

Impact of Code Transformation on Detection of Smart Contract Vulnerabilities

Cuong Tran Manh

Faculty of Information Technology
VNU University of Engineering
and Technology
Hanoi, Vietnam
tranmanhcuong@vnu.edu.vn

Hieu Dinh Vo

Faculty of Information Technology
VNU University of Engineering
and Technology
Hanoi, Vietnam
hieuvd@vnu.edu.vn

Abstract—While smart contracts are foundational elements of blockchain applications, their inherent susceptibility to security vulnerabilities poses a significant challenge. Existing training datasets employed for vulnerability detection tools may be limited, potentially compromising their efficacy. This paper presents a method for improving the quantity and quality of smart contract vulnerability datasets and evaluates current detection methods. The approach centers around semantic-preserving code transformation, a technique that modifies the source code structure without altering its semantic meaning. The transformed code snippets are inserted into all potential locations within benign smart contract code, creating new vulnerable contract versions. This method aims to generate a wider variety of vulnerable codes, including those that can bypass detection by current analysis tools. The paper experiments evaluate the method’s effectiveness using tools like Slither, Mythril, and CrossFuzz, focusing on metrics like the number of generated vulnerable samples and the false negative rate in detecting these vulnerabilities. The improved results show that many newly created vulnerabilities can bypass tools and the false reporting rate goes up to 100% and increases dataset size minimum by 2.5X.

Index Terms—smart contract, vulnerability, dataset, code transform

I. INTRODUCTION

The rapid proliferation of blockchain technology has spurred its exploration across a multitude of domains, encompassing finance, supply chain management, and value chain applications [1, 2]. However, alongside these promising advancements, inherent security risks associated with blockchain technology necessitate careful consideration [3]. Vulnerabilities in smart contracts are one of the most serious threats, in the 2016 DAO attack, the Reentrancy vulnerability was exploited causing a loss of 60 million US dollars. This incident underscores the urgent need for robust security measures to safeguard smart contracts, especially considering the vast sums of money they can potentially manage, often reaching tens of billions of US dollars [4].

Recently, many commercial and research tools have focused on detecting vulnerabilities in smart contracts, and new techniques such as applying machine learning and deep learning are also gradually being used [4, 5, 6, 7, 8]. Almost all tools have a certain false positive/negative rate [9].

Moreover, the approaches that use deep learning require a lot of data. Unfortunately, currently, the number of quality datasets is limited. All human-verified datasets are under 1000 vulnerability source codes [10]. Meanwhile, datasets with more vulnerability samples are only verified by static analysis tools, which is less reliable. For example, the SoliAudit dataset [11] consists of about 20 thousand smart contracts, which have been used to evaluate many recent research methods. However, this dataset uses tools for validation, which raises concerns about not detecting false results missed by these validation tools. Another widely used dataset is SolidiFI, which takes a different approach by using available vulnerability templates and then injecting these vulnerabilities into normal code [9] to create source code files containing software vulnerabilities. SolidiFI inserts vulnerabilities at all potential locations and is also considered helpful for finding corner cases that are difficult to detect in real-life scenarios [12, 2]. This method and dataset are highly reliable when these vulnerabilities have been verified. However, this dataset includes over 9,000 vulnerabilities, which is relatively small given that they represent only 7 distinct vulnerability types, with fewer than 50 samples per type.

Because of the aforementioned situations, this paper aims to enhance the quantity of vulnerability datasets for smart contracts to mitigate potential associated risks. Additionally, to enhance the quality, the dataset needs to have the capability to identify numerous cases of false reports that current state-of-the-art analysis tools may miss. Specifically, we propose an approach using *semantic-preserving code transformation* to create a new vulnerability template while retaining its original attributes. After that, these samples are inserted into potential locations within the source code to generate a contract containing the vulnerability.

To assess the effectiveness of our proposed approach, we conducted a comprehensive evaluation centered on two key aspects: the quantity of generated vulnerability snippets and their qualitative across various analysis tools. Specifically, we designed and executed experiments to analyze the false negative rates exhibited by prominent tools like Slither, Mythril, and CrossFuzz when evaluated against the new dataset.

In summary, this paper aims to address the two questions:

- **RQ1:** How does code transformation affect tools that detect vulnerabilities? To find out, we must assess the quantity and quality of vulnerable examples within the newly generated dataset.
- **RQ2:** Is the injection method at all potential locations in the previous research approaches more effective than just one location?

II. MOTIVATION

Semantic-preserving code transformation is a technique for modifying the structure of source code while ensuring the program’s functional behavior remains unaltered. For example, refactoring operations that rename variables exemplifies this principle, as program execution remains unaffected despite the change in variable names. Furthermore, leveraging the programming language’s specification, we can exploit alternative syntactic constructs that offer equivalent functionality, such as `for` and `while` loops. By adhering to this concept, it becomes possible to generate numerous variations of a source code base while preserving its underlying logic and intended outcomes. Consequently, a source code harboring vulnerabilities will continue to exhibit those vulnerabilities even after transforming, and conversely, a vulnerability-free source code will remain secure.

```

1 result = DOWN;
2 if(finalPrice > initialPrice){
3     result = UP;
4 }

```

Fig. 1. Simple original solidity source code

```

1 result = DOWN;
2 for(uint i = 0; finalPrice > initialPrice; i++){
3     result = UP;
4     break;
5 }

```

Fig. 2. Modified version of Fig. 1

Two code snippets, illustrated in Fig. 1 and 2, demonstrate the concept of semantic equivalence despite syntactic variations. Although both programs achieve the same outcome, they employ distinct constructs and syntaxes. Fig. 1 utilizes an `if` statement on Line 2 for conditional execution, while Fig. 2 leverages a `for` loop. However, the `for` loop in Fig. 2 iterates only once due to the presence of a `break` statement after the first iteration, effectively mimicking the behavior of an `if` statement. This semantic equivalence is further corroborated by the Control-Flow Graph (CFG) depicted in Fig. 3. Disregarding the `uint i = 0;` statement, which has no meaning to the program’s logic, both code snippets exhibit identical control flow. It’s noteworthy that the `break;` statement serves as a control flow construct, dictating program execution paths, while the `i++` statement within the loop becomes dead code as it’s never executed.

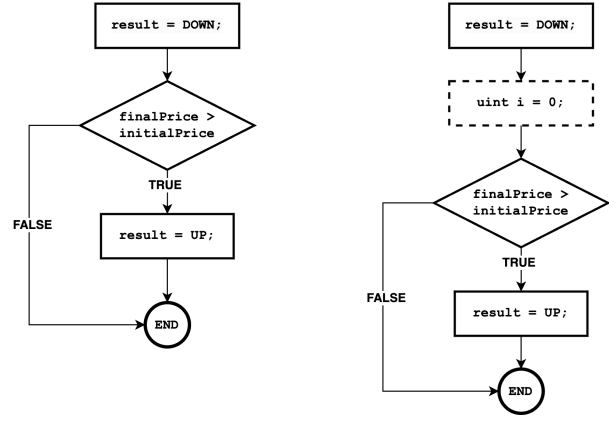


Fig. 3. CFG of two programs: Fig. 1 and Fig. 2

The variations of source code pose a significant challenge for vulnerability detection methods. Due to the existence of numerous code structures that can achieve the same program outcome or mean semantic equivalence, traditional analysis techniques may struggle to identify vulnerabilities across these variations [13]. This variability necessitates addressing two fundamental questions:

- **Capability of existing methods:** Can current vulnerability detection methods effectively transcend syntactic variations and pinpoint vulnerabilities even if they are expressed in a different code structure with equivalent semantics?
- **Generating variations for improvement:** How can we strategically generate additional code variations while preserving the program’s core logic? This ability to create diverse code sets would enable the development of more robust detection methods, equipping them to handle the broader spectrum of code structures that may harbor vulnerabilities that are not detected.

III. THE PROPOSED METHOD

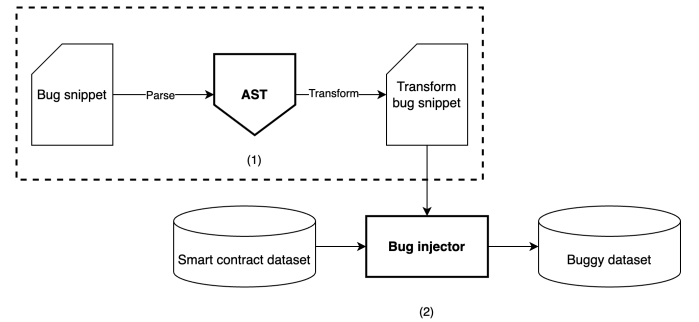


Fig. 4. The proposed method workflow

The process depicted in Fig. 4 consists of two main steps: (1) *transforming* the source code containing the vulnerability and (2) *injecting* the vulnerability snippet into the source code of the smart contract. In the first step, the source code containing the vulnerability is converted into an intermediate

TABLE I
ALL TRANSFORM OPERATORS IN THIS METHOD

Operator	Formula	Describe	Level
Variable renaming	Op_{rename}	Rename the variable declared in the vulnerability snippet to a different variable name.	Naming-level
Function renaming		Rename the function declared in the vulnerability snippet to a different function name.	
Permutation	$Op_{permutation}$ $Op_{subtraction}$ $Op_{division}$ $Op_{unequal}$	Swap the position of operands in a binary expression.	Expression-level
If branch swapping	Op_{if}	Invert the if statement condition and swap the position of both true/false statements.	Statement-level
If statement to for statement transforming	Op_{if2for}	Replace the if statement with a for loop.	
If statement to while statement transforming	$Op_{if2while}$	Replace the if statement with a while loop.	
Variable passing	Op_{tx}	Declare and use a new variable instead <code>tx.origin</code> expression.	

structure, Abstract Syntax Tree (AST), and then this AST is processed by a transform operator while ensuring the original semantics are preserved. The result of this transformation is a new source code snippet, which still contains vulnerabilities nature. Next, in the second step, the source code snippet containing these vulnerabilities is injected into all potential locations in a smart contract source code, to create new source code versions that have a vulnerability. This also ensures that the experimental results are not affected by the injection location. These source code versions are then syntactically checked and compiled testing to ensure validity. Each location where a vulnerability is added without causing a compilation error is considered a valid vulnerability called by *vulnerable location*.

The transform operators were chosen based on similar studies on semantic-preserving code transformation approaches [13, 14] and combined with the Solidity language specification. The operators were classified into three distinct levels based on the scope of their changing: naming-level, expression-level, and statement-level [14]. We provide a total of 7 specific stand-alone transform operators shown in Table I.

A. Naming-level

At this level, two operators are introduced: variable renaming and function renaming. The variable renaming operator systematically identifies and assigns new, valid Solidity identifiers to all declared variables within the code snippet, encompassing both initial declarations and subsequent uses. Importantly, this renaming is confined to the snippet to avoid unintended impact on the main program. Similarly, the function renaming operator identifies and renames declared functions and their calls within the snippet, adhering to Solidity's naming conventions while maintaining a restricted scope to ensure program integrity. These operators pave the way for generating diverse code variations with altered structures but equivalent functionality within the original snippets.

B. Expression-level

Permutation is employed as a code transformation technique to assess analysis tools' capability in comprehending an expression's value. Permutation is used for binary expression operations. This focuses on exchanging the positions of operands in summation, multiplication, equality comparison, and slightly more complex subtraction, division, and unequal, as illustrated by the following equations 1, 2, 3 and 4:

- Summation, multiplication, and equality comparison, this operator only swaps the positions of the two operands:

$$Op_{permutation} = \frac{E_1 \rightarrow E_2 \quad E_2 \rightarrow E_1}{E_1 \text{ op } E_2 \rightarrow E_2 \text{ op } E_1} \quad (1)$$

- For subtraction, in addition to reversing the two operands, add the - operator before the expression:

$$Op_{subtraction} = \frac{E_1 \rightarrow E_2 \quad E_2 \rightarrow E_1}{E_1 \text{ op } E_2 \rightarrow -(E_2 \text{ op } E_1)} \quad (2)$$

- For division, in addition to inverting the two operands, add the numerator 1/ and the inversion expression to the new denominator:

$$Op_{division} = \frac{E_1 \rightarrow E_2 \quad E_2 \rightarrow E_1}{E_1 \text{ op } E_2 \rightarrow 1/(E_2 \text{ op } E_1)} \quad (3)$$

- For unequal comparison operations, add a negation symbol after the permutation:

$$Op_{unequal} = \frac{E_1 \rightarrow E_2 \quad E_2 \rightarrow E_1}{E_1 \text{ op } E_2 \rightarrow \neg(E_2 \text{ op } E_1)} \quad (4)$$

Symbols in this paper: \rightarrow means transformation, E is an expression and S is a statement.

C. Statement-level

This level has four operators: if branch swapping, if statement to for/while statement transforming, and variable passing.

If branch swapping. This operator will reverse the position of both sides of the if statement, and also reverse the value of

the conditional expression. Two transformations are performed simultaneously, and the flow of the if statement will still be preserved. With this operator, the goal will be to check whether the tools understand the negative expression or not and evaluate whether the position of the conditional branch has any effect on the report result. Formula 5 represents this operator.

$$Op_{if} = \frac{E \quad S_T \rightarrow S_F \quad S_F \rightarrow S_T}{if(E)S_T \text{ else } S_F \rightarrow if(\neg E)S_F \text{ else } S_T} \quad (5)$$

If statement to for/while statement transforming. The essence of a loop is to check the condition before or after each iteration. Therefore, if the loop statement is executed only once, this is equivalent to checking the condition only once, similar to an if statement. However, the loop structure is much more complicated than the two branches of an if conditional. The goal of converting an if statement to a for or while loop is to make the control flow of the program more complicated while adding several statements that do not affect the meaning of the program. This can make it difficult for analysis tools not to handle the above loop structures strongly enough, or to understand how many loops there are. This operator only performs transform on conditional statements with only one branch. Two operator are represented by equations 6 and 7.

$$Op_{if2while} = \frac{E \quad S}{if(E) \quad S \rightarrow while(E)\{S; break;\}} \quad (6)$$

$$Op_{if2for} = \frac{E \quad S}{if(E) \quad S \rightarrow for(s_1; E; s_2)\{S; break;\}} \quad (7)$$

Equation 7 s_1 includes s_1 and s_2 as two optional statements. They have no semantic meaning in the program like the example in Fig. 2. At the same time, these statements are more complex and can affect gas consumption. According to the Solidity specification, we can use two empty statements at this location or declare a temporary variable.

Variable passing. This operator targets the `tx.origin` expression within Solidity smart contracts. The `tx.origin` variable holds the address of the original transaction initiator, which can be exploited by malicious contracts for deceptive purposes [15]. This operator will assign a variable using `tx.origin` to avoid using this expression directly. The analysis phase potentially ignores this vulnerability when there is no direct usage statement.

$$Op_{tx} = \frac{S(tx.origin) \rightarrow S'(v)}{S(tx.origin) \rightarrow v = tx.origin; S'(v)} \quad (8)$$

Variable passing operator for `tx.origin` is represented by Formula 8 with v as an added intermediate variable.

```

1 {
2   require (tx.origin == owner_txorigin);
3   receiver.transfer(amount);
4 }
```

Fig. 5. Simple source code use `tx.origin`

```

1 {
2   address tmpVar = tx.origin;
3   require (tmpVar == owner_txorigin);
4   receiver.transfer(amount);
5 }
```

Fig. 6. Simple source code use `tx.origin` after transform

Fig. 6 illustrates the transformed code derived from the original code presented in Fig. 5 that show `tx.origin` is replaced by `tmpVar`. This intermediary variable is subsequently employed in potentially vulnerable function calls, such as `require`.

IV. EXPERIMENTS

A. Experiment Setup

For the purposes of this evaluation, a subset of analysis tools was chosen from the original SolidiFI experiment setup. Specifically, we selected two prominent tools: *Slither*, which leverages static analysis techniques, and *Mythril*, which employs symbolic execution. The decision to exclude other tools was due to their discontinued status. Additionally, a state-of-the-art research approach, *CrossFuzz*, was incorporated to explore the effectiveness of fuzzing techniques. All tools were tested with their default configurations.

The original dataset is 272 Solidity code snippets from 7 vulnerability types: Overflow-underflow, Reentrancy, TOD, Timestamp dependency, Unchecked send, Unhandled exceptions and `tx.origin`.

B. Metric

A comprehensive evaluation of the generated dataset is crucial to ensure its effectiveness in facilitating research on smart contract vulnerability detection. This evaluation should encompass both quantitative and qualitative metrics. Quantitatively, we assess the volume of transformed bug snippets and the number of newly introduced vulnerable locations. Qualitatively, we examine the ratio of false negative reports.

Formular 9 represents the ratio of the number of false negative cases reported by the tool to the total number of bugs found in the dataset D . If this ratio on the generated dataset is higher than the original dataset, this means the newly created vulnerabilities are complex enough to bypass analysis tools. In this formula,

- n_{FN} is the number of false negative cases reported by the tool.
- N_D is the number of injected cases to all smart contracts by dataset D .

$$r_D = \frac{n_{FN}}{N_D} \quad (9)$$

C. Experiment result

To evaluate the quantity, the total number of new code snippets containing vulnerabilities has grown to **674**, representing a **248%** increase from the original 272. Initially, in terms of goals we see a clear improvement in the approach of creating variations using semantic-preserving code transformations.

TABLE II
EXPERIMENT RESULT

Operator	Reentrancy				Timestamp-Dependency			Overflow-Underflow			tx.origin		Unchecked send		TOD	
	CrossFuzz	1047	1148	-	-	930	939	964	1004		Mythril	968	1266	970	1336	1336
	Mythril	1122	1186	-	-	1031	1069	1079	1001		Slither	2	1369	210	1336	1336
	Slither	0	0	-	-	131	2	0	0		Vulnerable location	1336	1369	-	938	938
	Vulnerable location	1343	1346	-	-	1235	1235	1235	1078						938	938
Default																
Op_{rename}																
Permutation																
Op_{if2}																
Op_{if2for}																
$Op_{if2while}$																
Op_{tx}																

Additionally, if we combine the combinatorial transformations, except that Op_{if2for} and $Op_{if2while}$ are not used together, we have a total up to $\sum_{i=1}^7 C_7^i - \sum_{i=0}^5 C_5^i = 95$ operators. With this amount of data, we can use it for machine learning problems, especially new vulnerability detection methods to achieve greater efficiency.

Table II summarizes the experimental outcomes for all operators applied to each analysis tool. The results reveal most transformations significantly elevate the false negative rate r and can up to **100%** on Reentrancy, reported by Slither. Notably, certain vulnerabilities are entirely detectable in the original dataset. For instance, Slither successfully identifies all vulnerable locations associated with Reentrancy and `tx.origin`. However, after applying the transformation operators, Slither exhibits a substantial increase in false negatives. With the new dataset, we can detect more cases that analytics tools cannot identify and improve. An increase in the r ratio appears for all 3 tools in the experiment.

This experimental result shows that having solved the first research question (RQ1), the number of vulnerable increased significantly, and at the same time, they are more difficult to detect by current tools.

Permutation is shown to influence computations, leading to incorrect identification of underflow or overflow computations. Using an intermediate variable in Op_{tx} makes it difficult for tools to detect errors using the `tx.origin` expression. Statement transformations affect vulnerabilities related to multiple statements such as Reentrancy. Meanwhile, renaming does not have much effect, but in some cases, it is possible to inject some additional locations by renaming so that the snippets do not conflict with each other's function names.

We extracted random cases with false negative results and tested stand-alone this case. We consider making sure that there are no mistakes during testing. These cases that have code snippets after transform can bypass the detector of the above tools, while their original code snippet cannot. Notably, we observed that certain code transformations, such as statement reversal and the addition of negation symbols, could render the code undetectable by the tools, while the original code snippet was flagged appropriately. This is a limitation in the detection capabilities of these tools for specific code transform techniques.

V. RELATED WORK

SolidiFI injects vulnerabilities at all potential locations, instead of a single vulnerability [9]. They explain the method can find more corner cases that cannot be easily detected. However, the vulnerabilities injected into their dataset are independent, so they will have the same semantics in different locations and they still retain the characteristic of being vulnerable. In essence, a vulnerability at any location has the same semantics. Nevertheless, this methodology enables us to more precisely identify and assess the limitations of the analyzer.

To verify our findings and answer RQ2, we conducted a small experiment. We try to inject all Reentrancy vulnerabilities transformed by the Op_{if} operator and evaluate the ratio r

of Slither compared to injecting at all potential locations. The number of locations of each snippet is the same.

Table III shows that injecting vulnerability to only one position produces 10% of false negative rate. Meanwhile, injecting to all potential positions reaches 10.6%, presented in Table II. This suggests that Slither produced more false positives in contracts with multiple potential injection locations.

TABLE III
SLITHER RESULT ON TRANSFORMED REENTRANCY DATASET

Vulnerability type	Vulnerable location	False negative
Reentrancy	30	3

Indeed, as Slither reports, there are some cases where the detector of this tool doesn't cover loops or abstract/library syntax¹. Such bugs remain unidentified until user intervention. The lack of high-quality datasets makes it hard for tools to test their issues automatically. In this context, the proposed injection method targeting all potential insertion locations serves a two-fold purpose: facilitating the assessment and accurate evaluation of analysis tools and concurrently uncovering errors within the tools themselves, particularly those that might be easily missed.

VI. CONCLUSIONS

Reliable data is essential for research in software vulnerability detection problems. Unfortunately, the domain of smart contract security faces a significant hurdle due to the scarcity of such data. This paper proposes a novel method specifically designed to generate datasets tailored to smart contract vulnerabilities. This method leverages semantic-preserving code transformation techniques, demonstrably enhancing both the quantity and quality of existing datasets. The results show that the increased amount of data simultaneously increases the false negative rate on analysis tools. This also shows that the detection tools are sensitive to source code variants that have undergone transform operators.

These enriched datasets will pave the way for significant advancements in smart contract vulnerability detection accuracy. From there, it is possible to develop new detection methods such as machine learning and deep learning in this problem. Consequently, the potential for successful attacks on high-value blockchain systems can be substantially reduced.

ACKNOWLEDGMENT

This work has been supported by VNU University of Engineering and Technology under project number CN24.12.

REFERENCES

- [1] Dylan Yaga et al. "Blockchain technology overview". In: *arXiv preprint arXiv:1906.11078* (2019).
- [2] Lee Song Haw Colin et al. "An Integrated Smart Contract Vulnerability Detection Tool Using Multi-layer Perceptron on Real-time Solidity Smart Contracts". In: *IEEE Access* (2024).
- [3] Mohan Dhawan. "Analyzing safety of smart contracts". In: *Proceedings of the Conference: network and distributed system security symposium, San Diego, CA, USA*. 2017, pp. 16–17.
- [4] Zhenguang Liu et al. "Combining graph neural networks with expert knowledge for smart contract vulnerability detection". In: *IEEE Transactions on Knowledge and Data Engineering* 35.2 (2021), pp. 1296–1310.
- [5] Feng Mi et al. "VSCL: automating vulnerability detection in smart contracts with deep learning". In: *2021 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE. 2021, pp. 1–9.
- [6] Vu Trung Kien et al. "A Multimodal Deep Learning Approach for Efficient Vulnerability Detection in Smart Contracts". In: *GLOBECOM 2023-2023 IEEE Global Communications Conference*. IEEE. 2023, pp. 3421–3426.
- [7] Xiaobing Sun et al. "ASSBert: Active and semi-supervised bert for smart contract vulnerability detection". In: *Journal of Information Security and Applications* 73 (2023), p. 103423.
- [8] Zixian Zhen et al. "DA-GNN: A smart contract vulnerability detection method based on Dual Attention Graph Neural Network". In: *Computer Networks* (2024), p. 110238.
- [9] Asem Ghaleb and Karthik Pattabiraman. "How effective are smart contract analysis tools? evaluating smart contract static analysis tools using bug injection". In: *Proceedings of the 29th ACM SIGSOFT International Symposium on Software Testing and Analysis*. 2020, pp. 415–427.
- [10] Hanting Chu et al. "A survey on smart contract vulnerabilities: Data sources, detection and repair". In: *Information and Software Technology* (2023), p. 107221.
- [11] Jian-Wei Liao et al. "Soliaudit: Smart contract vulnerability assessment based on machine learning and fuzz testing". In: *2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*. IEEE. 2019, pp. 458–465.
- [12] Asem Ghaleb. "Towards effective static analysis approaches for security vulnerabilities in smart contracts". In: *Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering*. 2022, pp. 1–5.
- [13] Weiwei Zhang et al. "Challenging machine learning-based clone detectors via semantic-preserving code transformations". In: *IEEE Transactions on Software Engineering* (2023).
- [14] Thanh Le-Cong et al. "Evaluating Program Repair with Semantic-Preserving Transformations: A Naturalness Assessment". In: *arXiv arXiv:2402.11892* (2024).
- [15] Chihiro Kado et al. "An empirical study of impact of solidity compiler updates on vulnerabilities". In: *2023 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. IEEE. 2023, pp. 92–97.

¹<https://github.com/crytic/slither/pull/2419>