

# GasFuzz: Generating High Gas Consumption Inputs to Avoid Out-of-Gas Vulnerability

Fuchen Ma\*, Ying Fu\*, Meng Ren\*, Wanting Sun<sup>†</sup>, Zhe Liu<sup>‡</sup>, Yu Jiang\*, Jun Sun<sup>§</sup> and Jianguang Sun\*

\*Tsinghua University

<sup>‡</sup>Nanjing University of Aeronautics and Astronautics

<sup>§</sup>Singapore Management University

<sup>†</sup>Baidu Online Network Technology (Beijing) Co., Ltd

**Abstract**—The out-of-gas error occurs when smart contract programs are provided with inputs that cause excessive gas consumption, and would be easily exploited to make the DoS attack. Multiple approaches have been proposed to estimate the gas limit of a function in smart contracts to avoid such error. However, under estimation often happens when the contract is complicated. In this work, we propose GasFuzz, which could automatically generate inputs that maximizes the gas cost and reduce the under estimation cases. GasFuzz is designed based on feedback-directed mutational fuzz testing. First, GasFuzz builds the gas weighted control flow graph (CFG) of functions in smart contracts. Then, GasFuzz develops gas consumption guided selection and mutation strategies to generate the input that maximize the gas consumption. For evaluation, we implement GasFuzz based on js-evm, a widely used ethereum virtual machine written in javascript, and conduct experiments on 736 real-world transactions recorded on Ethereum. 44.02% of the transactions would have out-of-gas errors under the estimation results given by solc, means that the recorded real gas consumption for those recorded transactions is larger than the gas limit value estimated by solc. While GasFuzz could reduce the under estimation ratio to 13.86%. Compared with other well-known feedback-directed fuzzing engines such as PerFuzz and SlowFuzz, GasFuzz can generate a same or higher gas estimation value on 97.8% of the recorded transactions with less time, usually within 5 minutes. Furthermore, GasFuzz has exposed 25 previously unknown out-of-gas vulnerabilities in those widely-used smart contracts, 5 of which have been assigned unique CVE identifiers in the US National Vulnerability Database.

## I. INTRODUCTION

In order to avoid excessive consumption of resources, users will be charged with gas to execute smart contracts [22]. Gas cost is determined by the amount of resources consumed by computation and storage. [8][30]. Meanwhile, users may set a gas limit for each transaction to prevent potential malicious gas consumption. But if a transaction needs more gas than the limit set by the user, an out-of-gas error occurs and the transaction would fail. As a result, the transaction would be reverted and the gas spent on executing the transaction is wasted. Malicious users may provide specially crafted inputs to trigger a high-gas cost transactions to launch a Denial-of-Service(DoS) attack [1]. Indeed, the out-of-gas error happens frequently. We investigated transactions made from 2:00 to 3:00 on May 1st, 2019 on Ethereum and found that 21 of them have out-of-gas errors. Gas waste caused by these failed transactions accounts for 5.64% of all gas consumption.

Extensive research has focused on how to estimate the gas limit of smart contract functions to avoid this error. For

example, Solc [28], a widely-used compiler for smart contracts written in solidity language, provides a gas cost estimating tool to give a predication on the gas consumption for each function. It statically analyzes each contract function and provides an estimation based on the opcodes contained in the function. Marescotti et al. [45] present a symbolic execution method to estimate gas cost. The inputs triggering the gas cost could be achieved by solving the equations generated by the constraints for each program path. Although those works solve the problem to some extent, but in practice, their reference value is limited and users have to set the gas limit with manually evaluation. For instance, solc often estimates the gas consumption to be ‘infinite’ or under estimated values. The path conditions of symbolic based estimators are often complex to be solved and most of the results are incorrect.

In fact, finding out how many gas units used in a transaction is hard before the transaction is really executed. There are mainly two challenges. First, the gas consumption of a transaction comes from not only the gas charged by the opcodes, but also the gas used in data storage. It is usually not easy to estimate the data storage with static analysis and solve the complex data structure with symbolic techniques. Furthermore, a function which contains many branches may behave differently under different block states or with different inputs. It is often infeasible to generate some inputs manually which could consume some certain units of gas.

In this paper, we propose GasFuzz, a method that can automatically generate inputs which could lead to a high gas consumption of contract functions and reduce the under estimation ratio of existing works. GasFuzz is designed based on feedback-directed fuzz testing. Specifically, GasFuzz uses three steps to generate inputs: 1) Weighted control flow graph (W-CFG) generation, 2) Feedback-directed selection and mutation, 3) Contract execution. The first step generates a W-CFG for the contract, where each node in the CFG has a weight representing the gas cost of the node. The second step selects the seeds based on the feedback information including the coverage and gas consumption, and defines several mutators for selected seeds. The third step executes the functions with the mutated seeds and calculates the gas consumption of nodes and edges in W-CFG. The key idea behind GasFuzz is that the goal of finding out-of-gas vulnerabilities can be posed as an optimization problem whose goal is to find a set of inputs that maximizes gas consumption of a target smart contract function.

For evaluation, we implemented GasFuzz based on js-evm, one of the widely-used Ethereum Virtual Machine (EVM) for

the execution of smart contracts. We conduct experiments on 1000 real-world transactions recorded on Ethereum, where contracts used by 736 transactions could be compiled successfully. If we set the gas limit according to the limit value estimated by solc, 44.02% of the transactions will have out-of-gas errors, meaning that the recorded real gas consumption value is larger than the estimated value. There are also 33.43% of the transactions will be evaluated to be ‘infinite’ by solc, which may mislead the users to set a higher gas limit and drag the execution process of the transaction. While GasFuzz could reduce the under-estimation ratio to 13.86%, and can also expose meaningful reference value for all the ‘infinite’ excessive gas estimation situations by solc. We also compared GasFuzz with other well-known feedback-directed fuzzing tools such as SlowFuzz [52] and PerFuzz [40]. The results show that GasFuzz outperforms others. In particular, GasFuzz can generate a same or higher gas estimation on 97.8% of transactions with less time, usually within 5 minutes. Besides, GasFuzz found 25 previous unknown gas-related vulnerabilities in widely-used real world contracts which could lead to a Denial-of-Service attack, 5 of which have been assigned unique CVE identifiers in the US National Vulnerability Database and other 20 are appended for approval. In general, this work makes the following contributions:

- 1) We design and implement GasFuzz, a feedback-directed fuzzing approach that could generate inputs triggering a high-gas cost of smart contracts, and reduce the under estimation ratio of existing works.
- 2) We show two practical applications of GasFuzz. First, GasFuzz can be used to provide guidance on setting the gas limit to prevent potential out-of-gas errors. Second, GasFuzz could be used to reveal gas-related vulnerabilities of existing contracts and settings. We have found 25 confirmed vulnerabilities, and 5 of which with CVE IDs.

The rest of the paper is organized as follows. In Section 2, we introduce the background of blockchain and provide a high-level overview of GasFuzz with a motivating example. We formally describe the design of GasFuzz in Section 3. Section 4 demonstrates the evaluation results. Section 5 represents some limitations of GasFuzz. In Section 6, we introduce some other related works. Section 7 makes a conclusion.

## II. OVERVIEW

### A. Background of Blockchain and Gas

Blockchain is a distributed system widely used in all walks of life. This is a new application mode of computer technology including distributed data storage, point-to-point transmission, consensus mechanism, encryption algorithm and so on. Blockchain technology uses consensus mechanism to ensure that transactions are traceable. Meanwhile, the transactions on blockchain system could not be tempered with arbitrarily. In the aspect of privacy protection, blockchain technology uses many encryption algorithms to protect user’s transaction data. Ethereum is a successful practice of blockchain system. In the original Bitcoin system, users could only make system-specified transactions, such as buying Bitcoin. In the Ethereum system, the user’s freedom of action has been greatly improved. Developers can develop their own applications in Ethereum for publication. The applications deployed in the

Ethereum consists of smart contracts. Smart contract is written in Turing-completed language such as solidity. Ethereum also provides a virtual machine to execute the solidity code. The Ethereum Virtual Machine, also known as EVM, could recognize the binary code of the smart contracts as some opcodes, and execute the opcodes one by one.

Gas is a unit used to measure the workload of an operation in Ethereum. If an operation consumes more resources, it consumes more gas. This mechanism ensures that resources are not maliciously wasted. In the actual payment of gas costs, the amount of tokens consumed will be determined by the gas unit cost and the current price of gas. During the actual execution of the transaction, users will set up a maximum gas unit willing to pay for the transaction, which is called the gas limit. If the gas actually consumed by the transaction is larger than the gas limit, the execution of the entire transaction will fail, which is called an out-of-gas exception. Fig 1 represents a work flow of gas mechanism in Ethereum.

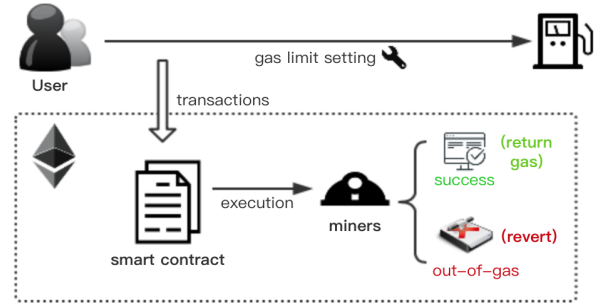


Fig. 1: Gas Mechanism in Ethereum, where miners will keep the gas though the transaction is not completed

It’s not good if an out-of-gas error happens, because the miners will still charge the users for gas though the transaction was not completed. However, it is also bad to set a very high gas limit. Due to the limitation of block gas limit, the miners preferred transactions with a low gas limit. A high gas limit will drag on the speed of completing the transaction. Besides, a high gas limit setting, which allows more opcodes executing, may give some attackers a chance to commit an attack.

### B. A Motivating Example

In this work, given a function in a smart contract, we aim to help users set gas limits more accurately, so as to avoid out-of-gas errors. There are two kinds of gas consumption mechanism in Ethereum, which are known as the execution gas cost and the transaction gas cost. The execution gas cost is based on the cost of computational operations which are executed as a result of the transaction. The transaction gas cost is based on the cost of sending data to the blockchain. There are 4 items which make up the full transaction cost: 1) the base cost of a transaction (which is 21,000 gas units); 2) the cost of a contract deployment (only charged on construct functions, 32,000 gas units); 3) the cost for every zero byte of data or code for a transaction; 4) the cost of every non-zero byte of data or code for a transaction. Given a trace of a smart contract, we can identify its gas cost by accumulating the gas consumption of each step. The problem is thus to identify the most gas-costing

traces. Given that the gas cost is path-based property, as far as we know, existing methods (e.g., fuzzing with the objective of maximizing branch coverage, static analysis or symbolic execution) are not effective on solving the problem. So we try to design trace-based feedback guided fuzzing technique.

In order to understand how our trace based technique works, let us consider a contract deployed on the Ethereum. The following code is intercepted from the contract [29] which is used to distribute tokens. The function we focus on is *distributeFixed*. This function transfers a fixed amount of ethers to each account that needs to be rewarded. Since this contract was deployed on Ethereum, 52 transactions have been made. Among these transactions, 80.77% of them are executed successfully, and 15.38% of the transactions failed because of out-of-gas error. All the out-of-gas transactions called the function *distributeFixed*. The function is shown as below:

```

1  contract DistributeTokens is Ownable{
2      ...
3      function distributeFixed
4      (address[] _addrs, uint _amountToEach) onlyOwner{
5          /* _addrs is an array of address input by the user */
6          for(uint i = 0; i < _addrs.length; ++i){
7              /* transfer some token to each address */
8              tokenReward.transfer(_addrs[i],_amountToEach);
9          }
10     }
11     ...
12 }

```

If we use the gas estimation tool provided by solc to estimate the gas cost of this function, the result is ‘infinite’. The main source of gas consumption in *distributeFixed* is the transfer at line 5. The transfer function is implemented based on the opcode CALL. CALL has a multi-part gas cost: 1) the basic gas cost is 700 units; 2) If the transferring value is non-zero, another 9000 gas units are charged; 3) If the destination account does not yet exist, another 25000 gas units are charged. This is the gas consumption of opcodes. In a real transaction, the memory used to store the transaction data on Ethereum also requires gas. We fetched 3 real transactions based on this function from Etherscan[37]. The inputs of the transactions and the results of gas used are shown in Fig. 2.

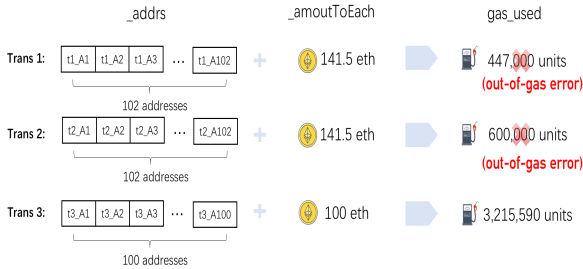


Fig. 2: Inputs for distributeFixed

The first transaction has 102 transfers. The gas limit set by users is 447,000 units. This transaction failed because of an out-of-gas error. The second transaction also has 100 ERC-20 transfers. The inputs of this transaction are the same as the first transaction. The user would like to try a higher gas limit for the transaction but significantly under-approximate the gas

consumption and, as a result, the transaction also failed. The third transaction has 100 ERC-20 transfers from the contract account to 100 other accounts, the amount in the transaction is 100 gwei. The gas cost of the transaction with this set of input is 3,215,590 units. This transaction is successfully executed. The first two transactions cost 1,047,000 units of gas in total. According to the setting of gas price in these two transactions, 31,857,000 gwei is charged by these transactions. **That means nearly six dollars are wasted in vain after these transactions. These out-of-gas errors would not occur if the user knows in advance how many gas will be consumed by this function and what are the inputs to trigger the gas consumption. GasFuzz addresses exactly this concern.**

### C. Gas Fuzzing Idea

Let us look at the original control flow graph of the function *distributeFixed* presented in Fig. 3. The transfer process needs many parameters, the complex sequences of opcodes are simplified in this figure to explain how GasFuzz works. Because of the long trace as shown in the CFG, it is infeasible to use existing symbolic execution based tools to get an accurate value of gas consumption of this function. As mentioned above, static analysis based tools result in an estimation of ‘infinite’. In contrast, GasFuzz applies a feedback-guided fuzzing techniques to identify an accurate gas estimation. The main idea of GasFuzz is to reserve the seeds that could cost more gas units on some of the statements in a given function to trigger a high gas consumption.

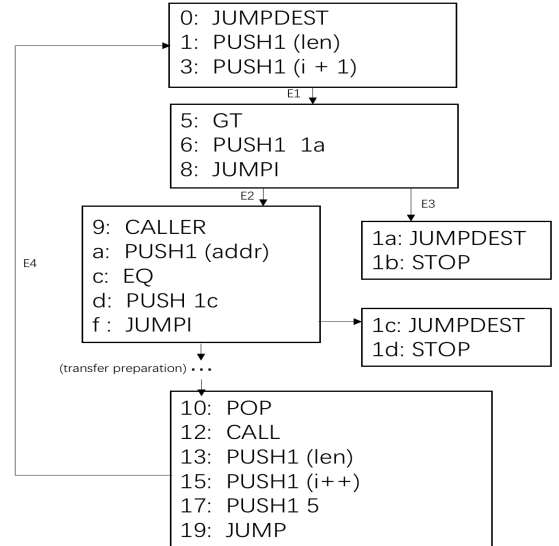


Fig. 3: The Control Flow Graph for function distributeFixed

We assume that the initial inputs of this function is the same as the third transaction in Fig. 2. GasFuzz would set up an initial environment and execute the function with the initial inputs. The initial inputs would execute edge E1, E2 and E4 for 81 times. GasFuzz will record the gas cost on each edge of the CFG as well as the total gas cost of the transaction. The initial inputs are reserved in the seeds pool for further mutation. At each iteration, GasFuzz picks an input from the seeds pool and randomly mutate it with several mutators.

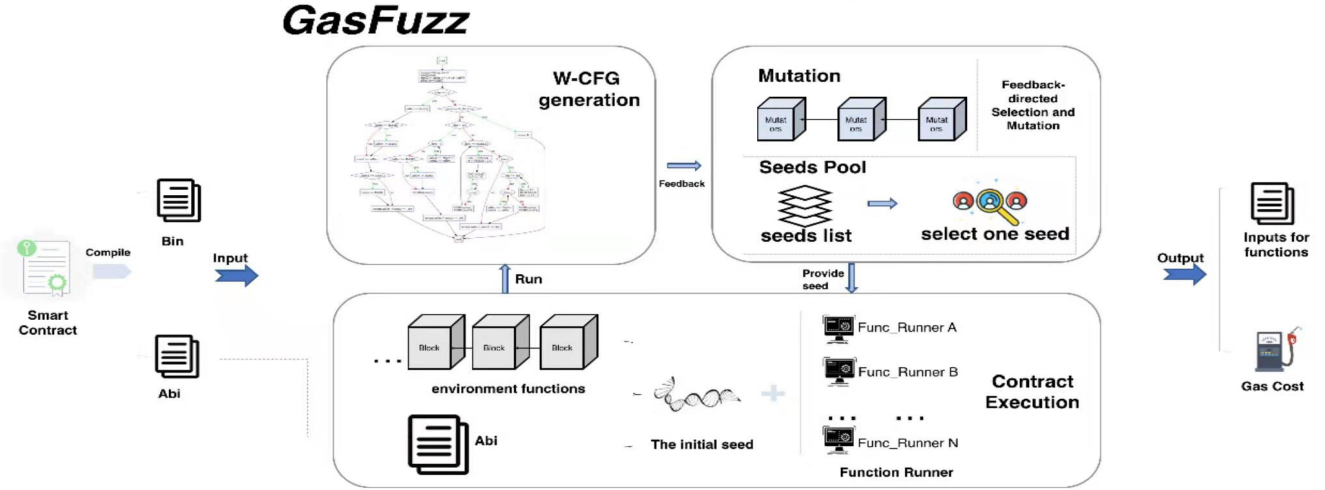


Fig. 4: GasFuzz architecture, which consists of: W-CFG Generation, Mutation and Environment Setup. The inputs of GasFuzz are the smart contract, while the outputs are the value of gas cost and the inputs that could trigger the cost.

As Table I shows, let us assume that the first mutated seed is *Input2*. GasFuzz will use the new input to run the function, the gas cost of this input only executes the edge E1, E2 and E4 for 32 times. The gas costs on these edges are not increased with the new input, and the total gas cost is not increased either, so it will be abandoned by GasFuzz. Let us further assume that the next input is *Input3*. With this input, the edges E1, E2 and E4 are executed for 100 times. The gas cost of these edges are increased, and this input will be reserved in the seeds pool by GasFuzz. The gas cost for each edge and the total gas cost are updated. As for *Input 4*, the edges are executed for 100 times as well, but if the total gas cost is increased, GasFuzz will keep it into the seeds pool as well. All of the reserved inputs will be mutated to generate new inputs. From all of the alternative inputs, GasFuzz will choose several ones which have the most gas cost of the whole function as the basic input to be mutated. In this way, GasFuzz always keeps the inputs that could trigger a local higher gas cost or a global higher gas cost and makes mutations based on these inputs.

TABLE I: Overview of the input mutation for GasFuzz

Input_Num	_addr.length	_amount	edge_exe_times
Input1	81	100 eth	81
<b>Input2</b>	<b>32</b>	<b>231 eth</b>	<b>32</b>
Input3	100	109 eth	32
Input4	100	31 eth	32

In this contract, if the length of the address array is larger, the transaction will make more transfer operations and consume more gas, eventually exceed the gas limit of a block which is 80,039,143 units [27]. GasFuzz generated sets of inputs that trigger a high gas cost which exceeds the gas limit. Malicious attackers may exploit this to implement an attack. The vulnerability has been assigned with a CVE Id CVE-XX-XX (hidden because of the double-blind review mechanism). A

more thorough analysis of GasFuzz's performance on various smart contracts will be presented in Section IV.

### III. GASFUZZ DESIGN

GasFuzz uses three steps to generate inputs to maximize the gas cost: 1) Weighted control flow graph (W-CFG) generation, 2) Feedback-directed mutation and selection, 3) Contract execution. The architecture of GasFuzz is shown in Fig4. GasFuzz starts the fuzzing process with a binary file (.bin) and application binary interface file (.abi) of the contract. GasFuzz extracts the types of parameters and randomly generates some initial inputs for contracts and generates some initial environment variables. Besides, GasFuzz deploys the contract and prepares some function runners for execution. GasFuzz uses the first input to run, and calculates the gas cost information for each edge in W-CFG. After the execution, the seed and feedback information are passed to the feedback-directed selection and mutation part. The feedback information determines whether the seed is saved or not. All the saved seeds are mutated under GasFuzz mutation strategies. After the mutation, GasFuzz selects one seed from the queue and pass it to the function runners. This is a cyclic process and Gasfuzz only stops when the user forces it to stop with time or value thresholds.

#### A. W-CFG Generation

GasFuzz defines a weighted control flow graph which gives a gas consumption weight to each node contained in the traditional CFG. Each node of the CFG is a sequential non-branching block of opcodes. If there is a JUMP opcode or JUMPI opcode, CFG will have a new branch. The weight of the node represents the gas cost of the node. As for the gas cost of some opcodes determined by the operands, a node may have an uncertain gas cost. If the node has an uncertain gas cost, weight of this node is defined as gas consumption



values excluding uncertainties. A simple W-CFG generated by GasFuzz is presented in Fig. 5.

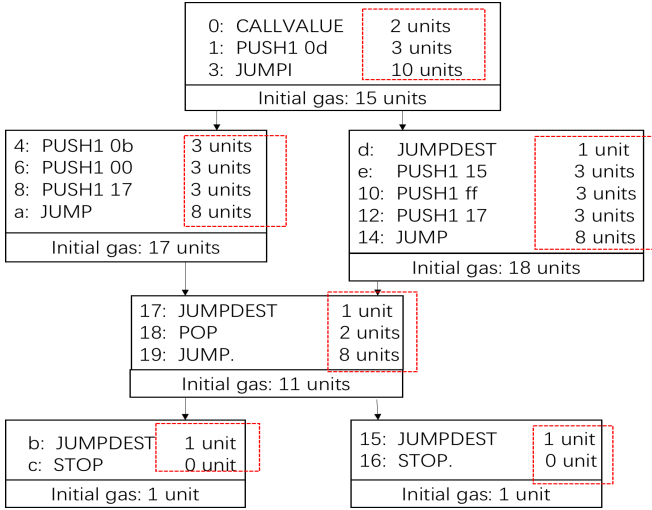


Fig. 5: Example of W-CFG, each block has an initial gas costs. When there is a JUMP or JUMPI opcode, there is a new branch.

As we could see from the W-CFG presented above, each node has some opcodes, and an initial gas consumption value. If the gas consumption of a node is not fixed, the initial gas consumption of the node will be the minimum value of the node. The W-CFG has a new branch when encounter with a JUMP or JUMPI opcode. JUMP opcode takes one parameter from the stack as the destination. JUMPI opcode takes two parameters. The first parameter is the destination and the second one is the condition. As for this example, the destination of the JUMPI in the first node is ‘0d’ which is pushed into the stack in the first PUSH opcode. The condition of this JUMPI is the msg.value which is pushed by CALLVALUE opcode. If msg.value is 0, JUMPI will add the program counter by 1. If msg.value is not 0, JUMPI will jump to address 17. Each edge in the W-CFG has an initial gas consumption of 0. The gas consumption of an edge is defined as the total units of gas cost of the root node of this edge in the execution process.

### B. Feedback-directed selection and mutation

Feedback-directed fuzzing is the kernel component of GasFuzz. It keeps collecting feedback information during test case execution and using the feedback to optimize the fuzzing process, by selecting the right seeds for generating new test cases. Seeds selection determines whether a seed is a good one, and seeds mutation determines the how the selected seeds change for the next execution process. Algorithm 1 outlines the feed-back directed selection and mutation for GasFuzz. GasFuzz will keep running until the duration time set by the user is up or the user would like to stop the process.

**Seeds Selection.** GasFuzz could generate an initial input according to the types of parameters in smart contracts’ functions. The initial inputs are put into the contract functions for the first execution. The feedback information of the execution,

### Algorithm 1: Feedback-directed selection and mutation process

---

**Input :** initial\_inputs

```

1 select_seed = initial_inputs;
2 old_feedBack = function.execute(select_inputs);
3 SeedsPool.push(initial_inputs);
4 repeat
5   foreach strategy of mutation_strategies do
6     newSeed = mutate(original_seed, strategy);
7     newSeed.change_Array();
8     new_seeds.push(newSeed);
9   end
10  foreach seed of new_seeds do
11    feedBack = function.execute(seed);
12    if feedBack.isInteresting(old_feedback) then
13      SeedsPool.push(seed);
14      old_feedback = feedBack;
15    end
16  end
17 until Duration_timeIsUp() || UserStop();
Output: Inputs that lead to a high gas cost or Warnings

```

---

which contains gas cost and the number of times each edge is exercised, and the total gas cost, is stored by GasFuzz. The initial seed is mutated then under several mutation operators. The mutated inputs are put into the function again. If the inputs could cost more gas of the function or hits more times of some edges in W-CFG, the inputs are defined as interesting seeds and would be reserved. The selection process could be represented as the following:

$$\{s|S, total(s) > totalcur\} \cup \{s|S, \exists e : E \text{ cost}(s, e) > costcur(e)\}$$

where  $E$  represents the set of edges in the W-CFG,  $s$  represents seeds, and  $total$  represents the total gas cost of the transaction. All interesting seeds are stored in a priority queue called seeds pool. Seeds that consume more gas are assigned with higher priority. Each iteration, GasFuzz picks a seed from the queue to mutate, and puts the mutated seed into the function again. GasFuzz repeats the above process iteratively.

For the local optimum problem, it occurs when a node does not has an growth on its gas consumption, which hinders the execution of its subsequent nodes which may lead to a high gas consumption. In order to tackle with this problem, GasFuzz uses an sampling algorithm similar to MCMC [56] in seed selector which is implemented in the function *isInteresting* in the algorithm above. In this way, GasFuzz saves those none-interesting seeds with a probability and successfully avoid the local optimal problems.

**Seeds Mutation.** Seeds mutation operators of GasFuzz could be divided in two groups: 1) Traditional mutation. For mutating common inputs of functions, we defined several mutators which are adopted from AFL [33]. Each selected seed is mutated using the mutation operators. The mutators we adopt include bit flipping, byte flipping, arithmetic increment and decrement of integer values, replacing of bytes with

interesting integer values (0, MAX\_INT), etc. 2) Gas-cost-specific mutation. The second group of the mutation operators are used for mutating some special type of variables such as arrays. The length of an array variable may affect the times of the execution of some loops in a function. More loops means more gas consumption somehow. An array variable is replaced by a new array with a different length and different elements. The elements of the array are generated randomly according to the type of the parameters, and the length is randomly selected.

The seed in GasFuzz is represented as a gene chain, which contains all bytes of the function inputs in one contract. GasFuzz defines a `gene_map` to match the inputs for each parameter of each function. Besides, smart contracts need some state information to run. GasFuzz defines state variables as new functions and connects them with the original functions in the contract. The inputs of the functions are state variables. GasFuzz will generate new state variables with the same strategies as other inputs and store them in the gene chain.

### C. Contract Execution

The input of GasFuzz is the abi file and binary file of a smart contract. Construction of an environment and several runners for a contract to run is essential. The environment of GasFuzz is a simulation one which contains the same features as the Ethereum environment. For each contract, the first step is to deploy the contract, that is, to execute the construct function of the contract. The construct process is done by executing the binary code compiled from the contract source code with on the Ethereum virtual machine. After one contract has been deployed, GasFuzz generates a runner for each function. This runner contains the signature of the function and each input for the function. With the help of the runner, GasFuzz will find the matching piece of bytecode from the bin file according to the signature of the function. The signature of a function is a string that only related to the name and the inputs of a function. The number of inputs and their types are extracted by GasFuzz from the abi file, and recorded in the runner of function. Ethereum virtual machine could run a transaction easily with the parameters provided by these drivers. To startup the fuzzing process, an initial seed is also needed. The initial seed is generated by the abi file of the contract. A map is used to record the mapping relation between the formal parameter and the actual parameter stored in the seed.

For detail implementation, GasFuzz is based on js-evm, and relays on several open-source programs such as solidity-types[24], to accomplish the whole task. types to generate a variable for a certain type in smart contracts. In the following, we mainly describe some implementation details for mutation part and environment variables definition part.

The seed is represented as a variable called gene in our implementation. A gene stores all of the inputs of the functions in one contract. Another structure named `gene_map` is used to fetch seeds for each input parameter of a function. For each input of the function defined in the contract, GasFuzz fetches the name of the function, the length of the inputs, the name of the input and the type of the input as the components to generate a key. The key is used to get the seed from the `gene_map`. For each input, we use an open-source tool called 'solidity-types' to generate a random value according to the

type of the parameter. The value is represented as a buffer and connected to the variable gene. The starting position and the ending position of the value in the gene are recorded in the `gene_map`. By this way, we generate random values for the inputs and maintain a structure to effectively get the values. As for the environment variables, we define the following variables in GasFuzz: the coinbase of a block as an address variable; the difficulty of a block as an integer; the block number as an integer; the block timestamp as an integer; the sender as an address; the original address of a transaction named `tx.origin` as an address.

## IV. EVALUATION

In this section, we evaluate GasFuzz to answer the following four questions: **a) Is out-of-gas a real threat? b) Is GasFuzz capable of generating inputs that approximate the gas limit for a given smart contract function? c) What advantages does GasFuzz have compared with other feedback-directed fuzzing tools? d) Is GasFuzz able to be applied in the practical use?**

All our experiments were performed on a machine with 126GB memory, equipped with an Intel(R) Xeon(R) Gold 6148 CPU @ 2.40GHz and running 64-bit Linux operating system. All of the 255 real world contracts used in our experiment are compiled by solc version 0.4.26. We took 1000 recorded transactions randomly from Etherscan[37] for all of the following experiments. We used solc to compile all these contracts and contracts used by 736 transactions could be compiled successfully. The other contracts could not be compiled because of the version of the compiler. It's necessary to mention here that some of the tools referred in related work are not available at present, so we will not compare GasFuzz with them in the following experiments, but with the most widely used solc and reward-directed fuzzing such as SlowFuzz and PerFuzz.

### A. Risks of out-of-gas error

Because of the fact that lots of users of smart contracts use the estimation tool of solc to set the gas limit when they made a transaction, we use the result of solc to investigate the current level of out-of-gas risk. Fig.6 shows a comparison between the gas used of these transactions and the gas estimation results of the tool in solc. We use 'diff' to express this contrast. As shown in the following formula:

$$diff = solc_{est}(function) - real\_cost(trans)$$

where 'diff' represents the difference between the gas estimation of solc and the gas units actually consumed by the recorded transaction. In Fig 6, the X-axis represents the interval where 'diff' is located, and the Y-axis represents the number of transactions in that interval. The red bar represents the case where the 'diff' is negative, and the blue bar represents the case where the 'diff' is positive. We use 0 to represent the 'infinite' estimation given by solc. **As the result shows, if the users use the estimation results of solc to set the gas limit, out-of-gas error is likely to happen. In particular, 44.02% of the transactions will be out-of-gas, and 33.43% of the transactions will be evaluated to be 'infinite' without meaningful guidance.**

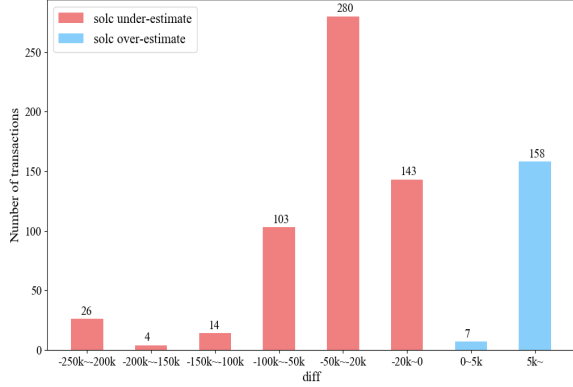


Fig. 6: Risks of out-of-gas error with the results of solc

We chose the most representative 8 functions from the 255 contracts involved in the transactions as illustrative examples to show the detail situations. The 8 functions are with the top most probability of out-of-gas error and are from different contracts. The functions' details and the gas estimation results of them are listed in Table II. The column *Trans* represents the number of transactions that called this function. *solc* column shows the estimation result of solc. Column *real* represents the recorded maximum cost of the executed transactions based on this function.

TABLE II: Chosen functions' details, and their estimation of gas cost by solc and recorded real gas cost.

Contract	Function	Trans	solc	real
BAToken	transfer	1	43517	21916
DSTokenBase	transfer	1	43919	37175
Dogg	transfer	2	45216	52835
AccessoryData	addSERAPHIM	1	41921	49816
ENIGMA	Activate	5	infinite	25689
KittyItemMarket	additem	8	infinite	113844
CommunityChest	send	8	infinite	32523
FutureIExchange	admin_token_withdraw	4	infinite	31484

If users set gas limit according to the results of solc, 93.3% transactions based on the 8 contracts will have out-of-gas problems or infinite limit. The under estimation is mainly because the gas estimation tool in solc uses static analysis to calculate the gas consumption of all opcodes. solc does not consider the gas consumption caused by data storage during the execution of transactions, which results in the estimation bias compared with the real situation. Furthermore, if the function has some complex loops which depends on the inputs for the function, the execution times of each opcode are uncertain. In this case, solc will give an estimate of 'infinite', and users have to set the gas limit with manually analysis, which is not correct. As mentioned in the later experiments, we detect out-of-gas vulnerabilities in 25 contracts out of the 255 tested contracts.

**Answer to RQ-a:** The results of current gas estimation

tools can lead to a large number of out-of-gas occurrences. Out-of-gas error is indeed a potential threat.

### B. GasFuzz Performance

In order to answer the second question, we evaluated GasFuzz with those functions and ran each function with 5 minutes. The results are shown in Figure 7. According to the relationship among the results of GasFuzz and solc and real gas cost, we divide the transactions into three intervals. As shown in Fig. 7, the first interval represents the results of GasFuzz are smaller than solc and real gas cost, the second interval represents the results of GasFuzz are between solc's result and real gas cost, and the third one represents the results of GasFuzz are larger than solc and real gas cost. The red bars in the figure represent that, within this interval, solc's result is smaller than the real gas consumption and would definitely result in out-of-gas error, the blue bar represents that although the solc's result is larger than real consumption, but would still be potentially attacked by the inputs generated by GasFuzz, especially in the third interval.

The figure shows that GasFuzz can exceed the actual gas cost in 634 transactions (493+17+124) of the 225 contracts, counts for 86.14% of all transactions. In these transactions, if user sets gas limit according to results of GasFuzz not the solc, out-of-gas problems might be successfully avoided. As we illustrated in the formal section, the results of solc exceed the real gas consumption on only 22.55% of the transactions. **If GasFuzz is used for gas limit settings, the risk of out-of-gas occurring in these transactions will be greatly reduced, from 44.02% to 13.86%.**

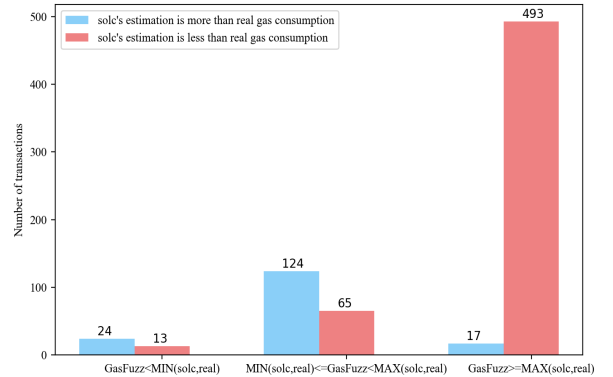


Fig. 7: GasFuzz results compared with results of solc and the real gas consumption in the transaction

We use the 8 illustrative examples in the formal subsection to observe the effectiveness of GasFuzz. The results are shown in Table III. Column GasFuzz represents the results of maximum gas given by GasFuzz in 5 minutes. Column Time shows how long it takes for GasFuzz to find the maximum gas cost. The last two columns show the increase ratio of gas cost given by GasFuzz compared with solc and the real situation.

We could find in the table that, GasFuzz could generate inputs to trigger a gas which is higher than the estimation

TABLE III: GasFuzz results on the 8 functions

Contract	GasFuzz	Time	/solc	/real
BAToken	25879	0.542	-40.5%	18.08%
DSTokenBase	53179	1.949	21.08%	43.05%
Dogg	55072	0.583	21.80%	4.23%
AccessoryData	49601	0.224	18.32%	-0.43%
ENIGMA	26949	8.496	/	4.90%
KittyItemMarket	116324	0.417	/	2.18%
CommunityChest	32884	0.173	/	1.11%
FutureIExchange	32991	0.389	/	4.79%

result of solc by 20% on average. Except for the contract AccessorData, GasFuzz's results are higher than the maximum value of real transactions. For the contract BAToken, GasFuzz gives a lower result than solc, we will explain the reason below. For the contract AccessoryData, the gas consumption is lower than the real value by only 0.43% which means GasFuzz's result is very closed to the real value. If the user sets the gas limit with the estimation result by solc, 7 of the contracts are in dangerous of out-of-gas error. However, with GasFuzz's result, only one of the contract may occur out-of-gas error.

Besides the real out-of-gas situations in these transactions, some of the transactions face potential out-of-gas errors. A potential out-of-gas means the limit of some successfully executing transactions is actually not proper in some other situations. Figure 8 shows the relationship between the results of GasFuzz on the 8 functions and fuzzing time. As the figure shows, GasFuzz could approach or even exceed the real gas used in the transaction, which means some of the functions have potential out-of-gas error. With the inputs generated by GasFuzz, out-of-gas errors may occur. We will use 3 functions as examples to illustrate the three types of inputs that GasFuzz generated within 5 minutes as below.

**GasFuzz exceeds the results of solc and the real gas cost in the transaction.** The first function we focused on is *transfer* taken from the contract Dogg. This function is used to transfer some ether to some other addresses. This function needs two parameters: *to* and *value*. The first is the destination address and the second is the amount of tokens to be transferred. The function is shown as the following:

```

1 function transfer
2 (address _to, uint256 _value) public returns (bool success)
3 {
4     _transfer(msg.sender, _to, _value);
5     return true;
6 }

```

'msg.sender' in function 'transfer' represents the caller of the transaction. This is a state variable used in Ethereum. GasFuzz finds three different inputs for this function. At 0.32 seconds, GasFuzz generates an input that could cost 25603 gas units. At 0.49 seconds, another input is generated cost 25607 units of gas. After 0.58 seconds, GasFuzz finds an input to cost 55072 units of gas. This result is higher than the real gas cost in the transaction based on this function, which is 52835 units. This means there may be potential out-of-gas errors in these transactions when given the inputs generated by GasFuzz.

This situation occurred 28 of the 30 transactions based on eight contracts mentioned above. As for all the transactions we used in our experiment, 69.29% of them have this situation.

In this situation, if the user used solc to set the gas limit of the transaction, there is a chance that out-of-gas error could occur. In this situation, user could just set the gas limit as a value which is a little larger than the result of GasFuzz.

**GasFuzz exceeds results of solc but is lower than the real gas cost in the transaction.** The second function we focused on is 'addSERAPHIM' from the contract 'AccessoryData'. The function is listed as the following. The function is used to add a new seraphim to the system. The 'onlyCREATOR' is a modifier used to limit the scope of the executor of the function. The function first check whether the new seraphim is a new one, then it add the number of the seraphims by one.

```

1 function addSERAPHIM(address _newSeraphim)
2 onlyCREATOR public {
3     if (seraphims[_newSeraphim] == false) {
4         seraphims[_newSeraphim] = true;
5         totalSeraphims += 1;
6     }
7 }

```

GasFuzz gives two gas consumption results for this function. At 0.135 second, GasFuzz find a set of inputs that could cost 48937 units of gas. At 0.224 second, GasFuzz gives another gas consumption result of 49601 units. Solc gives a result of 41921 gas units for this function. However, in the real transaction, 49816 units of gas have been consumed. GasFuzz find a result better than solc but lower than the real situation. The reason that GasFuzz may not find the worst case for gas consumption may be the time for fuzzing. If we ran the tool with more time, maybe it will give us a higher result. However, this result is higher than the one of solc, which means the risks of out-of-gas error are decreased with the help of GasFuzz.

**GasFuzz exceeds the real gas cost but is lower than the estimation result by solc.** The last function we focused on is *transfer* in BAToken.sol. This function is listed as follows.

```

1 function transfer
2 (address _to, uint256 _value) returns (bool success)
3 {
4     if (balances[msg.sender] >= _value && _value > 0)
5     {
6         ...
7         Transfer(msg.sender, _to, _value);
8         return true;
9     }
10    else
11        return false;
12 }

```

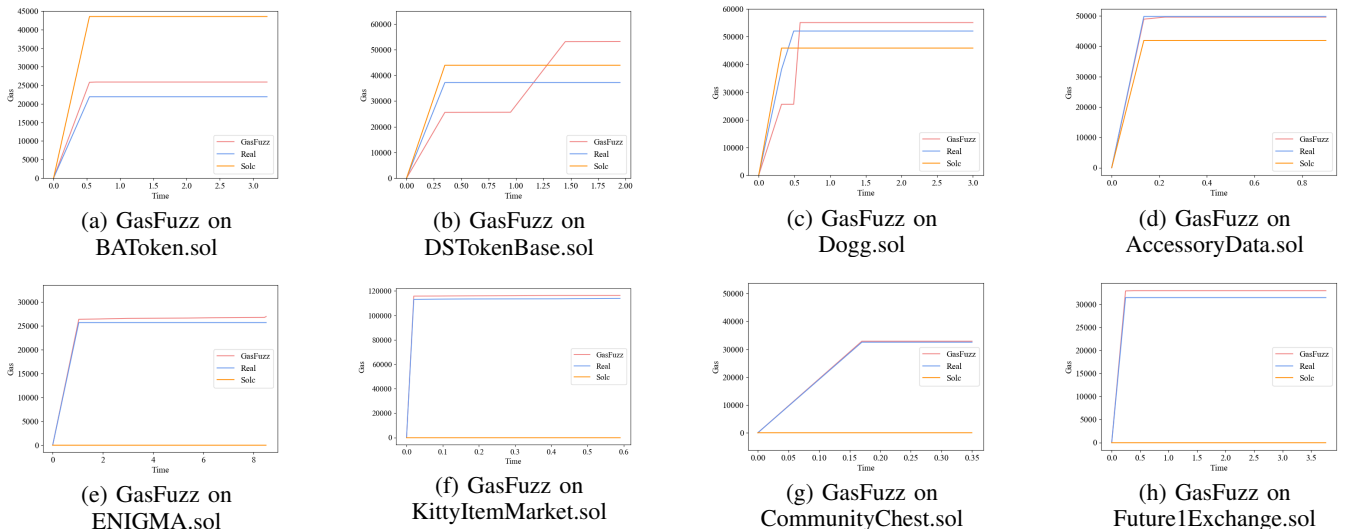
Solc predicts that 43517 gas units would be charged while GasFuzz gives a result of 25879 gas units which is lower than the estimation of solc. To find out the reason for the under estimation, we made a thorough analysis of this result. The following codes are extracted from the opcode sequences of the execution.

```

1 {"pc":1420,"op":85,"gas":"0xfffffaa6c","gasCost":"0x0",
2  "stack":["0xd356a28b","0x1f8","0x0","0x1","0x100000000000000",
3  "0000000000000000000000000000","0x0"],
4  "depth":0,"opName":"SSTORE"}

```





'SSTORE' is an opcode used to store a data in the memory. The opcode takes two parameters from the stack. The first element is the value that needs storing. The second element is the address that the value is stored at. As for this example, the value '0x1000000000000000000000000000000000000000000000000000000000000000' is the address and the value '0x0' is the value stored at this address. As stated in the Ethereum yellow paper[26], the SSTORE opcode will give a gas refund if it changes some non-zero value to a zero value in some address. The refund gas is half the gas used before, which could not exceed 15,000. In our example, the gas used before is 43,517. So the refund gas is 15,000. solc gas estimation did not consider the refunding, which leads to a higher gas estimation. In this situation, GasFuzz gives a more accurate result than solc which is also higher than the real transaction costs. As we mentioned in Section II, fewer gas limit settings mean faster execution. Besides, excessive gas estimates are also dangerous because attackers can exploit the gas loss.

### C. GasFuzz Engine

GasFuzz is a feedback-directed fuzzing method. However, some other feedback-directed fuzzing tools could also be easily used in this problem. We compared GasFuzz with two other fuzzing tools: SlowFuzz [52] and PerfFuzz [40]. SlowFuzz provides a fuzzing method to detect algorithmic complexity vulnerabilities for C/C++ programs. SlowFuzz uses resource-usage-guided evolutionary search techniques to automatically find inputs that maximize computational resource utilization for a given application. Similar to SlowFuzz, PerfFuzz generates inputs via feedback-directed mutational fuzzing. PerfFuzz

uses multi-dimensional feedback and independently maximizes execution counts for all program locations.

1) *Implementations for Gas Used Problem.*: In order to apply the fuzzing tools mentioned above, based on their open-source version, we customize their feedback mechanisms for fuzzing smart contracts. SlowFuzz retains the seeds that could maximize the total path of program execution. While PerfFuzz retains seeds that could lead to more execution of an edge in the program CFG. In our customization, we use the same CFG as GasFuzz. For SlowFuzz, we calculated the execution times for each edge in the CFG during an execution. And the total length of path is represented as the sum of execution times for all edges. As for PerfFuzz, we record the maximum execution times for each edge in the CFG. If a seed could execute some edges more than the times recorded as the maximum value, it will be reserved.

2) *Comparison among Different Fuzzing engines.*: We compared GasFuzz with the fuzzing tools as well as a random fuzzing engine. The random engine save all the seeds randomly without any feedback information. The comparison is committed on the 736 transactions we took. The result of the comparison is shown in TableIV.

TABLE IV: The number of transactions that each fuzzer can maximize the gas estimation compared with other engines

	PerFuzz	RnadamFuzz	SlowFuzz	GasFuzz
total	628	644	623	<b>722</b>
percentage	85.1%	87.3%	84.4%	<b>97.8%</b>

GasFuzz performs the best in 97.8% of the transactions compared with other fuzzing engines, which is the best among all of the engines. RandomFuzz performs better than the other two fuzzing engines with a ratio of 87.3%. The reason for this result may be that the directed strategies used in Perfuzz

TABLE V: Performance of each fuzzer on 8 smart contracts

	PerFuzz			RandomFuzz			SlowFuzz			GasFuzz		
	Result(unit)	Time(s)	GasRate	Result(unit)	Time(s)	GasRate	Result(unit)	Time(s)	GasRate	Result(unit)	Time(s)	GasRate
BAToken	25879	0.712	36347	25879	3.012	8592	25879	2.166	11948	25879	0.542	<b>47747</b>
DSTokenBase	53179	2.412	22048	53179	1.904	27930	53179	0.579	<b>91846</b>	53179	1.949	27285
Dogg	55072	4.457	12356	25667	2.287	11223	55008	2.549	21580	55072	0.583	<b>94463</b>
AccessoryData	49601	0.704	70456	49601	2.315	21426	49601	1.467	33811	49601	0.224	<b>221433</b>
ENIGMA	26949	52.246	516	26693	60.543	441	26565	3.751	<b>7082</b>	26949	8.496	3172
KittyItemMarket	116196	2.522	46073	116196	0.546	212813	116324	1.681	69199	116324	0.417	<b>278954</b>
CommunityChest	32884	0.426	77192	32884	158.368	208	32820	0.651	50415	32884	0.173	<b>190081</b>
Future1Exchange	32991	1.234	26735	32991	1.601	20606	32991	0.767	43013	32991	0.389	<b>84810</b>

and SlowFuzz are not suitable for this problem. Maximizing the execution times for each edge or the total length of the execution path may lead to a highest gas cost brought by the opcodes. However, the gas consumption of a transaction also contains the gas cost brought by data storing. For this reason, the fuzzing engines in PerFuzz and SlowFuzz are not ideal in this problem. In order to observe the difference among the four engines more specifically, we count the fuzz results of the four engines on the eight contracts we mentioned above. The results are shown in TableV. In the table, we introduce a new concept called 'GasRate'. Its definition is shown in the following formula:

$$GasRate = Result/Time$$

This variable is used to measure the effectiveness of the fuzz engine. If this tool can find higher gas consumption in less time, which means a higher GasRate, we think the effectiveness of this engine is better than others. As the table shows, GasFuzz has a highest GasRate on 6 contracts. GasFuzz finds the highest gas cost among four engines on all of the 8 contracts. GasFuzz can find the best results in a relatively short period of time. We now use the contract 'ENIGMA' to show the fuzzing process of these fuzzing engines. We chose the contract 'ENIGMA' because the 4 engines gives many gas consumption results of function 'Activate' in this contract which could clearly show the trend of the results of the 4 engines in the process of fuzzing. The results of 4 fuzzing engines are shown in Fig9.

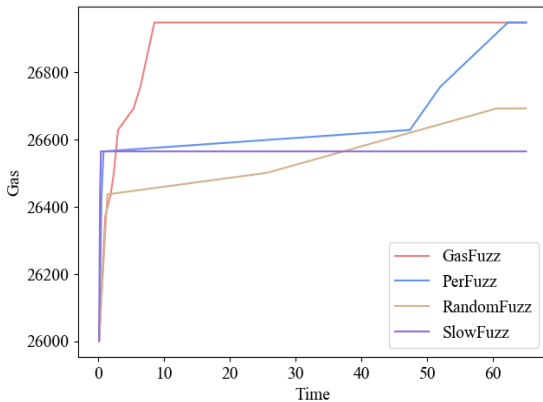


Fig. 9: Trend of results given by 4 fuzzing engines on function 'Activate' from contract 'ENIGMA'

As we could see from the figure, GasFuzz could reach a high result in a short time which is 8.5 seconds to be precisely. PerFuzz gives the same result as GasFuzz at 52.24 seconds. SlowFuzz and RandomFuzz performs worse than the other two engines even with more than 5 minutes.

**Answer to RQ-c:** GasFuzz performs better than some well used feedback-directed fuzzing engines in this problem, where the gas weighted control flow graph and the gas directed seed selection and mutation strategies contribute to the improvements.

#### D. Practical Applications

GasFuzz aims at generating a set of inputs that could trigger a high gas cost. It can be useful for different practical purposes, and we evaluate GasFuzz for two potential applications.

1) *Gas Estimation and Limit Setting:* Choosing a proper gas limit for each transaction is a tough problem. Many works introduced in Section 7 focus on solve this problem in different ways. GasFuzz could also be used to estimate the gas limit of a transaction. To our best knowledge, except for solc, we are the second publicity available tool <sup>1</sup>. The popular solc often generates 'infinite' estimation or under estimation. GasFuzz could give a more accurate result as presented in Table III.

2) *Gas Related Vulnerabilities Detection:* Another application of GasFuzz is to detect gas related vulnerabilities. A transaction with a gas cost that exceeds the gas limit will be reverted. Malicious attackers may exposure a set of inputs to trigger a loop to consume plenty of gas. As a result, the service provided by the contract is stopped. Such vulnerabilities may cause a loss of money. The following code

```

1 function transfer
2 (address[] _address, uint256[] _values) onlyOwner public {
3   require(_address.length == _values.length);
4   for (uint i = 0; i < _address.length; i += 1) {
5     token.transfer(_address[i], _values[i]);
6   }
7   AirDroppedTokens(_address.length);
8 }

```

is fetched from a real-world smart contract. This function is used to transfer some ether to some accounts. Attackers may create a bunch of address to cause a denial of service attack to this contract. The transfer function in the loop may be executed for many times and cost plenty of gas. As a result, the gas

<sup>1</sup><http://github-xxxx-xxx>, hidden for double blind review and would be open after review process

used exceed the gas limit of a block and the transaction would revert.

GasFuzz could generate a set of inputs which identifies this problem. This previously unknown issue has been assigned with a CVE ID. We have already detected 25 confirmed out-of-gas vulnerabilities with the help of GasFuzz, 5 of which have been assigned unique CVE identifiers in the US National Vulnerability Database and other 20 are appended for approval.

**Answer to RQ-d:** GasFuzz has practical values and is really helpful in gas estimation tasks as well as in looking for gas-related vulnerabilities.

## V. THREATS TO VALIDITY

**Mutation and selection strategies limitations.** The mutation strategies of GasFuzz are hybrid for the current version. They are designed based on the mutation strategies in AFL. GasFuzz uses some global functions to represent the environment variables. For mutations, these variables are given different values randomly on the basis of their types. This mutation strategy performed correctly for GasFuzz. However, more efficient mutation strategies for environment variables could be adapted to GasFuzz to build a more complex environment for the contract to run. A more complex environment means more accounts and more state information, it may lead to some different behaviours of smart contracts. Second, even we have implemented MCMC in the seed selection, we will still face the local optimization problem, which needs to be solved with more efficient algorithms or improve the reservation probability for none-interesting seed.

**Smart contracts support limitations.** The current prototype of GasFuzz could not deal with the smart contract functions that call another contract deployed on some address in Ethereum. The gas cost of these functions determined by the called contract, which could not be calculated by GasFuzz for now. This problem may be solved by fetching the called contract and making a relationship between these two contracts. We will try to make GasFuzz adapted to these kinds of contracts in our future work.

**Data collection limitations.** One of the subjective threats is the data collection of the experiment. The experimental process contains a large amount of data, and there are often some errors in the statistics and analysis phases. In order to solve this problem, we use fully automated script files to statistics and analyze the data. To a certain extent, script files could avoid the impact of the data collection process on the experimental results. Another possible subjective threats of this experiment is the time of fuzzing process. The result of fuzzing technique depends on the duration time sometimes. More fuzzing time may give a larger result of gas costs.

## VI. RELATED WORK

**Gas cost estimation.** Solc [28] provides an api to estimate the gas cost of the functions of contracts as showed in Section 5. However, the estimation way of solc has a lot of constraints which lead to many ‘infinite’ estimations and under estimations. GASTAP [2] infers sound gas upper bounds for all public functions in a smart contracts with complex transformation and analysis processes on the code. Matteo Marescotti [45]

have presented a solution to the problem of estimating the gas consumption for Ethereum smart contracts based on techniques inspired by bounded model-checking techniques. These works contribute lots of premiere ideas for further research, but none are publicly available except for Solc.

**Security of Ethereum and Smart contracts.** Security of Ethereum and smart contracts is essential to the development of blockchain technique. Some works make a survey on the attacks of Ethereum and smart contracts [5]. Plenty of research focus on Ethereum and smart contract security. For example, echidna [21] provides a fuzzing framework for developers to check if there is any problem in their contracts. It takes a list of invariants (properties that should always remain true) as input. For each invariant, it generates random sequences of calls to the contract and checks if the invariant holds. ContractFuzzer [38] defines some oracles to detect vulnerabilities in smart contracts such as dangerous delegate bugs. Vandal [11] presents a security analysis framework for Ethereum smart contracts. Another recent work[35] gives the first complete small-step semantics of EVM bytecode. This formalization relies on a combination of hyper and safety properties. MadMax [34] classifies and identifies gas-focused vulnerabilities, and present a static program analysis technique to automatically detect gas-focused vulnerabilities with very high confidence. EVMFuzz [32] uses differential testing to fuzz different versions of EVM and try to find out the vulnerabilities in EVM. Our another work EVM\* provides a framework to make a reinforcement on EVM [43] to test whether there is a dangerous operation during transaction execution and stop the dangerous transaction in time.

**Performance problems detection.** A lot of work has been done to detect the waste of gas and the optimization pattern for gas consumption. For example, a recent work [15] identifies 24 anti-patterns from the execution traces of real smart contracts. Besides, they design and develop GasReducer, the first tool to automatically detect all these anti-patterns from the bytecode of smart contracts and replace them with efficient code through bytecode-to-bytecode optimization. Their another work [14] proposes an emulation-based framework to automatically measure the resource consumptions of EVM operations.

**Fuzzing tools.** Traditional fuzzing techniques aims at finding memory safety problems. Fuzzing testing could be divided as mutation based fuzzing and generation based fuzzing. Mutation based fuzzing such as AFL[33], mutates seeds with mutation strategies for each fuzzing process. Generation based fuzzing generate inputs based on the domain knowledge of the program. Some works focus on the improvement of AFL to increase the fuzzing performance. Enfuzz[16] proposes a method to combine several fuzzers to work together and share the seeds with each other. PafI[41] utilizes efficient guiding information synchronization and task division to extend those existing fuzzing optimizations of single mode to industrial parallel mode. Saffl[55] focuses on the combination of symbolic execution and fuzzing testing tools to increase the coverage. Some other works focus on the techniques to find the weakness of fuzzing techniques, such as Anti-fuzz[25]. Their results show that it is relatively easy to implement and apply anti-fuzzing techniques that are able to completely mask crashes and, by extension, vulnerabilities from fuzzers. However this is hard to implement in smart contracts deployed on Ethereum,

because the contracts on Ethereum could not be changed since it is deployed.

**Main difference.** For the gas estimation tools, GasFuzz focuses on generating inputs to trigger a high gas consumption. Besides, in section V, we demonstrated GasFuzz could perform better than the related available tools. For the traditional fuzzing tools aimed at smart contracts, GasFuzz uses different mutation strategies and detect different vulnerabilities. Echidna and ContractFuzzer randomly mutates the inputs for each generation, while GasFuzz combines the strategies of AFL and the randomly methods for environment variables. Besides, GasFuzz aims at generating inputs trigger a high gas cost but not detecting the memory safety bugs. As for the performance bug detection tools, GasFuzz could automatically find the bottleneck of gas cost for a smart contract, either for gas limit reference or vulnerability detection. The previous works aim at C/C++ programs or artificially define some optimization or known patterns for smart contracts.

## VII. CONCLUSION

In this paper, we designed GasFuzz, a fuzzing approach to generate inputs that could lead to a high gas consumption, which can be used to reduce the under estimation ratio of existing tools and the detect previously unknown out-of-gas vulnerabilities. First, GasFuzz will generate a W-CFG for the function. Second, GasFuzz uses feedback-directed selection and mutation strategies. Finally, GasFuzz provides an environment for contract transaction. The result demonstrates that GasFuzz could successfully generate some inputs to trigger high gas cost in a short time. In 86.14% of the transactions, the gas cost given by GasFuzz is higher than the gas cost of real-world transactions or values estimated by solc, which could successfully decrease the risk of out-of-gas error, from 44.02% to 13.86% of all transactions. For the practical application of GasFuzz, we evaluate GasFuzz on more real-world contracts and detect 25 previously unknown gas-related vulnerabilities, and 5 of which are assigned with CVE Ids.

Our future work will focus on the optimization of GasFuzz. We will design some gas-related mutators, accelerate the fuzzing procedure with some static analysis such as the result of solc, and try to detect more gas-related vulnerabilities in smart contracts with GasFuzz. Besides, we would like to find some gas hot spots with the help of GasFuzz and try to design some patterns to optimize the gas cost of these spots.

## REFERENCES

- [1] Ethereum smart contract best practices. [https://consensys.github.io/smart-contract-best-practices/known\\_attacks/#gas-limit-dos-on-the-network-via-block-stuffing](https://consensys.github.io/smart-contract-best-practices/known_attacks/#gas-limit-dos-on-the-network-via-block-stuffing), 2019.
- [2] Elvira Albert, Pablo Gordillo, Albert Rubio, and Ilya Sergey. Gastap: A gas analyzer for smart contracts. *CoRR*, abs/1811.10403, 2018.
- [3] Maher Alharby and Aad van Moorsel. Blockchain-based smart contracts: A systematic mapping study. *arXiv preprint arXiv:1710.06372*, 2017.
- [4] Sidney Amani, Myriam Bégel, Maksym Bortin, and Mark Staples. Towards verifying ethereum smart contract bytecode in isabelle/hol. In *Proceedings of the 7th ACM SIGPLAN International Conference on Certified Programs and Proofs*, pages 66–77. ACM, 2018.
- [5] Nicola Atzei, Massimo Bartoletti, and Tiziana Cimoli. A survey of attacks on ethereum smart contracts. *IACR Cryptology ePrint Archive*, 2016:1007, 2016.
- [6] Nicola Atzei, Massimo Bartoletti, and Tiziana Cimoli. A survey of attacks on ethereum smart contracts (sok). In *International Conference on Principles of Security and Trust*, pages 164–186. Springer, 2017.
- [7] Alex Biryukov, Dmitry Khovratovich, and Sergei Tikhomirov. Findel: Secure derivative contracts for ethereum. In *International Conference on Financial Cryptography and Data Security*, pages 453–467. Springer, 2017.
- [8] BlockGeeks. Ethereum-gas. <https://blockgeeks.com/guides/ethereum-gas/>, 2019.
- [9] Andreas Bogner, Mathieu Chanson, and Arne Meeuw. A decentralised sharing app running a smart contract on the ethereum blockchain. In *Proceedings of the 6th International Conference on the Internet of Things*, pages 177–178. ACM, 2016.
- [10] Joseph Bonneau. Ethiks: Using ethereum to audit a coniks key transparency log. In *International Conference on Financial Cryptography and Data Security*, pages 95–105. Springer, 2016.
- [11] Lexi Brent, Anton Jurisevic, Michael Kong, Eric Liu, Francois Gauthier, Vincent Gramoli, Ralph Holz, and Bernhard Scholz. Vandal: A scalable security analysis framework for smart contracts. *arXiv preprint arXiv:1809.03981*, 2018.
- [12] Roberto Casado-Vara, Alfonso González-Briones, Javier Prieto, and Juan M Corchado. Smart contract for monitoring and control of logistics activities: Pharmaceutical utilities case study. In *The 13th International Conference on Soft Computing Models in Industrial and Environmental Applications*, pages 509–517. Springer, 2018.
- [13] Ting Chen, Xiaoqi Li, Xiapu Luo, and Xiaosong Zhang. Under-optimized smart contracts devour your money. In *2017 IEEE 24th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pages 442–446. IEEE, 2017.
- [14] Ting Chen, Xiaoqi Li, Ying Wang, Jiachi Chen, Zihao Li, Xiapu Luo, Man Ho Au, and Xiaosong Zhang. An adaptive gas cost mechanism for ethereum to defend against under-priced dos attacks. In *International Conference on Information Security Practice and Experience*, pages 3–24. Springer, 2017.
- [15] Ting Chen, Zihao Li, Hao Zhou, Jiachi Chen, Xiapu Luo, Xiaoqi Li, and Xiaosong Zhang. Towards saving money in using smart contracts. In *2018 IEEE/ACM 40th International Conference on Software Engineering: New Ideas and Emerging Technologies Results (ICSE-NIER)*, pages 81–84. IEEE, 2018.
- [16] Yuanliang Chen, Yu Jiang, Jie Liang, Mingzhe Wang, and Xun Jiao. Enfuzz: From ensemble learning to ensemble fuzzing. *arXiv preprint arXiv:1807.00182*, 2018.
- [17] Konstantinos Christidis and Michael Devetsikiotis. Blockchains and smart contracts for the internet of things. *Ieee Access*, 4:2292–2303, 2016.
- [18] Christopher D Clack, Vikram A Bakshi, and Lee Braine. Smart contract templates: essential requirements and design options. *arXiv preprint arXiv:1612.04496*, 2016.
- [19] ConsenSys. Mythril. <https://github.com/ConsenSys/mythril-classic>, 2019.
- [20] Shaen Corbet, Brian Lucey, and Larisa Yarovaya. Datestamping the bitcoin and ethereum bubbles. *Finance Research Letters*, 26:81–88, 2018.
- [21] crytic. echidna. <https://github.com/crytic/echidna/>, 2019.
- [22] Chris Dannen. *Introducing Ethereum and Solidity*. Springer, 2017.
- [23] Kevin Delmolino, Mitchell Arnett, Ahmed Kosba, Andrew Miller, and Elaine Shi. Step by step towards creating a safe smart contract: Lessons and insights from a cryptocurrency lab. In *International Conference on Financial Cryptography and Data Security*, pages 79–94. Springer, 2016.
- [24] duy tai. solidity-typesp. <https://github.com/duytai/solidity-types>, 2019.
- [25] Emil Edholm and David Göransson. Escaping the fuzz - evaluating fuzzing techniques and fooling them with anti-fuzzing. 2016.
- [26] Ethereum. Ethereum. <https://ethereum.github.io/yellowpaper/paper.pdf>, 2019.
- [27] Ethereum. gas\_limit. <https://ethstats.net/>, 2019.
- [28] Ethereum. Solc. <https://github.com/ethereum/solidity>, 2019.
- [29] etherscan. contract. <https://etherscan.io/address/0x2359ffa8e5b7a7f3d44f4aa9fbc03d061a4b5d0c>, 2019.

- [30] Ethos. What is ethereum gas. <https://www.ethos.io/what-is-ethereum-gas/>, 2019.
- [31] Gianni Fenu, Lodovica Marchesi, Michele Marchesi, and Roberto Tonelli. The ico phenomenon and its relationships with ethereum smart contract environment. In *2018 International Workshop on Blockchain Oriented Software Engineering (IWBOSE)*, pages 26–32. IEEE, 2018.
- [32] Ying Fu, Meng Ren, Fuchen Ma, Yu Jiang, Heyuan Shi, and Jiaguang Sun. Evmfuzz: Differential fuzz testing of ethereum virtual machine. *arXiv preprint arXiv:1903.08483*, 2019.
- [33] google. American fuzzing lop. <http://lcamtuf.coredump.cx/afl/>, 2019.
- [34] Neville Grech, Michael Kong, Anton Jurisevic, Lexi Brent, Bernhard Scholz, and Yannis Smaragdakis. Madmax: Surviving out-of-gas conditions in ethereum smart contracts. *Proceedings of the ACM on Programming Languages*, 2(OOPSLA):116, 2018.
- [35] Ilya Grishchenko, Matteo Maffei, and Clara Schneidewind. A semantic framework for the security analysis of ethereum smart contracts. In *International Conference on Principles of Security and Trust*, pages 243–269. Springer, 2018.
- [36] Everrett Hildenbrandt, Manasvi Saxena, Xiaoran Zhu, Nishant Rodrigues, Philip Daian, Dwight Guth, and Grigore Rosu. Kevm: A complete semantics of the ethereum virtual machine. Technical report, 2017.
- [37] <https://etherscan.io/>. Etherscan. <https://etherscan.io/>, 2019.
- [38] Bo Jiang, Ye Liu, and W. K. Chan. Contractfuzzer: Fuzzing smart contracts for vulnerability detection. 2018.
- [39] Ahmed Kosba, Andrew Miller, Elaine Shi, Zikai Wen, and Charalampos Papamanthou. Hawk: The blockchain model of cryptography and privacy-preserving smart contracts. In *2016 IEEE symposium on security and privacy (SP)*, pages 839–858. IEEE, 2016.
- [40] Caroline Lemieux, Rohan Padhye, Koushik Sen, and Dawn Xiaodong Song. Perffuzz: automatically generating pathological inputs. In *ISSTA*, 2018.
- [41] Jie Liang, Yu Jiang, Yuanliang Chen, Mingzhe Wang, Chijin Zhou, and Jiaguang Sun. Paf: extend fuzzing optimizations of single mode to industrial parallel mode. In *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pages 809–814. ACM, 2018.
- [42] Loi Luu, Duc-Hiep Chu, Hrishi Olickel, Prateek Saxena, and Aquinas Hobor. Making smart contracts smarter. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pages 254–269. ACM, 2016.
- [43] Fuchen Ma, Ying Fu, M Ren, Mingzhe Wang, Yu Jiang, Kaixiang Zhang, Huizhong Li, and Xiang Shi. Evm\*: From offline detection to online reinforcement for ethereum virtual machine. *2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pages 554–558, 2019.
- [44] Daniele Magazzeni, Peter McBurney, and William Nash. Validation and verification of smart contracts: A research agenda. *Computer*, 50(9):50–57, 2017.
- [45] Matteo Marescotti, Martin Blicha, Antti EJ Hyvärinen, Sepideh Asadi, and Natasha Sharygina. Computing exact worst-case gas consumption for smart contracts. In *International Symposium on Leveraging Applications of Formal Methods*, pages 450–465. Springer, 2018.
- [46] Bill Marino and Ari Juels. Setting standards for altering and undoing smart contracts. In *International Symposium on Rules and Rule Markup Languages for the Semantic Web*, pages 151–166. Springer, 2016.
- [47] Hartwig Mayer. Ecdsa security in bitcoin and ethereum: a research survey. *CoinFabrik*, June, 28:126, 2016.
- [48] Esther Mengelkamp, Benedikt Notheisen, Carolin Beer, David Dauer, and Christof Weinhardt. A blockchain-based smart grid: towards sustainable local energy markets. *Computer Science-Research and Development*, 33(1-2):207–214, 2018.
- [49] Ricky KP Mok, Edmond WW Chan, Xiapu Luo, and Rocky KC Chang. Inferring the qoe of http video streaming from user-viewing activities. In *Proceedings of the first ACM SIGCOMM workshop on Measurements up the stack*, pages 31–36. ACM, 2011.
- [50] Michael Mylrea and Sri Nikhil Gupta Gouriseti. Blockchain for smart grid resilience: Exchanging distributed energy at speed, scale and security. In *2017 Resilience Week (RWS)*, pages 18–23. IEEE, 2017.
- [51] Alex Norta. Designing a smart-contract application layer for transacting decentralized autonomous organizations. In *International Conference on Advances in Computing and Data Sciences*, pages 595–604. Springer, 2016.
- [52] Theofilos Petsios, Jason Zhao, Angelos D Keromytis, and Suman Jana. Slowfuzz: Automated domain-independent detection of algorithmic complexity vulnerabilities. 2017.
- [53] Petar Tsankov, Andrei Marian Dan, Dana Drachsler-Cohen, Arthur Gervais, Florian Buenzli, and Martin T. Vechev. Securify: Practical security analysis of smart contracts. In *ACM Conference on Computer and Communications Security*, 2018.
- [54] Dejan Vujčić, Dijana Jagodić, and Siniša Randić. Blockchain technology, bitcoin, and ethereum: A brief overview. In *2018 17th International Symposium INFOTEH-JAHORINA (INFOTEH)*, pages 1–6. IEEE, 2018.
- [55] Mingzhe Wang, Jie Liang, Yuanliang Chen, Yu Jiang, Xun Jiao, Han Liu, Xibin Zhao, and Jiaguang Sun. Saff: increasing and accelerating testing coverage with symbolic execution and guided fuzzing. In *Proceedings of the 40th International Conference on Software Engineering: Companion Proceedings*, pages 61–64. ACM, 2018.
- [56] Wiki. Mcmc. [https://en.wikipedia.org/wiki/Markov\\_chain\\_Monte\\_Carlo](https://en.wikipedia.org/wiki/Markov_chain_Monte_Carlo), 2019.
- [57] Gavin Wood et al. Ethereum: A secure decentralised generalised transaction ledger. *Ethereum project yellow paper*, 151:1–32, 2014.
- [58] Ronghua Xu, Yu Chen, Erik Blasch, and Genshe Chen. Blendcac: A smart contract enabled decentralized capability-based access control mechanism for the iot. *Computers*, 7(3):39, 2018.
- [59] Fan Zhang, Ethan Cecchetti, Kyle Croman, Ari Juels, and Elaine Shi. Town crier: An authenticated data feed for smart contracts. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pages 270–282. ACM, 2016.