

SSR: Safeguarding Staking Rewards by Defining and Detecting Logical Defects in DeFi Staking

Zewei Lin^{†§}, Jiachi Chen[†], Jingwen Zhang^{†§}, Zexu Wang^{†§}, Yuming Feng[§], Weizhe Zhang^{†§}, and Zibin Zheng^{†*}

[†]Sun Yat-sen University, {linzw3, zhangjw273, wangzx97}@mail2.sysu.edu.cn, {chenjch86, zhizibin}@mail.sysu.edu.cn

[‡]Harbin Institute of Technology, wzzhang@hit.edu.cn,

[§]Peng Cheng Laboratory, fengym@pcl.ac.cn

Abstract—Decentralized Finance (DeFi) staking is one of the most prominent applications within the DeFi ecosystem, where DeFi projects enable users to stake tokens on the platform and reward participants with additional tokens. However, logical defects in DeFi staking could enable attackers to claim unwarranted rewards by manipulating reward amounts, repeatedly claiming rewards, or engaging in other malicious actions. To mitigate these threats, we conducted the first study focused on defining and detecting logical defects in DeFi staking. Through the analysis of 64 security incidents and 144 audit reports, we identified six distinct types of logical defects, each accompanied by detailed descriptions and code examples. Building on this empirical research, we developed SSR (Safeguarding Staking Reward), a static analysis tool designed to detect logical defects in DeFi staking contracts. SSR utilizes a large language model (LLM) to extract fundamental information about staking logic and constructs a DeFi staking model. It then identifies logical defects by analyzing the model and the associated semantic features. We constructed a ground truth dataset based on known security incidents and audit reports to evaluate the effectiveness of SSR. The results indicate that SSR achieves an overall precision of 92.31%, a recall of 87.92%, and an F1-score of 88.85%. Additionally, to assess the prevalence of logical defects in real-world smart contracts, we compiled a large-scale dataset of 15,992 DeFi staking contracts. SSR detected that 3,557 (22.24%) of these contracts contained at least one logical defect.

Index Terms—DeFi Staking, Smart Contracts, Logical Defects Detection, LLM, Static Analysis.

I. INTRODUCTION

In recent years, Decentralized Finance (DeFi) has experienced rapid growth, with DeFi staking emerging as one of its most prominent applications. DeFi staking refers to the practice whereby DeFi projects enable users to stake tokens on the platform in order to increase the total value locked (TVL) [1] and reward participants with additional tokens [2]. This mechanism serves two primary purposes: for project developers, it helps increase the project's token TVL, thereby supporting token value and attracting investment; for users, it offers an opportunity to earn additional rewards.

Logical defects in DeFi staking can allow attackers to claim unwarranted rewards. Logical defects arise from errors or incomplete logic in the contract, while non-logical defects result from external environmental factors or improper operations [3], [4]. Logical defects in DeFi staking can be exploited to manipulate staking reward amounts, repeatedly

claim rewards, and perform other malicious actions [5], [6]. To date, logical defects in DeFi staking contracts have led to substantial financial losses. In October 2024, the DeFi staking project OTSea Staking suffered a loss of \$26,000 due to a logical defect that allowed manipulation of the variable used in calculating staking rewards [7].

To the best of our knowledge, no prior research has focused on the logical defects in DeFi staking. Existing studies on DeFi staking primarily focus on its tokenomics, which combines aspects of tokens and economics [8], including token transactions and distribution [9], [10]. Meanwhile, current methods for detecting contract defects are unable to detect logical defects in DeFi staking, as they cannot identify logic related to staking, such as the formulas used to calculate rewards [11], [12]. These logics are defined according to unique requirements, making them difficult to detect using predefined rules. Furthermore, security analyses of DeFi staking incidents often fail to identify the root causes. For instance, while some analyses note manipulated staking rewards, they neglect to examine the code leading to the manipulation [13].

To address this gap, an empirical study on logical defects in DeFi staking is conducted. 64 security incidents and 144 audit reports related to DeFi staking are collected and analyzed. Two researchers manually examined these incidents and reports using the open card sorting method. Based on the root causes identified, six types of logical defects in DeFi staking are defined: *Staking Logical Variables Manipulation (SVM)*, *Rewards without TimeDelay (RT)*, *Single Liquidity Pool Reliance (SLR)*, *Omission in Status Update (OSU)*, *Unsafe Verification (UV)*, and *Unauthorized Staking Asset Access (UAA)*.

Based on the empirical study, we developed a detection tool for logical defects in DeFi staking, named SSR (Safeguarding Staking Reward). Given the source code of a DeFi staking contract, SSR first employs a large language model (LLM) to extract basic staking-related information. It then constructs a comprehensive model of the DeFi staking logic based on the extracted data and the contract's control flow graph, and identifies relevant semantic features. Finally, leveraging the DeFi staking model and the extracted semantic features, SSR detects six types of logical defects using predefined rules.

To evaluate the effectiveness of SSR in detecting logical defects in DeFi staking, we constructed a ground truth dataset consisting of 40 DeFi staking smart contracts, derived from

* Zibin Zheng is the corresponding author.

the security incidents and audit reports used in the empirical study. On this dataset, SSR achieved an overall precision of 92.31%, a recall of 87.92%, and an F1-score of 88.85%. To further assess the prevalence of logical defects in real-world DeFi staking contracts, we developed a large-scale dataset comprising 15,992 open-source DeFi staking smart contracts. Our analysis revealed that 3,557 (22.24%) of these contracts contained at least one logical defect. We assessed SSR's performance on the large-scale dataset through random sampling and manual verification. The experimental results indicated that SSR achieved an overall precision of 89.41%.

The primary contributions of this work are as follows:

- We conducted the first study on logical defects in DeFi staking. By analyzing 64 security incidents and 144 audit reports, we identified and classified six types of logical defects, thereby expanding the existing classification [12]. We provided detailed descriptions and code examples of these defects to inform future security solutions.
- We proposed SSR, a tool for detecting logical defects in DeFi staking. It extracts fundamental information about DeFi staking using an LLM and constructs a DeFi staking model for logical defect detection. On the ground truth dataset, it achieved an overall precision of 92.31%, recall of 87.92%, and an F1-score of 88.85%.
- We collected a large-scale dataset of 15,992 real-world DeFi staking smart contracts to further assess the prevalence of logical defects. SSR revealed that 3,557 (22.24%) of these contracts contain at least one logical defect, achieving an overall precision of 89.41%.
- To support future research, we published the source code of SSR, along with all analysis results and datasets at <https://github.com/Zer0zw/SSR>.

II. BACKGROUND

A. Smart Contract and Decentralized Finance (DeFi)

A smart contract is a program deployed on a blockchain [14] that autonomously executes actions based on predefined conditions, without the need for intermediaries [15], [16]. Due to the inherent characteristics of smart contracts, defects within them can lead to significant economic losses.

Decentralized Finance (DeFi) refers to financial applications built on blockchains, such as Ethereum [17], using smart contracts. Its aim is to replicate the functions of traditional financial institutions, such as banks, exchanges, and lending platforms, in a decentralized manner [18]. Among DeFi applications, staking is one of the most prominent

B. DeFi Staking

DeFi staking refers to the practice in which DeFi projects allow users to stake tokens on the platform to increase the total value locked (TVL) and reward participants with additional tokens [5]. Tokenomics in DeFi staking refers to the design of blockchain-based economic systems, including the token distribution, utility, and rewards that influence user behavior [2].

In DeFi staking smart contracts, three core functionalities are typically implemented: staking, claiming rewards,

and unstaking [19]. Staking involves users depositing their cryptocurrency assets into a platform's staking pool or smart contract. Claiming rewards is the process through which users receive incentives, usually in the form of additional tokens, based on the proportion of their staked assets contributed to the DeFi project. Unstaking enables users to withdraw their deposited assets after completing the required lock-up period.

III. LOGICAL DEFECTS IN DeFi STAKING

To gain deeper insights into the logical defects of DeFi staking and to design effective detection methods, we conducted an empirical study. Based on the findings, we classified these defects into six distinct categories. This section presents the classification criteria and methodology, followed by precise definitions and illustrative examples for each category of logical defect in DeFi staking.

A. Data Collection

1) *DeFi Staking Security Incidents*: To identify the logical defects affecting DeFi staking contracts, we collected and analyzed security incidents involving DeFi staking projects. We searched for relevant security incidents using the keywords "stake" and "staking" on blockchain security platforms, including DeFiHackLab [20], SlowMist [21], Blocksec [22], Peckshield [23], and Certik [24]. As a result, we identified 127 security incidents associated with DeFi staking projects.

2) *Smart Contract Audit Reports*: To gain a more comprehensive understanding of the logical defects in DeFi staking, we also analyzed audit reports to evaluate which defects had been mitigated through audits. We collected audit reports from two sources: the DAppScan [25] dataset and smart contract audit platforms. DAppScan is a large-scale dataset of weaknesses from Decentralized Applications (DApps) [26], accompanied by audit reports from various security teams. We searched for audit reports within the DAppScan dataset using keywords. In addition, we collected audit reports from Cyberscope [27], CertiK [24], and other smart contract auditing platforms. In total, we collected 201 relevant audit reports.

B. Data Analysis

1) *Manually Filtering*: To conduct a more accurate analysis of the logical defects in DeFi staking, data unrelated to these defects is filtered out. Two researchers manually examined and filtered the collected security incidents and smart contract audit reports. For security incidents, we analyzed the root cause of the incident, excluding cases where the targeted contract was not a DeFi staking contract or where the incident's cause was unrelated to logical defects (e.g., private key leakage, Ponzi schemes). For audit reports, we excluded those that did not focus on DeFi staking contracts. After this filtering process, we identified 64 security incidents and 144 audit reports.

2) *Open Card Sorting*: To ensure accurate and comprehensive classification, we employed the open card sorting method [28], which is commonly used in empirical research [29], [30], to categorize the collected incidents and reports. First, cards were created for each incident and report,

each card containing the project name, a description of the defect, and the root cause (i.e., the code snippets that exhibited the defect). Then, two researchers, each with more than three years of experience in smart contract research, manually analyzed and classified these cards. The open card sorting process consisted of two rounds.

In the first round, 40% of the cards were randomly selected and jointly analyzed by both researchers to establish an initial classification. During the analysis, the researchers initially analyzed the defective code snippets guided by the descriptions of the defects. In the case of attack events, they also examined the specific attack processes. Subsequently, the cards were categorized based on the root cause of each defect. For instance, in cases of staking reward manipulation, the researchers investigated the precise reason for the manipulation, such as whether it involved altering a key variable in the reward calculation or relying on a single liquidity pool. This approach emphasized identifying the root cause of each defect, rather than broadly grouping cases under general labels like “staking reward manipulation”. Cards for which the root cause could not be identified were excluded from the classification.

In the second round, the remaining 60% of the cards were analyzed and classified independently by each researcher. The steps for analyzing the cards in this round were identical to those used in the first round. Discrepancies in classification were subsequently discussed collaboratively to finalize the categorization of logical defects in DeFi staking.

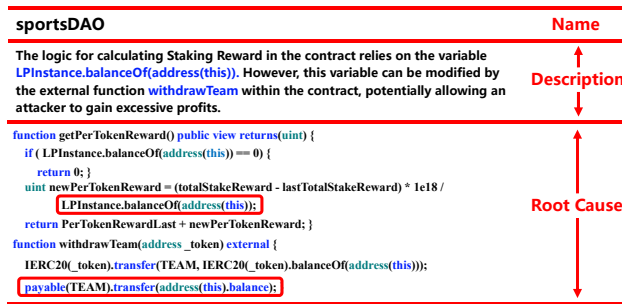


Fig. 1. Example of a card of DeFi staking security incident

Figure 1 illustrates a card example generated from a DeFi staking security incident. In this case, the targeted project is sportsDAO [31], where staking rewards are calculated based on the token balance within the contract. However, the *withdrawTeam* function allows any user to transfer all staking tokens to the contract owner, reducing the token balance to zero. As the token balance decreased significantly, the staking rewards available to the attacker increased accordingly. Based on root cause analysis, this issue is classified as a case of *Staking Logical Variables Manipulation*.

C. Definition of Logical Defects in DeFi Staking

Based on open card sorting, we identified six distinct types of logical defects in decentralized finance (DeFi) staking. These defects may allow attackers to exploit reward calculations, repeatedly claim rewards, or steal others' staked assets and rewards. Table I provides concise definitions of these six

defect types. Below, we provide detailed explanations and code examples for each defect.

It is important to note that this paper focuses exclusively on logical defects in DeFi staking and does not address non-logical defects, such as reentrancy. Furthermore, if a DeFi staking contract contains multiple logical defects, each defect can independently lead to a security vulnerability. While multiple defects may coexist within a single contract, our analysis of 64 security incidents revealed that each incident was ultimately caused by a specific defect.

TABLE I
DEFINITION OF LOGICAL DEFECTS IN DeFi STAKING.

Logical Defect	Definition
<i>Staking Logical Variables Manipulation (SVM)</i>	The values of certain critical variables related to DeFi staking logic can be arbitrarily modified by external functions.
<i>Rewards without Timedelay (RT)</i>	Staking rewards are calculated without considering the staking duration, and no temporal restriction is imposed on claiming rewards.
<i>Single Liquidity Pool Reliance (SLR)</i>	Staking rewards and minted staking tokens depend solely on a single liquidity pool.
<i>Omission in Status Update (OSU)</i>	The relevant staking state is not updated in a timely manner.
<i>Unsafe Verification (UV)</i>	A lack of proper validation checks or insufficient validation conditions when claiming rewards or unstaking tokens.
<i>Unauthorized Staking Asset Access (UAA)</i>	The existence of functions that allow unauthorized access to users' staking asset.

(1) **Staking Logical Variables Manipulation (SVM):** This defect arises because certain critical variables in the DeFi staking system can be arbitrarily modified. Depending on the role of these variables, such modifications can lead to various consequences. For example: manipulating the amount of staking rewards; enabling or disabling staking or reward-claiming functions; and altering fee rates. Such manipulations can be exploited by attackers to unjustly obtain rewards or alter the operational states. It should be noted that this defect does not include normal user staking or unstaking operations.

```

1 mapping(address => uint) public userStakeAmount;
2 uint256 public rewardRate
3 function claimReward() public {
4     uint _reward = userStakeAmount[_msgSender()].mul
      (rewardRate);
5     require(_reward > 0, "Stake Reward is 0");
6     _standardTransfer(address(this), _msgSender(),
      _reward); }
7 function setRewardRate(uint _newRewardRate) public {
8     rewardRate = _newRewardRate; }

```

Fig. 2. An example of Staking Logical Variables Manipulation defect.

Example: Figure 2 presents a simplified example of the *Staking Logical Variables Manipulation* defect. In this example, users claim rewards by invoking the *claimReward* function (lines 3–6), where the reward amount is calculated as the product of the staked token amount and *rewardRate* (line 4). However, the external function *setRewardRate* allows arbitrary modification of *rewardRate* (line 8). An attacker can exploit this defect by adjusting the reward rate before claiming rewards, thereby obtaining unjustified profits.

(2) **Rewards without Timedelay (RT):** This defect arises because the calculation of DeFi staking rewards does not consider the staking duration, and there are no time restrictions

on claiming them. Since the rewards are not tied to the staking duration, an attacker can repeatedly execute the sequence of stake, claim rewards, and unstake within a single block. Each iteration yields a reward, allowing the attacker to unjustly accumulate rewards multiple times.

```
1 mapping(address => uint) public lastClaimTime;
2 function claimReward() public {
3     uint stakeDuration = block.timestamp -
4       lastClaimTime[msg.sender()]
5     uint _reward = stakeAmount[msg.sender()].mul(
6       stakeDuration).mul(1e18).div(totalSupply());
7     lastClaimTime[msg.sender()] = block.timestamp
8     _standardTransfer(address(this), msg.sender(),
9       _reward); }
```

Fig. 3. An example of Rewards with Timedelay.

Example: Figure 3 illustrates a simplified example of reward calculation that considers staking duration. In this case, the staking rewards are calculated based on both the user's stake amount and the duration of the staking period (line 4). By incorporating stake duration into the reward calculation, rewards remain zero in the block where tokens are staked, thereby preventing exploitation through repeated cycles of staking, claiming rewards, and unstaking. Additionally, by updating the timestamp of the last reward claim (line 5), any subsequent claims within the same block yield zero rewards, thereby mitigating exploitation through multiple invocations of the *claimReward* function.

(3) Single Liquidity Pool Reliance (SLR): This defect arises from the calculation of staking rewards or the amount of staking tokens minted based on a single liquidity pool. An attacker can exploit this defect by initiating a flash loan attack to manipulate the token ratio within the liquidity pool and subsequently inflate the calculated rewards.

```
1 IPancakePair private _lp;
2 function rewardTokenPriceInToken() public view
3   returns (uint256) {
4     (uint256 lpReserve0, uint256 lpReserve1, uint256
5       lpTimestamp) = _lp.getReserves();
6     return (_tokenLpIndex == 0 ? lpReserve0 :
7       lpReserve1).mul(1e18).div(_tokenLpIndex == 0
8       ? lpReserve1 : lpReserve0); }
```

Fig. 4. An example of Single Liquidity Pool Reliance defect.

Example: Figure 4 presents a simplified example of the *Single Liquidity Pool Reliance* defect. In the function *rewardTokenPriceInToken* (lines 2–4), the staking reward is calculated based on the token ratio in the liquidity pool, obtained via *_lp.getReserves()* (line 3). Since the rewards rely solely on a single liquidity pool, an attacker can manipulate the token ratio within it and thereby inflate the rewards.

(4) Omission in Status Update (OSU): This defect refers to the failure to promptly update DeFi staking-related status. Specifically, it involves variables such as staking duration and account balance not being updated immediately after actions such as staking, claiming rewards, or unstaking.

Example: Figure 5 presents a simplified example of the *Omission in Status Update* defect. In this case, the *unStake* function enables users to unstake tokens and transfers a specified amount of staked tokens to the user (line 5). However, the

user stake amount has not been updated (line 4). As a result, an attacker can repeatedly invoke the *unStake* function to unjustly receive staked tokens multiple times.

```
1 function unStake(uint _amount) public {
2     uint _userBalances=userStakeAmount[msg.sender];
3     require(_userBalances >= _amount, "User Stake
4       Amount not enough");
5     // userStakeAmount[msg.sender] = _userBalances.
6       sub(_amount);
7     _standardTransfer(address(this), msg.sender,
8       _amount); }
```

Fig. 5. An example of Omission in Status Update defect.

(5) Unsafe Verification (UV): This defect refers to the lack of necessary validation conditions required for claiming rewards or unstaking tokens. For instance, this may occur when the contract fails to verify the staking status before executing an unstake or reward claim operation.

```
1 function claimReward(uint256 stakeIndex) external {
2     require(stakeIndex < stakers[msg.sender].length,
3       "Invalid stake index");
4     // require(stakers[msg.sender][stakeIndex].
5       isUnstake, "Already unstaked");
6     reward = calculateReward(msg.sender, stakeIndex);
7     require(reward > 0, "No rewards to claim");
8     stakeInfo.rewardsClaimed = stakeInfo.
9       rewardsClaimed.add(reward);
10    rewardToken.transfer(msg.sender, reward); }
```

Fig. 6. An example of Unsafe Verification defect.

Example: Figure 6 illustrates a simplified example of the *Unsafe Verification* defect. In this example, users can claim rewards for the staked NFTs using the function *claimReward*. However, the validation to verify whether the NFT has already been unstaked (line 3) is missing. As a result, an attacker could claim rewards even after the NFT has been unstaked.

(6) Unauthorized Staking Asset Access (UAA): This defect refers to the unauthorized transfer of staked assets or rewards belonging to other users. For instance, this can be accomplished by directly modifying users' balances, manipulating accounts to perform staking or unstaking operations, or unauthorizedly transferring staked tokens or rewards from any account. An attacker could exploit these defects to access and steal users' assets without their consent.

```
1 function forceTransfer(address from, address to,
2   uint _amount) public {
3     userStakeAmount[from] = userStakeAmount[from].
4       sub(_amount);
5     userStakeAmount[to] = userStakeAmount[to].add(
6       _amount);
7     _standardTransfer(from, to, _amount); }
```

Fig. 7. An example of Unauthorized Staking Asset Access defect.

Example: Figure 7 illustrates a simplified example of the *Unauthorized Staking Asset Access* defect. In this example, the *forceTransfer* function facilitates the transfer of staking tokens from the *from* account to a specified *to* account (lines 2–4). However, the function lacks appropriate permission checks to verify whether the caller is the *from* account or has the necessary authorization to access its tokens. As a result, an attacker can steal staking tokens from any account.

IV. METHODOLOGY

To aid smart contract developers in identifying logical defects within DeFi staking, we have developed a detection tool called SSR (Safeguarding Staking Rewards).

A. Design Decision

SSR utilizes static analysis to detect logical defects in DeFi staking contracts before their deployment on the blockchain, eliminating the need for transactional data or runtime information. However, the inherent complexity of DeFi staking mechanisms introduces several technical challenges.

The Complexity of DeFi Staking Logic. A thorough understanding of the staking logic requires substantial expertise. This complexity arises from several core functionalities, such as staking, claiming rewards, and unstaking, which distinguish DeFi staking contracts from other contract types. Moreover, the staking logic can vary significantly across different contracts. To address this challenge, SSR utilizes the code comprehension capabilities of LLMs to extract fundamental information related to DeFi staking. The extracted information includes variables such as the staking amount, duration, and functionalities such as staking and claiming rewards.

Identification of Transfer Amount Calculation. The calculation of transferred token amount, such as the amount of staking rewards, is a critical factor in DeFi staking contracts. Accurately identifying its calculation logic presents two main challenges: first, locating the variables related to the transfer amount; second, deriving the corresponding calculation formula. To address these challenges, SSR first identifies the variables that determine the transfer amount from LLM-extracted transfer statements. It then constructs a calculation dependency graph (CDG) based on these variables to identify the state variables involved in the transfer amount calculations.

B. Overview

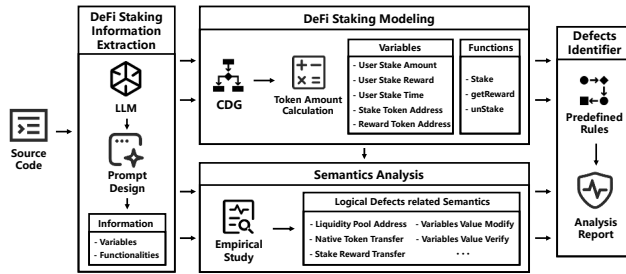


Fig. 8. Overview of SSR

Figure 8 illustrates the overall framework of SSR. Given the source code of a DeFi staking contract, SSR first employs LLMs to extract fundamental information related to the DeFi staking logic (Section IV-C). Based on the extracted information, SSR constructs a DeFi staking model, encompassing relevant variables, functions, and the transfer amount calculation (Section IV-D). Additionally, SSR extracts the semantics related to logical defects in DeFi staking based on the findings of prior empirical studies (Section IV-E). Ultimately, SSR detects six types of logical defects in DeFi staking based on the predefined rules (Section IV-F).

C. DeFi Staking Information Extraction

In this module, SSR leverages the open-source LLM DeepSeek-V3 [32] to extract fundamental information of DeFi staking, including relevant variables and functionalities.

Given the critical role of prompt design in determining the effectiveness of LLMs, we developed specialized prompt sets tailored to extract variables and functionalities relevant to DeFi staking. We adopted the prompt structure recommended by DeepSeek-V3 [33], which comprises two components: the system prompt (SP) and the user prompt (UP). The following presents a detailed discussion of the prompt design process.

System Prompt (SP). The system prompt (SP) defines the role of the LLM as a smart contract auditor. To improve the stability and accuracy of the model's responses, the LLM is instructed to generate multiple responses in the background and select the most frequent one. Additionally, the output format is restricted to *JSON*, excluding any explanatory text, thereby structuring the LLM's output to facilitate subsequent modeling automation. The SP used to extract variables and functionalities related to DeFi staking remains unchanged.

User Prompt (UP). The UP is responsible for guiding the LLM in analyzing and extracting variables and functionalities related to DeFi staking. The corresponding prompt templates are shown in Figures 9 and 10, respectively.

Prompt Template	
<p>System Prompt (SP): You are a smart contract auditor. You will be asked questions related to code semantic. You can mimic answering them in the background five times and provide me with the most frequently appearing answer. Furthermore, please strictly adhere to the output format specified in the question, there is no need to explain your answer.</p> <p>{%CONTRACT_SOURCE_CODE%}</p> <p>User Prompt (UP) – Variables: In this DeFi staking contract, which variable is used to record the { %VARIABLE_DESCRIPTION }? Please answer in a section starts with "{ %VARIABLE_TYPE }".</p> <p>Please answer in the following json format: { " %VARIABLE_TYPE " : ["Variable Name 1", "Variable Name 2"],</p> <p>Note: 1. When outputting the variable name, there is no need to include the contract name. If the obtained variable is a regular variable, please output the variable name as "Variable Name". If the obtained variable is an element of a struct in a contract, please output the variable name as "Struct Name", "Element Name", without including the index. 2. If the contract does not include the corresponding function, leave the "Variable name" and "Description" fields empty in the response. For instance, if the contract does not have a variable for recording the number of tokens staked by the user, the response should be formatted in JSON as follows: "User Stake Amount": [].</p>	
Variable Type	Variable Description
User Stake Amount	number of tokens staked by the user
User Stake Reward	rewards of the user
User Stake Time	user's staking duration
Stake Token Address	contract address of the tokens staked by the user
Reward Token Address	contract address of the tokens received as rewards by the user

Fig. 9. Prompt Template of DeFi Staking-related Variables

For variables, SSR identifies five types closely related to DeFi staking logic: *User Stake Amount*, *User Stake Reward*, *User Stake Time*, *Stake Token Address*, and *Reward Token Address*. SSR first provides a description of each variable's functionality and instructs the LLM to extract the corresponding variable names from the contract. The types and descriptions of these variables are illustrated in Figure 9. To reduce costs, SSR integrates the type names and descriptions of the five distinct variable types into a single prompt, enabling the simultaneous identification of these variables. Additionally, supplementary notes are appended to the end of the UP to further constrain the output format of the LLM, thereby enhancing the automation and accuracy of modeling.

For functionalities, SSR identifies three core operations in the DeFi staking: *Stake*, *getReward*, and *unStake*. For each functionality, SSR first provides a description and guides the LLM in extracting the functions and statements that implement the corresponding logic. The types and descriptions of these functionalities are illustrated in Figure 10. For example, for the *getReward* functionality, the LLM extracts the function that allows users to claim staking rewards, along with the specific statement that transfers the reward tokens to the user. Similar to the approach used for variables, different functionality types are integrated into a single prompt to reduce costs, with the output format explicitly constrained at the end.

Prompt Template		
<p>System Prompt (SP): You are a smart contract auditor. You will be asked questions related to code semantic. You can mimic answering them in the background five times and provide me with the most frequently appearing answer. Furthermore, please strictly adhere to the output format specified in the question, there is no need to explain your answer.</p> <p>{%CONTRACT_SOURCE_CODE%}</p> <p>User Prompt (UP) - Functionalities:</p> <p>In this DeFi Staking Contract, for each function employed to allow users to { %FUNCTION_DESCRIPTION }, which statement is used to { %STATEMENT_DESCRIPTION }? Please answer in a section starts with "{ %FUNCTION_TYPE } Calculation".</p> <p>Please answer in the following json format:</p> <p>{"FUNCTION_TYPE": [{"Function": "Function Name 1", "Node": "Transfer Statement"}], ("Function": "Function Name 2", "Node": "Transfer Statement")},</p> <p>Please ensure that the obtained statement is complete, but do not include the semicolon at the end of the sentence.</p>		
Function Type	Function Description	Statement Description
Stake	stake tokens	transfer equity tokens to users who have staked their tokens
getReward	claim rewards	transfer the staking rewards to the user
unStake	unstake tokens	transfer the tokens from unstaking to the user

Fig. 10. Prompt Template of DeFi Staking-related Functionalities

D. DeFi Staking Modeling

In this module, SSR constructs a DeFi staking model leveraging foundational information extracted by the LLM. This model provides a critical basis for subsequent semantic analysis and defect detection. The following details the model's structure and the methodologies used in its construction.

DeFi Staking Model Structure. The structure of the DeFi staking model consists of three components: DeFi staking-related variables, functions, and calculations of the transferred token amounts. Specifically, the variables include the *User Stake Amount*, *User Stake Reward*, *User Stake Duration*, *Stake Token Address*, and *Reward Token Address*. The functions encompass three core operations in DeFi staking: staking, claiming rewards, and unstaking. The calculations of the transferred token amounts involve determining the number of equity tokens acquired through staking, the stake rewards earned, and the number of tokens received upon unstaking.

The Model Construction Method. SSR employs the smart contract analysis tool Slither [34] to model the DeFi staking contract based on its source code.

In variable analysis, SSR first identifies relevant variables based on their names, which are extracted in the preceding module. It then reconstructs the internal model of a two-layer data structure, which is commonly employed to store user stake information in DeFi contracts. A typical example is *userInfo.stakeAmount*, where *userInfo* is a mapping from user addresses to a struct that contains individual staking data. Different keys within *userInfo* convey distinct logical

meanings, for instance, *userInfo.stakeTime* denotes the timestamp when the user initiated staking. To accurately identify the variable that stores a user's staked amount, it is crucial to determine both the mapping variable, *userInfo*, and the specific key, *stakeAmount*.

In function analysis, SSR first identifies functions related to the three core functionalities by their names, which are extracted in the preceding LLM-based module. It then further refines and filters additional relevant functions by analyzing invocation relationships and evaluating whether they can be externally invoked by users.

In analyzing the calculation of the transferred token amount, SSR first locates the transfer statement and then identifies the variable that represents the transferred amount within it. Subsequently, to identify the variables used for the transferred amount calculation, SSR constructs a calculation dependency graph (CDG). The nodes in the CDG represent the variables that the transferred amount depends on, while the edges capture the dependencies among them.

Algorithm 1 Construction of CDG

Require: Transferred amount variable Var_{amount}

Require: Smart Contract *contract*

Ensure: Calculation Dependency Graph CDG

```

1:  $CDG \leftarrow \emptyset$ 
2:  $queue \leftarrow [Var_{amount}]$ 
3:  $visited \leftarrow visited \cup \{Var_{amount}\}$ 
4: while  $queue$  is not empty do
5:    $Var_{cur} \leftarrow queue.dequeue()$ 
6:    $VarSet_{cal} \leftarrow GetCalculationDepend(Var_{cur}, contract)$ 
7:   for each  $Var_{cal}$  in  $Dep_{cal}$  do
8:     if  $Var_{cal} \notin visited$  then
9:        $queue.enqueue(Var_{cal})$ 
10:       $visited \leftarrow visited \cup \{Var_{cal}\}$ 
11:     end if
12:      $CDG.addCalculationEdge(Var_{cur}, Var_{cal})$ 
13:   end for
14:    $VarSet_{con} \leftarrow GetControlDepend(Var_{cur}, contract)$ 
15:   for each  $Var_{con}$  in  $Dep_{con}$  do
16:     if  $Var_{con} \notin visited$  then
17:        $queue.enqueue(Var_{con})$ 
18:        $visited \leftarrow visited \cup \{Var_{con}\}$ 
19:     end if
20:      $CDG.addControlEdge(Var_{cur}, Var_{con})$ 
21:   end for
22: end while

```

The algorithm for constructing the CDG is presented in Algorithm 1. Starting with a variable of the transferred token amount, the algorithm extracts all other variables on which it depends. This dependency analysis is recursively applied to the newly identified variables until no additional dependencies are found. There are two types of edges E_{cal} and E_{con} , which represent the dependency between variable values in calculations and the control flow relationship of variable modification operations, respectively. These edges are determined by analyzing the variables read by the nodes in the DFG and

CFG associated with the given variable. Once the CDG is constructed, the state variables on which the transferred token amount depends are identified. Variables that do not contribute to detecting defects, such as local variables, are excluded.

```

1 function pendingToken(address account) public view
  returns(uint) {
2   unit _rewardPerToken = getPerTokenReward()
3   unit rewardAmount = userLPStakeAmount[account]
4   * (_rewardPerToken - userRewardPerTokenPaid[
      account]) / (1e18) + (userRewards[
      account]);
5   return rewardAmount }
6 function getReward() public {
7   uint _reward = pendingToken(_msgSender());
8   require(_reward > min, "The stake reward is less
      than the minimum obtainable reward.");
9   userRewards[_msgSender()] = 0;
10  if (_reward > 0) {
11    _standardTransfer(address(this), _msgSender
      (), _reward);
12  }
  return ; }

```

Fig. 11. A concrete example for CDG Construction.

Taking the contract in Figure 11 as a concrete example, when a user calls function `getReward()` (line 6-10) to claim staking rewards, the contract transfers an amount, `_reward`, to the user (line 11). Since `_reward` is calculated using the function `pendingToken()` (lines 1-5), depending on `_rewardPerToken`, `userLPStakeAmount`, and others, the E_{cal} between `_reward` and these variables is extracted. Furthermore, since the transfer is conditioned on the `_reward` exceeding the minimum threshold `min` (line 8), the E_{con} between `_reward` and `min` is extracted. The process then recursively expands to capture additional dependencies, such as `_rewardPerToken`, which is calculated in the `getPerTokenReward()` function (line 2).

Additionally, since the balance of native tokens held by the contract is stored in a special variable, `this.balance`, SSR performs an additional analysis to determine whether the transfer amount depends on this token balance.

Finally, by constructing the DeFi Staking model described above, SSR extracts information related to the DeFi Staking logic. The names and descriptions of the extracted logic are presented in Table II.

TABLE II
EXTRACTED DeFi STAKING LOGIC

Name	Description
<i>Reward(var)</i>	Variable <i>var</i> denotes the DeFi staking reward.
<i>StakeTime(var)</i>	Variable <i>var</i> denotes the duration of time during which the user holds a stake in the asset.
<i>Amount(var)</i>	Variable <i>var</i> denotes the number of tokens that the user has staked.
<i>unStake(func)</i>	Function <i>func</i> is used to transfer the tokens from unstaking to the user.
<i>CalDepend(var₁, var₂)</i>	The value of the variable <i>var₁</i> is calculated based on the value of the variable <i>var₂</i> .
<i>DependonBalance(var)</i>	The value of the variable <i>var</i> is calculated based on the balance of native tokens in the contract.

E. Semantics Analysis

In this module, SSR extracts semantics related to logical defects in DeFi staking based on the DeFi staking model. This includes semantics concerning the meanings of variables

and the behaviors of functions. The names and descriptions of extracted semantics are presented in Table III.

TABLE III
SEMANTICS RELATED TO LOGICAL DEFECTS IN DeFi STAKING

Name	Description
<i>lpPool(con)</i>	Variable <i>con</i> denotes an external liquidity pool contract.
<i>ModifyVar(func, var)</i>	The value of the variable <i>var</i> is modified within the function <i>func</i> .
<i>VerifyVar(func, var)</i>	The execution of certain logic within the function <i>func</i> requires the verification of variable <i>var</i> .
<i>NaTokenTrans(func, to)</i>	In function <i>func</i> , the native token of the contract is transferred to the address <i>to</i> .
<i>RewardTrans(func, from, to)</i>	In function <i>func</i> , the staking reward of the address <i>from</i> is transferred to the address <i>to</i> .
<i>permissCheck(func, addr)</i>	The execution of certain logic within the function requires verifying whether the function caller is the address <i>addr</i> or possesses authorization for <i>addr</i> .

SSR analyzes the semantics of liquidity pool contracts by examining their features, including the presence of functions such as mint and burn, as well as the ability to determine the ratio between two distinct tokens within the pool. These features are inferred through the analysis of widely adopted third-party liquidity pool contracts, including Uniswap [35] and PancakeSwap [36].

ModifyVar and *VerifyVar* indicate whether a function modifies or verifies the value of a specific variable, respectively. For *VerifyVar*, SSR determines the verification of a variable's value by examining specific nodes in the CFG, such as modifiers and require statements. The logic controlled by these nodes depends on verifying the value of the variable. For *ModifyVar*, SSR determines the modification of a variable's value by checking whether any function performs a write operation on the variable. By modeling the two-layer data structure of variables in the *DeFi Staking Modeling* module, SSR can identify functions that modify or verify these variables, as well as the special keys that are modified or verified.

NaTokenTrans and *RewardTrans* represent the transfer of native tokens and staking rewards within the functions, respectively. SSR identifies token transfer based on the naming and parameter conventions defined in the ERC standards [37], [38] and determines whether the transferred token is a native token or a reward token. Specifically, it recognizes native tokens by detecting the special variable `this.balance` and identifies staking rewards using the DeFi staking model.

permissCheck represents the permission check performed within a function, including determining whether the caller is a specific address and verifying whether the caller has the necessary permissions, such as authorization to transfer another user's tokens. SSR first extracts the control-flow nodes that govern a function's execution from the contract's CFG, including special nodes like modifiers and require statements. It then analyzes whether these nodes contain logic that verifies if the function invoker corresponds to a specific address. If such logic is found, SSR considers it a permission check within the function.

F. Defects Identifier

In this module, SSR leverages the DeFi staking model and semantic features to identify six distinct types of logical

defects in DeFi staking. The following provides detailed explanations of the detection methods for each defect type.

(1) Staking Logical Variables Manipulation (SVM) SSR identifies two scenarios that could lead to this defect: dependence on ordinary variables and the native token balance. The detection logic for both scenarios is outlined in Rule 1 and Rule 2, respectively. Rule 1 identifies a defect when staking rewards depend on a variable that arbitrary external accounts can modify. Rule 2 identifies a defect when staking rewards depend on the native token balance, which arbitrary external accounts can transfer. In addition to the scenario involving staking rewards, SSR applies similar logic to detect cases related to tokens deposited during staking and unstaking.

$$isSVM = \text{Reward}(re) \wedge \text{CalDepend}(re, var) \wedge \text{ModifyVar}(func, var) \wedge \neg \text{permissCheck}(func, _) \quad (1)$$

$$isSVM = \text{Reward}(re) \wedge \text{DependonBalance}(re) \wedge \text{NaTokenTrans}(func, _) \wedge \neg \text{permissCheck}(func, _) \quad (2)$$

(2) Rewards without Timedelay (RT) The logic for detecting this defect is defined in Rule 3. A defect is identified when the calculation of staking rewards does not rely on any variable associated with the staking duration ($\neg \text{StakeTime}(var)$). In determining whether a variable is related to staking duration, SSR considers not only relevant variables in the DeFi staking model but also special time-related variables in Solidity [39] smart contracts, such as *block.timestamp*.

$$isRT = \text{Reward}(re) \wedge \neg \text{StakeTime}(var) \wedge \text{CalDepend}(re, var) \quad (3)$$

(3) Single Liquidity Pool Reliance (SLR) The logic for detecting this defect is defined in Rule 4. A defect is identified when the staking rewards depend solely on a single liquidity pool. In Rule 4, $\{con \mid \text{CalDepend}(re, con), \text{lpPool}(con)\}$ represents the set of liquidity pool contracts on which the calculation of staking reward depends.

$$isSLR = \text{Reward}(re) \wedge \#Pool == 1 \wedge Pool = \{con \mid \text{CalDepend}(re, con), \text{lpPool}(con)\} \quad (4)$$

(4) Omission in Status Update (OSU) This defect involves three scenarios: (i) whether the staked amount and duration are updated in the stake function; (ii) whether the staking rewards and duration are updated in the claim reward function; and (iii) whether the staked amount is updated in the unstake function. Taking the status update in the *unStake* function as an example, the detection logic is shown in Rule 5. If the staked tokens amount (*Amount(a)*) is not updated in the *unStake* function (*unStake(func)*), a defect is considered to exist.

$$isOSU = \text{unStake}(func) \wedge \text{Amount}(a) \wedge \text{ModifyVar}(func, a) \quad (5)$$

(5) Unsafe Verification (UV) This defect involves two scenarios: whether the staked amount is verified in the unstake function, and whether the staking reward is verified in the claim reward function. Taking the unstake function as an example, the detection logic is shown as Rule 6. If the amount of staked tokens is not verified ($\neg \text{verifyVar}(func, a)$) before the token transfer is executed, a defect is considered to exist.

$$isUV = \text{unStake}(func) \wedge \text{Amount}(a) \wedge \neg \text{verifyVar}(func, a) \quad (6)$$

(6) Unauthorized Staking Asset Access (UAA) This defect involves two scenarios: transferring the user's staked token or rewards through token transfer, or directly modifying the token balance. Taking staking rewards as an example, the detection logic for both scenarios is in Rule 7 and Rule 8, respectively. Rule 7 states that if a function permits the transfer of rewards from the address *from* to another address, but fails to validate whether the caller is *from* or check for the appropriate permissions, a defect is considered to exist. Rule 8 states that if a function can directly modify the staking rewards without permission checks, a defect is considered to exist.

$$isUAA = \text{RewardTrans}(func, from, _) \wedge \neg \text{perCheck}(func, from) \quad (7)$$

$$isUAA = \text{Reward}(re) \wedge \text{modifyVar}(func, re) \wedge \neg \text{perCheck}(func, from) \quad (8)$$

V. EVALUATION

A. Evaluation Setup

Research Questions. We focus on the following three research questions.

- **RQ1:** How effective is SSR in detecting logical defects in DeFi staking contracts?
- **RQ2:** What is the prevalence of logical defects in real-world DeFi staking contracts?
- **RQ3:** How accurate is SSR in modeling the logic associated with DeFi staking in contracts?

Dataset. To address the research questions outlined above, we constructed two datasets. First, to assess the effectiveness of SSR in detecting logical defects in DeFi staking, we created a ground truth dataset consisting of 40 DeFi staking contracts, collected based on the security incidents and audit reports. Second, to investigate the prevalence of logical defects in real-world DeFi staking contracts, we compiled a large-scale dataset comprising 15,992 such contracts. The details of these two datasets are described below.

For the ground truth dataset, we first collected open-source smart contracts associated with the security incidents and audit reports analyzed in the previous empirical study. We then manually verified whether each contract qualified as a DeFi staking contract and whether it could be successfully compiled. As a result, we obtained 40 DeFi staking contracts, comprising 21 positive cases and 19 negative cases.

TABLE IV
THE DATASET SCALE AT EACH PREPROCESSING STEP

	Original	Keyword Filtered	Deduplicated	LLM Filtered
Scale	1,030,719	137,642	69,444	15,992

For the large-scale dataset, we initially collected 1,030,719 open-source smart contracts from nine blockchains [17], [40]–[47], sourced from a widely used repository [48]. Data preprocessing was carried out in three stages. First, an initial filter was applied using the keyword “stak” to identify potential DeFi staking contracts. Next, duplicates were removed, and those that were not compilable were discarded. In the final stage, the LLM DeepSeek-V3 was used to verify whether these contracts are indeed DeFi staking contracts. As a result, we

obtained 15,992 real-world DeFi staking contracts. The dataset size at each preprocessing step is presented in Table IV.

The experiment was executed on a server running Ubuntu 22.04.1 LTS equipped with 32 Intel(R) Xeon(R) Platinum 8360H CPUs (3.00GHz) and 64 GB of memory.

B. RQ1: Effectiveness of SSR in the Ground Truth Dataset

To answer RQ1, we applied SSR to the ground truth dataset, which consists of 40 DeFi staking contracts. The results are summarized in Table V, which reports the number of each type of logical defect (Incs), along with the counts of true positives (TP), false positives (FP), and false negatives (FN). Precision, recall, and F1-score for identifying each defect were calculated using the following formulas: $\frac{\#TP}{\#TP+\#FP} \times 100\%$, $\frac{\#TP}{\#TP+\#FN} \times 100\%$, and $\frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$. Additionally, we calculated the overall evaluation metrics to assess the effectiveness of SSR. The overall precision is computed as $\frac{\sum_{i=1}^n p_{c_i} \times |c_i|}{\sum_{i=1}^n |c_i|}$, where p_{c_i} denotes the precision for detecting defect type i , and $|c_i|$ is the number of contracts with defect i . The overall recall and F1-score were computed using analogous formulas by substituting the respective metric for p_{c_i} . This evaluation approach is consistent with methodologies employed in prior studies [49], [50].

TABLE V
DETECTION RESULT IN GROUND TRUTH DATASET

Type	# Incs	# TP	# FP	# FN	Prec (%)	Rec (%)	F1 (%)
SVM	1	1	0	0	100.00	100.00	100.00
RT	5	4	0	1	100.00	80.00	88.89
SLR	3	2	0	1	100.00	66.67	80.00
OSU	5	4	0	1	100.00	80.00	88.89
UV	6	6	2	0	75.00	100.00	85.71
UAA	8	7	0	1	100.00	87.50	93.33
Total	/	/	/	/	92.31	87.92	88.85

The experimental results indicate that SSR achieves an overall precision of 92.31%, a recall of 87.92%, and an F1-score of 88.85% on the ground truth dataset.

Since the causes of FPs and FNs are consistent across both the ground truth and large-scale datasets, their analysis is presented collectively in response to RQ2.

C. RQ2: Prevalence of Logical Defects in DeFi Staking

To answer RQ2, we applied SSR to a large-scale dataset consisting of 15,992 unlabeled DeFi staking contracts. The detection results are presented in Table VI. The second and third columns in the table display the number of logical defects detected by SSR for each type, along with the corresponding percentage. The results indicate that 3,557 (22.24%) of the contracts in the dataset contain at least one type of defect.

To assess the effectiveness of SSR in detecting logical defects in DeFi staking in a large-scale dataset, we employed random sampling combined with manual verification. Specifically, to evaluate the precision of SSR, we conducted random sampling at a 95% confidence level with a 10% margin of error for each type of logical defect. This sampling methodology is consistent with prior studies [30], [49], [51]. The sample sizes for each defect type are provided in the fourth column of Table VI. Given the rarity of certain logical defects, we

performed separate random samplings for each defect type, rather than sampling directly from the entire dataset. For instance, contracts identified by SSR as exhibiting the *Single Liquidity Pool Reliance* (SLR) defect account for only 0.28% of the dataset. A random sample drawn from the entire dataset may fail to capture these rare defects, which could hinder an evaluation of SSR's performance for this defect.

TABLE VI
DETECTION RESULT IN LARGE-SCALE DATASET

Type	# Defects	Per (%)	# Sample	# TP	# FP	Prec (%)
SVM	430	2.69	79	68	11	86.08
RT	516	3.23	82	79	3	96.34
SLR	44	0.28	31	29	2	93.55
OSU	889	5.56	87	79	8	86.21
UV	1,315	8.22	90	80	10	88.89
UAA	846	5.29	87	79	8	90.81
Total	3,557	22.24	/	/	/	89.41

We manually verified the sampled datasets to determine the number of true positives (TP) and false positives (FP) detected by SSR for each type of logical defect, along with the corresponding precision (Prec), as presented in Table VI. The formulas used to calculate the precision for each defect type, as well as the overall precision, are consistent with those employed in the analysis of RQ1. The experimental results indicate that SSR achieved an overall precision of 89.41% in detecting logical defects in the large-scale dataset.

TABLE VII
FALSE NEGATIVES OF SSR IN LARGE-SCALE DATASET

Defect Type	# FN	FN rate (%)
Staking Logical Variables Manipulation (SVM)	1	1.04
Rewards without Timedelay (RT)	2	2.08
Single Liquidity Pool Reliance (SLR)	2	2.08
Omission in Status Update (OSU)	3	3.13
Unsafe Verification (UV)	2	2.08
Unauthorized Staking Asset Access (UAA)	4	4.17
Total	11	11.46

In addition, to validate the presence of logical defects that were not identified by SSR, we performed random sampling at a 95% confidence level with a 10% confidence interval on contracts for which SSR produced negative detection results. We randomly sampled and manually verified 96 such contracts in which no logical defects were detected by SSR. The experimental results are presented in Table VII. These results indicate that SSR exhibits a false negative rate of 11.46% on the large-scale dataset.

False Positives (FPs). For RT, OSU, and UV, the primary causes of FPs arise from limitations in the model's ability to capture complex staking logic during the DeFi staking modeling process.

```

1 mapping (address => uint) public stakingTime; //used
  for the unstaking locktime
2 mapping (address => uint) public firstTime; //used
  for the APY boost
3 mapping (address => uint) public lastClaimedTime; //
  used for the claiming locktime
4 mapping (address => uint) public progressiveTime; //
  used for the claiming locktime

```

Fig. 12. An example of Unauthorized Staking Asset Access defect.

For example, as illustrated in Figure 12, several variables related to users' staking times are defined. Although these variables share the same data structure, their semantics and func-

tionalties differ. Specifically, the variable *stakeTime* represents the time threshold at which a user can withdraw their staked assets; *firstTime* logs the timestamp of a user's initial staking activity, which is used to calculate bonuses for the annual percentage yield (APY); *lastClaimedTime* records the most recent time staking rewards were claimed, which is used to calculate the staking reward; and *progressiveTime* is used to compute progressive staking rewards.

SSR struggles to accurately identify the complex logic surrounding staking time in DeFi staking models. As a result, it fails to determine the staking time variables on which staking rewards depend. Furthermore, it cannot distinguish between which time variables need to be updated following a staking reward claim and which do not. In this case, *lastClaimedTime* needs to be updated after a staking reward claim, while *firstTime* does not. However, SSR incorrectly assumes that *firstTime* must also be updated, resulting in a false OSM defect report.

For SVM, SLR, and UAA, the primary source of FPs stems from inaccuracies in pattern-based permission analysis and variable identification. For example, due to incorrect recognition of permission controls, SSR mistakenly classifies functions intended for adjusting project strategies, such as modifying the transaction fee rate, which should be restricted to project developers, as callable by any external account. Because staking rewards depend on the transaction fee rate, this misidentification results in a false SVM defect report.

False Negatives (FNs). For RT, OSU, and UV, the primary cause of FNs lies in the limitations of the LLM, which lead to the incorrect identification of DeFi staking-related variables. For instance, if the LLM mistakenly identifies a variable unrelated to stake time as the stake time, it could result in the omission of an RT defect.

For SLR, SVM, and UAA, the primary cause of FNs stems from errors or omissions in the control flow graph (CFG) and data flow graph (DFG) generated by Slither. For example, inaccuracies in the DFG may hinder the identification of all variables involved in the staking reward calculation. This could result in missing a variable that may be subject to external manipulation, leading to the omission of an SVM defect.

Since the complex staking calculation is the primary cause of FPs and FNs, we performed an analysis of the prevalence of these complex calculation patterns in real-world contracts. A random sample was taken at a 95% confidence level with a 10% confidence interval from the large-scale dataset. The results show that 5.21% (5/96) of the samples contain the complex patterns, which SSR could not identify.

D. RQ3: Accuracy of the DeFi Staking Model

Since the accuracy of the DeFi staking model directly impacts the effectiveness of SSR in detecting logical defects, we further validated the accuracy of the DeFi staking model by manually modeling the contracts in the ground truth dataset and comparing them with the models constructed by SSR.

The accuracy of the DeFi staking model is calculated by the average of three component accuracies: variables, func-

tions, and calculations of transferred token amounts. And the accuracy of each component is determined by comparing the modeled variables and functions with the ground truth lists obtained through manual analysis. The calculation formula is as follows: $M_{total} = \frac{M_{Var} + M_{Func} + M_{Cal}}{3}$, where M_{Var} , M_{Func} , and M_{Cal} , represent the evaluation metrics of modeling accuracy for DeFi staking-related variables, functions, and token amount calculations, respectively. M_{total} denotes the overall modeling accuracy for the DeFi staking contract. These metrics include precision, recall, and F1-score.

Taking the accuracy of the model for variables as an example, five distinct types of variables are identified in the model. The accuracy of each variable type is evaluated individually, and their average is calculated to determine the overall accuracy of the variable modeling. A similar approach is applied to assess the accuracy of the other components.

TABLE VIII
THE ACCURACY OF DEFI STAKING MODELING

	Precision (%)	Recall (%)	F1-Score (%)
Variables	71.75	72.08	71.17
Functions	90.72	91.46	90.38
Calculations	91.25	94.05	91.00
Total	84.57	85.87	84.18

The accuracy of the DeFi Staking model is validated using the ground truth dataset, and the results are summarized in Table VIII. The DeFi Staking Modeling module of SSR achieved an overall precision of 84.57%, recall of 85.87%, and F1-score of 84.18%. These results indicate that, while SSR may encounter difficulties in accurately identifying certain complex DeFi staking logic, it remains effective in modeling the majority of staking logic with high accuracy.

TABLE IX
THE ACCURACY OF DEFI STAKING MODEL USING DIFFERENT LLM

Model	Precision (%)	Recall (%)	F1-Score (%)
DeepSeek-v3	84.57	85.87	84.18
DeepSeek-r1	80.52	79.94	78.53
GPT-4o	77.86	79.20	77.10

To evaluate the impact of various LLMs on DeFi staking modeling, Table IX presents the accuracy of DeFi staking models constructed using several commonly used LLMs. The results show no significant differences among the different LLMs. The primary factor influencing accuracy variations among LLMs is their differing ability to extract variables related to DeFi staking. In contrast, the accuracy of modeling both DeFi staking functions and the calculation of transferred amounts shows minimal variation. Therefore, considering both recognition accuracy and usage costs, SSR selects DeepSeek-V3 for DeFi staking modeling.

VI. THREATS TO VALIDITY

A. External Validity.

The *Defects Identifier* module of SSR is designed to detect six predefined types of logical defects. However, a limitation remains when new defect types are not included in the existing categories. To address this issue, we collected a broad range of real-world security incidents and audit reports related to DeFi staking. These materials were analyzed and categorized

using an open card sorting method to improve the comprehensiveness of the classification. Moreover, the SSR framework is inherently extensible, enabling future research to incorporate additional defect types.

In the evaluation phase, random sampling was used to assess the effectiveness of SSR, which could introduce sample bias. To mitigate this concern, we adopted a confidence-based random sampling method to reduce potential bias. This strategy has been employed in prior studies [30], [52] and offers a more statistically grounded approach to evaluation.

B. Internal Validity.

When constructing the ground truth dataset and labeling the data obtained through random sampling, we employed a manual labeling approach, which could be prone to errors. To mitigate this risk, the labeling was performed by two researchers, followed by a double-checking process. Both researchers were involved in the empirical study and possess significant expertise in logical defects in DeFi staking.

VII. RELATED WORK

A. DeFi Staking Analysis

Existing studies on DeFi staking primarily focus on the tokenomics, examining the trading and distribution of staked tokens. Based on these analyses, researchers can predict the future distribution of tokens and develop strategies.

Lin et al. [10] developed a continuous-time economic model featuring a tokenized digital network and analyzed how DeFi staking platforms manage their lifecycle by adjusting staking incentives. Carré et al. [9] developed a token trading model to examine the efficiency and security implications of liquid staking within the context of DeFi lending. Xiong et al. [53] investigated the opportunities and risks associated with leveraging staking through liquid staking derivatives within the Ethereum ecosystem. These studies focus on the analysis of tokenomics, whereas SSR specifically examines the security defects in DeFi staking smart contracts.

B. Smart Contract Defects Detection

Given the immutable nature of smart contracts, any defects can lead to substantial financial losses once deployed. Consequently, numerous studies have focused on detecting smart contract defects. These detection methods can be classified into three types: static analysis, dynamic analysis, and formal verification [54]. Static analysis methods are widely used in contract analysis, including ONENTE [55], ZEUS [56], Securify [57], Slither [34], Ethainter [58], and Defectchecker [11]. Dynamic analysis approaches, such as fuzzing, are also commonly used, as demonstrated by Contract-Fuzzer [59], Mythril [60], FunFuzz [61], and SmartReco [62]. Other studies, such as KEVM [63], Isabelle [64], and PropertyGPT [65], employ formal verification methods.

As the complexity of smart contracts increases, new defects continue to emerge. Due to the nature of these new defects, traditional detection tools are often unable to identify them. Consequently, numerous studies have focused on defining and

detecting these new defects. For example, NFTGuard [52], CrySol [30], and CDRipper [51] target the detection of defects related to NFTs, cryptographic defects, and centralization defects within smart contracts.

C. Analysis of Logical Defects in Smart Contracts

Since logical defects in smart contracts are closely tied to the specific logic within the contracts, they cannot be detected by traditional contract analysis tools. However, with the recent development of LLMs, the automatic identification of logic in contracts has become possible. As a result, many studies have focused on detecting logical defects within contracts by leveraging this advancement.

GPTScan [12] is the first tool to combine GPT with static analysis for detecting logical vulnerabilities in smart contracts. It utilizes GPT as a versatile code comprehension tool and identifies ten specific types of logical vulnerabilities. S'oley [66] trained a model to detect logical vulnerabilities based on contracts and related code changes from GitHub [67]. Liu et al. [3] collected contracts associated with logical vulnerabilities from 6,165 real audit reports and fine-tuned ChatGPT on this data to improve its ability to detect logical vulnerabilities. Since these studies do not account for the unique logic of DeFi staking, they are unable to detect the logical defects within it. SSR, on the other hand, focuses on the logic of DeFi staking by modeling it specifically and using the resulting model to detect logical defects.

VIII. CONCLUSION

In this paper, we present an empirical study aimed at identifying logical defects in DeFi staking. Through the analysis of 64 security incidents and 144 audit reports, we defined and classified six types of logical defects in DeFi staking. Based on this classification, we introduce SSR, a tool specifically designed to identify logical defects in DeFi staking. SSR first extracts fundamental information about DeFi staking using large language models (LLMs) and constructs the DeFi staking model. It then identifies logical defects by analyzing the model and related semantic features. The evaluation results demonstrate that SSR effectively detects logical defects in DeFi staking, achieving an overall precision of 92.31%, recall of 87.92%, and an F1-score of 88.85% on the ground truth dataset. Furthermore, SSR revealed that 3,557 (22.24%) of 15,992 real-world DeFi staking contracts contain at least one logical defect, highlighting the widespread occurrence.

ACKNOWLEDGMENT

This work was supported in part by the National Key Research and Development Program of China under Grant 2023YFB2704700, the National Natural Science Foundation of China (No. 62032025, 62302534), the Guangdong Basic and Applied Basic Research Foundation (Grant No. 2025A1515011632), the Major Key Project of Peng Cheng Laboratory under Grant PCL2025AS07.

REFERENCES

- [1] “Total value locked (tv) in staking protocols,” 2025. [Online]. Available: <https://www.stelareum.io/en/defi-tvl/staking.html>
- [2] “Tokenomics in defi staking,” 2023. [Online]. Available: <https://www.certik.com/zh-CN/resources/blog/tokenomics-in-defi-staking>
- [3] Q. Liu, M. Wu, J. Chen, T. Chen, X. Chen, R. Jiang, Y. Yang, Z. Lin, and Y. Cheng, “Exploring the potential of chatgpt in detecting logical vulnerabilities in smart contracts,” *Blockchain: Research and Applications*, p. 100294, 2025.
- [4] Z. Liao, S. Song, H. Zhu, X. Luo, Z. He, R. Jiang, T. Chen, J. Chen, T. Zhang, and X. Zhang, “Large-scale empirical study of inline assembly on 7.6 million ethereum smart contracts,” *IEEE Transactions on Software Engineering*, vol. 49, no. 2, pp. 777–801, 2022.
- [5] “2022 year in review - staking,” 2023. [Online]. Available: <https://www.certik.com/zh-CN/resources/blog/2022-year-in-review-staking>
- [6] “Top defi lending and staking protocol vulnerabilities and mitigations,” 2024. [Online]. Available: <https://metana.io/blog/top-defi-lending-and-staking-protocol-vulnerabilities-and-mitigations/>
- [7] “Blocksec: Attack transaction for otseastaking,” 2024. [Online]. Available: <https://app.blocksec.com/explorer/tx/eth/0x90b4fcf583444d44ef8625c6f253cfcb786d2f4eda7198bdab67a54108cd5f4>
- [8] “Tokenomics - wikipedia,” 2025. [Online]. Available: <https://en.wikipedia.org/wiki/Tokenomics>
- [9] S. Carré and F. Gabriel, “Liquid staking: When does it help?” *Available at SSRN*, 2024.
- [10] L. W. Cong, Z. He, and K. Tang, “The tokenomics of staking,” National Bureau of Economic Research, Tech. Rep., 2025.
- [11] J. Chen, X. Xia, D. Lo, J. Grundy, X. Luo, and T. Chen, “Defectchecker: Automated smart contract defect detection by analyzing evm bytecode,” *IEEE Transactions on Software Engineering*, vol. 48, no. 7, pp. 2189–2207, 2021.
- [12] Y. Sun, D. Wu, Y. Xue, H. Liu, H. Wang, Z. Xu, X. Xie, and Y. Liu, “Gptscan: Detecting logic vulnerabilities in smart contracts by combining gpt with program analysis,” in *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering*, 2024, pp. 1–13.
- [13] “Autoshark lightning loan attack: Price flash crash 99
- [14] S. Nakamoto, “Bitcoin whitepaper,” URL: <https://bitcoin.org/bitcoin.pdf> (-: 17.07. 2019), vol. 9, p. 15, 2008.
- [15] Z. Zheng, S. Xie, H.-N. Dai, W. Chen, X. Chen, J. Weng, and M. Imran, “An overview on smart contracts: Challenges, advances and platforms,” *Future Generation Computer Systems*, vol. 105, pp. 475–491, 2020.
- [16] Z. Zheng, S. Xie, H.-N. Dai, X. Chen, and H. Wang, “Blockchain challenges and opportunities: A survey,” *International journal of web and grid services*, vol. 14, no. 4, pp. 352–375, 2018.
- [17] “Etherscan,” 2025. [Online]. Available: <https://etherscan.io/>
- [18] S. Werner, D. Perez, L. Gudgeon, A. Klages-Mundt, D. Harz, and W. Knottenbelt, “Sok: Decentralized finance (defi),” in *Proceedings of the 4th ACM Conference on Advances in Financial Technologies*, 2022, pp. 30–46.
- [19] “9 best defi staking platforms in 2025,” 2025. [Online]. Available: <https://www.solulab.com/best-defi-staking-platforms/>
- [20] “Defi hacks reproduce - foundry,” 2025. [Online]. Available: <https://github.com/SunWeb3Sec/DeFiHackLabs/>
- [21] “Slowmist,” 2025. [Online]. Available: <https://cn.slowmist.com/>
- [22] “Blocksec,” 2025. [Online]. Available: <https://blocksec.com>
- [23] “Peckshield,” 2025. [Online]. Available: <https://peckshield.com>
- [24] “Certik: Elevate your web3 journey,” 2025. [Online]. Available: <https://www.certik.com>
- [25] Z. Zheng, J. Su, J. Chen, D. Lo, Z. Zhong, and M. Ye, “Dappscan: building large-scale datasets for smart contract weaknesses in dapp projects,” *IEEE Transactions on Software Engineering*, 2024.
- [26] “Introduction to dapps,” 2025. [Online]. Available: <https://ethereum.org/en/developers/docs/dapps/>
- [27] “Cyberscope,” 2025. [Online]. Available: <https://www.cyberscope.io/>
- [28] D. Spencer, *Card sorting: Designing usable categories*. Rosenfeld Media, 2009.
- [29] J. Chen, X. Xia, D. Lo, J. Grundy, X. Luo, and T. Chen, “Defining smart contract defects on ethereum,” *IEEE Transactions on Software Engineering*, vol. 48, no. 1, pp. 327–345, 2020.
- [30] J. Zhang, Y. Shen, J. Chen, J. Su, Y. Wang, T. Chen, J. Gao, and Z. Chen, “Demystifying and detecting cryptographic defects in ethereum smart contracts,” *arXiv preprint arXiv:2408.04939*, 2024.
- [31] “Sportsdao - bscscan,” 2025. [Online]. Available: <https://bscscan.com/address/0x66666625ab26131b490e7015333f97306f05bf816#code>
- [32] DeepSeek-AI, “Deepseek-v3 technical report,” 2024. [Online]. Available: <https://arxiv.org/abs/2412.19437>
- [33] “Prompt library,” 2025. [Online]. Available: <https://api-docs.deepseek.com/prompt-library>
- [34] J. Feist, G. Grieco, and A. Groce, “Slither: a static analysis framework for smart contracts,” in *2019 IEEE/ACM 2nd International Workshop on Emerging Trends in Software Engineering for Blockchain (WETSEB)*. IEEE, 2019, pp. 8–15.
- [35] “Uniswap docs: v2 protocol/smart contracts,” 2025. [Online]. Available: <https://docs.uniswap.org/contracts/v2/concepts/protocol-overview/smart-contracts>
- [36] “Pancakeswap: Router v2,” 2025. [Online]. Available: <https://docs.pancakeswap.finance/to-delete/smart-contracts/pancakeswap-exchange/v2-contracts/router-v2>
- [37] “Erc20 - openzeppelin docs,” 2025. [Online]. Available: <https://docs.openzeppelin.com/contracts/4.x/erc20>
- [38] “Erc-721 non-fungible token standard,” 2025. [Online]. Available: <https://ethereum.org/en/developers/docs/standards/tokens/erc-721/>
- [39] “Solidity language — solidity development,” 2025. [Online]. Available: <https://soliditylanguage.org/>
- [40] “Bnb smart chain (bnb) blockchain explorer,” 2025. [Online]. Available: <https://bscscan.com/>
- [41] “Celo explorer - celo foundation,” 2025. [Online]. Available: <https://celo.blockscout.com/>
- [42] “Tron — decentralize the web,” 2025. [Online]. Available: <https://tron.network/>
- [43] “Polygon.io - stock data api real-time historical api,” 2025. [Online]. Available: <https://polygon.io/>
- [44] “Arbitrum,” 2025. [Online]. Available: <https://arbitrum.io/>
- [45] “Explorer — avalanche,” 2025. [Online]. Available: <https://subnets.avax.network/>
- [46] “Fantom explorer,” 2025. [Online]. Available: <https://explorer.fantom.network/>
- [47] “Optimism.io,” 2025. [Online]. Available: <https://www.optimism.io/>
- [48] M. Ortner and S. Eskandari, “Smart contract sanctuary,” [Online]. Available: <https://github.com/tintinweb/smart-contract-sanctuary>
- [49] Z. Lin, J. Chen, J. Wu, W. Zhang, Y. Wang, and Z. Zheng, “Crypwarn: Warning the risk of contract-related rug pull in defi smart contracts,” *IEEE Transactions on Software Engineering*, 2024.
- [50] S. Yang, X. Lin, J. Chen, Q. Zhong, L. Xiao, R. Huang, Y. Wang, and Z. Zheng, “Hyperion: Unveiling dapp inconsistencies using llm and dataflow-guided symbolic execution,” *arXiv preprint arXiv:2408.06037*, 2024.
- [51] Z. Lin, J. Chen, J. Wu, W. Zhang, and Z. Zheng, “Definition and detection of centralization defects in smart contracts,” *arXiv preprint arXiv:2411.10169*, 2024.
- [52] S. Yang, J. Chen, and Z. Zheng, “Definition and detection of defects in nft smart contracts,” in *Proceedings of the 32nd ACM SIGSOFT International Symposium on Software Testing and Analysis*, 2023, pp. 373–384.
- [53] X. Xiong, Z. Wang, X. Chen, W. Knottenbelt, and M. Huth, “Leverage staking with liquid staking derivatives (lsds): Opportunities and risks,” *arXiv preprint arXiv:2401.08610*, 2023.
- [54] P. Qian, Z. Liu, Q. He, B. Huang, D. Tian, and X. Wang, “Smart contract vulnerability detection technique: A survey,” *arXiv preprint arXiv:2209.05872*, 2022.
- [55] L. Luu, D.-H. Chu, H. Olickel, P. Saxena, and A. Hobor, “Making smart contracts smarter,” in *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, 2016, pp. 254–269.
- [56] S. Kalra, S. Goel, M. Dhawan, and S. Sharma, “Zeus: analyzing safety of smart contracts,” in *Ndss*, 2018, pp. 1–12.
- [57] P. Tsankov, A. Dan, D. Drachler-Cohen, A. Gervais, F. Buenzli, and M. Vechev, “Securify: Practical security analysis of smart contracts,” in *Proceedings of the 2018 ACM SIGSAC conference on computer and communications security*, 2018, pp. 67–82.
- [58] L. Brent, N. Grech, S. Lagouvardos, B. Scholz, and Y. Smaragdakis, “Ethaider: a smart contract security analyzer for composite vulnerabilities,” in *Proceedings of the 41st ACM SIGPLAN Conference on Programming Language Design and Implementation*, 2020, pp. 454–469.

- [59] B. Jiang, Y. Liu, and W. K. Chan, "Contractfuzzer: Fuzzing smart contracts for vulnerability detection," in *Proceedings of the 33rd ACM/IEEE International Conference on Automated Software Engineering*, 2018, pp. 259–269.
- [60] T. Durieux, J. F. Ferreira, R. Abreu, and P. Cruz, "Empirical review of automated analysis tools on 47,587 ethereum smart contracts," in *Proceedings of the ACM/IEEE 42nd International conference on software engineering*, 2020, pp. 530–541.
- [61] M. Ye, Y. Nan, H.-N. Dai, S. Yang, X. Luo, and Z. Zheng, "Funfuzz: A function-oriented fuzzer for smart contract vulnerability detection with high effectiveness and efficiency," *ACM Transactions on Software Engineering and Methodology*, vol. 33, no. 7, pp. 1–20, 2024.
- [62] J. Zhang, Z. Zheng, Y. Nan, M. Ye, K. Ning, Y. Zhang, and W. Zhang, "Smartreco: Detecting read-only reentrancy via fine-grained cross-dapp analysis," *arXiv preprint arXiv:2409.18468*, 2024.
- [63] E. Hildenbrandt, M. Saxena, N. Rodrigues, X. Zhu, P. Daian, D. Guth, B. Moore, D. Park, Y. Zhang, A. Stefanescu *et al.*, "Kevm: A complete formal semantics of the ethereum virtual machine," in *2018 IEEE 31st Computer Security Foundations Symposium (CSF)*. IEEE, 2018, pp. 204–217.
- [64] S. Amani, M. Bégel, M. Bortin, and M. Staples, "Towards verifying ethereum smart contract bytecode in isabelle/hol," in *Proceedings of the 7th ACM SIGPLAN international conference on certified programs and proofs*, 2018, pp. 66–77.
- [65] Y. Liu, Y. Xue, D. Wu, Y. Sun, Y. Li, M. Shi, and Y. Liu, "Propertygpt: Llm-driven formal verification of smart contracts through retrieval-augmented property generation," *arXiv preprint arXiv:2405.02580*, 2024.
- [66] M. Soud, W. Nuutinen, and G. Liebel, "Soley: Identification and automated detection of logic vulnerabilities in ethereum smart contracts using large language models," *arXiv preprint arXiv:2406.16244*, 2024.
- [67] "Github - build and ship software on a single, collaborative platform," 2025. [Online]. Available: <https://github.com/>