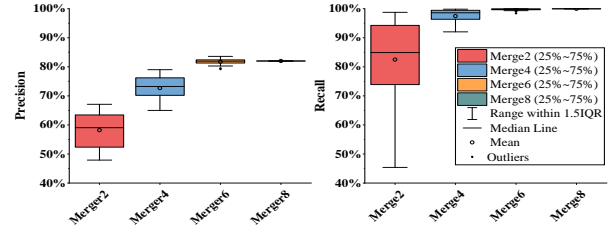


Fig. 14. Precision and recall for combinations of N *Origin Tools* (N = 2, 4, 6, 8)

Combining more tools with ReEP reduces the likelihood of missing vulnerabilities. Even though integrating a single tool with ReEP did not improve recall, recall improved as more tools were integrated. As shown in RQ2, the DB2 dataset contains 31 positive labels, and the maximum number of true positives (TP) detected by 8 tools is 29, detected by Slither. Combining ReEP with a single tool did not increase the TP count. However, when ReEP combined Slither with other tools, the TP count was at least 29. As $Recall = TP / (TP + FN)$, and the total $TP + FN$ equals 31 (the number of positive labels), the increased TP count led to an improved recall. As shown in Figure 14, recall consistently improves and becomes more stable as the number of tool combinations increases. Furthermore, as research on Reentrancy vulnerabilities advances, ReEP could leverage diverse detection scopes from multiple tools to enhance detection effectiveness, especially for tools with high false positive rates. By automating analysis, ReEP improves effectiveness and efficiency in Reentrancy vulnerability detection.

TABLE 3
The average number of reported functions and time for combinations of N *Origin Tools* (N = 2, 4, 6, 8).

	Merger2	Merger4	Merger6	Merger8
Avg. Func.	6	9	10	11
Avg. Time (s)	13.6	16.7	18.2	19.1

Additionally, we analyzed ReEP's efficiency when combining multiple tools. For direct analysis, we focused solely on the time T_R , that ReEP required to conduct subsequent analysis after obtaining the function names provided by *Origin Tools*. $T_R = T_{Overall} - T_{Ori}$, where $T_{Overall}$ is the overall running time for the entire process, and T_{Ori} is the time to run *Origin Tools*.

To analyze the cost of ReEP's combination analysis, we counted the average number of functions reported and time consumed for combinations of two, four, six, and

eight tools. The results are shown in Table 3. As ReEP integrates more tools, the average number of analyzed functions increases, leading to higher time consumption. Notably, when combining eight tools, the average number of analyzed functions is 11, but the corresponding time is only 19.1 seconds, which is relatively small. This highlights the importance of path pruning techniques for efficient analysis.

Answer to RQ3. Experimental results show that as ReEP combines more tools, vulnerability detection precision and recall improve, and results become more stable, demonstrating strong extensibility. Combining multiple tools achieved a peak precision of 83.6% and reduced the likelihood of missing vulnerabilities. Meanwhile, scaling multiple tools is moderately time-consuming, the findings highlight ReEP's extensibility and effectiveness for enhancing detection.

4.5 RQ4: Impact of Different Stages in ReEP

To evaluate the impact of each stage of ReEP on the overall effectiveness, we conducted comparative experiments on precision and time consumption for Merge6's *Best_combo* using three modes: *Stage1* (Target State Search), *Stage2* (Symbolic Execution Verification), and *Stage1&2*. In *Stage1*, ReEP integrated the FDG and SMC-CFG analysis to prune the path, while *Stage2* involved the verification of path accessibility through symbolic execution.

Table 4 presents a comparison of the detection results among the three modes, where *Avg. Tool* represents the average detection results from the *Origin Tools*. As evidenced by the table data, we can observe that *Stage1* detected all the 41 Reentrancy vulnerabilities (TP) but have 556 false positives (FP). Stage 1 does not significantly improve the precision (from 0.5% to 6.9%) due to a high number of false positives. In contrast, *Stage2*'s symbolic execution verification remarkably reduces the number of false positives (FPs) to only 51, which is less than one-tenth of the FPs identified in *Stage1*. Nonetheless, the number of true positives (TPs) detected by symbolic execution is 35, as symbolic execution may encounter path explosion, leading to the omission of certain TP contracts. The *Stage1&2* mode, which combines symbolic execution verification with path pruning, achieves the highest precision of 83.6%.

TABLE 4
Comparison of Detection Results for Different Modes

	Avg. Tool	Stage1	Stage2	Stage1&2
# TP	19	41	35	41
# FP	3870	556	51	8
Precision	0.5%	6.9%	40.7%	83.6%
Time (s)	8.2	1.3	23.8	15.6

Table 4 provides the average time consumption for the three modes. While *Stage2* takes an average of 23.8 seconds, the processing time is reduced to 15.6 seconds in *Stage1&2*. The pruning in *Stage1* plays a crucial role in enhancing ReEP's overall performance. Experimental observations indicate that *Stage1* combines FDG with SMC-CFG pruning, leading to reduced running time and mitigated symbolic execution path explosion. Furthermore, it facilitates targeted path analysis, enabling deeper path searches and improving the overall precision of vulnerability detection.

Answer to RQ4. The *Stage1&2* mode, by combining path pruning with symbolic execution verification, achieves

efficient path reachability validation, resulting in the highest precision of 83.6% when merging multiple tools.

5 THREATS TO VALIDITY

5.1 Internal Validity

The internal validity threats stem from the reliance on existing Reentrancy detection tools. While ReEP can improve precision by reducing false positives, it is not capable of detecting more Reentrancy cases than *Origin Tools* (as observed in the experimental results of RQ2). Consequently, the capabilities of ReEP might be constrained by the limitations of *Origin Tools*. Fortunately, a key strength of ReEP is its ability to merge multiple tools to enhance its detection capabilities. As demonstrated in RQ3, ReEP achieved a peak precision of 83.6% by merging six tools, underlining its impressive extensibility. Even though new Reentrancy attack patterns emerge, ReEP can integrate new detection tools, effectively extending its capabilities to adapt to evolving vulnerabilities.

5.2 External Validity

The external validity threats primarily stem from the inherent challenges associated with manual inspections of large datasets, which tend to be highly error-prone and time-consuming. The dataset *DB1* comprises 230,548 verified contracts obtained from Etherscan, among which 21,212 contracts were flagged for Reentrancy vulnerabilities by state-of-the-art automated detection tools. Manually verifying this large-scale real-world dataset can be error-prone and time-consuming. To ensure the dataset's accuracy, the authors adopted a rigorous approach involving 50 participants, including graduate students and PhDs with extensive experience in smart contract research. These participants conducted two rounds of thorough checks on the detection results, minimizing bias and ensuring the reliability of the manual examination of the reentrant contracts detected by the tools. Utilizing tool evaluation to enhance the accuracy of the results. Additionally, by employing ReEP and integrating eight state-of-the-art tools, we successfully identified 7 additional real-world Reentrancy vulnerability contracts that were initially missed. Therefore, this dataset represents real-world smart contracts deployed on the Ethereum with Reentrancy vulnerabilities.

6 RELATED WORK

6.1 Symbolic Execution.

Recent research on symbolic execution in smart contract security can be divided into two categories according to the effect: accuracy and efficiency.

To enhance the accuracy of symbolic execution in detecting security vulnerabilities, various tools such as Oyente [9], Mythril [14], Manticore [10], and Maian [21] conduct thorough analyses of contracts, and explore all possible paths to generate vulnerability reports. Sailfish [19] utilizes a storage dependency graph (SDG) to detect vulnerabilities in contracts. SmartDagger [7] constructs the cross-contract control flow graph from bytecodes, facilitating cross-contract vulnerability detection. However, blind path traversal can easily lead to imprecise results, especially in cross-contract analysis where different contracts sharing the same storage may cause data confusion. More importantly, many

existing tools face a significant limitation in their capacity to adapt and detect evolving vulnerability patterns. ReEP distinguishes itself by validating existing tool results to reduce false positives and integrating different tools to detect various vulnerability patterns.

To improve the efficiency of symbolic execution for faster detection of security vulnerabilities, tools like Oyente [9], MPro [22], and Mythril [14] analyze smart contract bytecode. They apply corresponding path pruning strategies to expedite path traversal and mitigate path explosion issues. Smartian [11], SmartDagger [7], and Smartest [12] have optimized speed by refining search strategies for symbolic execution paths or employing heuristic algorithms. Additionally, Park [23] proposes a parallel symbolic execution-based approach to accelerate detection. ReEP stands out by guiding symbolic execution based on the detection results of existing tools, achieving efficient path searching and traversal.

6.2 Dynamic and Static Detection.

Traditional static detection techniques for smart contracts [6]–[8], [24], [25] have several limitations. The public nature of the blockchain, along with diverse permission control in smart contracts, complicates static vulnerability detection, often resulting in false positives. Additionally, unclear call relationships between functions require further exploration for effective target state identification triggering vulnerabilities.

Dynamic detection methods [9]–[13], [22], [26]–[28], relying on program testing and verification, enhance result reliability. However, the inability to access global state information and perform cross-contract analysis leads to path explosion and high resource consumption. Resource constraints and vulnerability identification limitations can further contribute to detection failures or errors.

FuzzSlice [29] performs fuzz testing within a given time budget to eliminate potential false positives in static analysis. The difference is that ReEP validates suspicious vulnerability information from existing tools through symbolic execution, showcasing strong extensibility and avoiding detection failures due to the inability to generate effective inputs. During the *Target state search* phase, ReEP explores vulnerability target states to guide path pruning, eliminating irrelevant path branches to enhance path traversal efficiency. In the *Symbolic execution verification* phase, combined with program instrumentation to achieve vulnerability state-guided symbolic execution verification. ReEP combines the strengths of both static and dynamic methods, which enables efficient identification of critical states to improve precision.

7 CONCLUSION AND FUTURE WORKS

In this paper, we present ReEP, a tool that effectively improves the precision of existing Reentrancy vulnerability detection tools. ReEP evaluates results from existing tools and validates vulnerability likelihood to reduce false positives. When new vulnerability patterns emerge, integrating the corresponding detection tools ensures reliable Reentrancy detection. By analyzing the program dependency relationships of vulnerability functions in the detection report, ReEP guides path pruning and symbolic execution. Cross-contract symbolic execution is employed to verify the reachability of

vulnerability paths and confirm the existence of vulnerabilities. We implemented and validated our tool with eight state-of-the-art detection tools. After applying ReEP, the average precision of these eight tools increased significantly from 0.5% to 73%. Furthermore, by merging six tools, the precision further improved, reaching a maximum of 83.6%, while the best performance of the current state-of-the-art tools is only 31.8%. These results demonstrate that ReEP effectively unites the strengths of existing works, enhances the precision of Reentrancy vulnerability detection tools, and efficiently identifies Reentrancy vulnerabilities in real-world scenarios.

In future work, we aim to expand the scope and capabilities of detection by combining multiple technologies, covering a broader range of vulnerability types, and supporting the detection of bytecode, among other enhancements.

REFERENCES

- [1] PeckShield, "Web3 industry security report," 2022. [Online]. Available: <https://peckshield.com/static/pdf/2023.pdf>
- [2] SLOWMIST, "2022-blockchain-security-and-aml-analysis-annual-report(en)," 2023. [Online]. Available: <https://www.slowmist.com/report>
- [3] C. Ferreira Torres, M. Baden, R. Norvill, B. B. Fiz Pontiveros, H. Jonker, and S. Mauw, "Ægis: Shielding vulnerable smart contracts against attacks," in *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security*, 2020, pp. 584–597.
- [4] R. Ji, N. He, L. Wu, H. Wang, G. Bai, and Y. Guo, "Deposafe: Demystifying the fake deposit vulnerability in ethereum smart contracts," in *2020 25th International Conference on Engineering of Complex Computer Systems (ICECCS)*. IEEE, 2020, pp. 125–134.
- [5] Z. Zhang, N. Zhang, J. Su, Z. Zhong, M. Ye, and J. Chen, "Turn the rudder: A beacon of reentrancy detection for smart contracts on ethereum," in *2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)*, 2023, pp. 295–306.
- [6] J. Feist, G. Grieco, and A. Groce, "Slither: a static analysis framework for smart contracts," in *2019 IEEE/ACM 2nd International Workshop on Emerging Trends in Software Engineering for Blockchain (WETSEB)*. IEEE, 2019, pp. 8–15.
- [7] Z. Liao, Z. Zheng, X. Chen, and Y. Nan, "Smartdagger: a bytecode-based static analysis approach for detecting cross-contract vulnerability," in *Proceedings of the 31st ACM SIGSOFT International Symposium on Software Testing and Analysis*, 2022, pp. 752–764.
- [8] J. Ye, M. Ma, Y. Lin, Y. Sui, and Y. Xue, "Clairvoyance: Cross-contract static analysis for detecting practical reentrancy vulnerabilities in smart contracts," in *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering: Companion Proceedings*, 2020, pp. 274–275.
- [9] L. Luu, D.-H. Chu, H. Olickel, P. Saxena, and A. Hobor, "Making smart contracts smarter," in *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, 2016, pp. 254–269.
- [10] M. Mossberg, F. Manzano, E. Hennenfent, A. Groce, G. Grieco, J. Feist, T. Brunson, and A. Dinaburg, "Manticore: A user-friendly symbolic execution framework for binaries and smart contracts," in *2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE)*. IEEE, 2019, pp. 1186–1189.
- [11] J. Choi, D. Kim, S. Kim, G. Grieco, A. Groce, and S. K. Cha, "Smartian: Enhancing smart contract fuzzing with static and dynamic data-flow analyses," in *2021 36th IEEE/ACM International Conference on Automated Software Engineering (ASE)*. IEEE, 2021, pp. 227–239.
- [12] S. So, S. Hong, and H. Oh, "Smartest: Effectively hunting vulnerable transaction sequences in smart contracts through language model-guided symbolic execution." in *USENIX Security Symposium*, 2021, pp. 1361–1378.
- [13] C. F. Torres, J. Schütte, and R. State, "Osiris: Hunting for integer bugs in ethereum smart contracts," in *Proceedings of the 34th Annual Computer Security Applications Conference*, 2018, pp. 664–676.
- [14] ConsenSys, "Mythril," 2020. [Online]. Available: <https://github.com/ConsenSys/mythril>
- [15] J. Feist, G. Grieco, and A. Groce, "Slither Analyzer," Jun. 2023. [Online]. Available: <https://github.com/cryptic/slither>

- [16] Solidity, "Solidity compiler," 2013. [Online]. Available: <https://docs.soliditylang.org/en/latest/installing-solidity.html>
- [17] T. Durieux, J. F. Ferreira, R. Abreu, and P. Cruz, "Empirical review of automated analysis tools on 47,587 ethereum smart contracts," in *Proceedings of the ACM/IEEE 42nd International conference on software engineering*, 2020, pp. 530–541.
- [18] P. Tsankov, A. Dan, D. Drachsler-Cohen, A. Gervais, F. Buenzli, and M. Vechev, "Securify: Practical security analysis of smart contracts," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2018, pp. 67–82.
- [19] P. Bose, D. Das, Y. Chen, Y. Feng, C. Kruegel, and G. Vigna, "Sailfish: Vetting smart contract state-inconsistency bugs in seconds," in *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2022, pp. 161–178.
- [20] C. Schneidewind, I. Grishchenko, M. Scherer, and M. Maffei, "Ethor: Practical and provably sound static analysis of ethereum smart contracts," in *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 621–640. [Online]. Available: <https://doi.org/10.1145/3372297.3417250>
- [21] I. Nikolić, A. Kolluri, I. Sergey, P. Saxena, and A. Hobor, "Finding the greedy, prodigal, and suicidal contracts at scale," in *Proceedings of the 34th annual computer security applications conference*, 2018, pp. 653–663.
- [22] W. Zhang, S. Banescu, L. Pasos, S. Stewart, and V. Ganesh, "Mpro: Combining static and symbolic analysis for scalable testing of smart contract," in *2019 IEEE 30th International Symposium on Software Reliability Engineering (ISSRE)*. IEEE, 2019, pp. 456–462.
- [23] P. Zheng, Z. Zheng, and X. Luo, "Park: accelerating smart contract vulnerability detection via parallel-fork symbolic execution," in *Proceedings of the 31st ACM SIGSOFT International Symposium on Software Testing and Analysis*, 2022, pp. 740–751.
- [24] C. Schneidewind, I. Grishchenko, M. Scherer, and M. Maffei, "ethor: Practical and provably sound static analysis of ethereum smart contracts," in *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, 2020, pp. 621–640.
- [25] S. Tikhomirov, E. Voskresenskaya, I. Ivanitskiy, R. Takhaviev, E. Marchenko, and Y. Alexandrov, "Smartcheck: Static analysis of ethereum smart contracts," in *Proceedings of the 1st international workshop on emerging trends in software engineering for blockchain*, 2018, pp. 9–16.
- [26] F. Ma, Z. Xu, M. Ren, Z. Yin, Y. Chen, L. Qiao, B. Gu, H. Li, Y. Jiang, and J. Sun, "Pluto: Exposing vulnerabilities in inter-contract scenarios," *IEEE Transactions on Software Engineering*, vol. 48, no. 11, pp. 4380–4396, 2021.
- [27] J. Su, H.-N. Dai, L. Zhao, Z. Zheng, and X. Luo, "Effectively generating vulnerable transaction sequences in smart contracts with reinforcement learning-guided fuzzing," in *37th IEEE/ACM International Conference on Automated Software Engineering*, 2022, pp. 1–12.
- [28] T. D. Nguyen, L. H. Pham, J. Sun, Y. Lin, and Q. T. Minh, "sfuzz: An efficient adaptive fuzzer for solidity smart contracts," in *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering*, 2020, pp. 778–788.
- [29] A. Murali, N. S. Mathews, M. Alfarel, M. Nagappan, and M. Xu, "Fuzzslice: Pruning false positives in static analysis warnings through function-level fuzzing," in *2024 IEEE/ACM 46th International Conference on Software Engineering (ICSE)*. IEEE Computer Society, 2023, pp. 767–779.
- [30] G. Wood et al., "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum project yellow paper*, vol. 151, no. 2014, pp. 1–32, 2014.
- [31] Z. Zheng, S. Xie, H.-N. Dai, W. Chen, X. Chen, J. Weng, and M. Imran, "An overview on smart contracts: Challenges, advances and platforms," *Future Generation Computer Systems*, vol. 105, pp. 475–491, 2020.
- [32] J. Chen, X. Xia, D. Lo, J. Grundy, X. Luo, and T. Chen, "Defining smart contract defects on ethereum," *IEEE Transactions on Software Engineering*, vol. 48, no. 1, pp. 327–345, 2020.
- [33] S. So, M. Lee, J. Park, H. Lee, and H. Oh, "Verismart: A highly precise safety verifier for ethereum smart contracts," in *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2020, pp. 1678–1694.
- [34] smartbugs, "Smartbugs wild dataset," 2020. [Online]. Available: <https://github.com/smartbugs/smartbugs-wild>
- [35] B. Zhao, Z. Li, S. Qin, Z. Ma, M. Yuan, W. Zhu, Z. Tian, and C. Zhang, "{StateFuzz}: System {Call-Based}{State-Aware} linux driver fuzzing," in *31st USENIX Security Symposium (USENIX Security 22)*, 2022, pp. 3273–3289.
- [36] M. R. Parvez, "Combining static analysis and targeted symbolic execution for scalable bug-finding in application binaries," Master's thesis, University of Waterloo, 2016.
- [37] S. Arzt, S. Rasthofer, R. Hahn, and E. Bodden, "Using targeted symbolic execution for reducing false-positives in dataflow analysis," in *Proceedings of the 4th ACM SIGPLAN International Workshop on State of the Art in Program Analysis*, 2015, pp. 1–6.
- [38] Y. Xue, M. Ma, Y. Lin, Y. Sui, J. Ye, and T. Peng, "Cross-contract static analysis for detecting practical reentrancy vulnerabilities in smart contracts," in *Proceedings of the 35th IEEE/ACM International Conference on Automated Software Engineering*, 2020, pp. 1029–1040.
- [39] F. Contro, M. Crosara, M. Ceccato, and M. Dalla Preda, "Ethersolve: Computing an accurate control-flow graph from ethereum bytecode," in *2021 IEEE/ACM 29th International Conference on Program Comprehension (ICPC)*. IEEE, 2021, pp. 127–137.
- [40] X. Rival and K. Yi, *Introduction to static analysis: an abstract interpretation perspective*. MIT Press, 2020.
- [41] J. Krupp and C. Rossow, "teether: Gnawing at ethereum to automatically exploit smart contracts," in *27th {USENIX} Security Symposium ({USENIX} Security 18)*, 2018, pp. 1317–1333.
- [42] P. Godefroid, N. Klarlund, and K. Sen, "Dart: Directed automated random testing," in *Proceedings of the 2005 ACM SIGPLAN conference on Programming language design and implementation*, 2005, pp. 213–223.
- [43] F. Victor and A. M. Weintraud, "Detecting and quantifying wash trading on decentralized cryptocurrency exchanges," in *Proceedings of the Web Conference 2021*, 2021, pp. 23–32.
- [44] T. Chen, Z. Li, Y. Zhang, X. Luo, T. Wang, T. Hu, X. Xiao, D. Wang, J. Huang, and X. Zhang, "A large-scale empirical study on control flow identification of smart contracts," in *2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*. IEEE, 2019, pp. 1–11.
- [45] T. Chen, Y. Zhang, Z. Li, X. Luo, T. Wang, R. Cao, X. Xiao, and X. Zhang, "Tokenscope: Automatically detecting inconsistent behaviors of cryptocurrency tokens in ethereum," in *Proceedings of the 2019 ACM SIGSAC conference on computer and communications security*, 2019, pp. 1503–1520.
- [46] E. Coppa, H. Yin, and C. Demetrescu, "Symfusion: Hybrid instrumentation for concolic execution," in *37th IEEE/ACM International Conference on Automated Software Engineering*, 2022, pp. 1–12.
- [47] Inpluslab, "Smart contract dataset," 2021. [Online]. Available: <http://xblock.pro/#/>
- [48] etherscan, "A block explorer for ethereum," 2020. [Online]. Available: <https://etherscan.io/>
- [49] Remix, "Ethereum ide and tools for the web," 2023. [Online]. Available: <https://github.com/ethereum/remix>
- [50] E.Org, "Ethereum," 2023. [Online]. Available: <https://ethereum.org/en/>
- [51] P. Tsankov, A. Dan, D. Drachsler-Cohen, A. Gervais, F. Buenzli, and M. Vechev, "Securify: Practical security analysis of smart contracts," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2018, pp. 67–82.
- [52] C. F. Torres, A. K. Iannillo, A. Gervais, and R. State, "Confuzzius: A data dependency-aware hybrid fuzzer for smart contracts," in *2021 IEEE European Symposium on Security and Privacy (EuroS&P)*, 2021, pp. 103–119.