A Survey on Ethereum Systems Security: Vulnerabilities, Attacks and Defenses

Huashan Chen, Marcus Pendleton, Laurent Njilla, and Shouhuai Xu

Abstract—The blockchain technology is believed by many to be a game changer in many application domains, especially financial applications. While the first generation of blockchain technology (i.e., Blockchain 1.0) is almost exclusively used for cryptocurrency purposes, the second generation (i.e., Blockchain 2.0), as represented by Ethereum, is an open and decentralized platform enabling a new paradigm of computing — Decentralized Applications (DApps) running on top of blockchains. The rich applications and semantics of DApps inevitably introduce many security vulnerabilities, which have no counterparts in pure cryptocurrency systems like Bitcoin. Since Ethereum is a new, yet complex, system, it is imperative to have a systematic and comprehensive understanding on its security from a holistic perspective, which is unavailable. To the best of our knowledge, the present survey, which can also be used as a tutorial, fills this void. In particular, we systematize three aspects of Ethereum systems security: vulnerabilities, attacks, and defenses. We draw insights into, among other things, vulnerability root causes, attack consequences, and defense capabilities, which shed light on future research directions.

Index Terms—Blockchain, Ethereum, Security, Smart Contract, Vulnerabilities, Attacks, Defenses

I. INTRODUCTION

The notion of *blockchain* was implicitly introduced in 2008 as the key underlying technique of the cryptocurrency known as Bitcoin [1], which uses a *transaction-centered* model known as *unspent transaction outputs* (UTXO). In this model, a blockchain is a distributed and public ledger, which records the payment transactions between parties over a peer-to-peer (P2P) network. Unlike traditional digital cash systems [2], in which there is a trusted third party (e.g., bank), there is no trusted third party in a blockchain system in general and in Bitcoin in particular. Bitcoin is often referred to as Blockchain 1.0 because it only offers payment services. The innovation of the Bitcoin system is its consensus protocol, which allows mutually distrusting nodes in a P2P network to eventually reach a consensus on the outcome after executing payment transactions. Unlike traditional consensus protocols [3], the participants are from an open network and are incentivized by the payment of Bitcoins (or BTCs), which are "mined" through a clever cryptographic hash function known as Proof-of-Work (PoW), an idea originally proposed as an anti-spam technique [4].

Perhaps inspired by the success of Bitcoin as well as the need to support semantically richer (than just payment) applications, the notion of *smart contracts* has been introduced

H. Chen and S. Xu are with the Department of Computer Science, University of Texas at San Antonio. M. Pendleton and L. Njilla are with the U.S. Air Force Research Laboratory, Rome, NY. Correspondence: shxu@cs.utsa.edu

to represent autonomous programs, leading to a new paradigm of Decentralized Applications (DApps) that run on top of blockchains and consist of many interacting smart contracts. The Ethereum system was launched in 2015 to support smart contracts, while offering its inherent cryptocurrency known as Ether [5] and using an *account-centered* model (rather than the UTXO model mentioned above). Ethereum has become the de facto standard platform for DApps. At the moment of writing, the market value of Ethereum is over US$31B with approximately one million smart contracts executing on top of the Ethereum blockchain [6], [7]. The success of Ethereum ushers in Blockchain 2.0, which goes much beyond the payment-centered Blockchain 1.0.

While Ethereum facilitates semantically richer applications than Bitcoin, it also enlarges the threat surface, as evidenced by the many high-profile attacks. One example is the DAO [8] attack in year 2016, in which case an attacker exploited the so-called reentrancy vulnerability (which will be detailed later) to steal approximately US$60M. In July 2017, a vulnerability in the Parity wallet contract caused the loss of US$31M [9]. In April 2018, the MyEtherWallet wallet fell victim to a BGP and DNS hijacking attack, enabling the hacker to steal approximately US$17M [10]. These attacks highlight that our capabilities in securing the Ethereum system are limited. This should not be taken as a surprise because Ethereum is a new programming paradigm with DApps running on top of blockchains with many autonomous contracts.

The motivation of the present survey is in three-fold of *researchers*, *practitioners*, and *students*. From the standpoint of a *researcher* who wants to investigate Ethereum security, there is a need for a source of systematized treatment on the problems related to Ethereum security. Despite the fact that there have been some surveys, they did not offer a systematic and comprehensive view on Ethereum vulnerabilities, attacks, and defenses as we do. While referring to the related prior work in Section I-B for details, we mention the following: there is neither systematic understanding of the Ethereum vulnerabilities that have been discovered, nor systematic understanding of their root causes; this may explain why there are still a number of vulnerabilities that are completely open. From the standpoint of a *practitioner*, there is a need for a source of best practices and guiding principles. Industry has conducted due diligence in summarizing many best practices [11], which however may overwhelm practitioners. Therefore, it might be more useful to have a small number of guiding principles that are easier to adopt in practice. From the standpoint of a *student* who wants to learn about Ethereum security, there is a need for a succinct yet comprehensive and systematic source that also offer references to materials of greater details.

A. Our contributions

We provide a systematic and comprehensive survey on Ethereum systems security. It is systematic in the sense that *vulnerabilities*, *attacks*, and *defenses* as well as the relationships between them are accommodated. It is comprehensive in the sense that it covers both the Ethereum platform via a layered architecture and the environment in which the Ethereum platform operates. In terms of vulnerabilities, we enumerate 44 types of Ethereum vulnerabilities according to the layers of the Ethereum architecture and the environment in which Ethereum operates. Perhaps more importantly, we analyze and systematize the root causes of those vulnerabilities. This allows us to provide insights into how to prevent some Ethereum vulnerabilities and how to cope with the inevitable vulnerabilities, including those that are largely open. Some of our findings and insights are highlighted as follows:

1) Ethereum smart contracts introduce new kinds of vulnerabilities that do not have counterparts in traditional paradigm of applications.
2) It is important to design more secure programming languages and supporting tools for programmers to write more secure smart contracts.
3) The vulnerabilities caused by the design and implementation of the Ethereum blockchain are harder to cope with than vulnerabilities in the traditional paradigm.
4) The vulnerabilities in the Ethereum environment are largely caused by human, usability, and networking factors.
5) There are many vulnerabilities (e.g., *outsourceable puzzle*, *51% hashrate*, and *under-priced opcodes*) that must be tackled in order to adequately defend Ethereum or blockchain-based DApp systems in general.

In terms of attacks, we systematize 26 attacks against Ethereum according to the layers of the Ethereum architecture. Perhaps more importantly, we relate these attacks to the vulnerabilities and systematize the attack consequences. Some of our findings are highlighted as follows:

1) The largest single-incident financial loss thus far was caused by a denial-of-service (DoS) attack against the Parity wallet. This attack disabled a library that is used by many contracts and is therefore quite different from traditional DoS attacks.
2) DApps running on top of blockchains are not fully decentralized when they use centralized web interfaces, even if the underlying blockchain is fully decentralized.
3) Ethereum application-layer attacks have caused the largest financial losses.

In terms of defenses, we systematize 47 defenses into two classes: proactive defenses, which aim to prevent attacks as much as possible; and reactive defenses, which aim to cope with the hidden vulnerabilities (i.e., their existence may not be known to the defender). Perhaps more importantly, we present a deeper analysis of the defenses according to (i) their capabilities in mitigating the damages that can be incurred by exploiting vulnerabilities, and (ii) the defense effort that has been investigated to cope with each kind of vulnerabilities. Some of our findings are highlighted as follows:

1) Industry has come up with a significant set of best practices for guiding the development of smart contracts. These best practices, if adequately executed, can indeed avoid many vulnerabilities.
2) Existing proactive defenses can defend against attacks that exploit many vulnerabilities; in contrast, existing reactive defenses can only defend against attacks that exploit a few vulnerabilities.
3) There are large discrepancies between levels of effort and investment into different kinds of attacks (i.e., much more investment into defending against high-profile attacks than low-profile attacks).
4) Existing studies focus on defending against attacks that attempt to exploit vulnerabilities in the DApp back-end (i.e., smart contracts), but largely ignore the protection of the DApp front-end (i.e., browser) and the interactions between the front-end and the back-end.

Although the present paper focuses on the Ethereum system, the aforementioned findings related to vulnerabilities, attacks and defenses might be applicable to blockchain-based systems in general. Moreover, we discuss some fundamental research problems that must be adequately tackled in order to secure and defend Ethereum and blockchain-based systems, including:

1) There is a lack of deep understanding on the rigorously-defined properties of blockchain-based systems.
2) There is a lack of deep understanding on the rigorous analysis methodologies that are necessary and sufficient for analyzing the desired properties of blockchain-based systems.
3) There is a lack of deep understanding on the metrics that are necessary and sufficient for analyzing the security and risk of blockchain-based systems.

In order to improve readability and ease the reference to the large number of vulnerabilities, we denote the vulnerabilities by $\mathcal{V}_1, \ldots, \mathcal{V}_{44}$, respectively. Similarly, we respectively denote the attacks by $\mathcal{A}_1, \ldots, \mathcal{A}_{26}$ and the defenses by $\mathcal{D}_1, \ldots, \mathcal{D}_{47}$. Moreover, we use "$\mathcal{A}_i(\mathcal{V}_j, \cdots)$" to denote that attack $\mathcal{A}_i$ exploits vulnerabilities $\mathcal{V}_j$ and possibly others.

The preceding discussion justifies how the present paper would adequately serve the needs of students and researchers. For practitioners, we mention that we systematize the best practices into a small number of principles that may be easier to adopt in practice.

B. Related work

Table I highlights the relationship between the present work and related prior surveys, which accommodate some of vulnerabilities, attacks, defenses, and building-blocks. The most closely related survey is Atzei et al. [12], which discussed 12 types of vulnerabilities, 9 attacks, and 3 defenses in the context of Ethereum smart contracts (i.e. Blockchain 2.0). In contrast, we present a much more systematic treatment, by accommodating 44 types of vulnerabilities, 26 attacks, and 47 defenses. From a methodological standpoint, we further discuss the root causes of vulnerabilities (e.g., the root cause of the *unchecked call return value* vulnerability is the *inconsistent*

TABLE I: Comparison between surveys related to blockchain security and privacy.

Study	Vulnerabilities			Attacks			Defenses			Building-blocks	
	Blockchain 1.0	Blockchain 2.0	Insights	Blockchain 1.0	Blockchain 2.0	Insights	Blockchain 1.0	Blockchain 2.0	Insights	Cryptography	Consensus
[12]		12			9			3			
This work		44	✓		26	✓		47	✓		
[13]	20			6			5				
[14]	11			11			30				
[15]				22		✓	33				yes
[16]								10	✓		
[17]								27	✓		
[18]										yes	
[19], [20], [21], [22]											yes

exception handling of Solidity). Moreover, we draw insights from the perspectives of vulnerability, attack, and defense.

There are a number of surveys on blockchain security and privacy which however have different perspectives than ours. First, Li et al. [13] reviewed security of blockchain technologies through the lens of 20 types of vulnerabilities, 6 attacks, and 5 defenses, without distinction between Blockchain 1.0 and 2.0 aspects. Similarly, Zhu et al. [14] reviewed 11 smart contract vulnerabilities, 11 attacks against blockchain data, and 30 defenses. On the contrary, we focus on the Ethereum blockchain system by accommodating 44 types of vulnerabilities, 26 attacks, and 47 defenses.

Second, Saad et al. [15] explored blockchains' attack surface in terms of cryptographic constructions, distributed system architecture, and applications; they cover 22 attacks and 33 defenses in Blockchain 1.0 and 2.0, but not vulnerabilities. Their survey is orthogonal to ours because (i) we focus on the Ethereum system, rather than multiple implementations of blockchains; (ii) we discuss vulnerabilities and their root causes as well as the attacks exploiting them, rather than attack surface; (iii) we provide insights into, and contrast, the vulnerabilities at different layers of the Ethereum architecture.

Third, Harz et al. [16] discussed 10 smart contract verification tools. Similarly, Angelo et al. [17] discussed 27 tools for analyzing Ethereum smart contracts. On the contrary, we focus on the defenses of the Ethereum systems security, rather than purely on smart contracts.

Fourth, Zhang et al. [18] presented a comprehensive review on Bitcoin-like transactions and the underlying (cryptographic) mechanisms. Their review is geared towards the abstract blockchain model for Bitcoin-like transactions (i.e., Blockchain 1.0); in contrast, we focus on the Ethereum ecosystem, including the design and implementation of the blockchain platform and DApps. There are surveys that focus on Bitcoin and cryptocurrencies [23], [24], [25], which use a transaction-centered model known as *unspent transaction outputs* (UTXO). In contrast, we focus on Ethereum, which uses an account-centered model.

Fifth, Orthogonal to the purpose of the present survey and tutorial, there are several surveys purely on blockchain consensus protocols [19], [20], [21], [22]. In contrast, we focus on the overall Ethereum system.

C. Paper outline

Section II briefly reviews the Ethereum system and discusses the survey methodology. Section III presents the 44 Ethereum vulnerabilities according to the layers of the Ethereum architecture and a deeper analysis of their root causes. Section IV presents the 26 attacks against Ethereum and a deeper analysis on their consequences. Section V presents the 47 defenses, classified as proactive or reactive and a deeper analysis on their capabilities and investments. Section VI discuss future research directions towards securing blockchain-based systems, including Ethereum as a special case. Section VII conclude the present paper.

II. ETHEREUM REVIEW AND SURVEY METHODOLOGY

A. A brief review of the Ethereum system

Figure 1 highlights a 4-layer architecture of the Ethereum blockchain, operates across these 4 layers. At the *application layer*, Ethereum clients execute smart contracts, which are associated to Ethereum accounts, in EVM. The *data layer* contains the blockchain data structures. The *consensus layer* assures a consistent state of the blockchain. Note that Ethereum plans to replace its current use of Proof-of-Work (PoW) with a Proof-of-Stake (PoS). The *network layer* formulates a Ethereum peer-to-peer (P2P) network of *nodes* or *clients* such that a node can always get the updated state of the blockchain from some of the active nodes. The *environment* serves these 4 layers via a corresponding component: a web user interface to interact with applications; databases for storing blockchain data; cryptographic mechanisms for supporting the consensus protocols; and Internet service for the network layer.

1) The application layer: Ethereum supports two types of accounts: *externally owned accounts* (EOA) and *contract accounts*. An EOA is used to keep a user's funds in Wei, which is the smallest subdenomination of Ether and is worth 10^{-18} Ether. An EOA is associated with, and addressed by, a public key; access to an EOA is authenticated by showing the ownership of the corresponding private key. In contrast, a
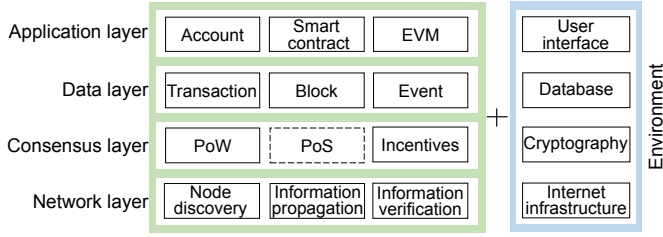
Fig. 1: Architecture of the Ethereum blockchain and its environment in which the Ethereum blockchain runs.

contract account is associated with a piece of executable byte-code (i.e., smart contracts), which defines some business logic of interest. An EOA or contract account has a dynamic state, which is defined by: (i) *nonce*, which tracks the number of *transactions* that have been initiated by the owner of the EOA or the number of contracts created by the contract account; (ii) *balance*, which is the amount of Wei (i.e., 10^{-18} Ether) owned by the EOA or contract account; (iii) *storageRoot*, which is the hash of the root of the account's storage data structure *trie* that records a contract's state variables associated to the corresponding piece of bytecode (i.e., not applicable to EOA); (iv) *codeHash*, which is the hash value of a contract account's bytecode (i.e., not applicable to EOA). The state of a blockchain is defined by the states of the accounts on the blockchain.

Smart contracts are building-blocks of decentralized applications (DApps) running on top of the Ethereum blockchain. A DApp often has a user interface as its front-end and some smart contracts as its back-end. At the moment of writing, 2,497 DApps are running on top of Ethereum, including finance, governance, gambling, exchange, and wallet applications [26]. Some DApps issue their own cryptocurrency, called *tokens*, for purposes like Initial Coin Offering (ICO) and exchanges. A Ethereum-based token is a special kind of smart contract (e.g., ERC-20 [27]).

Smart contracts execute in EVMs, which are quasi-Turing-complete machines using a stack-based architecture; the term "quasi" means that the execution is limited by the amount of gas offered by the transaction in question [28].

2) The data layer: A *transaction* is an interaction between an EOA (called *sender*) and anther EOA or contract account (called *recipient*). A transaction is specified by: (i) *nonce*, which is a counter for tracking the total number of transactions that have been initiated by the sender; (ii) *recipient*, which specifies a transaction's destination EOA or contract account; (iii) *value*, which is the amount of money (unit: Wei) to be transferred from the sender to the recipient (if applicable); (iv) *input*, which is the bytecode or data corresponding to the purpose of the transaction; (v) *gasPrice* and *gasLimit*, which respectively specify the unit price and the maximum amount of gas the sender is willing to pay the winning miner of a *block* containing the transaction; (vi) (v, r, s), which is the Elliptic Curve Digital Signature Algorithm (ECDSA) signature of the sender. The execution of a transaction updates the states of the accounts involved and therefore the state of the blockchain.

Figure 2 depicts the lifecycle of a transaction. (i) A sender constructs a transaction and digitally signs it. (ii) The sender submits the signed transaction to an Ethereum client via a JSON-RPC call. (iii) The client validates the received transaction and broadcasts it to the Ethereum P2P network. (iv) Any client that receives the transaction and is a miner adds the transaction to its transaction pool. (v) A miner executes a sequence of transactions chosen from its transaction pool, formulates a new block, and updates the state of the blockchain as follows. For a `money-transfer` transaction, the specified *value* is transferred from the sender's EOA to the recipient's EOA or contract account; for a `contract-creation` transaction where *input* is a piece of bytecode, a new contract account is created and is associated with the bytecode; for a `contract-invocation` transaction where *recipient* is a callee contract and *input* uniquely identifies the callee function (and possibly some associated arguments), the bytecode associated to the callee contract account is loaded into the EVM. (vi) The miner solves a PoW by finding a random nonce such that the hash value of the block metadata in question is smaller than a certain threshold, which reflects the difficulty of creating a block. Unlike Bitcoin's computation-intense PoW [1], Ethereum uses a memory-intense puzzle called "Ethash" [5]. (vii) Upon creating a block, the miner broadcasts it to the Ethereum P2P network so that other clients can validate the block. (viii) Upon validating a block, the client in question appends the block to the blockchain.

Trie [5] is the data structure for storing Ethereum blockchain data (e.g., account states). Like a Patricia tree [29], a trie stores $(key, value)$ pairs while facilitating search as follows: the path from the root to a leaf node corresponds to a *key* and the leaf node contains a *value* (e.g., the state of an account). As illustrated in Figure 3, a block header may point to a state trie, a transaction trie (for bookkeeping transaction data), and a receipt trie (for bookkeeping the data related to the execution of transactions). Each contract account corresponding to a leaf or branch node on the state trie uses a separate storage trie to bookkeep the persistent data of the contract; this storage trie also use a $(key, value)$ structure, where the position of each slot corresponds to a *key* and the contract's state variable in each slot corresponds to a *value*. Note that an Ethereum blockchain has a *single* state trie because the state of the blockchain dynamically evolves.

3) The consensus layer: At the moment of writing, Ethereum takes about 14 seconds to create a block, meaning that multiple miners could create valid blocks simultaneously and that there could be many *stale* blocks. Ethereum uses a variant of the GHOST consensus protocol [30] to select the "heaviest" branch as the *main chain* where the "heaviest" branch is the sub-tree rooted at the fork in question and has the highest cumulative block difficulty [31], [32], while noting that *stale* blocks are not on the main chain.

Ethereum rewards not only the *regular* blocks on the main chain, but also the stale blocks referred by a regular block. As illustrated in Figure 4, an *uncle* block is a stale block referenced by a regular block called *nephew* (via a dashed arrow). The distance between two blocks is their height difference on the tree. The miner of a regular block receives one unit of "static block reward," which is worthy of 2 Ethers
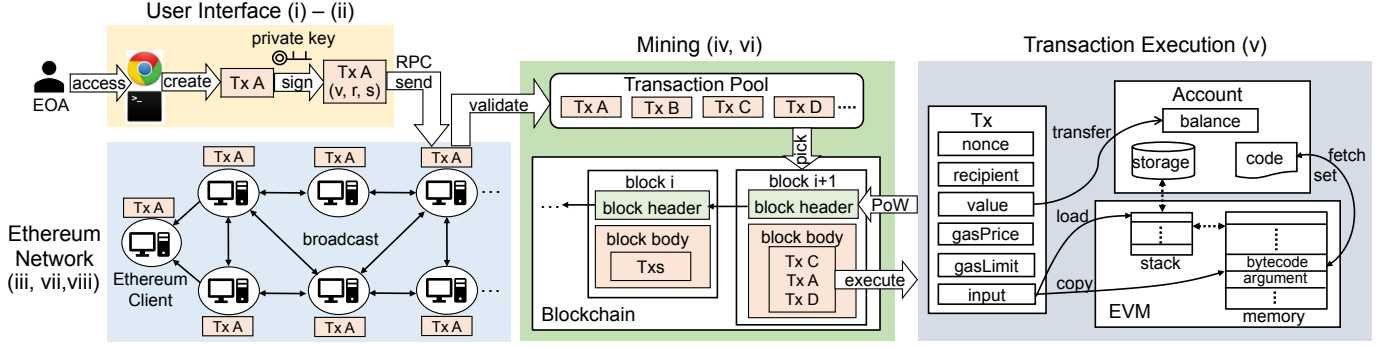
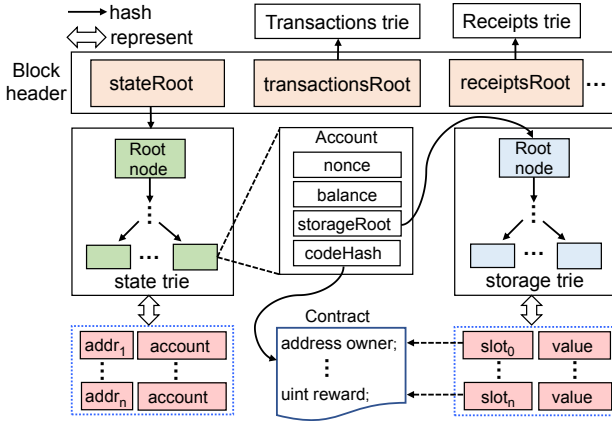Fig. 2: An illustration of the lifecycle of an Ethereum transaction.

Fig. 3: Storage structure of Ethereum blockchain.

at the time of writing. In order to incentivize referencing to uncle blocks, the miner of a nephew regular block further receives $1/32$ of the static block reward for a reference (and for up to 2 references). The miner of the referenced uncle block is rewarded with $1 - d/8$ of the static block reward, where $1 \leq d \leq 6$ is the distance between the uncle block and the referencing nephew block. Miners of regular blocks also receive the gas fee specified by the transactions in question.
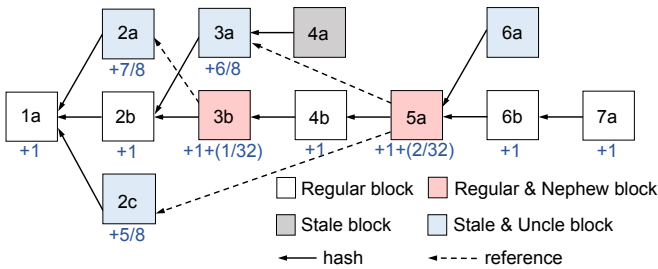

Fig. 4: A sub-tree of an Ethereum blockchain (adapted from [33]), where a value beneath a node is the total award to the miner of the node (unit: static block reward).

4) The network layer: The Ethereum network is a structured P2P network where each node (i.e., client) stores a copy of the entire blockchain. For node discovery and routing purposes, each node maintains a dynamic routing table of 160 buckets and each bucket contains up to 16 entries of other nodes' IDs,

IP address, UDP/TCP ports. Ethereum uses the RLPx protocol [34], [35] to discover target clients and uses the Ethereum Wire Protocol [36] to facilitate the exchange of Ethereum blockchain information (e.g., transactions, blocks) between clients.

5) The environment: The Ethereum blockchain runs in an environment, which naturally operates across the 4-layers to provide respective services, namely: a web interface for users to interact with the Ethereum blockchain; a database for Ethereum clients to store the blockchain data; cryptographic mechanisms for security purposes; and the Internet infrastructure to support blockchain networking and communication among Ethereum nodes. We separate the Ethereum blockchain architecture from the environment because attacks against the Ethereum blockchain may come from the environment and these attacks may be better addressed in the environment rather than by the Ethereum blockchain, leading to a clean and modular abstraction.

B. Survey methodology

1) Scope: Since our focus is on Ethereum security, we will systematize the vulnerabilities of, the attacks against, and the defenses for Ethereum. Since these aspects are related to the programming language for writing smart contracts and the client software, we focus on the widely-used Solidity and Geth [37] as well as Parity [38], respectively.

2) Methodology: Our methodology can be characterized as follows. First, we use a layered architecture to present the Ethereum ecosystem, from the application down to the data, the consensus, the network, and the environment layers. This layered view allows us to describe matters at the most appropriate layer.

Second, we consider security from three perspectives (i.e., *vulnerabilities*, *attacks*, and *defenses*) as well as the relationships between them. For each vulnerability, we discuss, among other things, its root cause and status (i.e., eliminated, can be eliminated by best practice, or still open). For each attack, we discuss, among other things, its cause, tactic, and direct impact. For each defense, we discuss its mechanism and the (hidden) vulnerabilities it aims to protect from exploitation. We aim to provide insights into the design and implementation of Ethereum as well as future research directions in order to address the range of open problems.

III. VULNERABILITIES

Figure 5 highlights our classification of Ethereum vulnerabilities based on their *location*, *cause*, and *status* (i.e., *eliminated* vs. *can be avoided by best practice* vs. *open*). For ease of reference, we denote the 44 types of vulnerabilities as $\mathcal{V}_1, \ldots, \mathcal{V}_{44}$, respectively. In what follows, we group them according to location.

A. Vulnerabilities at the application layer

1) Reentrancy ($\mathcal{V}_1$): This vulnerability was first observed from the DAO attack [39]; its variants were later reported in [40]. The vulnerability occurs when an external callee contract calls back to a function in the caller contract *before* the execution of the caller contract is completed (i.e., cyclic calls in a sense). This allows the attacker to bypass the due validity check until the caller contract is drained of Ether or the transaction runs out of gas. The vulnerability is caused by: (i) a contract's control-flow decision relies on some of its state variable(s) that should be, but are not, updated by the contract itself before calling another (i.e., an external) contract [40]; and (ii) there is no gas limit when handing the control-flow to another contract. The vulnerability can be prevented by one of the following methods [41]: (i) assuring that a contract's state variables are updated *before* calling another contract; (ii) introducing a mutex lock on the contract state to assure that only the lock owner can change the state; (iii) using the transfer method to send money to other contracts because this method only forwards 2,300 gas to the callee contract.

2) Delegatecall injection ($\mathcal{V}_2$): This vulnerability was first observed from an attack against the Parity wallet [42]. In order to facilitate code-reuse, EVM provides an opcode, DELEGATECALL, for embedding a callee contract's bytecode into the bytecode of the caller contract as if it's a piece of the latter's bytecode [43]. As a consequence, a malicious callee contract can directly modify (or manipulate) the state variables of the caller contract. This vulnerability is caused by the ability that the state variables of a caller contract can be updated by the bytecode of a callee contract. The vulnerability can be completely prevented by declaring a contract that is meant to be shared via the delegatecall as a library, which is stateless [44].

3) Frozen Ether ($\mathcal{V}_3$): This vulnerability was first observed from another attack against the Parity wallet [45]. The vulnerability results from the ability of users to deposit their money to their contract accounts with the inability to spend their money from those accounts, effectively freezing their money. The vulnerability is caused by [46]: (i) contracts not providing any function for spending money relying on the money-spending function of another contract (as a library) and (ii) the callee contract (i.e., the library) being killed accidentally or deliberately. The vulnerability can be prevented by assuring that mission-critical functions, or money-spending functions in this case, are not outsourced to another contract.

4) Upgradable contract ($\mathcal{V}_4$): This vulnerability was first discussed in [12]. The idea of contract upgrading was introduced to mitigate the problem that smart contracts, once deployed, cannot be modified even if they are later found to have vulnerabilities. In order to allow contract upgrading, there are two approaches: (i) splitting a contract into a *proxy contract* and a *logic contract* such that developers can upgrade the latter but not the former; and (ii) using a *registry contract* to bookkeep the updated contracts. While effective, these approaches introduce a new vulnerability: when the contract developer is malicious, the updated (logic) contract can be malicious. The vulnerability is caused by one contract being able to call another contract, which can be malicious. This vulnerability (i.e., unsecure contact updating) remains to be an open problem.

5) DoS with unexpected revert ($\mathcal{V}_5$): This vulnerability was first reported in [41]. It occurs either when a transaction is reverted due to a caller contract encountering a failure in an external call, or the callee contract deliberately performs the revert operation to disrupt the execution of the caller contract. This vulnerability is caused by the execution of a caller contract being reverted by a callee contract. This vulnerability can be prevented by letting a recipient invoke a transaction to "pull" the money that was set aside by a sender for the recipient, which effectively prevents a sender's transaction from being reverted [47].

6) Integer overflow and underflow ($\mathcal{V}_6$): This vulnerability was first observed from the attack against the BEC tokens [48]. It occurs when the result of an arithmetic operation falls outside of the range of a Solidity data type, causing (for example) unauthorized manipulation to the attacker's balance [48] or other state variables. The vulnerability is caused by that Solidity source code does not perform proper validation on numeric inputs, and that neither the Solidity compiler nor the EVM enforces integer overflow/underflow detection. This vulnerability can be prevented by using the *SafeMath* library [49] that handles these issues.

7) Manipulated balance ($\mathcal{V}_7$): This vulnerability was first reported in [50] and was also known as the "forcing Ether to contracts" vulnerability. This vulnerability occurs when a contract's control-flow decision relies on the value of this.balance or address(this).balance, which can be leveraged by an attacker to cause (for example) that only the attacker can obtain the money; see [44] for a detailed description. This vulnerability can be prevented by not using a contract's balance in any condition check [50].

8) Authentication through tx.origin ($\mathcal{V}_8$): This vulnerability was first discussed in [51]. The tx.origin is a global variable in Solidity and refers to the original EOA that initiates the transaction in question. This vulnerability occurs when a contract uses tx.origin for authorization, which can be compromised by a phishing attack. Figure 6 provides an example of this vulnerability (Line 6) in function withdrawAll, which uses tx.origin to confirm that it is the owner of contract UserWallet that is calling the function. However, if the owner of contract UserWallet is tricked to transfer money to a malicious AttackContract, the fallback function (Line 14) will be automatically executed, which in turn calls the withdrawAll() function. Since the condition check (Line 6) validates the EOA that initiated the transaction rather than the intermediate caller (i.e., AttackContract), the attacker can pass the authority check and steal the money in
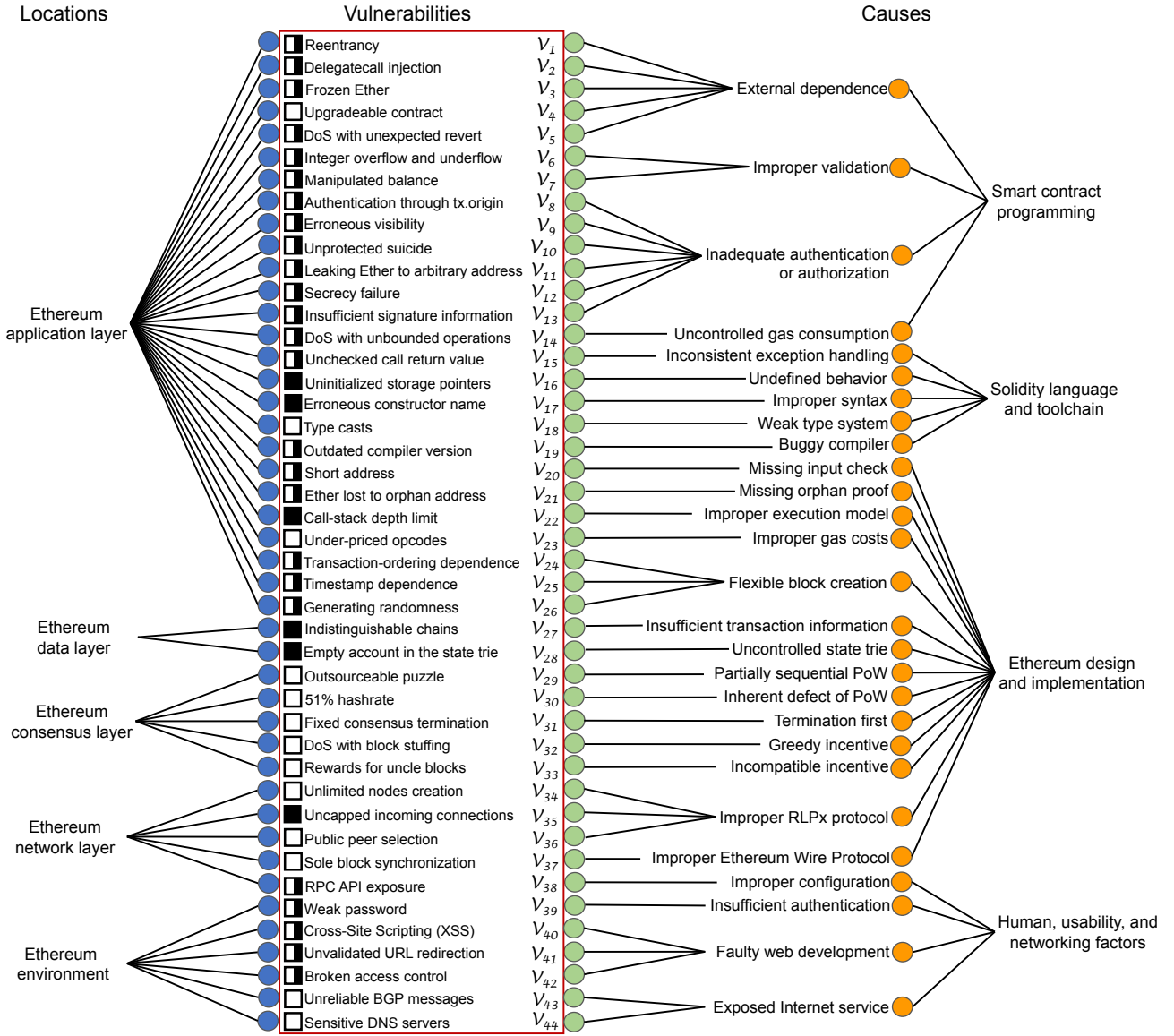
Locations   Vulnerabilities   Causes

Fig. 5: A classification of Ethereum vulnerabilities and their state-of-the-art treatments (■ means "eliminated already", ◨ means "can be avoided by best practice", and □ means open (i.e., has yet to be eliminated).

contract `UserWallet`. This vulnerability can be prevented by using `msg.sender`, instead of `tx.origin`, for authentication [52] because `msg.sender` returns the account that incurred the message.

9) Erroneous visibility ($\mathcal{V}_9$): This vulnerability was first observed in an attack against the Parity wallet [42]. It occurs when a function's visibility is incorrectly specified and thus permits unauthorized access. Specifically, Solidity provides four types of visibility to restrict access to a contract's functions, namely `public`, `external`, `internal`, and `private`, which respectively says a function can be called arbitrarily, only externally, only internally (i.e., within the contract and its derived contracts), or only within the contract. Functions that should not be called from an external contract should be specified as `private` or `internal`. However, Solidity makes functions as `public` by default, which allows attackers to directly call these improperly specified functions. Solidity

```
1  contract UserWallet { // vulnerable contract
2      address owner;
3      constructor() public { owner = msg.sender; }
4      ...
5      function withdrawAll(address addr) public {
6          require(tx.origin == owner);
7          addr.transfer(this.balance);
8      }
9  }
10
11 contract AttackContract { // malicious phishing contract
12     address attacker;
13     function () payable {
14         UserWallet(msg.sender).withdrawAll(attacker);
15     }
16 }
```

Fig. 6: The `tx.origin` vulnerability and its exploitations.

(starting version 0.5.0) mitigates the vulnerability by making it mandatory for programmers to explicitly specify function visibility [53]. Still, this vulnerability cannot be completely prevented unless programmers correctly specify their functions' visibilities.

10) Unprotected suicide ($\mathcal{V}_{10}$): This vulnerability was first observed from an attack against the Parity wallet [45]. A contract account can be killed by the contract's owner (or a trusted third party) using the suicide or self-destruct method. A contract may have an "owner" whose privileges are typically specified when creating the contract. When a contract account is killed, its associated contract bytecode and storage are deleted forever. This vulnerability is caused by the inadequate authentication enforced by a contract. The vulnerability can be mitigated by enforcing advanced authentication (e.g, multi-factor authentication) to assure that a suicide operation must be approved by multiple parties [54].

11) Leaking Ether to arbitrary address ($\mathcal{V}_{11}$): This vulnerability was first reported in [55]. The vulnerability occurs when a contract's funds can be withdrawn by any caller, who is neither the owner of the contract nor an investor who deposited funds to the contract. This vulnerability is caused by the failure in checking an caller's identity when the caller invokes a function to send Ether to an arbitrary address. This vulnerability can be prevented by enforcing an adequate authentication on the functions for sending funds.

12) Secrecy failure ($\mathcal{V}_{12}$): This vulnerability was first observed from a multi-player game in [56] and was also called *keeping secrets* in [12]. It can be exploited to benefit an attacker. In blockchain, restricting the visibility of a variable or function does not assure that the variable or function is secret, because of the public nature of blockchain (i.e., details of transactions are publicly known). Although restricting a state variable to be private can prevent other contracts from accessing it, any one can see the value of a state variable from the relevant transaction data. This vulnerability is caused by the lack of secrecy of sensitive data in an untrusted environment. A possible solution to preventing this vulnerability is to use cryptographic techniques, such as timed commitments [56], [12].

13) Insufficient signature information ($\mathcal{V}_{13}$): This vulnerability was first exploited in a replay attack against smart contracts [57]. The vulnerability occurs when a digital signature turns out to be valid for multiple transactions, which can happen when one sender (say Alice) sends money to multiple recipients through a proxy contract (instead of initiating multiple transactions) [58]. In the proxy contract mechanism, Alice can send a digitally signed message off-chain (e.g. via email) to the recipients, similar to writing personal checks in the real world, to let the recipients withdraw money from the proxy contract via transactions. In order to assure that Alice does approve a certain payment, the proxy contract verifies the validity of the digital signature in question. However, if the signature does not give the due information (e.g., nonce, proxy contract address), a malicious recipient can replay the message multiple times to withdraw extra payments. This vulnerability can be prevented by incorporating the due information in each message [59].

14) DoS with unbounded operations ($\mathcal{V}_{14}$): This vulnerability was first observed from the GovernMental contract [60] and its variants were later discussed in [61], [62]. Recall that each block has a "gas limit" field that specifies the maximum total amount of gas that can be consumed by the transactions in a block. This vulnerability occurs when the amount of gas that is required for executing a contract exceeds the block gas limit. This vulnerability is caused by improper programming with unbounded operations in a contract (e.g., loop over a large array). This vulnerability can be mitigated by assuring (i) contracts should not use loops over data structures, especially those data structures that can be operated by EOA; and (ii) when a contract has to use loop over data structure, the contract should keep track of the loop and resume the aborted execution when the sender of the transaction re-invokes the same contract (in order to finish the execution of the contract) [61].

15) Unchecked call return value ($\mathcal{V}_{15}$): This vulnerability was first discussed in [63], [54] and is also known as *mishandled exceptions*. It has two variants, called *gasless send* and *unchecked send* [12], [64]. Recall that Solidity provides two methods for a contract to call another: (i) directly referencing to a callee contract's instance; (ii) using one of the following four low-level methods: send, call, delegatecall and callcode. There is a discrepancy in Solidity's handling of exceptions occurring in the execution of callee contracts [63]: if an exception occurs in case (i), the exception is automatically propagated back to the caller and the transaction is reverted entirely; if an exception occurs in case (ii), the callee contract returns FALSE back to the caller contract. This discrepancy can lead to unintended transactions unless the caller contract carefully addresses the discrepancy. At the moment of writing, neither the Solidity compiler nor the EVM addresses the discrepancy. This vulnerability can be prevented by letting a caller contract check and address the discrepancy mentioned above.

16) Uninitialized storage pointer ($\mathcal{V}_{16}$): This vulnerability was first reported in [65]. Recall that in Solidity, the contract state variables are always laid out consecutively in storage, starting from slot 0. For a complicated local variable (e.g., struct, array, or mapping), a reference is assigned to an unoccupied slot in the storage to point to the state variable. If the local variable is not explicitly initialized, then the local variable's reference points to slot 0 by default, causing the content starting from slot 0 to be overwritten [66]. This vulnerability is caused by Solidity's treatment of uninitialized complicated local variables. This vulnerability has been eliminated by Solidity compiler, starting version 0.5.0, by reporting an error to contracts that contain uninitialized storage pointers [53].

17) Erroneous constructor name ($\mathcal{V}_{17}$): This vulnerability was first observed from the Rubixi contract [67], where the constructor function has an incorrect name that allows anyone to become the owner of the contract. Prior to Solidity version 0.4.22, a function declared with the same name as the contract's is considered as the contract's constructor, which is only executed one-time upon the creation of the contract. If the constructor's name is misspelled by a programmer for whatever reason, the intended constructor becomes a public, normal function that can be invoked by *any* EOA. This vulnerability

is caused by that Solidity does not provide a special syntax to distinguish a constructor function from a regular function [68]. The vulnerability has been eliminated, starting Solidity version 0.4.22, by introducing the new keyword `constructor` [69].

```
1  contract CounterLibrary { function add(uint) public returns (uint) }
2  contract CounterLib { function add(uint) public returns (uint) }
3  contract Game {
4      function play(CounterLibrary c) public {
5          c.add(1);
6      }
7  }
```

Fig. 7: The type casts vulnerability.

18) Type casts ($\mathcal{V}_{18}$): This vulnerability was first reported in [12]. Recall that a contract written in the Solidity language can call another contract by directly referencing to the callee contract's instance. As illustrated in Figure 7, contract `Game` calls the function add() in contract `CounterLibrary` by referencing to its instance c (Line 5). When function play() (Line 4) is invoked, the argument specifying the callee contract's address is cast to `CounterLibrary`. However, the Solidity compiler can only check whether or not the `CounterLibrary` contract declared function add(), but cannot check whether or not the address argument conforms to that of the `CounterLibrary` contract's. If the address associated to the `CounterLib` contract (Line 2) contains a function that is named add() and has the same declaration, then the add() function in contract `CounterLib` is executed, instead of the desired add() function in contract `CounterLibrary`. As a consequence, the EVM can be misled to run the attacker's contract. This vulnerability is caused by the incompetent type system of Solidity. Currently, there is no feasible way to avoid the vulnerability.

19) Outdated compiler version ($\mathcal{V}_{19}$): This vulnerability was first reported in [54]. It occurs when a contract uses an outdated compiler, which contains bugs and thus makes a compiled contract vulnerable. This vulnerability can be prevented by using an up-to-date compiler.

20) Short address ($\mathcal{V}_{20}$): This vulnerability was first discussed in [70]. Recall that in a `contract-invocation` transaction, the function selector and arguments are orderly encoded in the *input* field, where the first four bytes specify the callee function and the rest data arranges arguments in chunks of 32 bytes. However, if the length of the encoded arguments is shorter than expected, EVM will auto-pad extra zeros to the arguments to make up for 32 bytes. Consider function transfer(address addr, uint tokens) as an example. If the trailing one (i.e., last) byte of addr is left off, two extra hex zeros will be added to the end of tokens, which amplifies the number of tokens being sent. This vulnerability is caused by that EVM does not check the validity of addresses [71]. This vulnerability can be prevented by checking the length of a transaction's input (i.e., `msg.data`) [72].

21) Ether lost to orphan address ($\mathcal{V}_{21}$): This vulnerability was first reported in [12]. When transferring money, Ethereum only checks whether the length of the recipient's address is no greater than 160-bit but not the validity of the recipient's address. If money is sent to a non-existing orphan address, Ethereum automatically registers for the address than terminating the transaction. Since the address is not associated to *any* EOA or contract account, there is no way to withdraw the transferred money, which is effectively lost. This vulnerability is caused by that EVM is not orphan-proof. At the moment of writing, this vulnerability can only be prevented by manually assuring the correctness of the recipient's address.

22) Call-stack depth limit ($\mathcal{V}_{22}$): This vulnerability was first reported in [73]. Recall that in the original specification of Ethereum execution model [5], EVM's call-stack has a hard limit of 1024 frames. When a contract calls another contract, the call-stack depth of the transaction increases by one; when the number of nested calls exceeds 1024, Solidity throws an exception and aborts the call. An attacker can recursively call a contract, which may be deployed by the attacker, for 1023 times and then calls a victim contract to reach the stack depth limit, which fails any subsequent external call made by the victim contract. Since Solidity does not propagate exceptions in low-level external calls, the victim contract may not be aware of the failure. This vulnerability is caused by EVM's inadequate execution model, and has been eliminated by the hard fork for EIP-150, which re-defines the gas-consumption rules of external calls to make it impossible to reach 1024 in call stack depth [74].

23) Under-priced opcodes ($\mathcal{V}_{23}$): This vulnerability was first observed from two DoS attacks [75], [76], [77]. Recall that Ethereum uses the gas mechanism to prevent the abuse of computing resources (e.g., CPU, disk, network). This vulnerability occurs when a contract contains a lot of under-priced opcodes that consume a large amount of resources at a low gas cost, meaning that the execution of the contract wastes a lot of computing resources. This vulnerability is caused by the failure in properly setting the gas cost for consuming computing resources. In order to mitigate this vulnerability, Ethereum has raised the gas cost for the opcodes that were abused to launch the two DoS attacks described in [74]. However, it is not clearly whether or not the vulnerability can be completely prevented by this mechanism or not [77].

24) Transaction ordering dependence (a.k.a. front running; $\mathcal{V}_{24}$): This vulnerability was first discussed in [63]. It refers to the concurrency issue that the forthcoming state of blockchain depends on the execution order of transactions, which is however determined by the miners. Typically, miners group and order transactions into a new block based on the reward offered by the transactions. Since transactions are publicly broadcast to the network, a malicious EOA can offer a higher *gasPrice* to have its transactions assembled into blocks sooner than the others'. Moreover, a malicious miner can always pick up its own transactions regardless of the *gasPrice*. This vulnerability is caused by that the state of a contract depends on how minors select transactions to assemble blocks. This vulnerability can be mitigated by using a cryptographic commit-reveal scheme to hide the information (e.g., *gasPrice*, *value*) offered by transactions [78], [79], or by introducing a guard condition to assure that an invocation of a contract either returns the expected output or fails [63].

25) Timestamp dependence ($\mathcal{V}_{25}$): This vulnerability was first reported in [63]. It occurs when a contract uses the `block.timestamp` as a part of the triggering condition for executing a critical operation (e.g., money transfer) or as the source of randomness, which however can be manipulated by a malicious miner. The vulnerability is caused by that Ethereum only prescribes that a timestamp must be greater than the timestamp of its parent block and be within future 900 seconds of the current clock. If a contract uses a timestamp-based condition (e.g., `block.timestamp %25 == 0`) to determine whether or not to transfer money, a malicious miner can slightly shift the timestamp to satisfy the condition to benefit the attacker. This vulnerability can be prevented by not using `block.timestamp` in any decision-making conditions.

26) Generating randomness ($\mathcal{V}_{26}$): This vulnerability was first reported in [80]. Many gambling and lottery contracts select winners randomly, for which a common practice is to generate a pseudorandom number based on some initial private seed (e.g., `block.number`, `block.timestamp`, `block.difficulty` or `blockhash`). However, these seeds are fully controlled by miners, meaning that a malicious miner can manipulate these variables to make itself the winner [81], [82]. This vulnerability is caused by manipulable entropy sources. There are several proposals for addressing this problem, and each proposal has its own pros and cons. For example, the Oracle RNG proposal [83] uses existing external services to generate random numbers off-chain and then send back to the requesting contract, meaning that there is a single-point-of-failure in the Oracle RNG; the RANDAO proposal [84] uses a collaborative cryptographic commit-reveal scheme to generate a random number by multiple participants altogether, meaning that it achieves a limited throughput.

B. Vulnerabilities at the data layer

1) Indistinguishable chains ($\mathcal{V}_{27}$): This vulnerability was first observed from the cross-chain replay attack when Ethereum was divided into two chains, namely ETH and ETC [85]. Recall that Ethereum uses ECDSA to sign transactions. Prior to the hard fork for EIP-155 [86], each transaction consisted of six fields (i.e., *nonce*, *recipient*, *value*, *input*, *gasPrice* and *gasLimit*). However, the digital signatures were not chain-specific because no chain-specific information was even known back then. As a consequence, a transaction created for one chain can be reused for another chain. This vulnerability has been eliminated by incorporating and signing a chainID.

2) Empty account in the state trie ($\mathcal{V}_{28}$): This vulnerability was first observed from a DoS attack reported in [76], [77]. An *empty* account is an account that has zero nonce, zero balance, no code associated to it, and no storage associated to it. An empty account is functionally equivalent to a non-existing account, except that an empty account needs to be bookkept in the Ethereum state trie and thus increases the synchronization and transaction processing time. This means that an attacker can incur a large number of empty accounts to substantially increase the the synchronization and transaction processing time, effectively causing a DoS attack [76], [77]. An empty account can be incurred by an attacker using the SUICIDE opcode to transfer zero Ethers to a non-existing account. This vulnerability was caused by the lack of control over that an empty account should not be included in the state trie. This vulnerability has been eliminated by the hard fork for EIP-161 [87], which removes those empty accounts from the state trie and prevents any empty account from being stored in the state trie.

C. Vulnerabilities at the consensus layer

1) Outsourceable puzzle ($\mathcal{V}_{29}$): This vulnerability was reported in [32]. Recall that Ethereum adopts the PoW puzzle called Ethash, which was meant to be ASIC-resistant and be able to limit the use of parallel computing (owing to that the vast majority of a miner's effort will be reading a dataset via the limited memory bandwidth). However, a crafty miner can still divide the task of searching for a puzzle solution into multiple smaller tasks and then outsource them. This vulnerability is caused by that Ethash only makes the puzzle solution partially sequential in preimage search, rather than relying on an inherently sequential PoW. Several puzzles are proposed to cope with this problem [88], [89], which however have not been adopted by the Ethereum community.

2) 51% hashrate ($\mathcal{V}_{30}$): This vulnerability is inherent to PoW-based blockchains, where an attacker controlling a majority of the mining power can take over the blockchain [90], [91]. Such an attacker can reverse transactions by formulating blocks and performing double-spending at will. This vulnerability is inevitable to PoW-based consensus protocols.

3) Fixed consensus termination ($\mathcal{V}_{31}$): This vulnerability was first discussed in [92]. It refers to the fact that Ethereum's consensus protocol achieves a probabilistic agreement with a deterministic termination, meaning that a block is considered persistent in the main chain with a high probability if it is followed by a fixed number of m blocks. In other words, the consensus for block i will terminate when the chain depth reaches $i+m$ and all of the transactions contained in block i are committed (i.e., a merchant can take an external action, such as shipping goods). However, the probability of agreement can be affected by various factors [92]. For example, in the presence of communication delay, applications should wait for more blocks to be confirmed so as to achieve a higher security against double-spending. This vulnerability is caused by that no deterministic protocol can simultaneously guarantee *agreement*, *termination*, and *validity* in an asynchronous network [93].

4) DoS with block stuffing ($\mathcal{V}_{32}$): This vulnerability was first observed from the Fomo3D contract [94]. The vulnerability is that during a period of time, only the attacker's transactions are included in the newly-mined blocks and the other transactions are abandoned by the miners. This can happen when the attacker offers a higher *gasPrice* to incentivize the miners to select the attacker's transactions. This vulnerability is caused by the greedy mining incentive mechanism. At the moment of writing, there is no solution to prevent this vulnerability.

5) Rewards for uncle blocks ($\mathcal{V}_{33}$): This vulnerability was independently reported in [95], [33], [96]. The vulnerability

refers to that the uncle-rewarding mechanism reviewed above can lower the security of Ethereum when compared with Bitcoin. In particular, the uncle-rewarding mechanism reduces the risk for selfish miners that selectively release blocks to maximize their own profit because stale blocks may become uncle blocks and then receive rewards; this essentially incentivize selfish mining and double-spending. At the moment of writing, it is not clear how to eliminate this vulnerability.

D. Vulnerabilities at the network layer

1) Unlimited nodes creation ($\mathcal{V}_{34}$): This vulnerability was reported for the Geth client prior to its version 1.8 [35]. In the Ethereum network, each node is identified by a unique ID, which is a 64-byte ECDSA public key. An attacker could create an unlimited number of nodes on a single machine (i.e., with the same IP address) and use these nodes to monopolize the incoming and outgoing connections of some victim nodes, effectively isolating the victims from the other peers in the network. This vulnerability is caused by the weak restriction on the node generation process. This vulnerability can be eliminated by using a combination of IP address and public key as node ID. This countermeasure has not been adopted by the Geth developers who argue that it has a negative impact on the usability of the client.

2) Uncapped incoming connections ($\mathcal{V}_{35}$): This vulnerability was in the Geth client prior to its version 1.8 [35]. Each node can have a total number of `maxpeers` (with a default value 25) connections at any point in time, and can initiate up to $\lfloor(1+\texttt{maxpeers})/2\rfloor$ outgoing TCP connections with the other nodes. However, there was no upper limit on the number of incoming TCP connections initiated by the other nodes. This gives the attacker an opportunity to eclipse a victim by establishing an `maxpeers` many of incoming connections to a victim node, which has no outgoing connections. This vulnerability has been eliminated in Geth v1.8 by enforcing an upper limit on the number of incoming TCP connections to a node, with a default value $\lfloor\texttt{maxpeers}/3\rfloor = 8$.

3) Public peer selection ($\mathcal{V}_{36}$): This vulnerability was detected in Geth client prior to its version 1.8 [35]. Recall that the Ethereum P2P network uses a modified Kademlia DHT [97] for node discovery and that each node maintains a routing table of 256 *buckets* for storing information about the other nodes. The buckets are arranged based on the XOR distance between a node's ID and its neighboring node's ID [98]. When a node, say *A*, needs to locate a target node, *A* queries the 16 nodes in its bucket that are relatively close to the target node and asks each of these 16 nodes, say *B*, to return the 16 IDs of *B*'s neighbors that are closer to the target node. The process iterates until the target node is identified. However, the mapping from node IDs to buckets in the routing table is public, meaning that the attacker can freely craft node IDs that can land in a victim node's buckets and insert malicious node IDs into the victim node's routing table [35]. This vulnerability can be limited by making the "node IDs to buckets" mapping private. This countermeasure has not been adopted by the Geth developers who argue that it has a negative impact on the usability of the client.

4) Sole block synchronization ($\mathcal{V}_{37}$): This vulnerability was first reported in [99]. It allows an attacker to partition the Ethereum P2P network without monopolizing the connections of a victim client. Recall that each block header contains a *difficulty* field, which records the mining difficulty of the block. The total difficulty of the blockchain, denoted by *totalDifficulty*, is the sum of the difficulty of the blocks up to the present one. When a client, say *A*, receives from, say client *B*, a block of which the difference *totalDifficulty − difficulty* is greater than the *totalDifficulty* at the blockchain stored on client *A* (meaning that client *A* missed a number of blocks), *A* should start a block synchronization with *B*. Ethereum only allows a client to synchronize with one other client at a time (for network load considerations). This means that if client *B* is malicious and deliberately delays the synchronization in response to *A*'s request, the blockchain at client *A* is stalled and *A* rejects every subsequent block, which may facilitate double-spending and DoS attacks. This vulnerability can be mitigated by synchronizing *A* with multiple nodes, which however increases the network load.

5) RPC API exposure ($\mathcal{V}_{38}$): This vulnerability was first observed from the attack against the Geth and Parity clients [100]. The JSON-RPC of Ethereum clients provide various APIs for EOAs to communicate with the Ethereum network. For security purposes, the interface should only open locally and not be accessible from the internet. However, the standard port 8545 assigned to JSON-RPC can be accessed remotely in the Geth and Parity clients by default, which makes it possible for an attacker to call these remote clients via a JSON request [101]. Once having access to the remote client, the attacker can obtain sensitive data and perform certain unauthorized actions on the remote client. The vulnerability is caused by insecure API design and improper configuration. The vulnerability can be prevented by configuring the listening port (rather than using the default one) and adding access control to filter remote RPC calls.

E. Vulnerabilities in the Ethereum environment

We propose considering vulnerabilities in the environment with which Ethereum interacts because they also pose as threats to Ethereum.

1) Weak password ($\mathcal{V}_{39}$): This vulnerability was first observed from an attack against the DApp called Enigma [102]. The vulnerability refers to not only low-entropy passwords but also password reuse and insecure password storage. When the password of a DApp administrator is compromised, the attacker can manipulate the DApp's webpage at will to attack others.

2) Cross-Site Scripting (XSS; $\mathcal{V}_{40}$): This vulnerability was first observed from the attack against a cryptocurrency exchange, EtherDelta [103]. XSS is a browser-side vulnerability which allows an attacker to inject malicious JavaScript code into a HTTP webpage. To initiate a transaction via web-based DApp, an EOA may need to fetch its private key from a local file to the browser. An attacker can steal the private key by uploading a malicious JavaScript to the webserver, while claiming (for example) that the attacker is publishing information about exchanging some tokens. The webserver pushes the

attacker's token-exchange content to a victim's browser, which executes the malicious JavaScript code, causing the private key to be exposed to the attacker. This vulnerability can be prevented by letting a user sign transactions offline and then initiate transaction online, meaning that the private key is never loaded into the browser and cannot be compromised by the attack mentioned above.

3) Unvalidated URL redirection ($\mathcal{V}_{41}$): This vulnerability was first observed from an attack against the Ethereum wallet called MyEtherWallet [10]. Web applications frequently redirect users to other webpages. If the target URL is not authenticated, an attacker can replace the target URL with a phishing or malicious website, which renders the users vulnerable to attacks.

4) Broken access control ($\mathcal{V}_{42}$): This vulnerability was first observed from an attack against the decentralized application called CoinDash [104]. The access control failure at the server side causes an unauthorized access to the web application, which may result in the manipulation of the web page. Many DApps start with an Initial Coin Offering (ICO). A common practice is that participants obtain tokens by depositing Ether to a certain contract account, whose address is published on the website of the DApp. If an attacker breaks into the webserver and replaces the original contract address with its own, then the coin buyers will pay the attacker.

5) Unreliable Border Gateway Protocol (BGP) messages ($\mathcal{V}_{43}$): This vulnerability was first observed from an attack against MyEtherWallet [10]. Each BGP router maintains a routing table with routing paths between ASes. BGP peers broadcast IP prefixes they own and dynamically update their routing tables according to the IP prefixes they received from their neighbors. However, BGP has no mechanism for validating messages communicated between ASes, and therefore simply treats them as trusted [105]. An attacker can announce ownership of any prefix to its neighbor routers to redirect traffic towards the malicious ASes under its control. Then, the attacker can pretend to be a DNS server and redirect victims to malicious websites at IP addresses under the attacker's control.

6) Sensitive Domain Name System (DNS) servers ($\mathcal{V}_{44}$): This vulnerability was also observed from the attack against EtherDelta [106]. A wide range of vulnerabilities that can be exploited to incorrectly resolve DNS queries and redirect users to malicious websites, which can be achieved by hijacking DNS through cache poisoning, stale records, and compromising the resolver configuration file. [107].

F. Further analysis of vulnerability causes

Now we present a taxonomy of the root causes of the vulnerabilities reviewed above. As highlighted in Figure 5, the vulnerabilities are caused by incompetence or flaws in *smart contract programming*, *solidity language and toolchain*, *Ethereum design and implementation*, and *human, usability, and networking factors*.

1) Smart contract programming: These causes can be further divided into four sub-causes: *external dependence*, meaning a contract's execution relies on the behavior of an external contract; *improper validation*, meaning a failure in checking a condition allows the passing of an invalid input; *inadequate authentication or authorization*, causing failures in checking a caller's identity or privilege when the caller attempts to access a protected data item or functionality; and *uncontrolled gas consumption*, meaning a failure in gas allocation permits a DoS attack. These causes led to 14 types of vulnerabilities. Among these 14, only 4 (i.e., *Integer overflow and underflow ($\mathcal{V}_6$), erroneous visibility ($\mathcal{V}_9$), secrecy failure ($\mathcal{V}_{12}$), DoS with unbounded operations ($\mathcal{V}_{14}$)*) exist in traditional software, meaning that the other 10 are unique to Ethereum smart contract programming. This leads to:

Insight 1: Ethereum smart contracts introduce new kinds of vulnerabilities.

Insight 1 is interesting because it says that a new programming paradigm demands adequate training for programmers' secure coding practice. Indeed, among the 26 types of application-layer vulnerabilities, 20 can be prevented by using best practices, highlighting the importance of engineering best practices.

2) Solidity language and tool chain: These causes can be further divided into five sub-causes, among which four sub-causes are related to the improper design of the Solidity language (i.e., *inconsistent exception handling* between direct call and low-level calls; *undefined behavior* of uninitialized storage pointers; *improper syntax* of constructor function; and *weak type system* with flexible typing rules) and the other sub-cause is *buggy compiler* (i.e., insufficient tool chain support). These five causes contributed 5 types of vulnerabilities. This leads to the following insight, which is clearly interesting and important.

Insight 2: It is important to design more secure programming languages and have more secure supporting tools for programmers to write secure smart contracts.

3) Ethereum design and implementation: These causes can be further divided into 14 sub-causes, which are related to, among other things, EVM, blockchain, PoW consensus, incentive mechanism, and P2P protocol. These causes cut across the application layer, the data layer, the consensus layer, and the network layer. The root causes related to EVM include: (i) *missing input check*, meaning no check on the validity of a transaction's data; (ii) *missing orphan proof*, meaning no check on a nonexistent recipient address; (iii) *improper execution model*, meaning the behavior of EVM is not properly specified; and (iv) *improper gas costs*, meaning the gas costs of EVM opcodes are not properly specified.

The root causes related to blockchain include: (i) *flexible block creation*, meaning there are no restrictions on miners when they creating blocks, allowing them to create blocks to favor themselves; (ii) *insufficient transaction information*, meaning that a transaction can be accepted by multiple Ethereum blockchains (e.g, ETH, ETC, Ropsten), rather than a specific blockchain, owing to the lack of information specifying a unique blockchain; and (iii) *uncontrolled state trie*, meaning that there are no restrictions on the accounts that can be stored in the state trie.

The root causes related to the PoW consensus protocol include: (i) *partially sequential PoW*, meaning that Ethereum's PoW puzzle can still be outsourced to different miners; (ii)

inherent defect of PoW, meaning that a miner with 51% hashrate can extend the main chain at will; and (iii) *termination first*, meaning that the Ethereum consensus protocol prefers termination to agreement.

The root causes related to incentive mechanism include: (i) *greedy incentive*, meaning that a miner always selects transactions with higher *gas Price*; and (ii) *incompatible incentive*, meaning that a miner can get a higher payoff by deviating from the consensus protocol.

The root causes related to the P2P protocol include: (i) *improper RLPx protocol*, meaning that the node discovery and routing algorithms are not properly designed; and (ii) *improper Ethereum wire protocol*, meaning that the blockchain synchronization algorithm is not properly designed.

The root causes mentioned above contributed 18 vulnerabilities, cutting across the application layer, the data layer, the consensus layer, and the network layer; 9 out of these 18 vulnerabilities are still largely open (i.e., having yet to be eliminated). In summary, we draw:

Insight 3: The vulnerabilities caused by the design and implementation of the Ethereum blockchain are harder to cope with than the other vulnerabilities.

Insight 3 is interesting because it highlights the difficulty in designing and implementing a secure blockchain platform, despite (as we will see later) that the vulnerabilities in blockchain may not cause as big losses as the vulnerabilities at the application layer.

4) Human, usability and networking factors: These causes can be further divided into four sub-causes: *improper configuration*, meaning that an Ethereum client is installed with incorrect permissions; *insufficient authentication*, meaning the use of weak or exposed passwords; *faulty web development*, meaning that a DApp's web interface is vulnerable; and *exposed Internet service*, meaning attacks coming from the Internet are not blocked. These causes led to seven vulnerabilities. In summary, we draw the following interesting insight.

Insight 4: Vulnerabilities in the Ethereum environment are largely caused by human, usability, and networking factors, which however consist of a smaller fraction of the vulnerabilities when compared with the vulnerabilities that are inherent to the Ethereum blockchain and smart contracts.

5) Cross-cutting analysis: We observe that there are 13 vulnerabilities that are largely open. In order to draw insights into the approaches that may be able to eliminate them, we look deeper into their root causes (than what were mentioned above). First, we consider the vulnerabilities that are inherent to the Ethereum blockchain design, including: *under-priced opcodes* ($\mathcal{V}_{23}$), which are inherent to the EVM gas cost mechanism; *outsourceable puzzle* ($\mathcal{V}_{29}$), *51% hashrate* ($\mathcal{V}_{30}$), and *fixed consensus termination* ($\mathcal{V}_{31}$), which are inherent to the PoW consensus mechanism; and *DoS with block stuffing* ($\mathcal{V}_{32}$) and *rewards for uncle blocks* ($\mathcal{V}_{33}$), which are inherent to the Ethereum incentive mechanism. In order to eliminate or mitigate these vulnerabilities, alternate gas cost mechanisms, consensus mechanisms, and incentive mechanisms may have to be used.

Second, we consider the vulnerabilities that are inherent to the implementation of Ethereum clients, including *unlimited*

nodes creation ($\mathcal{V}_{34}$), *public peer selection* ($\mathcal{V}_{36}$), and *sole block synchronization* ($\mathcal{V}_{37}$). In order to eliminate or mitigate these vulnerabilities, alternate network protocols may have to be used without sacrificing platform usability or network performance.

Third, we consider the vulnerabilities that are inherent to the insecure environment in which Ethereum runs, including *unreliable BGP messages* ($\mathcal{V}_{43}$) and *sensitive DNS servers* ($\mathcal{V}_{44}$). In order to eliminate or mitigate these vulnerabilities, extra effort must be made to defend the environment regardless how Ethereum is designed and implemented.

Fourth, we consider the vulnerabilities that are inherent to the application-layer flexibility, including *upgradable contract* ($\mathcal{V}_4$) and *type casts* ($\mathcal{V}_{18}$). In order to eliminate or mitigate these vulnerabilities, alternate contract design pattern and type system may have to be used.

Summarizing the preceding discussion, we draw:

Insight 5: There are many vulnerabilities that are yet to be tackled in order to adequately defend Ethereum (or blockchain-based DApp systems in general), cutting across of every layer of the architecture, from the application layer, the data layer, the consensus layer, the network layer, down to the environment layer.

IV. ATTACKS

Corresponding to the presentation of vulnerabilities, we group the 26 attacks we consider according to the locations of the vulnerabilities they exploit. We describe each attack from the following perspective: its *history*, *cause*, *tactic*, and *direct impact* (i.e., the consequence when the attack succeeds). For ease of reference, we denote the 26 attacks by $\mathcal{A}_1, \ldots, \mathcal{A}_{26}$, respectively. Note that some of these 26 attacks may correspond to the same type of attacks (i.e., sharing the same name), but they exploit different vulnerabilities or vulnerability combinations and/or cause different consequences. For example, both $\mathcal{A}_2$ and $\mathcal{A}_3$ correspond to the *parity multisignature wallet attack*; however, $\mathcal{A}_2$ exploits the *delegatecall injection* vulnerability ($\mathcal{V}_2$) and the *erroneous visibility* vulnerability ($\mathcal{V}_9$) to cause the consequence of *unauthorized code execution*, while $\mathcal{A}_3$ exploits the *frozen ether* vulnerability ($\mathcal{V}_3$) and the *unprotected suicide* vulnerability ($\mathcal{V}_{10}$) to cause the consequence of *DoS*.

A. Attacks at the application layer

1) The DAO attack ($\mathcal{A}_1$): The contract `DAO` is a financial application running on top of Ethereum. In June 2016, it was attacked to cause the loss of US$60M [8]. The financial application is that investors (i.e., token holders) vote on investment proposals for spending their money (i.e., "investment crowdsourcing"). Once a proposal is approved by a majority, the amount of money approved by the supporting investors is moved to the proposer's account, and the amount of money owned by the minority that opposes the proposal is respectively "refunded" to each of them via some newly created contract accounts. This mechanism was implemented in the splitDAO() function in Figure 8, which shows how the attack exploits the *reentrancy* vulnerability ($\mathcal{V}_1$). Specifically,

the attack proceeds as follows: When a minority investor who opposes the proposal requests for a "refund", the `DAO` contract creates a new `DAO` contract account (Step 1-2) and transfers the requesting investor's money to the new `DAO` contract account (Step 3-4). The requesting investor may receive a reward for its past contribution or activity (Step 5-6). The vulnerability is that the number of tokens "refunded" to the requesting investor depends on the state variables `balances[msg.sender]` and `totalSupply`, which are updated in the end of the splitDAO() function after the external call msg.sender.call.value(). This allowed the attacker to recursively call the splitDAO() function (Step 7) before the state variables are updated, causing a malicious investor to draw more money than it deserves [39].
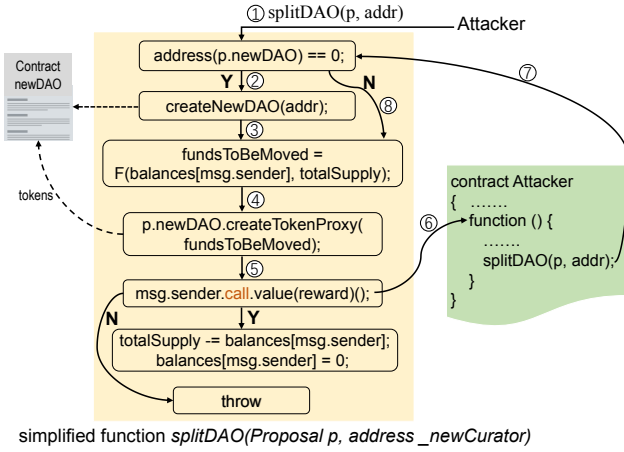

simplified function *splitDAO(Proposal p, address _newCurator)*

Fig. 8: The DAO attack exploiting the reentrancy vulnerability ($\mathcal{V}_1$).

2) Parity multisignature wallet attacks ($\mathcal{A}_2$ and $\mathcal{A}_3$): In Ethereum, a multisignature wallet is a smart contract which requires multiple private keys to unlock a wallet in order to safeguard Ether or tokens. As shown in Figure 9, a multisignature wallet supported by the Parity client consists of two contracts: (i) a library contract called `WalletLibrary`, which implements all of the core functions of a wallet; and (ii) an actual `Wallet` contract, which holds a reference (i.e., `_walletLibrary`) that forwards all of the unmatched function calls to the library contract via delegatecall (Line 7). The Parity multisignature wallet was compromised twice in 2017, which are briefly reviewed below.

The first attack, denoted by $\mathcal{A}_2$, exploited the *delegatecall injection* vulnerability ($\mathcal{V}_2$) and the *erroneous visibility* vulnerability ($\mathcal{V}_9$) to drain Ethers approximately worthy of US\$31M. The attacker took over the ownership of contract `Wallet` by sending a transaction to the contract with `msg.data` containing initWallet() as the callee function (Line 13). Since contract `Wallet` did not provide a function named initWallet(), the contract's fallback function was triggered to delegate the wallet initialization task to the function initWallet() in `WalletLibrary`, which replaced the original multi-owner of contract `Wallet` with the attacker's address specified in `msg.data`. This attack succeeds when the following four conditions are satisfied simultaneously [42], [108]: (i) the function initWallet() in the library was not specified as an *internal* one,

```
1  contract Wallet {
2      address _walletLibrary = new WalletLibrary();
3      address owner;
4      ...
5      function() payable {
6          if (msg.data.length > 0)
7              _walletLibrary.delegatecall(msg.data);
8      }
9  }
10
11 contract WalletLibrary {
12     ...
13     function initWallet(address[] _owners, uint _required, uint
                ↪ _daylimit){
14         initDaylimit(_daylimit);
15         initMultiowned(_owners, _required);
16     }
17 }
```

Fig. 9: Simplified vulnerable Parity multisignature wallet.

meaning that it can be externally called via delegatecall; (ii) the `WalletLibrary` was actually a stateful contract, meaning that it can change the state of `Wallet`; (iii) the function initWallet() did not check whether the `Wallet` contract had already been initialized (if so, no more initialization should be done); (iv) the `Wallet`'s fallback function did not check the function being called, but unconditionally forwarded any unmatched calldata to `WalletLibrary`, allowing unintended invocations.

The second attack, denoted by $\mathcal{A}_3$, exploited the *unprotected suicide* vulnerability ($\mathcal{V}_{10}$) and the *frozen Ether* vulnerability ($\mathcal{V}_3$), freezing approximately US\$280M in the affected wallets forever [109]. In response to the first attack, the Parity developers added a modifier, `only_uninitialized`, to protect function initWallet() such that a re-initialization of `Wallet` via delegatecall will throw an exception and be rejected by the modifier [110]. However, the shared `WalletLibrary` itself was left uninitialized, which allowed an attacker to bypass the `only_uninitialized` modifier and set himself as the owner of the `WalletLibrary` [111]. Once taking over the library, the attacker invoked the suicide method to kill the library, causing all of the `Wallet` contracts relying on the library unusable.

3) BECToken attack ($\mathcal{A}_4$): BECToken, an ERC-20 contract, was attacked in April 2018 by an exploitation of the *integer overflow* vulnerability ($\mathcal{V}_6$), causing a massive amount of tokens stolen and a temporary shutdown of token trading at exchange [112]. The vulnerability was in the function batchTransfer() shown in Figure 10, and the function was meant for users to transfer tokens to multiple recipients via two arguments, one specifying the array of the recipients' addresses and the other specifying the respective number of tokens. The statement at Line 3 calculates the total number of tokens the sender should pay for a particular transaction, but may have the following integer overflow: By setting `_value` to 2^{255} and `_receivers` to two accounts controlled by the attacker, the attack overflows the 256-bit variable `amount` and makes it zero [113]. As a consequence, the attack bypasses the two checks at Lines 4-5 and causes sending the two receivers extremely large numbers of tokens.

```
1  function batchTransfer(address[] _receivers, uint256 _value) public
       ↪ whenNotPaused returns (bool) {
2      uint cnt = _receivers.length;
3      uint256 amount = uint256(cnt) * _value;
4      require(cnt > 0 && cnt <= 20);
5      require(_value > 0 && balances[msg.sender] >= amount);
6      ... // transfer to specified recipients
7  }
```

Fig. 10: The vulnerable function in BECToken.

4) GovernMental attacks ($\mathcal{A}_5$, $\mathcal{A}_6$, $\mathcal{A}_7$, and $\mathcal{A}_8$): The contract `GovernMental` was an array-based pyramid Ponzi scheme, where the last participant wins a jackpot if no one joins the scheme within 12 hours after the last participant [114]. The contract has four vulnerabilities [12], which allowed the following four attack tactics and explains why $\mathcal{A}_5$, $\mathcal{A}_6$, $\mathcal{A}_7$, and $\mathcal{A}_8$ belong to the same type of attacks. The first attack, denoted by $\mathcal{A}_5$, exploits the *DoS with unbounded operations* vulnerability ($\mathcal{V}_{14}$) that when the array bookkeeping the number of participants becomes too large, the amount of gas required for operating on the array will go beyond the maximum gas that is permitted for assembling a block. This effectively halts the transaction and the winner cannot receive the 1,100 ETH jackpot. The second attack, denoted by $\mathcal{A}_6$, exploits the *unchecked call return value* vulnerability ($\mathcal{V}_{15}$) that the contract does not check the returned value when sending profits to the winner and the *call-stack depth limit* vulnerability ($\mathcal{V}_{22}$). As a consequence, the owner of contract `GovernMental` to fail the payment as follows: (i) Calling 1024 contracts before calling the target callee contract of the payee, which causes the callee contract to return FALSE to the caller contract, meaning that the callee contract does not receive the payment. (ii) The caller contract is supposed to check this returned value and then proceeds correspondingly but it does not. This causes the callee contract to lose money, which now belongs to the owner of the caller contract. The third attack, denoted by $\mathcal{A}_7$, exploits the *transaction-ordering dependence* vulnerability ($\mathcal{V}_{24}$) that a malicious miner can abandon some transactions related to `GovernMental` or reorder transactions to make itself the last player (i.e., winner) in each round. The fourth attack, denoted by $\mathcal{A}_8$, exploits the *timestamp dependence* vulnerability ($\mathcal{V}_{25}$) that a malicious miner can manipulate `block.timestamp` so that its own block appears to be the last block to make itself win.

5) HYIP attack ($\mathcal{A}_9$): The contract `HYIP` was another Ponzi scheme, which pays existing investors from funds contributed by new investors at the end of each day. This mechanism is implemented by the function **performPayouts()** highlighted in Figure 11, which contains the *DoS with unexpected revert* vulnerability ($\mathcal{V}_5$) (Line 7) [114]. The attack proceeds as follows: (i) The attacker, say Alice, writes an exploitation contract, named `Mallory`, in which the attacker invests and throws an exception in the fallback function (Line 14). (ii) When function **performPayouts()** is called to pay the investors, the fallback function is invoked and throws an exception, causing a reversion of the money transfer (Line 7) and thus DoS to contract `HYIP`. (iii) The attacker can blackmail

`HYIP` to pay a ransom for halting its attack, by undoing the **throw** operation (Line 15) via function **stopAttack** (Line 18), which can only be done by the contract owner, Alice.

```
1   contract HYIP {
2       Investor[] private investors;
3       ...
4       function performPayouts() {
5           for(uint idx = investors.length; idx−− > 0; ) {
6               uint payout=(investors[idx].amount*33)/1000;
7               if(!investors[idx].addr.send(payout)) throw;
8           }
9   }}
10
11  contract Mallory {
12      bool private attack = true;
13      ...
14      function() payable {
15          if (attack) throw;
16      }
17      function stopAttack() {
18          if(msg.sender == owner) attack = false;
19      }
20  }
```

Fig. 11: Simplified HYIP contract and attack.

6) Fomo3D attacks ($\mathcal{A}_{10}$ and $\mathcal{A}_{11}$): The contract `Fomo3D` was an extremely popular Ponzi game in year 2018, where the last participant who buys a key before the timer runs out won the jackpot. The price of keys gradually grows with the number of buyers. When a key was sold, the countdown extends for 30 seconds. In addition to the jackpot winner, Fomo3D implemented an airdrop lottery to attract participants. For each purchase over 0.1 ETH, the participant (i.e., buyer) had a random chance to be picked up for a tiny profit from the prize pool. These two incentive mechanisms can be attacked as follows [115].

The first attack, denoted by $\mathcal{A}_{10}$, is against the airdrop mechanism. Specifically, the attack exploited the *generating randomness* vulnerability ($\mathcal{V}_{26}$). As shown in Figure 12, function **airdrop()** generates a random seed by performing a deterministic computation on the current block state (i.e., `block.timestamp`, `block.difficulty`, etc.) and the address of `msg.sender` (Line 2-8). If the seed satisfies a certain condition (Line 10), the current key buyer wins an airdrop. However, since the block information is predictable, an attacker can simply pre-compute the addresses of new contracts and brute-forces for the winning seed (Line 2).

The second attack, denoted by $\mathcal{A}_{11}$, is against the winning mechanism [115]. Specifically, the attack exploited the *DoS with block stuffing* vulnerability ($\mathcal{V}_{32}$) and causes the attacker to win the prize of approximately US\$3M [94]. The attack proceeds as follows: When the timer of the game reached about three minutes, the attacker bought a key and then sent multiple transactions to his own accounts with high enough *gasPrice*. Owing to the choice of miners, these transactions were first assembled into blocks. Since the maximum amount of gas consumption for a block is limited, any transactions related to `Fomo3D` were not assembled into blocks. By congesting the network until the game was over, the attacker succeeded in becoming the last player.

```
1  function airdrop() private view returns(bool) {
2      uint256 seed= uint256(keccak256(abi.encodePacked(
3          (block.timestamp).add
4          (block.difficulty).add
5          ((uint256(keccak256(abi.encodePacked(block.coinbase)))) / (
                ↪ now)).add
6          (block.gaslimit).add
7          ((uint256(keccak256(abi.encodePacked(msg.sender)))) / (
                ↪ now)).add
8          (block.number)
9      )));
10     if((seed−((seed/1000)*1000)) < airDropTracker_)
11         return(true);
12     else
13         return(false);
14     }
15 }
```

Fig. 12: A snippet source code of Fomo3D.

7) ERC-20 signature replay attack ($\mathcal{A}_{12}$): This attack, which was first reported in [116], exploits the *insufficient signature information* vulnerability ($\mathcal{V}_{13}$). When a user transfers ERC-20 tokens, the user must have enough Ether to pay the transaction fee, which can be inconvenient when the user does not own any Ether. In order to alleviate the problem, the proxy-transfer method is introduced such that a user can authorize a proxy to carry out a transaction and pay the proxy some extra tokens as its service fee. As shown in Figure 13, when Alice is to transfer 100 MTC tokens to Bob, she can send a signed message off-chain to a proxy (Step 1) such that the proxy launches a transaction to transfer 100 tokens to Bob and receives 3 token from Alice for the service (Step 2). A signature is verified using function transferProxy(), which uses the Solidity function ecrecover() (Line 5) to identify Alice's account address that issued the signature. However, Alice's off-chain message may not provide her token contract address, which should be bound to her signature. As a consequence, the signature can be accepted as valid with respect to any token contract address (e.g., MTC, UGToken, and GGoken), meaning that Bob can replay the signed message to other kinds of token contracts, such as UGToken (Step 4), to obtain extra money from Alice (Step 5) [57].

8) Rubixi attack ($\mathcal{A}_{13}$): The Rubixi contract is a Ponzi scheme which contains the *erroneous constructor name* vulnerability ($\mathcal{V}_{17}$). The contract was originally named DynamicPyramid and was later renamed by the developer to Rubixi. However, the contract's constructor name was not updated accordingly, allowing anyone that calls the public function DynamicPyramid() to become the owner of the contract and therefore steal the funds of the contract.

B. Attacks at the data layer

1) ETH and ETC cross-chain replay attack ($\mathcal{A}_{14}$): Ethereum had a hard fork after the DAO attack, splitting into ETH and ETC that share the same transaction history. This means that a transaction that was validated by the ETH network could also be accepted by the ETC network when the recipient immediately rebroadcast the transaction on the ETC network, and vice versa [117]. Since both ETC and ETH networks did not implement any defense against this attack
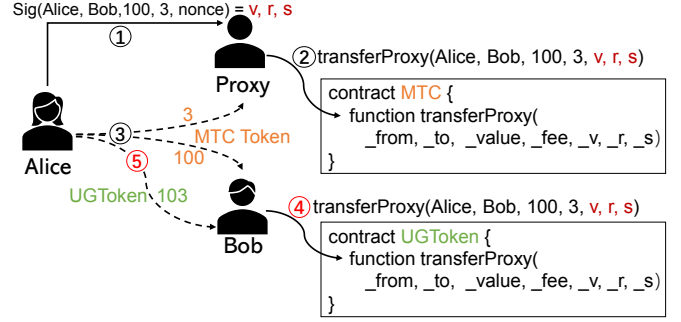

```
1  function transferProxy(address _from, address _to, uint256 _value,
      ↪ uint256 _fee, uint8 _v,bytes32 _r, bytes32 _s) returns (bool){
2      if(balances[_from] < _fee + _value) throw;
3      uint256 nonce = nonces[_from];
4      bytes32 h = sha3(_from, _to, _value, _fee, nonce);
5      if(_from != ecrecover(h, _v, _r, _s)) throw;
6      if(balances[_to] + _value < balances[_to] || balances[msg.sender] +
          ↪ _fee < balances[msg.sender]) throw;
7      balances[_to] += _value;
8      Transfer(_from, _to, _value);
9      balances[msg.sender] += _fee;
10     Transfer(_from, msg.sender, _fee);
11     balances[_from] −= _value + _fee;
12     nonces[_from] = nonce + 1;
13     return true;
14 }
```

Fig. 13: Cross-contract replay attack via transferProxy().

that exploited the *indistinguishable chains* vulnerability ($\mathcal{V}_{27}$), exchanges participating in both chains (e.g., Coinbase and Yunbi) lost a large amount of money [85].

2) Under-priced DDoS attacks ($\mathcal{A}_{15}$ and $\mathcal{A}_{16}$): These attacks, which were reported in [75], [76], [77], exploited both application-layer and data-layer vulnerabilities.

Specifically, the first attack, denoted by $\mathcal{A}_{15}$, exploited the *under-priced opcodes* vulnerability ($\mathcal{V}_{23}$) owing to the improper gas cost of EVM's EXTCODESIZE opcode. Prior to the EIP 150 hard fork, the EXTCODESIZE opcode only charged 20 gas for reading a contract's bytecode from disk and then deriving its length. As a consequence, the attacker can repeatedly send transactions to invoke a deployed smart contract with many EXTCODESIZE opcodes to cause a 2-3x slower block creation rate [75].

The second attack, denoted by $\mathcal{A}_{16}$, exploited the *under-priced opcodes* vulnerability ($\mathcal{V}_{23}$) owing to the improper gas cost of EVM's SUICIDE opcode and the *empty account in the state trie* vulnerability ($\mathcal{V}_{28}$). The SUICIDE opcode (renamed to SELFDESTRUCT after EIP 6) is meant to remove a deployed contract from the blockchain and send the remaining balance of Ether to the account designated by the caller. When the target account does not exist, a new account is created even though no Ether may be transferred; this consumes merely 90 gas [76]. Since an existing empty account is stored in the Ethereum state trie, the attacker created 19 million new empty accounts via the SUICIDE opcode at a low gas consumption, which wasted considerable disk space while increasing the synchronization and transaction processing time.

C. Attacks at the consensus layer

1) Ethereum Classic (ETC) 51% attack ($\mathcal{A}_{17}$): In January 2019, ETC suffered from a 51% attack that exploited the *51% hashrate* vulnerability ($\mathcal{V}_{30}$), in which the attacker carried out double-spending transactions against several exchanges, causing an estimated loss of US$1.1M [118]. Since 2018, ETC's mining hashrate had dropped significantly due to the declining in its price, which lowered the amount of computing resources that are required for launching a 51% attack. Moreover, cloud mining services (e.g., NiceHash) make it even easier to launch 51% attacks. The attack was disrupted when exchanges increased the number of blocks that are required for transaction confirmation and limited the participation of malicious addresses in ETC trade.

2) Selfish-mining attack ($\mathcal{A}_{18}$): The selfish-mining attack, which exploits the *rewards for uncle blocks* vulnerability ($\mathcal{V}_{33}$), means that miners may withhold their newly mined blocks and selectively publish some of their blocks to earn an unfair share of reward. A selfish miner continuously monitors the situation on a blockchain's public branches, estimates its advantages, and reveals its private blocks accordingly. When the public branches are shorter than the selfish-miner's private branch, the honest miners will switch to latter, rendering their previous mining effort useless and making the selfish-miners receive a higher reward. Ritz et al. [96] conducted a Monte Carlo simulation to emulate the block generation process in Ethereum and quantified the impact of uncle rewards on selfish-mining. Their simulation results showed that the uncle-block reward mechanism not only lowered the threshold of computational power at which selfish-mining becomes profitable, but also weakened the overall resilience against other attackers such as *double-spending*. Niu et al. [33] developed a mathematical analysis on a selfish-mining strategy through a 2-dimensional Markov process model, showing that Ethereum is more vulnerable to selfish-mining than Bitcoin.

3) Balance attack ($\mathcal{A}_{19}$): This attack exploits the *fixed consensus termination* vulnerability ($\mathcal{V}_{31}$) owing to the probabilistic PoW consensus in the presence of communication delays (or asynchronous networks) and the *unreliable BGP messages* vulnerability ($\mathcal{V}_{43}$). This attack was first reported in [119] via a theoretic analysis and testnet demonstration. The attack is to transiently partition the network into multiple subgroups of similar mining power so as to launch the double-spending attack at a subgroup of lower mining power. This allows the attacker to initiate transactions with merchants in one subgroup, while mining blocks in another subgroup to make its subtree outweighs the subtree mined in the victim group. After transactions in the victim subgroup get confirmed, the attacker reconnects the network. Since the mining power is roughly equally distributed among the subgroups, the subtree broadcast by the attacker has a good chance to be selected as the main chain, meaning that the attacker can double-spend in the victim subgroup. This attack was later deemed only theoretically possible because partitioning a public Ethereum network (e.g., using BGP-hijacking) may not be feasible in practice [120].

D. Attacks at the network layer

1) Account hijacking attack ($\mathcal{A}_{20}$): This attack exploits the *RPC API exposure* vulnerability ($\mathcal{V}_{38}$). In order to sign transactions, an EOA must first decrypt its private key that is stored on the local host and encrypted with a passphrase. This can be achieved by using the unlockAccount() API of an Ethereum client, which uses the passphrase to obtain the private key and loads it into the memory for 300 seconds (by default). The private key in memory can be accessed by any Ethereum API without authentication. This can be exploited as follows. Ethereum clients (e.g., Geth, Parity) typically use the default ports 8545 (HTTP) and 8546 (WebSocket) for the JSON-RPC interface. However, these client software neither configure those ports as local-only by default, nor adopt precautions (e.g., disabling remote calls). This allows an attacker to scan open Ethereum nodes and invoke eth_sendTransaction() API to transfer victims' money to the attacker's account. Once a victim types its passphrase to unlock its account, the eth_sendTransaction() API will be successfully executed. By the time the attack was observed in March 2018 [100], attackers had already stolen around US$20M from exposed Ethereum clients.

2) Eclipse attacks ($\mathcal{A}_{21}$ and $\mathcal{A}_{22}$): This attack allows an attacker, who can hijack the connections of some victim nodes in the P2P network, to create a virtual partition to isolate those victim nodes from the rest of the network. Victim nodes' connections can be hijacked at least by connection monopolization and poisoning victims' routing tables [35]. The connection monopolization attack, denoted by $\mathcal{A}_{21}$, exploits the *unlimited nodes creation* vulnerability ($\mathcal{V}_{34}$) and the *uncapped incoming connections* vulnerability ($\mathcal{V}_{35}$). When a client reboots, it has no incoming or outgoing connections. An attacker can create plenty of node IDs in advance and initiate enough incoming connections to the victim node immediately after its reboot. A node is eclipsed when its connection slots (25 by default) are occupied by incoming connections from the attacker. Despite the response to the aforementioned connection monopolization attack in imposing an upper limit on the number of incoming TCP connections, the following attack, denoted by $\mathcal{A}_{22}$, still can succeed. Specifically, an attacker still can exploit the *public peer selection* vulnerability ($\mathcal{V}_{36}$) to poison the victim nodes' routing tables when these tables are reboot and reset (e.g., the attacker could craft fake nodes and insert them into those routing tables to make the victim nodes' outgoing TCP connections point to the fake nodes controlled by the attacker). The attacker can further occupy the victim nodes' remaining connection slots by initiating connections to the victim nodes.

E. Attacks against the environment

1) EtherDelta attack ($\mathcal{A}_{23}$): EtherDelta is a popular exchange for users to trade ERC-20 tokens. It suffered from a code injection or EtherDelta attack that exploited the *Cross-Site Scripting* vulnerability ($\mathcal{V}_{40}$) in September 2017, causing a loss in cryptocurrency worth of thousands of dollars [103]. The attacker constructed a new token contract, which contains a piece of malicious JavaScript code in the token's name. Recall that EtherDelta extracts a newly created token's name

from the token contract's code and displays the name on the exchange's website. Recall also that when performing a token-trade transaction, a user needs to load its private key and account address to the web browser for signing the transaction. As a consequence, when the name of a newly created token was displayed on the user's browser, the malicious JavaScript code was executed to steal the user's private key from the browser, causing the loss of money protected by the private key.

2) CoinDash attack (A_{24}): CoinDash, a portfolio management platform, was compromised due to an exploitation of the *broken access control* vulnerability (V_{42}) during the course of its ICO in Jul 2017, causing the loss of US$7M worth of Ether within a few minutes. The attack proceeds by getting into the infected web application that hosts CoinDash's webpage, and replacing the ICO contract address with one that is controlled by the attacker [104].

3) Enigma attack (A_{25}): Enigma, a decentralized investment platform, was attacked prior to its ICO in August 2017 [102], owing to the exploitation of the *weak password* vulnerability (V_{39}). Through social engineering means, the attacker successfully stole an Enigma founder's password, which was disclosed in an unrelated attack and reused in Enigma. As a consequence, the attacker took control of the company's Slack channel, email lists, and Google account hosting the ICO's presale. The attacker replaced the official ICO contract address with the attacker's own address and sent messages to solicit buyers in fake presales.

4) MyEtherWallet attack (A_{26}): MyEtherWallet, a popular web-based wallet, was attack in April 2018 owing to the exploitation of the *unreliable BGP messages* vulnerability (V_{43}), the *sensitive DNS* vulnerability (V_{44}), and the *unvalidated URL redirection* vulnerability (V_{41}). The attack exploits a joint BGP and DNS hijacking to mislead users to a fake version of the website and compromised the victims' private keys [10]. The hacker then emptied the victims' wallets and stole approximately US$17M. The attacker first exploited a routing weakness in BGP and hijacked the traffic to Amazon's Route 53 servers, which provided DNS service to MyEtherWallet. When users visit their MyEtherWallet, their requests were rerouted to the fake DNS servers controlled by the attacker, which returned IP addresses to direct users to a phishing website. The unsuspecting users, who neglected the TLS/SSL certificate warning sign, proceeded to login to the phishing site, causing their passphrases and private keys exposed.

F. Further analysis of attack consequences

Now we present a taxonomy of the attack consequences mentioned above: *unauthorized code execution*, *DoS*, *unfair income*, *double-spending*, *private key leakage*, and *webpage manipulation*, which are elaborated below. Figure 14 highlights the relationship between these attack consequences, attacks, and vulnerabilities, where the vulnerabilities and attacks are ordered in such a way that allows us to clearly present the relationships between them (i.e., neither the vulnerabilities nor the attacks are ordered as they are discussed in the text). For ease of reference, we use "$A_i(V_j, \cdots)$" to denote that attack A_i exploits vulnerabilities V_j and possibly others.

1) Unauthorized code execution: This consequence occurs in one of the following four scenarios. (i) The attacker uses the DAO attack (A_1) to exploit the *reentrancy* vulnerability (V_1), namely $A_1(V_1)$. (ii) The attacker uses the first Parity multisignature wallet attack (A_2) to exploit the *delegatecall injection* vulnerability (V_2) and the *erroneous visibility* vulnerability (V_9), namely $A_2(V_2, V_9)$. (iii) The attacker uses the BECToken attack (A_4) to exploit the *Integer overflow* (V_6), namely $A_4(V_6)$. (iv) The attacker wages the Rubixi attack (A_{13}) to exploit the *erroneous constructor name* (V_{17}), namely $A_{13}(V_{17})$. The biggest financial loss caused by a single unauthorized code execution is US$60M, owing to the DAO attack (A_1) [8].

2) DoS: This consequence can be further divided into two sub-consequences: *DoS against smart contracts* and *DoS against Ethereum network*. The former sub-consequence occurs in one of the following four scenarios: (i) The attacker wages the second Parity multisignature wallet attack (A_3) to exploit the *frozen Ether* vulnerability (V_3) and the *unprotected suicide* vulnerability (V_{10}), namely $A_3(V_3, V_{10})$. (ii) The attacker uses the HYIP attack (A_9) to exploit the *DoS with unexpected revert* vulnerability (V_5), namely $A_9(V_5)$. (iii) The attacker uses the GovernMental attack (A_5) to exploit the *DoS with unbounded operations* vulnerability (V_{14}), namely $A_5(V_{14})$. (iv) The attacker uses the Fomo3D attack (A_{11}) to exploit the *DoS with block stuffing* vulnerability (V_{32}), namely $A_{11}(V_{32})$. The latter sub-consequence occurs in one of the following scenarios: (i) The attacker uses the under-priced DDoS attack (A_{15}) to exploit the *under-priced opcodes* vulnerability (V_{23}), namely $A_{15}(V_{23})$. (ii) The attacker uses the under-priced DDoS attack (A_{16}) to exploit the *under-priced opcodes* vulnerability (V_{23}) and the *empty account in the state trie* vulnerability (V_{28}), namely $A_{16}(V_{23}, V_{28})$. The biggest financial loss caused by a single DoS attack is US$280M, owing to the second attack against the Parity wallet (A_3) [109].

3) Unfair income: This consequence occurs in one of the following scenarios: (i) A malicious contract owner uses the GovernMental attack (A_6) to exploit the *unchecked call return value* vulnerability (V_{15}) and the *call-stack depth limit* (V_{22}) vulnerability to hold the money that should be transferred to a callee contract or EOA, namely $A_6(V_{15}, V_{22})$. (ii) A malicious miner wages the GovernMental attack (A_7) to exploit the *transaction-ordering dependence* vulnerability (V_{24}), namely $A_7(V_{24})$. (iii) A malicious miner wages the GovernMental attack (A_8) to exploit the *timestamp dependence* vulnerability (V_{25}), namely $A_8(V_{25})$. (iv) The attacker wages the Fomo3D attack (A_{10}) to exploit the *generating randomness* vulnerability (V_{26}) to make itself the winner, namely $A_{10}(V_{26})$. (v) An attacker uses the ERC-20 signature replay attack (A_{12}) to exploit the *insufficient signature information* vulnerability (V_{13}), namely $A_{12}(V_{13})$. (vi) The attacker uses the ETH and ETC cross-chain replay attack (A_{14}) to exploit the *indistinguishable chains* vulnerability (V_{27}) to repeatedly withdraw money, namely $A_{14}(V_{27})$. (vii) A malicious miner wages the selfish mining attack (A_{18}) to exploit the *rewards for uncle blocks* vulnerability (V_{33}) to get a higher mining reward, namely $A_{18}(V_{33})$. The biggest financial loss caused by a single unfair income attack is US$0.5M, owing to an ETH and ETC

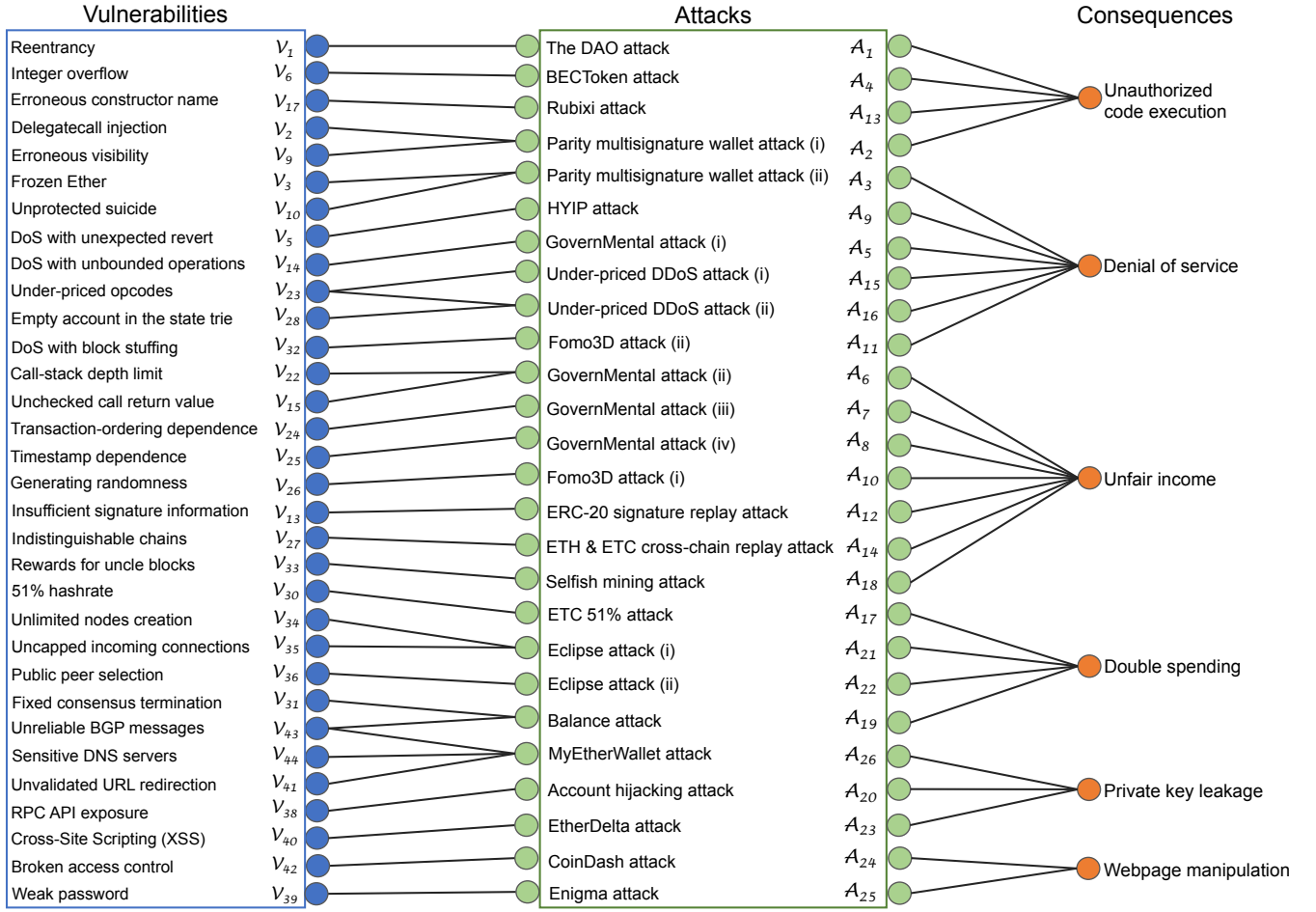Fig. 14: The relation between vulnerabilities, attacks and attack consequences.

cross-chain replay attack ($\mathcal{A}_{14}$) [85].

4) Double-spending: This consequence occurs in one of the following four scenarios: (i) The attacker uses the 51% attack ($\mathcal{A}_{17}$) to exploit the *51% hashrate* vulnerability ($\mathcal{V}_{30}$), namely $\mathcal{A}_{17}(\mathcal{V}_{30})$. (ii) The attacker uses the eclipse attack ($\mathcal{A}_{21}$) to exploit the *unlimited nodes creation* vulnerability ($\mathcal{V}_{34}$) and the *uncapped incoming connections* vulnerability ($\mathcal{V}_{35}$), namely $\mathcal{A}_{21}(\mathcal{V}_{34}, \mathcal{V}_{35})$. (iii) The attacker uses the eclipse attack ($\mathcal{A}_{22}$) to exploit the *public peer selection* vulnerability ($\mathcal{V}_{36}$), namely $\mathcal{A}_{22}(\mathcal{V}_{36})$. (iv) The attacker uses the balance attack ($\mathcal{A}_{19}$) to exploit the *fixed consensus termination* vulnerability ($\mathcal{V}_{31}$) and the *unreliable BGP messages* vulnerability ($\mathcal{V}_{43}$), namely $\mathcal{A}_{19}(\mathcal{V}_{31}, \mathcal{V}_{43})$. The biggest financial loss caused by a single double-spending attack is about US$1.1M, owing to the ETC 51% attack ($\mathcal{A}_{17}$) [118].

5) Private key leakage: This consequence occurs in one of the following three scenarios. (i) The attacker uses the MyEtherWallet attack ($\mathcal{A}_{26}$) to exploit the *unvalidated URL redirection* vulnerability ($\mathcal{V}_{41}$), the *unreliable BGP messages* vulnerability ($\mathcal{V}_{43}$), and the *sensitive DNS* vulnerability ($\mathcal{V}_{44}$) to direct the victim to a malicious website, namely $\mathcal{A}_{26}(\mathcal{V}_{41}, \mathcal{V}_{43}, \mathcal{V}_{44})$. (ii) A user's private key is stolen by an attacker that uses the account hijacking attack ($\mathcal{A}_{20}$) to exploit the *RPC API exposure* vulnerability ($\mathcal{V}_{38}$) to hijack the victim's

account from a remote Ethereum client, namely $\mathcal{A}_{20}(\mathcal{V}_{38})$. (iii) The attacker uses the EtherDelta attack ($\mathcal{A}_{23}$) to exploit the *Cross-Site Scripting* vulnerability ($\mathcal{V}_{40}$) to hijack the victim's browser session, namely $\mathcal{A}_{23}(\mathcal{V}_{40})$. The biggest financial loss caused by a private key leakage is US$20M, owing to an account hijacking attack ($\mathcal{A}_{20}$) [100].

6) Webpage manipulation: This consequence occurs in one of the following two scenarios. (i) A DApp's website interface is compromised by an attacker that wages the CoinDash attack ($\mathcal{A}_{24}$) to exploit the *broken access control* vulnerability ($\mathcal{V}_{42}$), namely $\mathcal{A}_{24}(\mathcal{V}_{42})$. (ii) The attacker uses the Enigma attack ($\mathcal{A}_{25}$) to exploit the *weak password* vulnerability ($\mathcal{V}_{39}$), namely $\mathcal{A}_{25}(\mathcal{V}_{39})$. These vulnerabilities reside in, or are inherited from, the Ethereum environment. The biggest financial loss caused by a webpage manipulation is US$7M, owing to an attack against CoinDash ($\mathcal{A}_{24}$) [104].

7) Summary and insights: Table II summarizes the aforementioned attacks, the layers at which the vulnerabilities they exploit reside, attack consequences, and financial losses incurred by those attacks. This allows us to draw the following insights:

Insight 6: The biggest financial loss in the Ethereum system is $280M, which is caused by a DoS attack against the Parity wallet that disabled a library that is used by many contracts.

Insight 6 is interesting at least from two points of view. First,

TABLE II: Overview of attacks and financial losses.

Attack targets	Vulnerability location	Real-world attacks	Attack consequences	Financial losses
DApp	Application	The DAO ($\mathcal{A}_1$)	Unauthorized code exec.	US$60M
		Parity (i) ($\mathcal{A}_2$)	Unauthorized code exec.	US$31M
		Parity (ii) ($\mathcal{A}_3$)	DoS	US$280M
	Environment	CoinDash ($\mathcal{A}_{24}$)	Webpage manipulation	US$7M
		Enigma ($\mathcal{A}_{25}$)	Webpage manipulation	US$0.5M
		MyEtherWallet ($\mathcal{A}_{26}$)	Private key leakage	US$17M
Ethereum	Consensus	Fomo3D (ii) ($\mathcal{A}_{11}$)	DoS	US$3M
		ETC 51% ($\mathcal{A}_{17}$)	Double spending	US$1.1M
	Data	ETH & ETC replay ($\mathcal{A}_{14}$)	Unfair income	US$0.5M
	Network	Account hijacking ($\mathcal{A}_{20}$)	Private key leakage	US$20M

DoS often causes *indirect* economic losses incurred by the unavailability of services. The DoS attack against Ethereum manifests *direct* economic losses. Second, the DoS attack is caused by code-reuse, which is a widely adopted practice in the traditional paradigm of programming. However, in the paradigm of DApps, code-reuse may impose a higher risk than ever, highlighting the importance of security auditing for popular contracts as well as the libraries they use.

Insight 7: DApps running on top of the Ethereum blockchain are not fully *decentralized* when they use centralized web interfaces.

Insight 7 is interesting because it points out one scenario that manifests that blockchain-based systems are *not* necessarily decentralized, despite the popular view. Indeed, vulnerabilities in those web interfaces can cause large financial losses. This highlights the importance of decentralizing the web interfaces to the blockchain-based systems.

Insight 8: Application-layer attacks have caused the largest financial losses among the attacks against Ethereum.

Insight 8 is interesting because a vulnerability at a lower layer would often cause a larger damage than a vulnerability at a higher layer. This rule of thumb does not necessarily apply to Ethereum or blockchain-based systems in general. This can be (at least partly) attributed to the fact that DApps operate directly on the digital assets, and highlights the importance of assuring application-layer security.

V. DEFENSES

In this section, we describe the 47 defenses that have been proposed for securing the Ethereum ecosystem, which are denoted by $\mathcal{D}_1, \ldots, \mathcal{D}_{47}$, respectively. Unlike vulnerabilities and attacks which naturally correspond to some layer of the Ethereum architecture, defenses are by no means geared towards the layers. Therefore, we propose categorizing them into two classes: proactive defenses and reactive defenses.

A. Proactive defenses

We categorize existing proactive defense mechanisms into the following five sub-classes based on the respective focuses of these mechanisms: *contract programming language, contract development, contract analysis, contract and Ethereum enhancement,* and *consensus protocols.*

1) Language-based security: Programming language approaches to securing smart contracts can be divided into two categories: high-level languages for developing more secure smart contracts and intermediate-level languages for facilitating contract formal analysis.

a) High-level languages: Vyper [121] ($\mathcal{D}_1$) removes a number of functionalities provided by Solidity (e.g., recursive calling, infinite loops, modifiers) and adds several new features (e.g., bounds and overflow checking) to eliminate vulnerabilities, such as the *DoS with unbounded operations* vulnerability ($\mathcal{V}_{14}$), the *erroneous visibility* vulnerability ($\mathcal{V}_9$), and the *Integer overflow and underflow* vulnerability ($\mathcal{V}_6$). Bamboo [122] ($\mathcal{D}_2$) uses polymorphism to mitigate the *transaction-ordering dependence* vulnerability ($\mathcal{V}_{24}$), while eliminating the *reentrancy* vulnerability ($\mathcal{V}_1$) and the *DoS with unbounded operations* vulnerability ($\mathcal{V}_{14}$). Obsidian [123] ($\mathcal{D}_3$) models smart contracts as finite state machines and tracks contracts' states to eliminate the *reentrancy* vulnerability ($\mathcal{V}_1$). Flint [124] ($\mathcal{D}_4$) uses an `Asset` type to assure the atomicity of operations and introduces restrictions on callers' capabilities to protect contract functions from unauthorized access, while aiming to eliminate the *reentrancy* vulnerability ($\mathcal{V}_1$), the *erroneous visibility* vulnerability ($\mathcal{V}_9$), the *Integer overflow and underflow* vulnerability ($\mathcal{V}_6$), and the *timestamp dependency* vulnerability ($\mathcal{V}_{25}$).

b) Intermediate-level languages: Simplicity [125] ($\mathcal{D}_5$) is an intermediate representation between high-level languages and low-level virtual machine. It provides formal semantics using the proof-assistant Coq [126], thus allowing formally analysis of contracts properties (e.g., safety and liveness). Scilla [127] ($\mathcal{D}_6$) distinguishes in-contract computation from inter-contract communication to disentangle contract-specific effects from each other.

2) Contract development: Since smart contracts are a new programming paradigm, these defenses can help developers in avoiding or mitigating common vulnerabilities.

a) Principles and best practices: A number of Solidity best practices were recommended in [11], such as *check, update, then interact* (i.e., checking conditions first, then updating state variables, and finally interacting with other contracts); *favor pull over push for external calls* (i.e., letting a recipient withdraw or "pull" the money set by the sender, rather than letting the sender directly transfer or "push" the money to the recipient). In order to make it easier for practitioners to adopt those best practices, we systematize 19 best practices ($\mathcal{D}_7$) according to the vulnerabilities towards which they are geared, while noting that 15 of these 19 best practices were scattered in [11], [54], [44]. Our systematization leads to 4 general principles or 6 specific sub-principles in total.

Figure 15 highlights these (sub-)principles. Intuitively, a programmer should be conscious of four kinds of controls: *call control, data control, resource control,* and *tool control.* (i) The *call control* principle says that a programmer should secure the interactions between smarts contracts and the interactions between EOAs and smart contracts. This principle
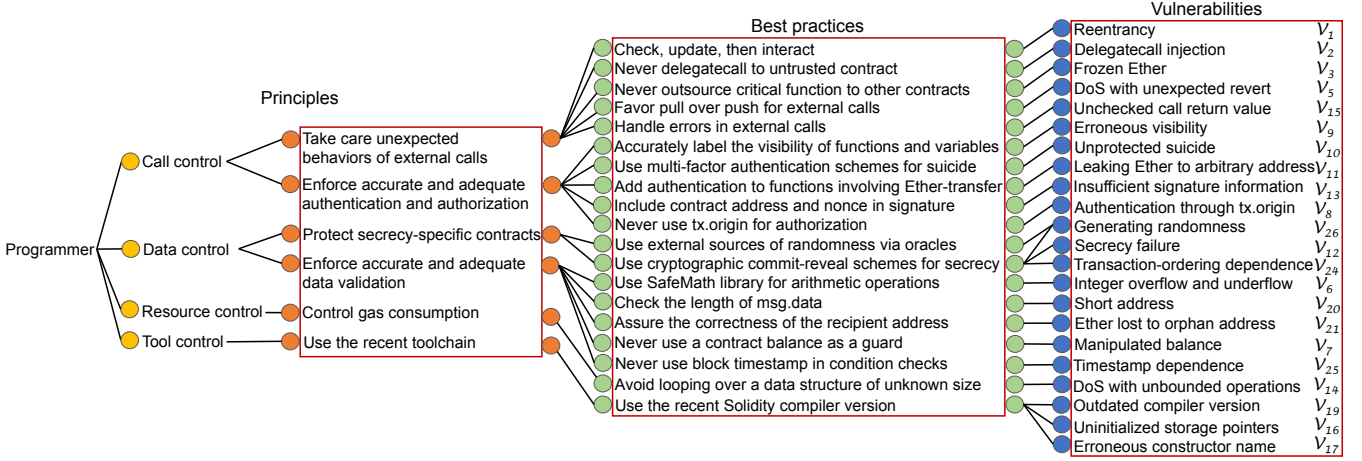
Fig. 15: Best practices and principles for guiding contract development.

is further divided into two sub-principles: one coping with the callee's unexpected behaviors and the other coping with the caller's access control. (ii) The *data control* principle says that a programmer should secure the data flow of a contract. This principle is further divided into two sub-principles: one dealing with the protection of sensitive data and the other dealing with the prevention of malformed data from entering a smart contract. (iii) The *resource control* principle says that a programmer should mitigate the exhaustion of available gas in Ethereum. (iv) The *tool control* principle says that a programmer should use updated tools (e.g., compiler, debugger) to eliminate known vulnerabilities.

b) Software engineering mechanisms: In order to defend against attacks that may exploit unknown vulnerabilities, several blockchain-oriented software engineering mechanisms were introduced [128] ($\mathcal{D}_8$), such as: the *rate limit* mechanism, which restricts the number of consecutive actions incurred by an EOA or restricts the amount of Ether sent by a contract within a period of time; the *balance limit* mechanism, which regulates the maximum amount of funds that can be held by a contract; the *speed bump* mechanism, which postpones some potentially-damaging operations.

3) Smart contract analysis: These defenses aim to enhance security of smart contracts via the following approaches: symbolic execution, abstract interpretation, formal verification, fuzzing and model-based vulnerability detection.

a) Symbolic execution: It works on a program' control-flow graph (CFG) and traverses all of the feasible execution paths on the graph to identify vulnerabilities [129]. This approach achieves neither soundness (i.e., zero false-negatives) nor completeness (i.e., zero false-positives), owing to the omission of some execution paths and the exploration of unreachable paths in real executions. Oyente [63] ($\mathcal{D}_9$) can detect four types of vulnerabilities — the *reentrancy* vulnerability ($\mathcal{V}_1$), the *mishandled exceptions* vulnerability ($\mathcal{V}_{15}$), the *transaction-ordering dependence* vulnerability ($\mathcal{V}_{24}$), and the *timestamp dependence* vulnerability ($\mathcal{V}_{25}$) — but incurs a high false-positive rate [64]. Maian [55] ($\mathcal{D}_{10}$) extends Oyente by considering multiple invocations of a contract, rather than a single invocation. It can detect three types of

vulnerabilities: the *frozen Ether* vulnerability ($\mathcal{V}_3$), the *leaking Ether to arbitrary address* vulnerability ($\mathcal{V}_{11}$), and the *unprotected suicide* vulnerability ($\mathcal{V}_{10}$). Mythril [130] ($\mathcal{D}_{11}$) uses "concolic analysis", which integrates symbolic and concrete execution of a smart contract, to detect 8 types of vulnerabilities, such as the *manipulated balance* vulnerability ($\mathcal{V}_7$), the *authentication through tx.origin* vulnerability ($\mathcal{V}_8$), and the *generating randomness* vulnerability ($\mathcal{V}_{26}$). The teEther tool [43] ($\mathcal{D}_{12}$) can detect vulnerabilities like the *erroneous visibility* vulnerability ($\mathcal{V}_9$) and the *erroneous constructor name* vulnerability ($\mathcal{V}_{17}$). The sCompile tool [131] ($\mathcal{D}_{13}$) reduces the number of suspicious execution paths that are not related to money-transfer. It can detect 4 types of vulnerabilities: the *reentrancy* vulnerability ($\mathcal{V}_1$), the *frozen Ether* vulnerability ($\mathcal{V}_3$), the *unprotected suicide* vulnerability ($\mathcal{V}_{10}$), and the *Ether lost to orphan address* vulnerability ($\mathcal{V}_{21}$). ECF [132] ($\mathcal{D}_{14}$) focuses on detecting callback-related vulnerabilities, such as the *reentrancy* vulnerability ($\mathcal{V}_1$).
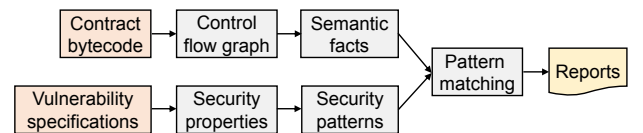

Fig. 16: Workflow of vulnerability detector Securify.

b) Abstract interpretation: Abstract interpretation aims to over-approximate the semantics of a program so as to achieve soundness in program analysis [133]. Securify [134] ($\mathcal{D}_{15}$) defines a set of *compliance* and *violation* patterns to characterize how contract comply and violate security properties extracted from some known vulnerabilities, such as the *reentrancy* vulnerability ($\mathcal{V}_1$), the *delegatecall injection* vulnerability ($\mathcal{V}_2$), and the *frozen Ether* vulnerability ($\mathcal{V}_3$). Figure 16 highlights how these patterns can be used to detect vulnerabilities. Zeus [64] ($\mathcal{D}_{16}$) defines a set of correctness and fairness policies and then embeds them as `assert` statements into the source code of contracts. It can detect six types of vulnerabilities, such as the *reentrancy* vulnerability ($\mathcal{V}_1$) and the *unchecked call return value* vulnerability ($\mathcal{V}_{15}$). FSolidM [135] ($\mathcal{D}_{17}$) abstracts smart contracts as finite state

machines and can detect the *reentrancy* vulnerability ($\mathcal{V}_1$) and the *transaction-ordering dependence* vulnerability ($\mathcal{V}_{24}$). MadMax [62] ($\mathcal{D}_{18}$) disassembles EVM bytecode into an intermediate representation and then leverages both data-flow analysis and context-sensitive flow analysis to detect gas-related vulnerabilities, such as the *DoS with unbounded operations* vulnerability ($\mathcal{V}_{14}$). Vandal [136] ($\mathcal{D}_{19}$) translates EVM bytecode into logic relations and use them to detect a number of vulnerabilities, such as the *reentrancy* vulnerability ($\mathcal{V}_1$), the *authentication through tx.origin* vulnerability ($\mathcal{V}_8$), the *unprotected suicide* vulnerability ($\mathcal{V}_{10}$), and the *unchecked call return value* vulnerability ($\mathcal{V}_{15}$).

In order to facilitate the aforementioned high-level contract analysis, several reverse engineering tools have been developed to convert EVM bytecode to source code or intermediate presentation. Porosity [137] ($\mathcal{D}_{20}$) is a decompiler for producing Solidity source code from EVM bytecode. Erays [138] ($\mathcal{D}_{21}$) lifts EVM bytecode to a high-level pseudocode by recovering the control-flow structure and transforming EVM from a stack-based machine to a register-based machine. EthIR [139] ($\mathcal{D}_{22}$) decompiles EVM bytecode to a high-level rule-based representation, which can then be fed into an automated static analyzer to infer high-level properties of the EVM bytecode.
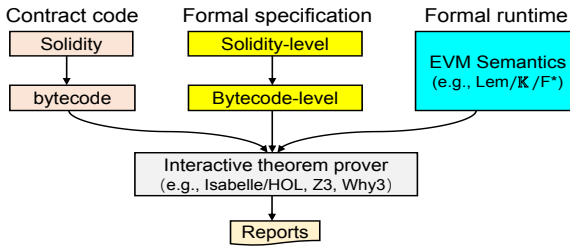

Fig. 17: Theorem proving for verifying smart contracts.

c) Formal verification: Formal verification proves the correctness of contract implementation with respect to a specification. This approach assures the completeness (i.e., no false positives). Figure 17 illustrates how to use theorem-proving to verify smart contracts at the EVM bytecode-level. Hirai [140] ($\mathcal{D}_{23}$) took the first step towards formalizing the EVM semantics, which can be accommodated by interactive theorem provers like Isabelle/HOL [141] to prove invariants and safety properties of smart contracts. Amani et al. [142] ($\mathcal{D}_{24}$) extend the work in [140] by splitting a contract into some basic blocks and then using a Hoare-style program logic to reason about semantic properties of contracts from properties of its parts. Hildenbrandt et al. [143] ($\mathcal{D}_{25}$) present a complete semantics of the EVM, referred to as KEVM, using the $\mathbb{K}$ framework [144] to achieve language-independent program verification. Park et al. [145] ($\mathcal{D}_{26}$) optimize the KEVM verifier by introducing EVM-specific abstractions and lemmas to avoid non-tractable reasoning in the underlying theorem prover. Grishchenko et al. [146] ($\mathcal{D}_{27}$) define a complete small-step semantics of EVM bytecode and formalize most of the semantics in the proof assistant F* [147]. Grishchenko et al. [148], [149] ($\mathcal{D}_{28}$) leverage the complete small-step semantics of EVM bytecode to build *EtherTrust*, which is the first sound and automated static analyzer to achieve formal security

related to the reachability properties of EVM bytecode. Other early-stage investigations include [150], [151], [152].

d) Fuzzing: Fuzz testing has been used to detect vulnerabilities in smart contracts. ContractFuzzer [46] ($\mathcal{D}_{29}$) can detect five types of smart contract vulnerabilities, such as the *reentrancy* vulnerability ($\mathcal{V}_1$) and the *unchecked call return value* vulnerability ($\mathcal{V}_{15}$). It generates inputs by crawling the ABI interfaces of smart contracts to extract their function selectors and data types of each argument, and instruments EVM to log contract execution behaviors for inspection. Re-Guard [153] ($\mathcal{D}_{30}$) aims to detect the *reentrancy* vulnerability ($\mathcal{V}_1$) in smart contracts by transforming smart contracts to semantically equivalent C++ program and generating random transactions via a fuzzing engine to check the execution traces of the C++ program.

e) Model-based vulnerability detection: Tann et al. [154] ($\mathcal{D}_{31}$) use sequence learning to detect three types of vulnerabilities, namely the *frozen Ether* vulnerability ($\mathcal{V}_3$), the *leaking Ether to arbitrary address* vulnerability ($\mathcal{V}_{11}$), and the *unprotected suicide* vulnerability ($\mathcal{V}_{10}$). Tikhomirov et al. [155] ($\mathcal{D}_{32}$) propose SmartCheck to detect vulnerabilities in Solidity contracts by translating Solidity source code into an XML-based parse-tree and checks it against specific XPath patterns; they can detect 10 types of vulnerabilities, such as the *DoS with unexpected revert* vulnerability ($\mathcal{V}_5$), the *manipulated balance* vulnerability ($\mathcal{V}_7$), but incur a high false-positives rate.

4) Contract and Ethereum enhancement: Kosba et al. proposed the Hawk [156] ($\mathcal{D}_{33}$) framework for incorporating cryptographic mechanisms to hide transaction data, effectively allowing contract developers to build privacy-preserving smart contracts without asking them to implement any cryptography. This framework can be used to defend against attacks that exploit the *secrecy failure* vulnerability ($\mathcal{V}_{12}$) and the *transaction-ordering dependence* vulnerability ($\mathcal{V}_{24}$). Zhang et al. [157] ($\mathcal{D}_{34}$) designed an authenticated data feed system, dubbed Town Crier, for smart contracts that require access to external data. This system can be used to mitigate the *generating randomness* vulnerability ($\mathcal{V}_{26}$). Adler et al. [158] ($\mathcal{D}_{35}$) extended Town Crier by implementing a voting-based decentralized oracle to address the centralized point-of-failure that is inherent to Town Crier. Chen [77] ($\mathcal{D}_{36}$) proposed an adaptive gas cost mechanism, which dynamically adjusts the cost of EVM operations according to their execution times, so as to defend against DoS attacks that exploit the *under-priced opcodes* vulnerability ($\mathcal{V}_{23}$). In order to harden the Ethereum network against eclipse attacks, Marcus et al. [35] ($\mathcal{D}_{37}$) proposed a set of countermeasures for eliminating some complicated artifacts of the Kademlia protocol used by Ethereum. Wang et al. [101] ($\mathcal{D}_{38}$) proposed some countermeasures for defending Ethereum clients against the attacks that may exploit the *RPC API exposure* vulnerability ($\mathcal{V}_{38}$).

5) New blockchain protocols: In order to tackle the *outsourceable puzzle* vulnerability ($\mathcal{V}_{29}$), Miller et al. [88] ($\mathcal{D}_{39}$) formalized the notion of non-outsourceable puzzles and employed Merkle-tree-based proofs for puzzle design. The basic idea underlying non-outsourceable puzzles is: If a pool operator outsources the mining task other miners, the miners can collect the full credit while the pool operator gets nothing,

which effectively disincentivizes pool operators from outsourcing their mining tasks. Zeng et al. [159] (D_{40}) extended the work of [88] by proposing a non-outsourceable puzzle that is compatible with the GHOST protocol used by Ethereum. Daian et al. [89] (D_{41}) designed a two-stage non-outsourceable puzzle where the outer puzzle relies on the solution to the inner puzzle.

In order to mitigate the *51% hashrate* vulnerability (V_{30}), Eyal et al. [160] (D_{42}) proposed the two-phase proof of work (2P-PoW) mechanism to disincentivize large mining pools, by incorporating two separate puzzles (instead of one) for miners to solve. Luu et al. [161] (D_{43}) implemented a new decentralized pooled mining protocol to defend against mining centralization, by replacing the traditional mining pool operator with an Ethereum smart contract.

B. Reactive defenses

Reactive defenses aim to react to potential exploitations of (unknown) vulnerabilities during the contract runtime to mitigate the damage. A runtime verification method monitors the execution traces to detect and possibly react to suspicious activities that may violate certain properties. DappGuard [162] (D_{44}) actively monitors the incoming transactions to a smart contract and leverages the aforementioned tool Oyente [63] to decide whether or not an incoming transaction can cause a security violation and if so, a counter transaction can be invoked to kill the contract in question. ContractLarva [163] (D_{45}) generates a new Solidity contract from the original contract and its specification, checks the original contract's runtime behaviors against this new contract's, and takes appropriate actions in the case of any discrepancy. Sereum [40] (D_{46}) uses taint analysis to monitor runtime data flows during smart contract execution for detecting and preventing the *reentrancy* vulnerability (V_1).

When detecting a violation, various mechanisms (D_{47}) have been proposed for mitigating the damage: (i) disabling the vulnerable smart contract or sensitive functionalities by using (for example) the *emergency stop* mechanism [164]; (ii) adopting a *stake-placing* mechanism to assure that any invocation, which potentially violates some properties, should pay a stake of compensation before running the contract and returns the stake back to the caller after the contract terminates normally; (iii) replacing vulnerable contracts with secure ones using the *virtual upgrade* mechanism [128], which can be realized by using a registry contract to hold the address of the latest version of a contract or by introducing a proxy contract to delegate calls to the latest version of a contract.

C. Further analysis based on defense capabilities

Now we present an analysis of defenses from the perspective of *defense capabilities*, meaning which defense can defend against the attacks that exploit certain vulnerabilities.

Figure 18 plots the Vern diagram of the 6 kinds of defenses discussed above, including 5 kinds of proactive defenses (i.e., *alternate language*, *contract analyzer*, *security enhancement*, *contract best practices*, *blockchain protocols*) and 1 kind of reactive defenses (i.e., *runtime verification*). For proactive
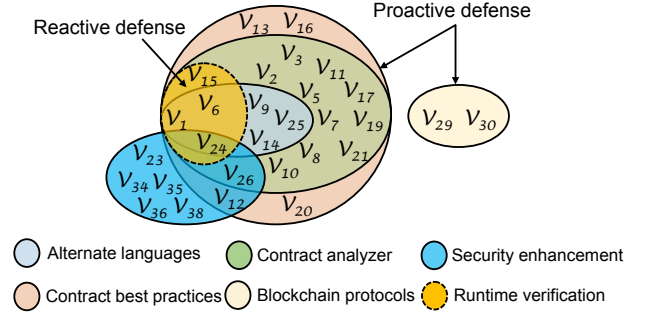

Fig. 18: Vern diagram representation of defenses against attacks that exploit some of 29 vulnerabilities; note that not all of the 44 vulnerabilities are relevant here because some vulnerabilities are already eliminated while other vulnerabilities are yet to be coped with.

defense, we observe the following: (i) using *contract best practices* in the course of developing contracts can prevent or mitigate attacks that attempt to exploit 22 types of application-layer vulnerabilities; (ii) using *smart contract analyzer* can detect or mitigate attacks that attempts to exploit 18 types of application-layer vulnerabilities; (iii) using *security enhancement* can prevent or mitigate attacks that attempts to exploit 8 types of vulnerabilities, including 4 application-layer vulnerabilities and 4 network-layer ones; (iv) using better-designed contract programming language (i.e., *alternate languages* for short) can prevent or mitigate attacks that attempt to exploit 6 types of application-layer vulnerabilities; (v) using better-designed *blockchain protocols* can mitigate attacks that attempt to exploit 2 types of consensus-layer vulnerabilities. For reactive defense, we observe that using *runtime verification* can mitigate attacks that attempt to exploit 4 types of application-layer vulnerabilities. In total, proactive defenses can defend against attacks that attempt to exploit 29 types of vulnerabilities, whereas reactive defenses can defend attacks that attempt to exploit 4 types of vulnerabilities, which are also covered by proactive defenses. Moreover, we observe that the vulnerabilities that can be defended by *alternate languages* and *runtime verification* are subsets of the vulnerabilities that can be defended by *contract analyzers*, which are in turn a subset of the vulnerabilities that can be defended by *contract development best practices*. We draw the following insights.

Insight 9: Industry has come up with a significant set of best practices for guiding the development of smart contracts.

Insight 9 highlights the importance of adopting best practices in the process of developing software. Nevertheless, these best practices alone are not adequate in assuring security.

Insight 10: Proactive defenses can defend against attacks that attempt to exploit many vulnerabilities. In contrast, reactive defenses can defend against attacks that attempt to exploit a few vulnerabilities.

Insight 10 reflects the state-of-the-art. Nevertheless, reactive defenses are still important because they may be able to defend attacks that attempt to exploit vulnerabilities that survived proactive defenses.

D. Further analysis based on defense investment

Now we present an analysis of defenses from the perspective of *defense investment*, meaning how much effort has been invested in designing defense against attacks that exploit a certain vulnerability. However, we note that some defenses are *not* geared towards any specific vulnerabilities or attacks; for example, software engineering mechanisms (i.e., D_8) are neither geared towards any specific vulnerability nor geared towards any specific attack.
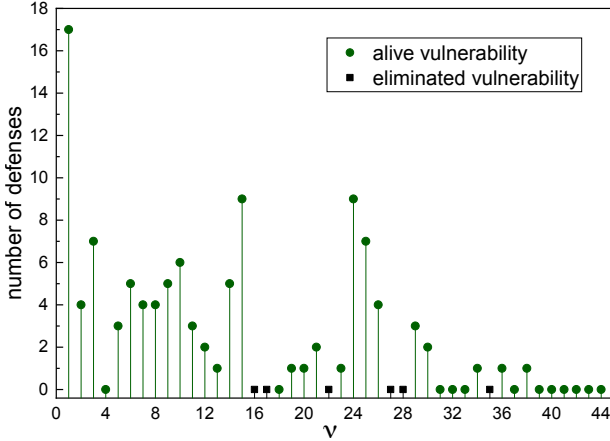

Fig. 19: The number defenses with respect to each vulnerability.

Figure 19 plots the summary of the number of defense mechanisms with respect to individual vulnerabilities. Table III elaborates which vulnerabilities may be protected by which defenses. For example, Figure 19 shows that V_1 can be defended by 17 defenses; Table III shows which these 17 defenses are. On one hand, we observe that the most extensively investigated vulnerability is the *reentrancy* vulnerability (V_1), which can be mitigated by 17 kinds of proactive and reactive defense mechanisms. Other vulnerabilities that have been substantially investigated include: the *unchecked call return value* vulnerability (V_{15}) and the *transaction-ordering dependence* vulnerability (V_{24}), each of which can be defended by 9 kinds of proactive and reactive defense mechanisms; the *frozen Ether* vulnerability (V_3) and the *timestamp dependence* vulnerability (V_{25}) can be defended by 7 kinds of proactive defense mechanisms. It appears that the vulnerabilities that have been thoroughly investigated are (i) the ones that have caused large financial losses, (ii) the ones that are inherent to the design of the Solidity language, and (iii) the ones that are inherent to the profit-making factor for assembling blocks. This is so because the *reentrancy* vulnerability (V_1) has caused the loss of US\$60M in the DAO attack (i.e., attack A_1 in Figure 14), the *frozen Ether* vulnerability (V_3) has caused the loss of US\$280M in the Parity wallet attack (i.e., attack A_3), the *unchecked call return value* vulnerability (V_{15}) is inherent to the exception handling mechanism in the Solidity language, the *transaction-ordering dependence* vulnerability (V_{24}) is inherent to the unpredictable nature of the Ethereum blockchain, and the *timestamp dependence* vulnerability (V_{25}) is inherent to the manipulable block in-

formation of the Ethereum blockchain. On the other hand, we observe 19 vulnerabilities that have zero or one defense mechanism. Most of these vulnerabilities are either caused by the Ethereum design and implementation, or reside in the Ethereum environment. This leads to:

Insight 11: There is a large discrepancy between the effort that has been invested to defend against attacks that exploit different vulnerabilities.

Insight 11 is interesting because it seems that the defense effort has been driven by the consequential financial loss incurred by the exploitation of certain vulnerabilities. This prioritization strategy is *not* adequate because it suggests in a sense that the defender is always chasing behind the attacker, who detects and exploits an vulnerability that then becomes known to the defender.

Insight 12: Existing studies focus on defending against attacks that attempt to exploit vulnerabilities in the DApp back-end (i.e., smart contracts), but largely ignore the protection of the DApp front-end (i.e., browser) and the interactions between the front-end and the back-end.

Insight 12 says that in order to adequately defend DApps, a good solution should consider the front-end interface, the back-end contracts, and their interactions. It is worth mentioning that attacks A_{23}, A_{24}, A_{25}, and A_{26} have exploited either the front-end interface or the front-end and back-end interactions to cause the loss of over US\$24M in total.

VI. DISCUSSION: GOING BEYOND ETHEREUM

The preceding discussion is geared towards the Ethereum system and offers a number of open problems for future research. While the findings and (some of) the open problems might be applicable to blockchain-based systems in general, in what follows we discuss future research directions that are equally applicable to Ethereum and blockchain-based systems.

A. Rigorous definition of properties

We observe that in order to adequately defend Ethereum and blockchain-based systems, there is an urgent need to understand the desirable security properties, which are however extremely difficult to formalize for a complex system like Ethereum. Some informal properties have been discussed in [165], which represents a very preliminary first step towards the ultimate goal. This leads to:

Insight 13: There is a lack of deep understanding on the rigorously-specified desirable security properties that should be possessed by Ethereum and blockchain-based systems.

B. Rigorous analysis methodologies

Having defined the rigorous properties that should be satisfied by Ethereum and blockchain-based systems, we need principled and rigorous methodologies to analyze that the desirable properties are indeed satisfied. For properties that are geared towards building-blocks, cryptography and formal methods have been two successful approaches, while noting that they have their own limitations. Since blockchain-based systems, such as Ethereum, as indeed complex systems, it is

TABLE III: Elaboration of the defenses with respect to vulnerabilities summarized in Figure 19.

	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9	V_{10}	V_{11}	V_{12}	V_{13}	V_{14}	V_{15}	V_{16}	V_{17}	V_{18}	V_{19}	V_{20}	V_{21}	V_{22}	V_{23}	V_{24}	V_{25}	V_{26}	V_{27}	V_{28}	V_{29}	V_{30}	V_{31}	V_{32}	V_{33}	V_{34}	V_{35}	V_{36}	V_{37}	V_{38}	V_{39}	V_{40}	V_{41}	V_{42}	V_{43}	V_{44}	total
D_1						✓			✓					✓																															3
D_2	✓													✓										✓																					3
D_3	✓																																												1
D_4	✓					✓			✓																✓																				4
D_7	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	✓	✓																			22
D_9	✓													✓										✓	✓																				4
D_{10}		✓								✓	✓																																		3
D_{11}	✓						✓	✓		✓					✓									✓	✓	✓																			8
D_{12}									✓									✓																											2
D_{13}	✓		✓							✓											✓																								4
D_{14}	✓																																												1
D_{15}	✓	✓	✓												✓									✓																					5
D_{16}	✓					✓	✓								✓									✓	✓																				6
D_{17}	✓																							✓																					2
D_{18}													✓																																1
D_{19}	✓						✓	✓		✓					✓																														5
D_{29}	✓	✓	✓												✓									✓																					5
D_{30}	✓																																												1
D_{31}			✓							✓	✓																																		3
D_{32}	✓	✓	✓		✓		✓	✓	✓					✓	✓										✓																				10
D_{33}												✓													✓																				2
D_{34}																										✓																			1
D_{35}																										✓																			1
D_{36}																							✓																						1
D_{37}																																		✓	✓	✓									3
D_{38}																																						✓							1
D_{39}																								✓																					1
D_{40}																								✓																					1
D_{41}																								✓																					1
D_{42}																									✓																				1
D_{43}																									✓																				1
D_{44}	✓					✓									✓									✓																					4
D_{46}	✓																																												1
total	17	4	7	0	3	5	4	4	5	6	3	2	1	5	9	–	–	0	1	1	2	–	1	9	7	4	–	–	3	2	0	0	0	1	–	1	0	1	0	0	0	0	0	0	

probably infeasible to prevent all attacks, meaning that attacks are inevitable and their risks must be adequately understood manageable. This calls for rigorous analysis methodologies from a holistic perspective. Towards this ultimate goal, the recently proposed approach of *cybersecurity dynamics* [166], [167], [168], [169], [170], [171], [172], [173], [174] has a great potential, although we are not aware of any results that have been published in the literature. This leads to:

Insight 14: There is a lack of deep understanding on the rigorous analysis methodologies that are necessary and sufficient for analyzing the desired properties of blockchain-based systems.

C. Metrics

It is a notoriously difficult problem to rigorously define metrics to systematically measure security properties of interest. However, given the high likelihood that blockchain-based systems will become digital infrastructures for the future society, if not already, there is an urgent need to define metrics to measure their security and risk. We are not aware of any systematic study on this matter, despite substantial efforts [175], [176], [177], [178], [179], [180], [181]. This leads to:

Insight 15: There is a lack of deep understanding on the metrics that are necessary and sufficient for quantifying the security and risk of blockchain-based systems.

VII. CONCLUSION

We have presented a systematic survey on the security of the Ethereum system, including its application, data, consensus, and network layers. The survey considered three perspectives, namely vulnerabilities, attacks and defenses, while correlating them. We discussed not only the locations of the vulnerabilities, but also their root causes. We systematized the attacks against, and the defenses for, the Ethereum system. We further systematized the best practices proposed by industry into a small number of guiding principles, which might be easier to adopt by practitioners. We provide insights into the state-of-the-art and into future research directions.

REFERENCES

[1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.
[2] D. Chaum, "Blind signatures for untraceable payments," in *Advances in Cryptology: Proceedings of CRYPTO '82, Santa Barbara, California, USA, August 23-25, 1982.*, pp. 199–203, 1982.

[3] M. J. Fischer, N. A. Lynch, and M. Paterson, "Impossibility of distributed consensus with one faulty process," *J. ACM*, vol. 32, no. 2, pp. 374–382, 1985.

[4] C. Dwork and M. Naor, "Pricing via processing or combatting junk mail," in *Advances in Cryptology - CRYPTO '92, 12th Annual International Cryptology Conference, Santa Barbara, California, USA, August 16-20, 1992, Proceedings*, pp. 139–147, 1992.

[5] G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum project yellow paper*, vol. 151, pp. 1–32, 2014.

[6] "Ethereum (eth) blockchain explorer." https://etherscan.io/.

[7] "Smart contract analytics." https://stat.bloxy.info/superset/dashboard/smart_contracts/?standalone=true.

[8] "Dao at v1.0." https://github.com/slockit/DAO/tree/v1.0.

[9] "The parity wallet hack explained – zeppelin blog." https://blog.zeppelin.solutions/on-the-parity-wallet-multisig-hack-405a8c12e8f7.

[10] "Bgp leaks and cryptocurrencies." https://blog.cloudflare.com/bgp-leaks-and-crypto-currencies/.

[11] "Solidity recommendations - ethereum smart contract best practices." https://consensys.github.io/smart-contract-best-practices/recommendations/.

[12] N. Atzei, M. Bartoletti, and T. Cimoli, "A survey of attacks on ethereum smart contracts (sok)," in *Principles of Security and Trust*, pp. 164–186, Springer, 2017.

[13] "A survey on the security of blockchain systems," *Future Generation Computer Systems*, 2017.

[14] L. Zhu, B. Zheng, M. Shen, S. Yu, F. Gao, H. Li, K. Shi, and K. Gai, "Research on the security of blockchain data: A survey," *CoRR*, vol. abs/1812.02009, 2018.

[15] M. Saad, J. Spaulding, L. Njilla, C. A. Kamhoua, S. Shetty, D. Nyang, and A. Mohaisen, "Exploring the attack surface of blockchain: A systematic overview," *CoRR*, vol. abs/1904.03487, 2019.

[16] D. Harz and W. Knottenbelt, "Towards safer smart contracts: A survey of languages and verification methods," *arXiv preprint arXiv:1809.09805*, 2018.

[17] M. Di Angelo and G. Salzer, "A survey of tools for analyzing ethereum smart contracts," in *2019 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPCON)*, IEEE, 2019.

[18] R. Zhang, R. Xue, and L. Liu, "Security and privacy on blockchain," *CoRR (to appear in ACM Computing Survey)*, vol. abs/1903.07602, 2019.

[19] C. Cachin and M. Vukolic, "Blockchain consensus protocols in the wild," *CoRR*, vol. abs/1707.01873, 2017.

[20] S. Bano, A. Sonnino, M. Al-Bassam, S. Azouvi, P. McCorry, S. Meiklejohn, and G. Danezis, "Consensus in the age of blockchains," *CoRR*, vol. abs/1711.03936, 2017.

[21] W. Wang, D. T. Hoang, Z. Xiong, D. Niyato, P. Wang, P. Hu, and Y. Wen, "A survey on consensus mechanisms and mining management in blockchain networks," *CoRR*, vol. abs/1805.02707, 2018.

[22] Y. Xiao, N. Zhang, W. Lou, and Y. T. Hou, "A survey of distributed consensus protocols for blockchain networks," *CoRR*, vol. abs/1904.04098, 2019.

[23] J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll, and E. W. Felten, "Sok: Research perspectives and challenges for bitcoin and cryptocurrencies," in *2015 IEEE Symposium on Security and Privacy*, pp. 104–121, May 2015.

[24] M. Conti, E. Sandeep Kumar, C. Lal, and S. Ruj, "A survey on security and privacy issues of bitcoin," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 3416–3452, 2018.

[25] F. Tschorsch and B. Scheuermann, "Bitcoin and beyond: A technical survey on decentralized digital currencies," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 2084–2123, 2016.

[26] "State of the dapps — dapp statistics." https://www.stateofthedapps.com/stats.

[27] "Erc-20 token standard | ethereum improvement proposals." https://eips.ethereum.org/EIPS/eip-20.

[28] S. Omohundro, "Cryptocurrencies, smart contracts, and artificial intelligence," *AI matters*, vol. 1, no. 2, pp. 19–21, 2014.

[29] "Patricia tree." https://xlinux.nist.gov/dads/HTML/patriciatree.html.

[30] Y. Sompolinsky and A. Zohar, "Secure high-rate transaction processing in bitcoin," in *International Conference on Financial Cryptography and Data Security*, pp. 507–527, Springer, 2015.

[31] J. Garay, A. Kiayias, and N. Leonardos, "The bitcoin backbone protocol: Analysis and applications," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 281–310, Springer, 2015.

[32] W. Wang, D. T. Hoang, P. Hu, Z. Xiong, D. Niyato, P. Wang, Y. Wen, and D. I. Kim, "A survey on consensus mechanisms and mining strategy management in blockchain networks," *IEEE Access*, vol. 7, pp. 22328–22370, 2019.

[33] J. Niu and C. Feng, "Selfish mining in ethereum," *arXiv preprint arXiv:1901.04620*, 2019.

[34] S. Kim, *Measuring Ethereum's peer-to-peer network*. PhD thesis, 2017.

[35] Y. Marcus, E. Heilman, and S. Goldberg, "Low-resource eclipse attacks on ethereum's peer-to-peer network.," *IACR Cryptology ePrint Archive*, vol. 2018, p. 236, 2018.

[36] "devp2p/eth.md at master · ethereum/devp2p · github." https://github.com/ethereum/devp2p/blob/master/caps/eth.md.

[37] "Github - ethereum/go-ethereum: Official go implementation of the ethereum protocol." https://github.com/ethereum/go-ethereum.

[38] "Github - paritytech/parity-ethereum: The fast, light, and robust evm and wasm client.." https://github.com/paritytech/parity-ethereum.

[39] "Analysis of the dao exploit." http://hackingdistributed.com/2016/06/18/analysis-of-the-dao-exploit/.

[40] M. Rodler, W. Li, G. O. Karame, and L. Davi, "Sereum: Protecting existing smart contracts against re-entrancy attacks," *arXiv preprint arXiv:1812.05934*, 2018.

[41] "Ethereum smart contract best practices." https://consensys.github.io/smart-contract-best-practices/.

[42] "A hacker stole $31m of ether — how it happened, and what it means for ethereum." https://medium.freecodecamp.org/a-hacker-stole-31m-of-ether-how-it-happened-and-what-it-means-for-ethereum-9e5dc29e33ce.

[43] J. Krupp and C. Rossow, "teether: Gnawing at ethereum to automatically exploit smart contracts," in *27th {USENIX} Security Symposium ({USENIX} Security 18)*, pp. 1317–1333, 2018.

[44] "Comprehensive list of known attack vectors and common anti-patterns." https://github.com/sigp/solidity-security-blog.

[45] "Security alert | parity technologies." https://www.parity.io/security-alert-2/.

[46] B. Jiang, Y. Liu, and W. Chan, "Contractfuzzer: Fuzzing smart contracts for vulnerability detection," in *Proceedings of the 33rd ACM/IEEE International Conference on Automated Software Engineering*, pp. 259–269, ACM, 2018.

[47] "Safety · ethereum/wiki wiki · github." https://github.com/ethereum/wiki/wiki/Safety#favor-pull-over-push-for-external-calls.

[48] "Nvd - cve-2018-10299." https://nvd.nist.gov/vuln/detail/CVE-2018-10299.

[49] "Safemath." https://github.com/OpenZeppelin/openzeppelin-solidity/blob/master/contracts/math/SafeMath.sol.

[50] "How to secure your smart contracts: 6 solidity vulnerabilities and how to avoid them (part 2)." https://medium.com/loom-network/how-to-secure-your-smart-contracts-6-solidity-vulnerabilities-and-how-to-avoid-them-part-2-730db0aa4834.

[51] "Tx.origin and ethereum oh my!." https://vessenes.com/tx-origin-and-ethereum-oh-my/.

[52] "Security considerations — solidity 0.5.10 documentation." https://solidity.readthedocs.io/en/develop/security-considerations.html#tx-origin.

[53] "Solidity v0.5.0 breaking changes — solidity 0.5.0 documentation." https://solidity.readthedocs.io/en/v0.5.0/050-breaking-changes.html.

[54] "Smart contract weakness classification and test cases." https://smartcontractsecurity.github.io/SWC-registry/.

[55] I. Nikolić, A. Kolluri, I. Sergey, P. Saxena, and A. Hobor, "Finding the greedy, prodigal, and suicidal contracts at scale," in *Proceedings of the 34th Annual Computer Security Applications Conference*, pp. 653–663, ACM, 2018.

[56] K. Delmolino, M. Arnett, A. Kosba, A. Miller, and E. 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*, pp. 79–94, Springer, 2016.

[57] "Replay attacks on ethereum smart contracts." https://github.com/nkbai/defcon26/tree/master/docs.

[58] "Solidity by example — solidity 0.5.3 documentation." https://solidity.readthedocs.io/en/v0.5.3/solidity-by-example.html#micropayment-channel.

[59] "Program the blockchain | signing and verifying messages in ethereum." https://programtheblockchain.com/posts/2018/02/17/signing-and-verifying-messages-in-ethereum/.

[60] "Governmental's 1100 eth jackpot payout is stuck because it uses too much gas : ethereum." https://www.reddit.com/r/ethereum/comments/4ghzhv/governmentals_1100_eth_jackpot_payout_is_stuck/.

[61] "Known attacks - ethereum smart contract best practices." https://consensys.github.io/smart-contract-best-practices/known_attacks/.

[62] N. Grech, M. Kong, A. Jurisevic, L. Brent, B. Scholz, and Y. Smaragdakis, "Madmax: Surviving out-of-gas conditions in ethereum smart contracts," *Proceedings of the ACM on Programming Languages*, vol. 2, no. OOPSLA, p. 116, 2018.

[63] 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*, pp. 254–269, ACM, 2016.

[64] S. Kalra, S. Goel, M. Dhawan, and S. Sharma, "Zeus: Analyzing safety of smart contracts," NDSS, 2018.

[65] "Storage allocation exploits in ethereum smart contracts." https://medium.com/cryptronics/storage-allocation-exploits-in-ethereum-smart-contracts-16c2aa312743.

[66] "Storage pointers in solidity." https://blog.b9lab.com/storage-pointers-in-solidity-7dcfaa536089.

[67] "Ethereum accounts, addresses and contracts." https://etherscan.io/address/0xe82719202e5965Cf5D9B6673B7503a3b92DE20be#contracts.

[68] "'constructor' modifier · issue #3196 · ethereum/solidity · github." https://github.com/ethereum/solidity/issues/3196.

[69] "Release version 0.4.22 · ethereum/solidity · github." https://github.com/ethereum/solidity/releases/tag/v0.4.22.

[70] "How to find $10m just by reading the blockchain." https://medium.com/golem-project/how-to-find-10m-by-just-reading-blockchain-6ae9d39fcd95.

[71] "The erc20 short address attack explained." https://vessenes.com/the-erc20-short-address-attack-explained/.

[72] "Worry-some bug / exploit with erc20 token transactions from exchanges : ethereum." https://www.reddit.com/r/ethereum/comments/63s917/worrysome_bug_exploit_with_erc20_token/dfwmhc3/.

[73] "Ethereum analysis: Gas economics and proof of work." https://github.com/LeastAuthority/ethereum-analyses.

[74] "Eips/eip-150.md at master · ethereum/eips · github." https://github.com/ethereum/EIPs/blob/master/EIPS/eip-150.md.

[75] "Transaction spam attack: Next steps." https://blog.ethereum.org/2016/09/22/transaction-spam-attack-next-steps/.

[76] "A state clearing faq." https://www.reddit.com/r/ethereum/comments/5es5g4/a_state_clearing_faq/?st=iw2e1mwo&sh=fa7768&depth=1.

[77] T. Chen, X. Li, Y. Wang, J. Chen, Z. Li, X. Luo, M. H. Au, and X. Zhang, "An adaptive gas cost mechanism for ethereum to defend against under-priced dos attacks," in *International Conference on Information Security Practice and Experience*, pp. 3–24, Springer, 2017.

[78] "How can i securely generate a random number in my smart contract? - ethereum stack exchange." https://ethereum.stackexchange.com/questions/191/how-can-i-securely-generate-a-random-number-in-my-smart-contract.

[79] "To sink frontrunners, send in the submarines." http://hackingdistributed.com/2017/08/28/submarine-sends/.

[80] "Predicting random numbers in ethereum smart contracts." https://blog.positive.com/predicting-random-numbers-in-ethereum-smart-contracts-e5358c6b8620.

[81] "Breaking randomness in the ethereum universe [part 1] - security boulevard." https://securityboulevard.com/2018/06/breaking-randomness-in-the-ethereum-universe-part-1/.

[82] "Randomness." https://github.com/fravoll/solidity-patterns/blob/master/docs/randomness.md.

[83] "Oracle." https://github.com/fravoll/solidity-patterns/blob/master/docs/oracle.md.

[84] "Randao: A dao working as rng of ethereum." https://github.com/randao/randao.

[85] "Rise of replay attacks intensifies ethereum divide - coindesk." https://www.coindesk.com/rise-replay-attacks-ethereum-divide.

[86] "Eips/eip-155.." https://github.com/ethereum/EIPs/blob/master/EIPS/eip-155.md.

[87] "Eip-161, state trie clearing." https://github.com/ethereum/EIPs/blob/master/EIPS/eip-161.md.

[88] A. Miller, A. Kosba, J. Katz, and E. Shi, "Nonoutsourceable scratch-off puzzles to discourage bitcoin mining coalitions," in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pp. 680–691, ACM, 2015.

[89] P. Daian, I. Eyal, A. Juels, and E. G. Sirer, "(short paper) piecework: Generalized outsourcing control for proofs of work," in *International Conference on Financial Cryptography and Data Security*, pp. 182–190, Springer, 2017.

[90] X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen, "A survey on the security of blockchain systems," *Future Generation Computer Systems*, 2017.

[91] M. Saad, J. Spaulding, L. Njilla, C. Kamhoua, S. Shetty, D. Nyang, and A. Mohaisen, "Exploring the attack surface of blockchain: A systematic overview," *arXiv preprint arXiv:1904.03487*, 2019.

[92] V. Gramoli, "From blockchain consensus back to byzantine consensus," *Future Generation Computer Systems*, 2017.

[93] M. J. Fischer, N. A. Lynch, and M. S. Paterson, "Impossibility of distributed consensus with one faulty process.," tech. rep., MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR COMPUTER SCIENCE, 1982.

[94] "The $3 million winner of fomo3d is still playing to win - longhash." https://www.longhash.com/news/the-3-million-winner-of-fomo3d-is-still-playing-to-win.

[95] A. Gervais, G. O. Karame, K. Wüst, V. Glykantzis, H. Ritzdorf, and S. Capkun, "On the security and performance of proof of work blockchains," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pp. 3–16, ACM, 2016.

[96] F. Ritz and A. Zugenmaier, "The impact of uncle rewards on selfish mining in ethereum," in *2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pp. 50–57, IEEE, 2018.

[97] P. Maymounkov and D. Mazieres, "Kademlia: A peer-to-peer information system based on the xor metric," in *International Workshop on Peer-to-Peer Systems*, pp. 53–65, Springer, 2002.

[98] S. K. Kim, Z. Ma, S. Murali, J. Mason, A. Miller, and M. Bailey, "Measuring ethereum network peers," in *Proceedings of the Internet Measurement Conference 2018*, pp. 91–104, ACM, 2018.

[99] K. Wüst and A. Gervais, "Ethereum eclipse attacks," tech. rep., ETH Zurich, 2016.

[100] "Billions of tokens theft case cause by eth ecological defects." https://mp.weixin.qq.com/s/ia9nBhmqVEXiiQdFrjzmyg.

[101] X. Wang, X. Zha, G. Yu, W. Ni, R. P. Liu, Y. J. Guo, X. Niu, and K. Zheng, "Attack and defence of ethereum remote apis," in *2018 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, IEEE, 2018.

[102] "Hackers nab $500,000 as enigma is compromised weeks before its ico." https://techcrunch.com/2017/08/21/hack-enigma-500000-ico/.

[103] "How one hacker stole thousands of dollars worth of cryptocurrency with a classic code injection...." https://hackernoon.com/how-one-hacker-stole-thousands-of-dollars-worth-of-cryptocurrency-with-a-classic-code-injection-a3aba5d2bff0.

[104] "Coindash tge hack findings report." https://blog.coindash.io/coindash-tge-hack-findings-report-15-11-17-9657465192e1.

[105] A. Mitseva, A. Panchenko, and T. Engel, "The state of affairs in bgp security: A survey of attacks and defenses," *Computer Communications*, vol. 124, pp. 45–60, 2018.

[106] "Cryptocurrency exchange etherdelta hacked in dns hijacking scheme." https://www.ccn.com/cryptocurrency-exchange-etherdelta-hacked-in-dns-hijacking-scheme/.

[107] A. R. Kang, J. Spaulding, and A. Mohaisen, "Domain name system security and privacy: old problems and new challenges," *arXiv preprint arXiv:1606.07080*, 2016.

[108] "An in-depth look at the parity multisig bug." http://hackingdistributed.com/2017/07/22/deep-dive-parity-bug/.

[109] "The $280m ethereum's parity bug. – comae technologies." https://blog.comae.io/the-280m-ethereums-bug-f28e5de43513.

[110] "The parity wallet hack reloaded - zeppelinos blog." https://blog.zeppelinos.org/parity-wallet-hack-reloaded/.

[111] G. Destefanis, M. Marchesi, M. Ortu, R. Tonelli, A. Bracciali, and R. Hierons, "Smart contracts vulnerabilities: a call for blockchain software engineering?," in *2018 International Workshop on Blockchain Oriented Software Engineering (IWBOSE)*, pp. 19–25, IEEE, 2018.

[112] "Batchoverflow." http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2018-10299.

[113] A. Li and F. Long, "Detecting standard violation errors in smart contracts," *arXiv preprint arXiv:1812.07702*, 2018.

[114] M. Bartoletti, S. Carta, T. Cimoli, and R. Saia, "Dissecting ponzi schemes on ethereum: identification, analysis, and impact," *arXiv preprint arXiv:1703.03779*, 2017.

[115] "What we learned from fomo3d." https://medium.com/@martinderka.

[116] "Def con® 26 hacking conference speakers." https://www.defcon.org/html/defcon-26/dc-26-speakers.html#Bai2.

[117] L. Kiffer, D. Levin, and A. Mislove, "Stick a fork in it: Analyzing the ethereum network partition," in *Proceedings of the 16th ACM Workshop on Hot Topics in Networks*, pp. 94–100, ACM, 2017.

[118] "Etc 51 % attack." https://bravenewcoin.com/insights/etc-51-attack-what-happened-and-how-it-was-stopped.

[119] C. Natoli and V. Gramoli, "The balance attack or why forkable blockchains are ill-suited for consortium," in *2017 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pp. 579–590, IEEE, 2017.

[120] P. Ekparinya, V. Gramoli, and G. Jourjon, "Impact of man-in-the-middle attacks on ethereum," in *2018 IEEE 37th Symposium on Reliable Distributed Systems (SRDS)*, pp. 11–20, IEEE, 2018.

[121] "Vyper — vyper documentation." https://vyper.readthedocs.io/en/latest/?badge=latest#.

[122] "Bamboo." https://github.com/pirapira/bamboo.

[123] M. Coblenz, "Obsidian: a safer blockchain programming language," in *Proceedings of the 39th International Conference on Software Engineering Companion*, pp. 97–99, IEEE Press, 2017.

[124] F. Schrans, S. Eisenbach, and S. Drossopoulou, "Writing safe smart contracts in flint," in *Conference Companion of the 2nd International Conference on Art, Science, and Engineering of Programming*, pp. 218–219, ACM, 2018.

[125] R. O'Connor, "Simplicity: A new language for blockchains," in *Proceedings of the 2017 Workshop on Programming Languages and Analysis for Security*, pp. 107–120, ACM, 2017.

[126] "Welcome! | the coq proof assistant." https://coq.inria.fr/.

[127] I. Sergey, A. Kumar, and A. Hobor, "Scilla: a smart contract intermediate-level language," *arXiv preprint arXiv:1801.00687*.

[128] "Software engineering techniques." https://consensys.github.io/smart-contract-best-practices/software_engineering/.

[129] J. C. King, "Symbolic execution and program testing," *Communications of the ACM*, vol. 19, no. 7, pp. 385–394, 1976.

[130] "Mythril." https://github.com/ConsenSys/mythril.

[131] J. Chang, B. Gao, H. Xiao, J. Sun, and Z. Yang, "scompile: Critical path identification and analysis for smart contracts," *arXiv preprint arXiv:1808.00624*, 2018.

[132] S. Grossman, I. Abraham, G. Golan-Gueta, Y. Michalevsky, N. Rinetzky, M. Sagiv, and Y. Zohar, "Online detection of effectively callback free objects with applications to smart contracts," *Proceedings of the ACM on Programming Languages*, vol. 2, no. POPL, p. 48, 2017.

[133] P. Cousot and R. Cousot, "Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints," in *Proceedings of the 4th ACM SIGACT-SIGPLAN symposium on Principles of programming languages*, pp. 238–252, ACM, 1977.

[134] P. Tsankov, A. Dan, D. D. Cohen, A. Gervais, F. Buenzli, and M. Vechev, "Securify: Practical security analysis of smart contracts," *arXiv preprint arXiv:1806.01143*, 2018.

[135] A. Mavridou and A. Laszka, "Designing secure ethereum smart contracts: A finite state machine based approach," *arXiv preprint arXiv:1711.09327*, 2017.

[136] L. Brent, A. Jurisevic, M. Kong, E. Liu, F. Gauthier, V. Gramoli, R. Holz, and B. Scholz, "Vandal: A scalable security analysis framework for smart contracts," *arXiv preprint arXiv:1809.03981*, 2018.

[137] M. Suiche, "Porosity: A decompiler for blockchain-based smart contracts bytecode," *DEF CON*, vol. 25, p. 11, 2017.

[138] Y. Zhou, D. Kumar, S. Bakshi, J. Mason, A. Miller, and M. Bailey, "Erays: Reverse engineering ethereum's opaque smart contracts," in *USENIX Security*, 2018.

[139] E. Albert, P. Gordillo, B. Livshits, A. Rubio, and I. Sergey, "Ethir: A framework for high-level analysis of ethereum bytecode," *arXiv preprint arXiv:1805.07208*, 2018.

[140] Y. Hirai, "Defining the ethereum virtual machine for interactive theorem provers," in *International Conference on Financial Cryptography and Data Security*, pp. 520–535, Springer, 2017.

[141] T. Nipkow, L. C. Paulson, and M. Wenzel, *Isabelle/HOL: a proof assistant for higher-order logic*, vol. 2283. Springer Science & Business Media, 2002.

[142] 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*, pp. 66–77, ACM, 2018.

[143] 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)*, pp. 204–217, IEEE, 2018.

[144] G. Roşu and T. F. Şerbănuţă, "An overview of the k semantic framework," *The Journal of Logic and Algebraic Programming*, vol. 79, no. 6, pp. 397–434, 2010.

[145] D. Park, Y. Zhang, M. Saxena, P. Daian, and G. Roşu, "A formal verification tool for ethereum vm bytecode," in *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pp. 912–915, ACM, 2018.

[146] I. Grishchenko, M. Maffei, and C. Schneidewind, "A semantic framework for the security analysis of ethereum smart contracts," in *International Conference on Principles of Security and Trust*, pp. 243–269, Springer, 2018.

[147] N. Swamy, C. Hriţcu, C. Keller, A. Rastogi, A. Delignat-Lavaud, S. Forest, K. Bhargavan, C. Fournet, P.-Y. Strub, M. Kohlweiss, *et al.*, "Dependent types and multi-monadic effects in f," in *ACM SIGPLAN Notices*, vol. 51, pp. 256–270, ACM, 2016.

[148] I. Grishchenko, M. Maffei, and C. Schneidewind, "Ethertrust: Sound static analysis of ethereum bytecode," *Technische Universität Wien, Tech. Rep*, 2018.

[149] I. Grishchenko, M. Maffei, and C. Schneidewind, "Foundations and tools for the static analysis of ethereum smart contracts," in *International Conference on Computer Aided Verification*, pp. 51–78, Springer, 2018.

[150] K. Bhargavan, A. Delignat-Lavaud, C. Fournet, A. Gollamudi, G. Gonthier, N. Kobeissi, N. Kulatova, A. Rastogi, T. Sibut-Pinote, N. Swamy, *et al.*, "Formal verification of smart contracts: Short paper," in *Proceedings of the 2016 ACM Workshop on Programming Languages and Analysis for Security*, pp. 91–96, ACM, 2016.

[151] "Why3." http://why3.lri.fr/.

[152] "Formal verification for solidity contracts." https://forum.ethereum.org/discussion/3779/formal-verification-for-solidity-contracts.

[153] C. Liu, H. Liu, Z. Cao, Z. Chen, B. Chen, and B. Roscoe, "Reguard: finding reentrancy bugs in smart contracts," in *Proceedings of the 40th International Conference on Software Engineering: Companion Proceeedings*, pp. 65–68, ACM, 2018.

[154] A. Tann, X. J. Han, S. S. Gupta, and Y.-S. Ong, "Towards safer smart contracts: A sequence learning approach to detecting vulnerabilities," *arXiv preprint arXiv:1811.06632*, 2018.

[155] S. Tikhomirov, E. Voskresenskaya, I. Ivanitskiy, R. Takhaviev, E. Marchenko, and Y. Alexandrov, "Smartcheck: Static analysis of ethereum smart contracts," in *2018 IEEE/ACM 1st International Workshop on Emerging Trends in Software Engineering for Blockchain (WETSEB)*, pp. 9–16, IEEE, 2018.

[156] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts," in *2016 IEEE symposium on security and privacy (SP)*, pp. 839–858, IEEE, 2016.

[157] F. Zhang, E. Cecchetti, K. Croman, A. Juels, and E. Shi, "Town crier: An authenticated data feed for smart contracts," in *Proceedings of the 2016 aCM sIGSAC conference on computer and communications security*, pp. 270–282, ACM, 2016.

[158] J. Adler, R. Berryhill, A. Veneris, Z. Poulos, N. Veira, and A. Kastania, "Astraea: A decentralized blockchain oracle," *arXiv preprint arXiv:1808.00528*, 2018.

[159] G. Zeng, S. M. Yiu, J. Zhang, H. Kuzuno, and M. H. Au, "A nonoutsourceable puzzle under ghost rule," in *2017 15th Annual Conference on Privacy, Security and Trust (PST)*, pp. 35–358, IEEE, 2017.

[160] "How to disincentivize large bitcoin mining pools." http://hackingdistributed.com/2014/06/18/how-to-disincentivize-large-bitcoin-mining-pools/.

[161] L. Luu, Y. Velner, J. Teutsch, and P. Saxena, "Smartpool: Practical decentralized pooled mining," in *26th {USENIX} Security Symposium ({USENIX} Security 17)*, pp. 1409–1426, 2017.

[162] T. Cook, A. Latham, and J. H. Lee, "Dappguard: Active monitoring and defense for solidity smart contracts," *Retrieved July*, vol. 18, p. 2018, 2017.

[163] J. Ellul and G. J. Pace, "Runtime verification of ethereum smart contracts," in *2018 14th European Dependable Computing Conference (EDCC)*, pp. 158–163, IEEE, 2018.

[164] M. Wohrer and U. Zdun, "Smart contracts: security patterns in the ethereum ecosystem and solidity," in *2018 International Workshop on Blockchain Oriented Software Engineering (IWBOSE)*, pp. 2–8, IEEE, 2018.

[165] V. Chia, P. H. Hartel, Q. Hum, S. Ma, G. Piliouras, D. Reijsbergen, M. van Staalduinen, and P. Szalachowski, "Rethinking blockchain security: Position paper," *CoRR*, vol. abs/1806.04358, 2018.

[166] S. Xu, "Cybersecurity dynamics," in *Proc. Symposium on the Science of Security (HotSoS'14)*, pp. 14:1–14:2, 2014.

[167] S. Xu, "Emergent behavior in cybersecurity," in *Proceedings of the 2014 Symposium on the Science of Security (HotSoS'14)*, pp. 13:1–13:2, 2014.

[168] S. Xu, "Cybersecurity dynamics: A foundation for the science of cybersecurity," in *Proactive and Dynamic Network Defense* (Z. Lu and C. Wang, eds.), vol. 74, pp. 1–31, Cham: Springer International Publishing, 2019.

[169] S. Xu, W. Lu, and L. Xu, "Push- and pull-based epidemic spreading in arbitrary networks: Thresholds and deeper insights," *ACM Transactions on Autonomous and Adaptive Systems (ACM TAAS)*, vol. 7, no. 3, pp. 32:1–32:26, 2012.

[170] Y. Han, W. Lu, and S. Xu, "Characterizing the power of moving target defense via cyber epidemic dynamics," in *Proc. 2014 Symposium on the Science of Security (HotSoS'14)*, vol. 10, pp. 1–12, 2014.

[171] R. Zheng, W. Lu, and S. Xu, "Active cyber defense dynamics exhibiting rich phenomena," in *Proc. 2015 Symposium on the Science of Security (HotSoS'15)*, pp. 2:1–2:12, 2015.

[172] M. Xu, G. Da, and S. Xu, "Cyber epidemic models with dependences," *Internet Mathematics*, vol. 11, no. 1, pp. 62–92, 2015.

[173] R. Zheng, W. Lu, and S. Xu, "Preventive and reactive cyber defense dynamics is globally stable," *IEEE Trans. Network Science and Engineering*, vol. 5, no. 2, pp. 156–170, 2018.

[174] Z. Lin, W. Lu, and S. Xu, "Unified preventive and reactive cyber defense dynamics is still globally convergent," *IEEE/ACM Trans. Netw.*, vol. 27, no. 3, pp. 1098–1111, 2019.

[175] D. M. Nicol, W. H. Sanders, and K. S. Trivedi, "Model-based evaluation: From dependability to security," *IEEE Trans. Dependable Sec. Comput.*, vol. 1, no. 1, pp. 48–65, 2004.

[176] M. Pendleton, R. Garcia-Lebron, J.-H. Cho, and S. Xu, "A survey on systems security metrics," *ACM Comput. Surv.*, vol. 49, pp. 62:1–62:35, Dec. 2016.

[177] A. Ramos, M. Lazar, R. H. Filho, and J. J. P. C. Rodrigues, "Model-based quantitative network security metrics: A survey," *IEEE Communications Surveys Tutorials*, vol. 19, no. 4, pp. 2704–2734, 2017.

[178] S. Noel and S. Jajodia, *A Suite of Metrics for Network Attack Graph Analytics*, pp. 141–176. Cham: Springer International Publishing, 2017.

[179] H. Chen, J. Cho, and S. Xu, "Quantifying the security effectiveness of firewalls and dmzs," in *Proceedings of the 5th Annual Symposium on Hot Topics in the Science of Security (HoTSoS'2018)*, pp. 9:1–9:11, 2018.

[180] J. D. Mireles, E. Ficke, J.-H. Cho, P. Hurley, and S. Xu, "Metrics towards measuring cyber agility." IEEE Transaction on Information Forensics & Security, 2019 (accepted for publication).

[181] J.-H. Cho, S. Xu, P. M. Hurley, M. Mackay, T. Benjamin, and M. Beaumont, "Stram: Measuring the trustworthiness of computer-based systems," *ACM Comput. Surv.*, vol. 51, no. 6, pp. 128:1–128:47, 2019.