Validity Is Not Enough: Uncovering the Security Pitfall in Chainlink's Off-Chain Reporting Protocol

Di Zhai[† ‡ §], Jiashuo Zhang[*¶], Jianbo Gao[* † ‡§], Tianhao Liu[† ‡ §], Tao Zhang[†‡], Jian Wang[* † ‡], Jiqiang Liu[* † ‡]
[†]Beijing Key Laboratory of Security and Privacy in Intelligent Transportation, Beijing Jiaotong University
[†]School of Cyberspace Science and Technology, Beijing Jiaotong University
[§]Beijing Advanced Innovation Center for Future Blockchain and Privacy Computing
[¶]School of Computer Science, Peking University
dizhai@bjtu.edu.cn, zhangjiashuo@pku.edu.cn, {gao, leolth, taozh, wangjian, jqliu}@bjtu.edu.cn

***Abstract*—Blockchain oracles play a crucial role in delivering price data from off-chain exchanges to smart contracts, enabling automated financial services. Chainlink, the dominant oracle service provider, employs Decentralized Oracle Networks (DONs) to provide price feeds. In Chainlink's DON, multiple oracle nodes independently observe the price of a cryptocurrency and run the Off-Chain Reporting (OCR) protocol to determine a unique price from their observation values. Price deviations originating from the OCR protocol will pose security risks. To prevent arbitrary price deviations induced by Byzantine oracle nodes, OCR's validity property guarantees that the determined price is bounded by honest observation values. However, this bound in real-world settings remains unclear, and it is unknown how much price deviation Byzantine behaviors can still induce.**

In this paper, we conduct an in-depth study of the potential impacts of Byzantine behaviors on the determined price in the OCR protocol, through both empirical and theoretical analyses. First, our empirical analysis reveals that, in real-world settings, Byzantine behaviors still have ample space to sway the determined price in the OCR protocol. We then detail Byzantine behaviors that strategically sway the determined price and formally model their impacts. Furthermore, we evaluate the impacts of these Byzantine behaviors using Chainlink's real-world price data. Our experimental results show that the price deviation induced by Byzantine behaviors can reach up to 8.47% of the ETH price. Our case studies further indicate that the downstream financial impacts of a price value swayed by Byzantine behaviors can be on the order of 10^5 USD, and the cumulative impacts of such price values may reach millions of USD. In summary, this work uncovers that Byzantine behaviors can still cause non-negligible impacts on the determined price in the OCR protocol, even under the validity guarantee. We have ethically reported our findings to Chainlink, aiming to support the security of the OCR protocol.

I. Introduction

Since smart contracts cannot directly access off-chain information, blockchain oracles play a crucial role in feeding real-world data into smart contracts [1], [2], primarily price data into Decentralized Finance (DeFi) contracts. Given the inherent centralization issue [3] of a single oracle node, current oracle service providers widely adopt Decentralized Oracle Networks (DONs) [4]–[10]. A DON consists of multiple independent oracle nodes, each of which observes the price of a certain cryptocurrency from its own data source, with discrepancies among oracle nodes' observation values. These oracle nodes run a distributed oracle protocol to determine a unique representative value, which is then transmitted on-chain. Since automated financial services depend on this determined price value, deviation in it can lead to security risks, such as arbitrage and incorrect execution of DeFi contracts [11], [12]. To prevent malicious oracle nodes (i.e., Byzantine oracle nodes) from causing arbitrary deviation in the determined price value, the distributed oracle protocol is required to satisfy the validity property. **The validity property guarantees that the determined price value lies within the range bounded by the minimum and maximum observation values of honest oracle nodes**, and this range is referred to as the ***honest range***, which specifies the allowable degree of deviation.

Chainlink is the dominant oracle service provider and maintains the most widely used DONs currently [13], [14]. As of May 2025, Chainlink's DONs secure 68% of the total value across all oracle services in the entire blockchain ecosystem, and 78% on the Ethereum mainnet [15]. Chainlink's DON relies on the Off-Chain Reporting (OCR) protocol [16] to determine a unique price value that is transmitted on-chain. Thus, unlike other distributed oracle protocols [17]–[22], OCR has already seen deployment in real-world applications. However, the security of OCR, especially in terms of price deviations that may be induced by Byzantine oracle nodes, is not yet well understood. In this regard, the only guarantee provided by OCR is the validity property. OCR allows any price deviation

[*]Corresponding authors

Network and Distributed System Security (NDSS) Symposium 2026
23-27 February 2026, San Diego, CA, USA
ISBN 979-8-9919276-8-0
https://dx.doi.org/10.14722/ndss.2026.240458
www.ndss-symposium.org

within the honest range, yet how wide this range is in real-world settings remains unclear. The width of the honest range is treated as a negligible parameter in related works [21], [22] and has not been thoroughly investigated so far. Moreover, the extent of price deviations that can be induced by Byzantine behaviors in the OCR protocol remains unknown.

To fill the gap, we conduct an in-depth study on price deviations that can be induced by Byzantine behaviors in the OCR protocol. These price deviations present exploitable opportunities for adversaries to gain unfair advantages, posing threats to the fairness of price feeds and causing financial impacts on other downstream applications. Since the OCR protocol has seen long-term practical use, our study combines empirical and theoretical analyses. First, we thoroughly investigate the honest range of price observation values in Chainlink's DON through an empirical analysis. Our results show that the width of the honest range can reach up to 13.13% of the ETH price, which is far beyond prior assumptions [21], [22]. A wide honest range implies that Byzantine behaviors still have ample space to sway the determined price value, posing a potential risk despite the validity guarantee. Motivated by this finding, we conduct a theoretical analysis of Byzantine behaviors' impacts on the determined price value. Specifically, we detail each type of Byzantine behavior that can occur in the OCR protocol and how they sway the determined price value. These Byzantine behaviors strategically exploit the OCR protocol to cause an observation value, which should not have been the representative value, to be finally adopted as such. We model the impacts of these Byzantine behaviors and propose metrics to characterize the impacts. Furthermore, we evaluate the real-world price deviations that Byzantine behaviors may induce. To this end, we conduct an empirical analysis by simulating the impacts of Byzantine behaviors using Chainlink's historical price feed instances. Our experimental results show that the price deviation induced by Byzantine behaviors in the OCR protocol can reach up to 8.47% of the ETH price. Additionally, we conduct case studies on Chainlink's primary downstream applications, demonstrating that the associated financial impacts of Byzantine behaviors can be on the order of 10^5 to 10^6 USD.

Our study uncovers that merely guaranteeing validity is insufficient, as the sway of Byzantine behaviors can still induce non-negligible price deviations, constituting a security pitfall in the OCR protocol. Due to factors such as cryptocurrency price volatility, the honest range under OCR's validity property in real-world settings can be considerably wide. The potential impacts of Byzantine behaviors on the determined price value should be further bounded, and we discuss possible mitigation strategies in Section VII. Our study offers new insights into distributed oracle protocols in real-world application contexts, highlighting that assessing the impacts of Byzantine behaviors remains essential despite the validity guarantee.

We summarize our contributions as follows.

- We conduct the first thorough investigation of the honest range in real-world settings. Through an empirical analysis of 84,097 Chainlink's historical price feed instances, we reveal that, despite the validity guarantee, Byzantine behaviors can still have ample space to sway the determined price value in the OCR protocol.
- We formally model the impacts of Byzantine behaviors on the determined price value in the OCR protocol. We introduce two types of Byzantine behaviors that can strategically sway the determined price value, and we conduct a theoretical analysis of their impacts under two different scenarios of the OCR protocol.
- We evaluate the price deviations that Byzantine behaviors can induce using real-world data. We conduct an empirical analysis encompassing 72,711 filtered price feed instances, 13 downstream liquidations, and 4,465,424 downstream charge transactions to assess the impacts of Byzantine behaviors in OCR. Our experimental results demonstrate that these impacts are non-negligible.
- We open-source our datasets and code to facilitate future research. The artifact is available at https://github.com/Zhai-Di/ocr-security. Additionally, we discuss possible mitigation measures to further bound the potential impacts of Byzantine behaviors in the OCR protocol.

II. Background

A. Decentralized Oracle Networks (DONs)

Blockchain Oracle. Smart contracts built on top of the blockchain cannot directly access off-chain data, this inherence causes the blockchain and smart contracts to function as an enclosed system. However, the execution of DeFi contracts inevitably relies on price data from exchanges. To bridge the connection between the blockchain and the external world, blockchain oracles are employed to provide price feeds for DeFi contracts. Blockchain oracles also introduce security issues [23], and oracle exploits have emerged as a root cause of many DeFi incidents [24].

Distributed Oracle Problem. Centralized oracles rely on the trustworthiness of a single entity, which introduces the risk of a single point of failure [25]. To address this risk, mainstream oracle service providers have adopted DONs for price feeds. In a DON, Byzantine oracle nodes may provide falsified observation values, and even the observation values of honest oracle nodes may exhibit discrepancies. Thus, oracle nodes in the DON must agree on a single representative value that is transmitted on-chain. This is referred to as the distributed oracle problem [17], [21], [22]. In the literature, it is regarded as a specialized instance of the convex agreement problem [26], [27], and distributed oracle protocols are designed to address the distributed oracle problem. As with classical approximate agreement protocols [28], [29], distributed oracle protocols satisfy the validity property, which ensures that the representative value is always in the convex hull [26] of the honest values. For scalar values, this property means that the representative value lies within the range bounded by the minimum and maximum honest values.

B. Chainlink's Price Feed Scheme

Chainlink's price feed scheme comprises both off-chain and on-chain components, as shown in Figure 1.

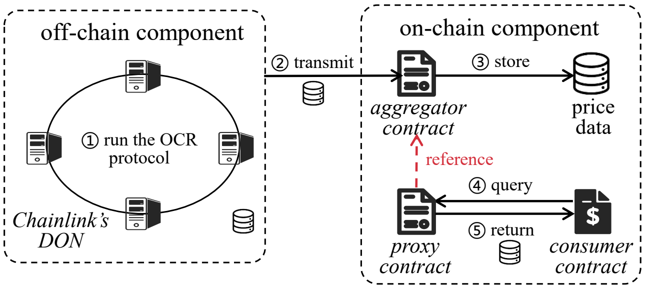


Fig. 1: Chainlink's price feed scheme

The off-chain component refers to Chainlink's DON. For a given blockchain, such as Ethereum, Chainlink deploys a DON for each asset pair it currently supports to provide price feeds. The number of oracle nodes in Chainlink's DON is typically set to 16, 19, or 31, with these oracle nodes designated by Chainlink. Chainlink previously hosted oracle competitions and selected winners to join specific DONs. However, Chainlink's DONs for price feeds currently do not allow external nodes to join. Each oracle node in Chainlink's DON has a unique cryptographic key pair, and receives bounties from Chainlink by providing observation values. The oracle nodes in Chainlink's DON run the OCR protocol [16] to determine a representative price value.

The on-chain component includes the aggregator contract, the proxy contract, and the consumer contracts. The aggregator contract interacts directly with the off-chain DON. Specifically, this contract records the on-chain addresses of each oracle node in the DON and receives price observation values transmitted from the DON. These observation values are stored in the aggregator contract. Furthermore, the aggregator contract provides a set of functions for retrieving price data, which return the median of the observation values from multiple oracle nodes. The proxy contract references a specific aggregator contract and ensures the continuity of Chainlink's price feed services when the underlying aggregator contract is upgraded, thereby providing significant convenience to users. A consumer contract refers to any contract that consumes Chainlink's price feeds, such as a DeFi contract. When querying the price of a cryptocurrency, the consumer contract calls one of the exposed functions of the proxy contract, without having to concern the underlying aggregator contract.

III. Review of Chainlink's OCR Protocol

In this section, we review the OCR protocol. We first introduce the system model, upon which the OCR protocol is built. We then present a protocol overview and the proof of OCR's validity property.

A. System Model

The off-chain DON consists of n oracle nodes, among which f are Byzantine oracle nodes [30]. Byzantine oracle nodes can arbitrarily deviate from the protocol, subject only to the underlying cryptographic assumptions. The remaining oracle nodes, referred to as honest oracle nodes, are assumed to behave honestly and conform to the protocol. The system satisfies the condition $n \geq 3f+1$, and the current configuration adopts $n = 3f + 1$ to achieve optimal resilience. The connections between oracle nodes are authenticated and encrypted, enabling them to send point-to-point messages in a network under the partial synchrony assumption [31], i.e., there exists an unknown Global Stabilization Time (GST), after which the message transmission delay between any two honest nodes has a known upper bound.

B. OCR Protocol Overview

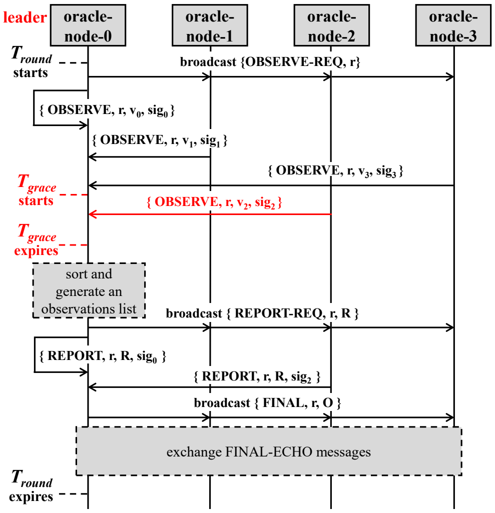
We describe the three sub-protocols that constitute OCR, namely, *pacemaker*, *report generation*, and *transmission*, with the **median** adopted as the representative value. Notably, our study primarily focuses on *report generation*.

1) Pacemaker: Pacemaker periodically initiates a new *report generation* instance. Specifically, *pacemaker* is responsible for triggering a designated leader to initiate a *report generation* instance. Furthermore, when the leader fails to make sufficient progress, *pacemaker* is capable of aborting the current *report generation* instance.

2) Report Generation: A *report generation* instance proceeds in multiple rounds, with a new round beginning every T_{round} time units. In each round, a unique oracle node is selected as the **leader**, while the remaining oracle nodes are referred to as **non-leaders**. *Report generation* generates one report per round. At the start of round r, the leader initiates the process by broadcasting OBSERVE-REQ messages to all oracle nodes, including itself. Upon receiving the OBSERVE-REQ message, the oracle node moves to round r, obtains a new observation value (e.g., ETH/USD price) from its own data source, signs it, and sends it back to the leader in an OBSERVE message. When the leader has gathered OBSERVE messages from $2f + 1$ distinct oracle nodes, it enters a grace period of duration T_{grace}, during which delayed observation values are awaited. The configuration of this grace period is motivated by Chainlink's reward mechanism, which incentivizes oracle nodes for their participation as observers and allows slightly delayed oracle nodes to remain eligible for rewards. Moreover, the resulting report is enriched with a broader set of observation values.

Upon the expiration of the grace period, the leader sorts the received observation values in non-decreasing order (allowing for duplicates), forming an ordered list referred to as the **observations list**. The observations list, together with the signatures from all observers, constitutes a report. The leader then includes this report in a REPORT-REQ message and broadcasts it to all oracle nodes. Upon receiving the REPORT-REQ message, the oracle node first verifies all signatures in the report. It then checks whether the time interval (e.g., the update interval for ETH/USD price is 1h under normal circumstances) is sufficient, or the price has fluctuated beyond a certain threshold (e.g., 0.5% for ETH/USD price) compared to the last price feed. If either condition is met, the oracle node signs the report, indicating that the oracle node validates this report. The oracle node subsequently sends a REPORT message containing its signed report to the leader.

Once a report has been validated by at least $f+1$[1] distinct oracle nodes, the leader assembles this report along with the signatures of the oracle nodes that validated it, generating an **attested report**. The leader then includes the attested report in a FINAL message, which is broadcast to all oracle nodes. Upon receiving the FINAL message, the oracle node verifies the signatures in the corresponding attested report, and then sends a FINAL-ECHO message containing the attested report to all other oracle nodes. To this end, the attested report is disseminated, ensuring that oracle nodes who have not yet received the FINAL message can quickly obtain it. When at least $f+1$[1] oracle nodes confirm receipt of the attested report, round r is completed, and the attested report is passed to *transmission*, from where it will be sent to the aggregator contract. The process that we primarily focus on is illustrated in Figure 2.



leader
oracle-node-0
oracle-node-1
oracle-node-2
oracle-node-3
T_{round} starts
broadcast {OBSERVE-REQ, r}
{ OBSERVE, r, v_0, sig_0 }
{ OBSERVE, r, v_1, sig_1 }
{ OBSERVE, r, v_3, sig_3 }
T_{grace} starts
{ OBSERVE, r, v_2, sig_2 }
T_{grace} expires
sort and generate an observations list
broadcast { REPORT-REQ, r, R }
{ REPORT, r, R, sig_0 }
{ REPORT, r, R, sig_2 }
broadcast { FINAL, r, O }
exchange FINAL-ECHO messages
T_{round} expires

Fig. 2: The illustrated DON consists of four oracle nodes, i.e., $n=4$, $f=1$, and the leader is oracle-node-0 in the current round. Here, sig_i denotes the signature of oracle-node-i, and v_i denotes the observation value of oracle-node-i. r denotes the current round number, R denotes the report generated by the leader, and O denotes the attested report.

3) Transmission: Transmission encapsulates the process by which oracle nodes send the attested report to the on-chain aggregator contract. Oracle nodes wait for the first report accepted by the aggregator contract. The consensus mechanism of the blockchain ensures that all honest oracle nodes see the same representative value (i.e., the median of the observations list) from this first accepted report and reach an agreement.

4) Representative Value: **The median of the observations list in the attested report that is first accepted by the aggregator contract will serve as the representative price value.** When a consumer contract requests cryptocurrency price data, the proxy contract simply returns this median. Since integer division rounds towards zero, the median is essentially the value at index $\lfloor \texttt{observations\_list.length}/2 \rfloor$. Recall that the leader enters the grace period after receiving $2f+1$ OBSERVE messages from distinct oracle nodes in *report generation*, and it waits for any potentially delayed observation values. Thus, the length of an observations list is not fixed, and the only guarantee is that it is at least $2f+1$. The OCR protocol satisfies the validity property, as detailed in Property 1.

Property 1 (Validity)**.** *Let L be the observations list in an attested report. L is sorted in non-decreasing order. Let m be the median of L, and let V_h denote the set of observation values from all the honest oracle nodes. It follows that* $\min(V_h) \leq m \leq \max(V_h)$.

Proof: We prove this property by contradiction. Without loss of generality, we assume that $m > \max(V_h)$. Since L contains at least $2f+1$ observation values, there are at least $f+1$ values in L that are greater than $\max(V_h)$. These $f+1$ observation values are all from Byzantine oracle nodes. However, according to the system model, there can be at most f Byzantine oracle nodes in the DON, which contradicts the assumption. Similarly, we can prove that $m \geq \min(V_h)$.

IV. Potential Risk despite the Validity Guarantee

According to the validity property, any price deviation within the honest range is allowed by the OCR protocol. However, how wide the honest range actually is remains unclear. To measure the honest range of price observation values in Chainlink's DON, we conduct a large-scale empirical study. To the best of our knowledge, we are the first to thoroughly investigate the honest range based on long-term price observation values in real-world settings. Our empirical analysis reveals that the honest range may still leave ample space for Byzantine behaviors to sway Chainlink's price data, posing a potential risk despite the validity guarantee.

A. Dataset

According to the code of Chainlink's aggregator contract [32], it emits a `NewTransmission` event upon receiving an attested report generated by the OCR protocol, and the parameters of `NewTransmission` event include the observations list contained in the report. Thus, we retrieved all `NewTransmission` events from block 12016450 (approximately when Chainlink began adopting OCR in March 2021) to block 22648562 (June 2025) through a full node on the Ethereum mainnet [33]. We parsed the `data` field of each `NewTransmission` event to extract the observations list.

With reference to the addresses of proxy contracts listed on Chainlink's official website [34], we use the database of Dune [35] to compute the cumulative number of calls made to the proxy contracts for each pairs on the Ethereum mainnet.

[1] Oracle nodes reach agreement on the representative value by leveraging the consensus mechanism of the blockchain they provide price feeds to.

As of 9 June 2025, we observe that 60.61% of the total calls were directed to the proxy contract for ETH/USD pair, indicating that ETH/USD price feed is the most demanded on the Ethereum mainnet. Therefore, we primarily analyze the ETH/USD price observation values, which includes 84,097 observations lists, a dataset we consider to be highly representative. Additionally, we collect BNB/USD price observation values using the same approach and extend our empirical analysis to the BNB/USD pair in Appendix D.

B. Measuring the Honest Range

Since the Chainlink's DON for ETH/USD price feeds consists of 31 oracle nodes (i.e., $n = 31$, $f = 10$), an observations list should contain 31 observation values under ideal conditions, where all 31 oracle nodes honestly conform to the OCR protocol and submit their observations to the leader before the grace period expires. Under such ideal conditions, any differences among the observation values should merely reflect **natural variations**, which arise due to differences in the data sources used by the oracle nodes and slight inconsistencies in the times at which they retrieve prices. However, we observed the following phenomena in the collected observation values.

- 13.46% of the observations lists, which we refer to as **incomplete lists**, contain fewer than 31 observation values. To the best of our knowledge, Chainlink has not published any reports specifically focused on this phenomenon. We infer that the underlying causes may include: ❶ some oracle nodes becoming unresponsive due to transient benign faults (e.g., crashes); ❷ high network latency or slow responses from queried data sources, resulting in the oracle node's OBSERVE message arriving at the leader after the grace period expired; and ❸ certain oracle nodes exhibiting Byzantine behaviors.
- Some observation values, which we refer to as **anomalies**, exhibit significant deviations from the other values within the same observations list, which cannot reasonably be attributed to natural variations. Table I presents several illustrative examples.

TABLE I: Examples of Anomalies.

Anomalies	Other Observations	Corresponding Event Logs
ob_1: 1654.79 ob_2: 1689.83	min: 3284.27 max: 3290.16	0x8252......1537/Logs
ob_1: 1639.27	min: 3273.99 max: 3278.50	0x394d......38d1/Logs
ob_{30}: 4810.35 ob_{31}: 4813.19	min: 3843.08 max: 3872.82	0x2fdd......d6d8/Logs
ob_{30}: 27047822019 ob_{31}: 27047822020	min: 3861.71 max: 3891.84	0xf9a3......608f/Logs
ob_{30}: 21759587641 ob_{31}: 21759587647	min: 3862.99 max: 3891.01	0x633f......c030/Logs
ob_{30}: 4591094991 ob_{31}: 4591095006	min: 3851.12 max: 3886.11	0xf289......9819/Logs

ob_i denotes the i-th value in an observations list.

These phenomena indicate that historical observation values should not all be regarded as honest and unaffected by other faults. However, identifying the root causes of incomplete lists and anomalies in the historical observation values is nearly impossible, as doing so would require access to the oracle nodes' internal execution logs corresponding to those time periods. Moreover, even values from honest oracle nodes, which are referred to as **honest observation values**, may exhibit a certain degree of natural variations. It is both challenging and of little practical significance to exactly determine whether each observation value is honest. Therefore, we first estimate the historical maximum of natural variations between honest observation values and then exclude lists that may contain anomalies based on this estimate, thereby enabling the remaining lists to be treated as consisting entirely of honest observation values. The corresponding rationale and experiments are detailed as follows.

1) Estimating the Maximum Natural Variation: We first exclude all incomplete lists to avoid the potential influence of Byzantine behaviors. For each complete list, let ξ denote the minimum difference between the largest and smallest values in any $2f + 1$-*subset* of a complete list. Since there are at least $2f + 1$ honest observation values in a complete list, the maximum natural variation is no smaller than ξ. The detailed proof is provided in Theorem 2 (Appendix A). **In the following, we use m to denote the median of an observations list.** We calculate ξ for each complete list, and, considering ETH price volatility, we normalize ξ relative to m (i.e., ξ/m). Among the 72,777 complete lists, the maximum value of ξ/m is 0.0851, indicating that historically there existed a natural variation equal to or exceeding 8.51% of the ETH price at that time.

2) Excluding Anomalies: We analyze the characteristics exhibited by an observations list that contains anomalies, as detailed in Theorem 1.

Theorem 1. *Let L denote an observations list consisting of $3f + 1$ values. Let ob_i denote the i-th value in L, where $1 \leq i \leq 3f + 1$. If there exists an anomaly in L, which is abnormally large (resp. small) that the difference between it and any honest observation value in L exceeds ϵ, then $ob_{3f+1} - ob_{2f+1} > \epsilon$ (resp. $ob_{f+1} - ob_1 > \epsilon$).*

Proof: Without loss of generality, assume a large anomaly ob_{al} exists in L that the difference between it and any honest observation value exceeds ϵ. Since L contains at most f observation values from Byzantine oracle nodes, there exists at least one honest observation value among the largest $f + 1$ values in L, which is denoted as ob_{hl}. As L is sorted in non-decreasing order, we have $ob_{hl} \geq ob_{2f+1}$ and $ob_{3f+1} \geq ob_{al}$. Because $ob_{al} - ob_{hl} > \epsilon$, it follows that $ob_{3f+1} - ob_{2f+1} > \epsilon$. Similarly, when there exists a small anomaly ob_{as} in L, we can prove that $ob_{f+1} - ob_1 > \epsilon$.

We apply the characteristics revealed by Theorem 1 to set filtering conditions, excluding lists potentially containing anomalies. Specifically, we define the parameter ϵ in Theorem 1 to represent the maximum acceptable natural variation,

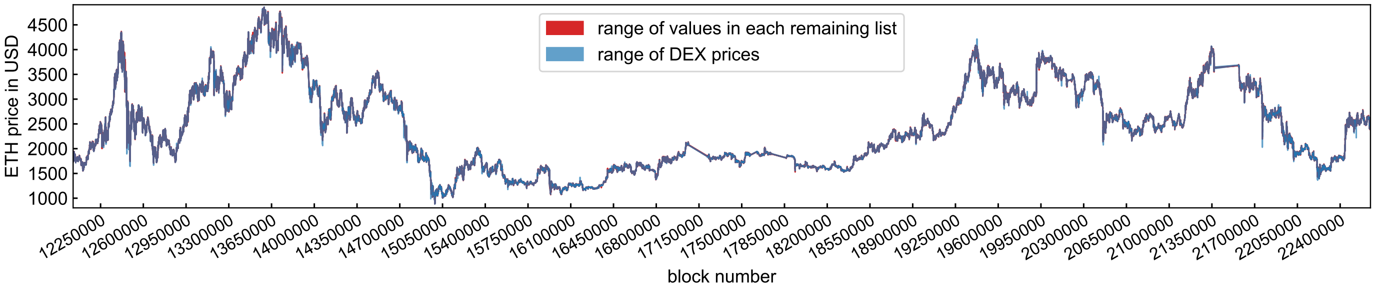


Fig. 3: Alignment between the remaining lists after filtering and the DEX prices

and exclude all lists that satisfy either $ob_{3f+1} - ob_{2f+1} > \epsilon$ or $ob_{f+1} - ob_1 > \epsilon$. Given that historically the maximum natural variation is at least 8.51% of m, we set $\epsilon = 0.0851 \times m$. The sensitivity analysis of ϵ, presented in Appendix B, demonstrates that our parameter choice has no significant effect on the subsequent experimental results. After our filtering, there remain 72,711 lists.

Furthermore, we align each remaining list with DEX prices to reference market moves. For each historical list, we take the block in which the price feed event occurs as the center block, and set a 6-block lead window and a 6-block lag window[2]. Within each 13-block window, we collected WETH/USDT price data from both Uniswap [36] and Sushiswap [37]. Thus, we aligned the range of values in each remaining list with the range of DEX prices within the corresponding 13-block window, and the results are presented in Figure 3. It can be observed that the remaining lists after filtering exhibit an overall trend consistent with the DEX price movements. Moreover, the range of values in each remaining list (represented by the red-shaded area in Figure 3) is almost covered by the range of DEX prices within the corresponding 13-block window (represented by the blue-shaded area in Figure 3). The results indicate that, after our filtering, the differences among observation values in each remaining list are primarily attributable to market volatility rather than oracle noise, and thus the remaining 72,711 lists are reasonably regarded as not containing any anomalies, with all observation values considered honest.

For each remaining list, we calculate the difference between its maximum and minimum values, which represents the width of the honest range, and normalize the results relative to m. The results are shown in Figure 4, which reveals that the width of the honest range in real-world settings can far exceed the related assumptions made in existing studies [21], [22]. For instance, there exists an honest range [3076.76, 3510.31][3], whose width reaches 433.55 USD, accounting for 13.13% of the ETH price (i.e., the median) at that moment. Within the 13-block window corresponding to this price feed, the market indeed experienced significant volatility. Even a single data source, Sushiswap, exhibits an ETH price change of 10.59% within this window, and diverse data sources used by Chainlink's oracle nodes result in larger price differences.

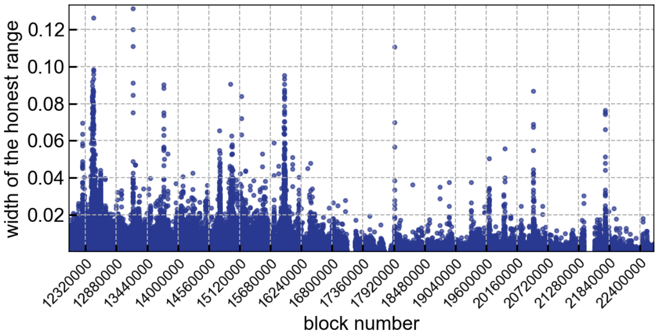


Fig. 4: Width of the honest range in historical price observation values

C. Our Finding

Our empirical study suggests that the honest range of real-world price observation values may have been overlooked, and the wide honest range introduces a potential risk. The validity property of OCR merely guarantees that Byzantine oracle nodes cannot push the determined price value outside the honest range. However, with a wide honest range, Byzantine behaviors can sway Chainlink's price data to a large extent. For instance, when the honest range is [3076.76, 3510.31][3], the ETH price determined by the OCR protocol without the presence of Byzantine behaviors is 3302.30 USD, whereas Byzantine behaviors may shift this price to 3510.31 USD or to 3076.76 USD. Since price deviations affect the transaction funds of downstream applications, Byzantine behaviors' sway on price data undermines the fairness of price feeds, providing adversaries with exploitable opportunities. Therefore, even under the guarantee of the validity property, it remains necessary to concretely assess the extent to which Byzantine behaviors in the OCR protocol can sway Chainlink's price data.

V. Byzantine Behaviors' Sway: Modeling Its Impacts on Price Data

Motivated by our finding (Section IV-C), we further conduct a theoretical analysis of the extent to which Byzantine behaviors can sway the determined price value in the OCR

[2] Following the default configuration of the OCR protocol, we estimate the duration of one OCR round to be 60s and adopt 6 blocks as the lead/lag window.

[3] The corresponding event log is in 0x80b9019cb90645cd4451dd2c19fc
17d1cbe20955941b9e282d195e244c96dfd4/Logs

protocol. Specifically, we first present the threat model and then propose the metrics for characterizing the impacts of Byzantine behaviors on the determined price value. Building on this, we model the impacts of Byzantine behaviors under two different scenarios of OCR.

A. Threat Model

Based on Section III-A, we consider a DON comprising n oracle nodes, among which f are Byzantine, where $n = 3f+1$. Our threat model assumes that Byzantine oracle nodes collude to maximize (or minimize) the price value determined by the OCR protocol for creating unfair advantages (e.g., arbitrage, frontrunning [12], [38]). This assumption is widely adopted in related studies [21], [22], [27], [28], [39], [40]. According to the OCR protocol, Byzantine oracle nodes cannot learn about any honest oracle node's observation values before receiving the observations list generated by the leader. We consider attack approaches that aim to intercept or infer the honest observation values to fall outside the scope of this paper. In each OCR round, a Byzantine oracle node acts as either the leader or a non-leader, depending on the output of a pseudo-random function. The capabilities of a Byzantine leader and a Byzantine non-leader are distinguished by their respective Byzantine behaviors, which are detailed below.

Byzantine non-leader's behavior:

Falsify the observation value. In *Report Generation*, upon receiving an OBSERVE-REQ message from the leader, a Byzantine oracle node first obtains a price observation value from data sources as normal. This correctly obtained observation value is referred to as the **original observation value**. However, this Byzantine oracle node then falsifies the original observation value to a sufficiently large (resp. small) value[4] so that the value will move to the end (resp. beginning) of the observations list, thereby swaying the median. Since such falsification is performed without knowledge of the distribution of other honest observation values, it is realistic to assume that a Byzantine oracle node typically falsifies a value significantly larger (or smaller) than its original observation value (e.g., an original observation value of 3702.42 USD might be falsified to 4302.42 USD). This Byzantine oracle node includes the falsified observation value in an OBSERVE message, signs it, and sends it to the leader. Upon verifying the signature of the sending oracle node, the leader then generates an observations list using this falsified observation value along with the other received observation values.

Byzantine leader's behavior:

Select a subset of observation values. In *Report Generation* of the OCR protocol, upon the expiration of the grace period and receiving observation values from more than $2f+1$ distinct oracle nodes, the Byzantine leader does not include all received values in the observations list as required. Instead, aiming to maximally sway the median, the Byzantine leader selects the $2f+1$ largest (resp. smallest) values from the received observation values to generate the observations list, discarding the rest, which exactly satisfies the threshold required by the OCR protocol. Such selection leads to a larger (resp. smaller) resulting median. Notably, such selection would go undetected, as a Byzantine leader could deceitfully claim that the $2f+1$ observation values finally included are all that it originally received.

B. Metrics

Byzantine behaviors can change the observations list (whose original length is denoted by l), resulting in a value that was not originally the median (i.e., its index is not equal to $\lfloor l/2 \rfloor$) becoming the final median. This is the essential reason why Byzantine behaviors in the OCR protocol can sway Chainlink's price data. Since an observations list is sorted in non-decreasing order, we leverage indices in the list to reflect changes in the median and propose metrics to characterize the impacts of Byzantine behaviors. The relevant definitions are as follows.

Definition 1 (Original Observations List: L_{ori})**.** Assuming that no Byzantine behaviors occur, each Byzantine oracle node conforms to the OCR protocol just like an honest oracle node. The observations list that will be formed under this ideal situation is referred to as the **original observations list**, denoted by L_{ori}, and it is sorted in non-decreasing order. Notably, when Byzantine behavior occurs, L_{ori} serves as an ideal reference list that does not actually exist.

Definition 2 (Final Observations List: L_{fin})**.** In *Report Generation* of the OCR protocol, Byzantine non-leaders falsify their observation values or the Byzantine leader selects a subset of observation values to generate the observations list. Under the aforementioned Byzantine behaviors, the finally generated observations list L_{fin} is referred to as the **final observations list**, which is sorted in non-decreasing order.

Definition 3 (**Metric:** Byzantine-Induced Price Deviation (ΔP) and Byzantine-Induced Index Deviation (ΔI))**.** The median of the final observations list L_{fin} is denoted as M_{fin}. Assuming all other conditions remain unchanged, while there is no Byzantine behavior, the original observations list is L_{ori}, with the median of L_{ori} being M_{ori}. Let ΔP denote the price deviation induced by Byzantine behaviors, referred to as the ***Byzantine-induced price deviation***. Then we have

$$\Delta P = |M_{fin} - M_{ori}|. \tag{1}$$

Let I_{ori} denote the index of M_{ori} in L_{ori}. Notably, in the scenarios we analyze below, M_{fin} also exists in L_{ori}. Thus, let I_{fin} be the index of M_{fin} in L_{ori}. Let ΔI denote the index deviation induced by Byzantine behaviors, referred to as the ***Byzantine-induced index deviation***. Then we have

$$\Delta I = |I_{fin} - I_{ori}|. \tag{2}$$

We use ΔP as the metric for characterizing the price deviation under the sway of Byzantine behaviors. Clearly, a larger ΔP indicates a greater impact of Byzantine behaviors

[4] Due to price volatility, the OCR protocol does not set thresholds on observation values, and Table I demonstrates that significantly deviating observation values are permitted.

on the determined price value. Since ΔP is correlated with the concrete positions in L_{ori} occupied by Byzantine oracle nodes' original observation values, we adopt the mathematical expectation $\mathbb{E}(\Delta P)$ to capture the average price deviation that can be induced by Byzantine behaviors. Given that L_{ori} is sorted in non-decreasing order, we leverage ΔI to reflect ΔP in our theoretical analysis.

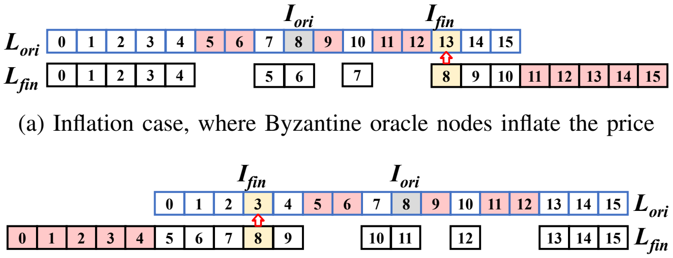
Definition 4 (**Metric:** Maximum Uncertain Price Deviation)**.** In *Report Generation* of the OCR protocol, f Byzantine oracle nodes exhibit the aforementioned Byzantine behaviors. Given a certain original observations list, if the f Byzantine oracle nodes aim to inflate the price, they can shift the median from M_{ori} to M'_{fin}; conversely, if they aim to deflate it, they can shift the median to M''_{fin}. Let $U_{max} = \max(M'_{fin} - M''_{fin})$. We define U_{max} as the ***maximum uncertain price deviation***.

We use U_{max} as the metric for characterizing the maximum deviation between the prices resulting from Byzantine behaviors with different aims. From the perspective of price feed users, it is unknown whether the potential f Byzantine oracle nodes aim to inflate or deflate the price. Byzantine behaviors with unknown aims induce uncertainty in the determined price value, and this metric captures the maximum deviation under such uncertainty. Thus, this metric should be effectively bounded **(when necessary, should be much less than the width of the honest range)**. Similarly, since L_{ori} is sorted in non-decreasing order, $\max(I'_{fin} - I''_{fin})$ can reflect $\max(M'_{fin} - M''_{fin})$.

C. Impacts of Byzantine Behaviors under Two Scenarios

We model the impacts of Byzantine behaviors under two possible scenarios, which are distinguished by whether the leader in the current OCR round is Byzantine. In Scenario I, the leader is honest, whereas in Scenario II, the leader is Byzantine. The OCR protocol uses a pseudo-random function to select the unique leader for several (a configurable parameter) consecutive rounds. Under the static adversary model, if one of the Byzantine oracle nodes is selected as the leader (with probability approximately 1/3), Scenario II occurs. Otherwise, Scenario I occurs. Under the adaptive adversary model, given that the selected leader may be subsequently corrupted, Scenario II occurs more frequently. Overall, these two scenarios are applicable under either adversary model.

1) Scenario I (Honest Leader): Under Scenario I, Byzantine oracle nodes can merely exhibit falsifying behavior, while selecting behavior does not occur. With the goal of maximally inflating (resp. deflating) the price, all f Byzantine oracle nodes falsify their respective original observation values into sufficiently large (resp. small) values. These falsified values will be moved to the last (resp. first) f positions of the final observations list. According to Property 1, the price (i.e., M_{fin}) does not take any of these falsified values. However, by using the original observations list as a reference, the Byzantine behaviors shift the median index from I_{ori} to I_{fin}, as illustrated in Figure 5.



(a) Inflation case, where Byzantine oracle nodes inflate the price

(b) Deflation case, where Byzantine oracle nodes deflate the price

Fig. 5: **Scenario I.** The illustrated DON consists of 16 oracle nodes, 5 of which are Byzantine and perform falsification. The blue-bordered boxes indicate the positions of each observation value in L_{ori}, while the black-bordered boxes indicate the positions of each observation value in L_{fin}. The numbers inside the boxes denote the indices of the observation values. The original observation values obtained by the falsifying oracle nodes should have fallen into the red-marked positions in L_{ori}; however, after falsification, they instead fall into the red-marked positions in L_{fin}.

Since Byzantine oracle nodes cannot control honest oracle nodes' values, their f original observation values occupy random positions (i.e., indices) in the ordering of a list. As a result, the index deviation exhibits randomness. For instance, Figure 6 shows a case with index deviation 0, where Byzantine behaviors cause no price deviation.

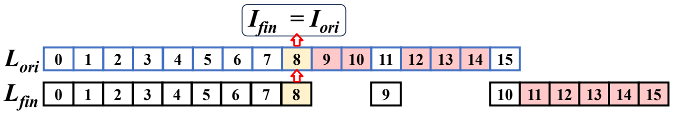


Fig. 6: Due to the particular positions of those falsifying oracle nodes' original observation values in L_{ori}, the median remains unchanged, i.e., $I_{fin} = I_{ori}$.

As a foundation for evaluating the metric *Byzantine-induced price deviation* under Scenario I, we use classical probability theory to derive the general regularity of the index deviation. In the inflation (resp. deflation) case, f values in L_{ori} are moved to the end (resp. beginning) of L_{fin}, which may cause the position of the median to shift backward (resp. forward). To calculate the probability distribution of ΔI (i.e., the index deviation induced by Byzantine behaviors), we introduce a discrete random variable X that describes the random positions of the falsifying oracle nodes' original observation values in L_{ori}. Specifically, X $(0 \leq X \leq f)$ denotes that, among the f original observation values, exactly X have indices positioned before I_{fin}. Let l denote the length of L_{ori} (where $l = 3f+1$), such that $I_{ori} = \lfloor l/2 \rfloor$. Eq. 3 presents the probability that X takes the value x. Eq. 4 (or Eq. 5) shows that ΔI is a random variable function with respect to X, allowing us to calculate

TABLE II: The probability distribution of ΔI in Scenario I when $l = 31$, where $\Pr[\Delta I = \Delta i]$ denotes the corresponding probability, which is identical in both the inflation and deflation cases.

ΔI	0	1	2	3	4	5	6	7	8	9	10
$\Pr[\Delta I = \Delta i]$	0.00007	0.00072	0.00395	0.01457	0.04038	0.08809	0.15416	0.21533	0.23216	0.17688	0.07370

the probability distribution of ΔI using $\Pr[X = x]$.

$$\Pr[X = x] = \frac{\binom{I_{fin}}{x} \cdot \binom{l-1-I_{fin}}{f-x}}{\binom{l}{f}}. \tag{3}$$

$$\begin{aligned} I_{fin} &= \left\lfloor \frac{l}{2} \right\rfloor + X, \quad \text{(inflation case)} \\ \Delta I &= \left| \left\lfloor \frac{l}{2} \right\rfloor + X - \left\lfloor \frac{l}{2} \right\rfloor \right| = X. \end{aligned} \tag{4}$$

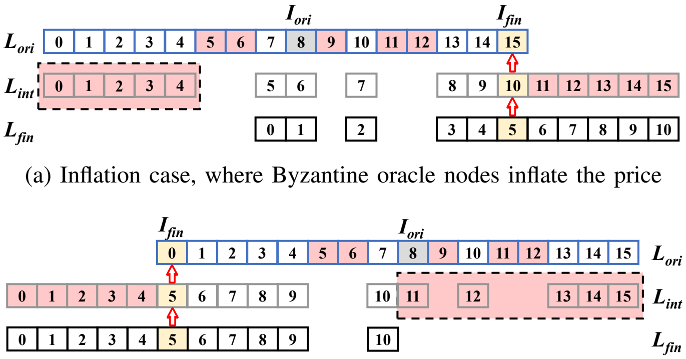
$$\begin{aligned} I_{fin} &= \left\lfloor \frac{l}{2} \right\rfloor - f + X, \quad \text{(deflation case)} \\ \Delta I &= \left| \left\lfloor \frac{l}{2} \right\rfloor - f + X - \left\lfloor \frac{l}{2} \right\rfloor \right| = f - X. \end{aligned} \tag{5}$$

For instance, when $l = 31$ (i.e., the size of Chainlink's DON for ETH/USD price feeds), the probability distribution of ΔI is shown in Table II. According to Table II, the mathematical expectation of ΔI (i.e., $\mathbb{E}(\Delta I)$) can be calculated, and the result is $\mathbb{E}(\Delta I) = 7.27$. Therefore, under Scenario I, Byzantine behaviors cause the median to shift approximately 7 positions in L_{ori} on average.

As a foundation for evaluating the metric *maximum uncertain price deviation* under Scenario I, we analyze the maximum uncertain index deviation. For the f falsifying oracle nodes, when they aim to inflate (resp. deflate) the price and shift the median to M'_{fin} (resp. M''_{fin}), corresponding to index I'_{fin} (resp. I''_{fin}) in L_{ori}, it always holds that $I'_{fin} - I''_{fin} \leq 2f$. In the illustration shown in Figure 5, $I'_{fin} - I''_{fin} = 2f$. Thus, under Scenario I, $\max(I'_{fin} - I''_{fin}) = 2f$. Furthermore, a more general conclusion can be drawn: when the observation values from Byzantine oracle nodes in L_{ori} are indexed within the interval $[\lfloor l/2 \rfloor - f + 1,\ \lfloor l/2 \rfloor + f - 1]$, then under Scenario I, $I'_{fin} - I''_{fin}$ reaches its maximum value of $2f$.

2) Scenario II (Byzantine Leader): Under Scenario II, both selecting and falsifying behaviors can occur, and these Byzantine behaviors can be divided into two steps for a clearer explanation. In the first step, similar to Scenario I, all f Byzantine oracle nodes falsify their respective original observation values into sufficiently large (resp. small) values. These falsified values are sent to the Byzantine leader, who sorts them into a non-decreasing list. We refer to this intermediate list as L_{int}, with the f falsified values occupying the first (resp. last) f positions.

In the second step, to further inflate (resp. deflate) the price, the Byzantine leader selects the last (resp. first) $2f+1$ values in L_{int} and discards the remaining observation values. Such selection results in the final observations list consisting of f falsified values and $f+1$ additional largest (resp. smallest) values, thereby making the final median (i.e., M_{fin}) as large (resp. small) as possible. By using the original observations list as a reference, the Byzantine oracle nodes shift the median index from I_{ori} to I_{fin}, as illustrated in Figure 7.



(a) Inflation case, where Byzantine oracle nodes inflate the price

(b) Deflation case, where Byzantine oracle nodes deflate the price

Fig. 7: **Scenario II.** The gray box represents L_{int}. The observation values at the red-marked positions in L_{ori} are falsified and moved to the red-marked positions in L_{int}. The observation values inside the red dashed box are discarded by the Byzantine leader.

In Scenario II, since the first step mainly involves f Byzantine non-leaders performing falsification, the index deviation induced by Byzantine behaviors also exhibits a certain degree of randomness. Notably, in the second step, the Byzantine leader has access to all observation values sent to it and can always select the largest (resp. smallest) $2f+1$ values. As a result, Byzantine behaviors under Scenario II always achieve $\Delta I > 0$.

As a foundation for evaluating the metric *Byzantine-induced price deviation* under Scenario II, we calculate the probability distribution of ΔI by introducing a discrete random variable Y, which describes the random positions of the falsifying oracle nodes' original observation values in L_{ori}. Specifically, Y $(0 \leq Y \leq f)$ denotes that, among the f original observation values, exactly Y have indices positioned before I_{fin}. Analogous to the analysis under Scenario I, l denotes the length of L_{ori}. Eq. 6 presents the probability that Y takes the value y. Eq. 7 (or Eq. 8) shows that ΔI is a random variable function with respect to Y, and we calculate the probability distribution of ΔI using $\Pr[Y = y]$.

$$\Pr[Y = y] = \frac{\binom{I_{fin}}{y} \cdot \binom{l-1-I_{fin}}{f-y}}{\binom{l}{f}}. \tag{6}$$

$$\begin{aligned} I_{fin} &= \left\lfloor \frac{2f+1}{2} \right\rfloor + f + Y = 2f + Y, \\ \Delta I &= 2f + Y - \left\lfloor \frac{l}{2} \right\rfloor. \quad \text{(inflation case)} \end{aligned} \tag{7}$$

$$
\begin{aligned}
I_{fin} &= \left\lfloor \frac{2f+1}{2} \right\rfloor - f + Y = Y, \\
\Delta I &= \left\lfloor \frac{l}{2} \right\rfloor - Y. \quad \text{(deflation case)}
\end{aligned}
\tag{8}
$$

Table III presents the probability distribution of ΔI under Scenario II for the setting $l = 31$ (i.e., the size of Chainlink's DON for ETH/USD price feeds). Furthermore, we calculate the mathematical expectation of ΔI, and the result is $\mathbb{E}(\Delta I) = 14.55$. This result indicates that, under Scenario II, Byzantine behaviors shift the median by approximately half the length of L_{ori} on average, nearly shifting it to the maximum (resp. minimum) value of L_{ori}.

As a foundation for evaluating the metric *maximum uncertain price deviation* under Scenario II, we analyze the maximum uncertain index deviation. For the f falsifying oracle nodes, when they aim to inflate (resp. deflate) the price and shift the index of the median to I'_{fin} (resp. I''_{fin}) in L_{ori}, it always holds that $I'_{fin} - I''_{fin} \leq 3f$. In the illustration shown in Figure 7, $I'_{fin} - I''_{fin} = 3f$. Thus, under Scenario II, $\max(I'_{fin} - I''_{fin}) = 3f$. Furthermore, a more general conclusion can be drawn: when the observation values from Byzantine oracle nodes in L_{ori} are indexed within the interval $[1, 3f-1]$, then under Scenario II, $I'_{fin} - I''_{fin}$ reaches its maximum value of $3f$. This indicates the potential existence of f Byzantine oracle nodes that can sway the determined price to the maximum value as well as the minimum value in L_{ori}, thereby introducing significant uncertainty to downstream transaction funds.

VI. Evaluation

Following the theoretical analysis, we evaluate the real-world impacts of Byzantine behaviors in the OCR protocol on Chainlink's price data and associated downstream applications. Specifically, we conduct an empirical analysis to answer the following three Research Questions (RQs).

- **RQ1.** How much deviation can Byzantine behaviors in the OCR protocol induce in Chainlink's real-world price data?
- **RQ2.** To what degree can Byzantine behaviors in the OCR protocol induce uncertainty in Chainlink's real-world price data?
- **RQ3.** To what degree can price deviations induced by Byzantine behaviors impact the transaction funds in the downstream applications?

We conduct experiments using the dataset described in Section IV-A to answer these RQs. Notably, this section builds on the definitions in Section V-B. In the following experiments, each of the 72,711 observations lists obtained after filtering in Section IV-B2 is treated as an original observations list (L_{ori}) without Byzantine behaviors, and l denotes the length of L_{ori}.

A. RQ1: Deviations in Chainlink's Real-World Price Data Induced by Byzantine Behaviors in the OCR Protocol

To answer RQ1, we evaluate the metric *Byzantine-induced price deviation* using real-world price observation values from Chainlink's DON. For a given L_{ori}, the positions occupied by the f values of Byzantine oracle nodes exhibit randomness, and the corresponding price deviation may therefore vary. Recall that our theoretical analysis provides the probability distribution of the index deviation (Table II and Table III). We align the values in each historical list according to their indices with the specified probabilities to calculate the mathematical expectation $\mathbb{E}(\Delta P)$, which captures the average *Byzantine-induced price deviation*. Specifically, the calculation in our simulation is based on Eq. 9 and Eq. 10, where v_j denotes the observation value at index j in L_{ori}. Clearly, $I_{ori} = \lfloor l/2 \rfloor$, and $M_{ori} = v_{\lfloor l/2 \rfloor}$.

$$
\begin{aligned}
&\text{inflation case:} \\
&\mathbb{E}(\Delta P) = \mathbb{E}(M_{fin} - M_{ori}) = \mathbb{E}(M_{fin}) - M_{ori}, \\
&\mathbb{E}(M_{fin}) = \sum_{j=0}^{l-1} p_j \cdot v_j, \\
&\text{where} \quad p_j = \Pr[I_{fin} = j] = \Pr[\Delta I = j - I_{ori}].
\end{aligned}
\tag{9}
$$

$$
\begin{aligned}
&\text{deflation case:} \\
&\mathbb{E}(\Delta P) = \mathbb{E}(M_{ori} - M_{fin}) = M_{ori} - \mathbb{E}(M_{fin}), \\
&\mathbb{E}(M_{fin}) = \sum_{j=0}^{l-1} p_j \cdot v_j, \\
&\text{where} \quad p_j = \Pr[I_{fin} = j] = \Pr[\Delta I = I_{ori} - j].
\end{aligned}
\tag{10}
$$

We simulate the price deviation induced by Byzantine behaviors under Scenario I for the 72,711 historical price feed instances. The results are sorted in descending order, and $\mathbb{E}(\Delta P)$, as well as $\mathbb{E}(\Delta P)/M_{ori}$ (the normalized value), at specific top fractions are shown in Table IV (inflation case) and Table V (deflation case).

According to our experimental results under Scenario I's inflation case (resp. deflation case), $\mathbb{E}(\Delta P)$ and $\mathbb{E}(\Delta P)/M_{ori}$ reach up to 120.86 USD (resp. 105.77 USD) and 3.57% (resp. 3.07%). This suggests that each 1 ETH involved in transactions relying on the swayed price feed is, in expectation, overvalued by as much as 120.86 USD, or 3.57% above the original ETH/USD price. A similar interpretation applies to the deflation case. Although the occurrence of significant $\mathbb{E}(\Delta P)$ or $\mathbb{E}(\Delta P)/M_{ori}$ may not seem very frequent, we will demonstrate through case studies that such price deviations can have a non-negligible impact on downstream applications.

Similarly, we simulate the price deviation induced by Byzantine behaviors under Scenario II for the 72,711 historical price feed instances. The experimental results for the inflation and deflation cases are presented in Table VI and Table VII, respectively.

According to our experimental results under Scenario II's inflation case (resp. deflation case), $\mathbb{E}(\Delta P)$ and $\mathbb{E}(\Delta P)/M_{ori}$ reach up to 209.24 USD (resp. 230.04 USD) and 8.47% (resp. 6.79%). It can be observed that the price deviations induced by Byzantine behaviors under Scenario II are more significant than those under Scenario I, highlighting the critical influence of the Byzantine leader in the OCR protocol.

TABLE III: The probability distribution of ΔI in Scenario II when $l = 31$, where $\Pr[\Delta I = \Delta i]$ denotes the corresponding probability, which is identical in both the inflation and deflation cases.

ΔI	5	6	7	8	9	10	11	12	13	14	15
$\Pr[\Delta I = \Delta i]$	2.3E-8	4.7E-7	5.2E-6	0.00004	0.00024	0.00120	0.00519	0.02002	0.07008	0.22581	0.67742

TABLE IV: Values of *Byzantine-induced price deviation* corresponding to the top fractions of results under Scenario I's inflation case.

Top fraction	0.005%	0.01%	0.1%	1%	10%
$\mathbb{E}(\Delta P)$ (USD)	96.21	70.09	29.86	14.02	4.29
$\mathbb{E}(\Delta P)/M_{ori}$	2.99%	2.65%	1.23%	0.55%	0.17%

TABLE V: Values of *Byzantine-induced price deviation* corresponding to the top fractions of results under Scenario I's deflation case.

Top fraction	0.005%	0.01%	0.1%	1%	10%
$\mathbb{E}(\Delta P)$ (USD)	64.37	50.81	26.59	13.00	4.08
$\mathbb{E}(\Delta P)/M_{ori}$	2.60%	2.07%	1.18%	0.53%	0.17%

TABLE VI: Values of *Byzantine-induced price deviation* corresponding to the top fractions of results under Scenario II's inflation case.

Top fraction	0.005%	0.01%	0.1%	1%	10%
$\mathbb{E}(\Delta P)$ (USD)	201.12	199.14	126.29	37.60	11.83
$\mathbb{E}(\Delta P)/M_{ori}$	6.60%	6.38%	4.77%	1.58%	0.49%

TABLE VII: Values of *Byzantine-induced price deviation* corresponding to the top fractions of results under Scenario II's deflation case.

Top fraction	0.005%	0.01%	0.1%	1%	10%
$\mathbb{E}(\Delta P)$ (USD)	198.00	123.95	71.06	30.95	10.35
$\mathbb{E}(\Delta P)/M_{ori}$	6.12%	4.78%	2.85%	1.29%	0.43%

Chainlink triggers a new price feed when the ETH/USD price fluctuates by more than 0.5%. Overall, compared to this deviation threshold, our experimental results demonstrate that the *Byzantine-induced price deviation* is non-negligible. Notably, an adversary may exploit Byzantine oracle nodes to sway the price toward its own interests, threatening the fairness of the price feed mechanism and causing downstream financial impacts.

B. RQ2: Uncertainty in Chainlink's Real-World Price Data Induced by Byzantine Behaviors in the OCR Protocol

To answer RQ2, we evaluate the metric *maximum uncertain price deviation* using real-world price observation values from Chainlink's DON. From the perspective of a downstream user, the price data injected into its contract may be either inflated or deflated due to potential Byzantine behaviors in the OCR protocol. We leverage the metric *maximum uncertain price deviation* to evaluate the maximum uncertainty present in Chainlink's real-world price data when it is swayed by Byzantine behaviors. As analyzed in Section V-C, under Scenario I (resp. Scenario II), the uncertainty induced by Byzantine behaviors reaches its maximum when Byzantine oracle nodes' original observation values occupy indices within the interval [6, 24] (resp. [1, 29]) of L_{ori}. Thus, we calculate the corresponding *maximum uncertain price deviation* (i.e., U_{max}) for each of the 72,711 historical price feed instances. The results are sorted in descending order, and U_{max}, as well as U_{max}/M_{ori} (the normalized value) at specific top fractions are shown in Table VIII(Scenario I) and Table IX(Scenario II).

According to our experimental results, under Byzantine behaviors in Scenario I, 37.92% of Chainlink's historical price feed instances exhibit a *maximum uncertain price deviation* that exceeds half of the width of the corresponding honest range, and 15.52% exceed two-thirds. Under Byzantine behaviors in Scenario II, the *maximum uncertain price deviation* equals the full width of the corresponding honest range, and the theoretical analysis supporting this result is given in Section V-C2. **Overall, the OCR protocol allows the metric *maximum uncertain price deviation* to approach the full width of the honest range. As the honest range of price observation values may be considerably wide in real-world settings, we argue that the metric *maximum uncertain price deviation* should be effectively bounded below the width of the honest range, in order to limit the uncertainty induced by Byzantine behaviors.**

C. RQ3: Downstream Financial Impacts of the Price Deviations Induced by Byzantine Behaviors

To answer RQ3, we conduct case studies to demonstrate the downstream financial impacts of the price deviations induced by Byzantine behaviors. We conduct the case studies from two different perspectives. The first focuses on the potential impacts of a single price feed swayed by Byzantine behaviors, while the second focuses on the potential cumulative impacts of a large number of price feeds swayed by Byzantine behaviors.

1) Downstream Impacts of a Single Price Feed Swayed by Byzantine Behaviors: DeFi lending platform Aave [41] is one of the primary downstream applications that rely on Chainlink's price feeds. During liquidations, Aave relies on these price feeds to calculate the amount of collateral that the liquidator receives for repaying a given debt, along with the liquidation bonus (a certain portion of the collateral). As a concrete instance, we consider a historical price feed

TABLE VIII: Values of *maximum uncertain price deviation* corresponding to the top fractions of results under Scenario I

Top fraction	0.005%	0.01%	0.1%	1%	10%
U_{max} (USD)	192.05	144.98	66.63	31.65	11.61
U_{max}/M_{ori}	5.34%	4.98%	2.77%	1.27%	0.47%

TABLE IX: Values of *maximum uncertain price deviation* corresponding to the top fractions of results under Scenario II

Top fraction	0.005%	0.01%	0.1%	1%	10%
U_{max} (USD)	337.66	303.56	209.55	69.91	22.68
U_{max}/M_{ori}	11.10%	9.58%	7.45%	2.89%	0.94%

(corresponding to the event log in footnote[5]) and demonstrate the impacts that price deviations induced by Byzantine behaviors can have on the liquidations executed on Aave. During the period when this price feed was valid, i.e., from block 13179376 to block 13179389, we identified 13 liquidations whose collateral asset was WETH, and each of them is listed in Table XVII (Appendix E). At that time, Aave relied on Chainlink's ETH/USD price data to value WETH, and the liquidation bonus rate for WETH is 5%. If Byzantine oracle nodes inflate the USD price of ETH, the amount of WETH received by the liquidator for repaying a given debt (including the liquidation bonus) will decrease, causing a loss to the liquidator. Conversely, if Byzantine oracle nodes deflate the USD price of ETH, this will cause a gain to the liquidator. In this OCR round, we simulate, under Scenario I and Scenario II, the average price that Byzantine behaviors can inflate (resp. deflate), i.e., $\mathbb{E}(M_{fin})$. We then compare these results with those in the absence of Byzantine behaviors to calculate the total decrease (resp. increase) in the amount of WETH received by all the liquidators, which is referred to as liquidators' average loss (resp. gain). When f Byzantine oracle nodes' original observation values occupy particular indices in a list (as detailed in Section VI-B), the potential inflation (resp. deflation) of the price reaches its maximum. Accordingly, we also calculate the total decrease (resp. increase) in the amount of WETH received by all the liquidators under this condition, which is referred to as liquidators' maximum loss (resp. gain). The results are presented in Table X.

TABLE X: Liquidators' loss or gain caused by Byzantine behaviors, where Byzantine behaviors in the inflation case can cause liquidators' loss, whereas Byzantine behaviors in the deflation case can cause liquidators' gain.

Liquidators' loss or gain (WETH)	Average loss	Maximum loss	Average gain	Maximum gain
Scenario I	-69.41	-101.22	+46.52	+77.10
Scenario II	-136.51	-141.49	+152.29	+175.02

We estimate the price of one WETH using 3230.32 USD, which is the ETH/USD price of the next OCR round in Chainlink's historical record, and the liquidators' loss or gain is on the order of 10^5 USD. These results demonstrate that **even a single price feed swayed by Byzantine behaviors can cause non-negligible financial impacts for the downstream application**.

2) Cumulative Downstream Impacts of Price Feeds Swayed by Byzantine Behaviors: Using Dune's database [35], we observe that the two contracts with the highest number of calls to the ETH/USD proxy contract both belong to the Ethereum Name Service (ENS). The two contracts obtain the latest ETH/USD price via Chainlink's proxy contract and convert the USD-denominated rental fee into ETH, which is then charged on the Ethereum mainnet.

To extract the actual ETH amounts in each ENS charge, we filtered and parsed 4,465,424 relevant function calls from block 14678295 (April 2022) to block 22648260 (June 2025). We simulate the sway of Byzantine behaviors on the OCR rounds underlying these 4,465,424 function calls. For each of these OCR rounds, we focus solely on the average price deviation that Byzantine behaviors can induce (i.e., $\mathbb{E}(\Delta P)$), because the likelihood of always inducing the maximum deviation is negligible. In an OCR round, if Byzantine oracle nodes inflate (resp. deflate) the USD price of ETH, a USD-denominated rental fee converts to less (resp. more) ETH than it would without Byzantine behaviors, resulting in a loss (resp. gain) for ENS's revenue. In all these OCR rounds, if the Byzantine behaviors are always under Scenario II, the total loss (or gain) for ENS's revenue is maximized, since $\mathbb{E}(\Delta P)$ under Scenario II exceeds that under Scenario I. Conversely, if Byzantine behaviors are always under Scenario I, the total loss (or gain) for ENS's revenue is minimized. We simulate all these OCR rounds under Scenario I and under Scenario II separately, and the results are presented in Table XI, with the loss or gain in USD calculated based on an ETH price of 2425.31 USD (June 25, 2025).

TABLE XI: ENS's loss or gain caused by Byzantine behaviors, where Byzantine behaviors in the inflation (resp. deflation) case can cause ENS's loss (resp. gain).

ENS's loss or gain		Total loss	Total gain
Scenario I	in ETH	-159.50	+179.66
	in USD	-386,837.63	+435,742.38
Scenario II	in ETH	-800.45	+1356.79
	in USD	-1,941,339.53	+3,290,630.66

It can be observed that the cumulative impacts of a large number of price feeds swayed by Byzantine behaviors on ENS's revenue are on the order of 10^5 to 10^6 USD. **Overall, our case studies demonstrate that the downstream financial impacts of the price deviations induced by Byzantine behaviors can reach at least on the order of 10^5 USD.**

[5] 0x80b9019cb90645cd4451dd2c19fc17d1cbe20955941b9e282d195e244c96dfd4/Logs

VII. Mitigation

In this section, we discuss possible mitigations and their trade-offs.

To counter the Byzantine leader's selecting behavior, a simple and direct approach is to eliminate the grace period, at the cost of reducing the reward opportunities for slower oracle nodes. Fundamentally avoiding the Byzantine leader's selecting requires integrating certain cryptographic techniques into the current OCR protocol, such that the leader only verifies signatures while the actual observation values remain hidden from it. However, this is a complex task, and we leave it for future research.

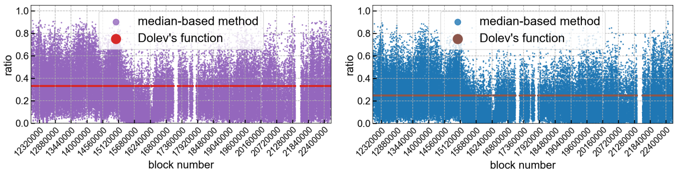
Regarding the falsifying behaviors, the goal should be to mitigate the impacts that falsifying imposes on the determined price, since it is difficult to identify which observation value is falsified. The metric *Byzantine-induced price deviation* will decrease as the DON's f/n ratio is reduced. Using the dataset filtered in Section IV-B2, we simulate two DON configurations, $n = 5f + 1$ and $n = 6f + 1$[6]. In each round, we assume f observation values are uniformly absent from the corresponding list because of delays, to simulate the elimination of the grace period. When Byzantine oracle nodes inflate (resp. deflate) the price, the maximum values of this metric under the two simulated configurations are 2.83% (resp. 2.36%) and 1.96% (resp. 2.34%), which are lower than those under the current configuration, 7.85% (resp. 5.87%). Therefore, for highly volatile cryptocurrencies such as ETH, it is necessary to further reduce the f/n ratio.

More importantly, we present an approach that bounds the metric *maximum uncertain price deviation* under the falsifying behaviors, i.e., using the approximation function introduced by Dolev [28] (defined in Eq. 11) to determine the unique price value.

$$d_{f,f}(L) = \mathrm{mean}(\mathrm{select}_f(\mathrm{reduce}^f(L))). \tag{11}$$

As given in Eq. 11, we set both parameters of the function to f, where f denotes the number of Byzantine oracle nodes in the DON. The observations list L is sorted in non-decreasing order, and the operation $\mathrm{reduce}^f(L)$ removes the f largest and f smallest values from L. By removing these values, we exclude potential values that fall outside the honest range, thereby ensuring the validity property. The output of $\mathrm{reduce}^f(L)$ is then sampled using the operation $\mathrm{select}_f(\cdot)$, which starts from the minimum value and selects every f-th value. The mean of the sampled values is taken as the determined price.

When the DON configuration $n = \alpha f + 1$ satisfies $\alpha \geq 4$, the metric *maximum uncertain price deviation* is bounded by a width of $1/(\alpha - 2)$ of the honest range. The proof is given in Appendix F. Notably, under falsifying behaviors, the current median-based method is unable to provide a clear bound on this metric, even if the f/n ratio decreases. We simulate the falsifying behaviors of f Byzantine oracle nodes and measure the metric *maximum uncertain price deviation* relative to the width of the honest range. Under the DON configurations $n = 5f + 1$ and $n = 6f + 1$, the comparisons between Dolev's approximation function and the median-based method are shown in Figure 8a and Figure 8b, respectively. Although the honest range is inherently determined by market volatility and thus beyond our control, the application of Dolev's function further bounds the uncertain price deviations induced by Byzantine behaviors, and the bounds it provides are explicit. In contrast, the median-based method allows Byzantine-induced uncertain price deviations to approach the full width of the honest range. The trade-off of adopting Dolev's function is an increase in the computation cost of the aggregator contract. We simulate the on-chain computation of both the median-based method and Dolev's function on the Sepolia testnet over a list containing 25 observation values[7]. The results show that the median-based method consumes 44,961 gas, while Dolev's function consumes 47,338 gas. Under our experimental setup, the additional computation cost introduced by Dolev's function is acceptable.



(a) $n = 31,\ f = 6$

(b) $n = 31,\ f = 5$

Fig. 8: Ratio of the *maximum uncertain price deviation* to the width of the honest range.

VIII. Related Work

Solutions to the distributed oracle problem. The distributed oracle problem is, to some extent, analogous to abstract sensor fusion [27]. The solutions [28], [40], [42]–[44] are designed to aggregate multiple values in a distributed setting to obtain a result that is correct in some sense. However, the focus of these solutions varies. Jin *et al.* [40] design and implement a sensor fusion system that provably preserves the privacy of sensor inputs and defends against pollution attacks. In the context of DONs, Chakka *et al.* [21] focus on the scalability issue and propose a distributed oracle protocol based on an honest simple majority, rather than the conventional requirement of an honest super majority. Bandarupalli *et al.* [22] focus on the complexity of the algorithm and propose a distributed oracle protocol with low communication complexity. The trade-off in both [21] and [22] is relaxing the condition of the validity property. Xiao *et al.* [17] focus on the trustworthiness of data sources and propose a decentralized oracle architecture that integrates truth discovery techniques with a Byzantine fault tolerance consensus protocol. Based on the prior work [17], Zeng *et al.* [18] further propose an approach that allows oracle nodes to dynamically join and leave the DON. However, neither [17] nor [18] has been deployed in real-world applications.

[6] These two configurations fully utilize the 31 values in each list.

[7] With the elimination of the grace period, f lagging observation values are absent from the list.

Blockchain oracles in real-world applications. Currently, mainstream oracle service providers [4]–[6] all adopt DONs as their underlying infrastructure. How to determine a unique price value from multiple price observation values is a critical problem for them. Excluding Chainlink, we briefly introduce how two other major blockchain oracles address this critical problem. In Chronicle's DON [5], a certain number of trusted nodes act as validators, whose multisignatures are used to ensure the validity of the observations set. Similar to Chainlink, Chronicle takes the median of the observations set as the representative value. Pyth [6] employs a standalone blockchain called Pythnet. Each price data point from multiple sources includes a corresponding confidence interval, and Pythnet's consensus mechanism aggregates and validates the price data based on the confidence intervals. However, these oracle services are not as widely used as Chainlink.

Security issues related to oracles. Given the crucial role of oracles in the DeFi ecosystem, attacks exploiting oracles cause monetary losses, thereby posing threats to the security of the DeFi ecosystem [24], [45]–[48]. Deng *et al.* [49] design a framework to automatically analyze DeFi contract behaviors under price deviations from oracles. Mo *et al.* [50] propose a framework to automatically detect unsafe transactions associated with deviated price feeds. To mitigate the uncertainty existing in price data from on-chain Decentralized Exchanges (DEXs), Park *et al.* [51] propose an adaptive conformal consensus algorithm to derive a consensus set of price data from multiple DEXs. Closely related to our study, Gansäuer *et al.* [11] reveal that Chainlink's price data on the Ethereum mainnet exhibit large deviations under market volatility and identify associated arbitrage implications. However, their study focuses solely on Chainlink's configuration of the time intervals and volatility thresholds of price updates, without conducting an in-depth analysis of the OCR protocol.

IX. Conclusion

In this paper, we present an in-depth study of the security of Chainlink's OCR protocol. We first reveal that the honest range of price observation values in real-world settings can reach 13.13% of the ETH price, which has been overlooked. We then formally model the impacts of Byzantine behaviors on price data, and further evaluate the price deviations induced by Byzantine behaviors and their downstream financial impacts using real-world data. Our experimental results show that Byzantine behaviors can induce price deviations of up to 8.47% of the ETH price, and the downstream financial impacts can be on the order of 10^5 to 10^6 USD. Finally, we discuss possible mitigation strategies. Our study offers new insights into distributed oracle protocols in real-world application contexts, demonstrating that the validity property alone is insufficient, and this limitation exposes a gap between the protocol's theoretical design and its practical deployment. Even if the validity property is satisfied, assessing and further bounding the impacts of Byzantine behaviors remain essential for distributed oracle protocols deployed in real-world applications.

Ethics Considerations

We have responsibly disclosed our findings to Chainlink before the publication of this paper. We aim for our findings to contribute constructively and to inform potential improvements in related protocols. Since our experiments use historical data for simulation, they have no effect on the actual functioning of Chainlink's price feeds. Additionally, all data involved in this study are publicly accessible and do not contain any private or sensitive information.

Acknowledgment

This work is supported by National Key R&D Program of China (2023YFB2703700), National Natural Science Foundation of China (62202011, 62402029), Beijing Municipal Natural Science Foundation(L251062, L251041), National Cryptologic Science Fund of China(2025NCSF02030), and Beijing Advanced Innovation Center for Future Blockchain and Privacy Computing. We also sincerely thank the NDSS reviewers for the helpful suggestions.

References

[1] V. Buterin, "A next-generation smart contract and decentralized application platform," Tech. Rep., 2014, [Online]. Available: https://ethereum.org/content/whitepaper/whitepaper-pdf/Ethereum_Whitepaper_-_Buterin_2014.pdf, Accessed Jul. 2025.

[2] F. Zhang, E. Cecchetti, K. Croman, A. Juels, and E. Shi, "Town crier: An authenticated data feed for smart contracts," in *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2016, pp. 270–282.

[3] ethereum.org, "Oracles," [Online]. Available: https://ethereum.org/en/developers/docs/oracles/, Accessed Jul. 2025.

[4] L. Breidenbach, C. Cachin, B. Chan, A. Coventry, S. Ellis, A. Juels, F. Koushanfar, A. Miller, B. Magauran, D. Moroz *et al.*, "Chainlink 2.0: Next steps in the evolution of decentralized oracle networks," Chainlink Labs, Tech. Rep., 2021, [Online]. Available: https://research.chain.link/whitepaper-v2.pdf, Accessed Jul. 2025.

[5] Chronicle, "Chronicle Docs," [Online]. Available: https://docs.chroniclelabs.org/, Accessed Jul. 2025.

[6] P. Network, "Blockchain oracle for market data," [Online]. Available: https://www.pyth.network/, Accessed Jul. 2025.

[7] WINkLink, "Introduction to WINkLink," [Online]. Available: https://doc.winklink.org/v1/doc/en/, Accessed Jul. 2025.

[8] A. S. De Pedro, D. Levi, and L. I. Cuende, "Witnet: A decentralized oracle network protocol," Witnet Project, Tech. Rep., 2017, [Online]. Available: https://arxiv.org/abs/1711.09756, Accessed Jul. 2025.

[9] J. Peterson, J. Krug, M. Zoltu, A. K. Williams, and S. Alexander, "Augur: a decentralized oracle and prediction market platform," 2015, [Online]. Available: https://arxiv.org/abs/1501.01042, Accessed Jul. 2025.

[10] P. Sztorc, "Truthcoin," [Online]. Available: https://www.truthcoin.info/papers/truthcoin-whitepaper.pdf, Accessed Jul. 2025.

[11] R. Gansäuer, H. B. Aoun, J. Droll, and H. Hartenstein, "Price oracle accuracy across blockchains: A measurement and analysis," in *International Workshop on Cryptoasset Analytics (CAAW)*, 2025.

[12] R. McLaughlin, C. Kruegel, and G. Vigna, "A large scale study of the ethereum arbitrage ecosystem," in *USENIX Security Symposium (USENIX Security)*. USENIX Association, 2023, pp. 3295–3312.

[13] Chainlink, "The year in Chainlink 2021: 7 pillars of momentum," [Online]. Available: https://blog.chain.link/the-year-in-chainlink-2021/, Accessed Jul. 2025.

[14] M. Kaleem and W. Shi, "Demystifying pythia: A survey of chainlink oracles usage on ethereum," in *International Conference on Financial Cryptography and Data Security (FC) Workshops*. Springer, 2021, pp. 115–123.

[15] Binance, "Chainlink dominance in DeFi oracles!" [Online]. Available: https://www.binance.com/en/square/post/24683741288217, Accessed Jul. 2025.

[16] L. Breidenbach, C. Cachin, A. Coventry, A. Juels, and A. Miller, "Chainlink off-chain reporting protocol," Chainlink Labs, Tech. Rep., 2021, [Online]. Available: https://research.chain.link/ocr.pdf, Accessed Jul. 2025.

[17] Y. Xiao, N. Zhang, W. Lou, and Y. T. Hou, "A decentralized truth discovery approach to the blockchain oracle problem," in *IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2023, pp. 1–10.

[18] H. Zeng, H. Cui, C. Wang, B. Zhang, Z. Yu, and B. Guo, "SenFEED: Dynamic decentralized oracle services for accurate and real-time sensor data," in *IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2025, pp. 1–10.

[19] L. Lys and M. Potop-Butucaru, "Distributed blockchain price oracle," in *International Conference on Networked Systems*. Springer, 2022, pp. 37–51.

[20] J. Dong, C. Song, Y. Sun, and T. Zhang, "DAON: A decentralized autonomous oracle network to provide secure data for smart contracts," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 18, pp. 5920–5935, 2023.

[21] P. Chakka, S. Joshi, A. Kate, J. Tobkin, and D. Yang, "Oracle agreement: From an honest super majority to simple majority," in *IEEE International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2023, pp. 714–725.

[22] A. Bandarupalli, A. Bhat, S. Bagchi, A. Kate, C.-D. Liu-Zhang, and M. K. Reiter, "Delphi: Efficient asynchronous approximate agreement for distributed oracles," in *International Conference on Dependable Systems and Networks (DSN)*. IEEE, 2024, pp. 456–469.

[23] A. Pasdar, Y. C. Lee, and Z. Dong, "Connect api with blockchain: A survey on blockchain oracle implementation," *ACM Computing Surveys (CSUR)*, vol. 55, no. 10, pp. 1–39, 2023.

[24] L. Zhou, X. Xiong, J. Ernstberger, S. Chaliasos, Z. Wang, Y. Wang, K. Qin, R. Wattenhofer, D. Song, and A. Gervais, "SoK: Decentralized finance (DeFi) attacks," in *IEEE Symposium on Security and Privacy (SP)*. IEEE, 2023, pp. 2444–2461.

[25] A. Beniiche, "A study of blockchain oracles," 2020, [Online]. Available: https://arxiv.org/abs/2004.07140, Accessed Jul. 2025.

[26] A. Constantinescu, D. Ghinea, R. Wattenhofer, and F. Westermann, "Meeting in a convex world: Convex consensus with asynchronous fallback," [Online]. Available: https://eprint.iacr.org/2023/1364, 2023, Accessed Jul. 2025.

[27] B. Ao, Y. Wang, L. Yu, R. R. Brooks, and S. Iyengar, "On precision bound of distributed fault-tolerant sensor fusion algorithms," *ACM Computing Surveys (CSUR)*, vol. 49, no. 1, pp. 1–23, 2016.

[28] D. Dolev, N. A. Lynch, S. S. Pinter, E. W. Stark, and W. E. Weihl, "Reaching approximate agreement in the presence of faults," *Journal of the ACM (JACM)*, vol. 33, no. 3, pp. 499–516, 1986.

[29] I. Abraham, Y. Amit, and D. Dolev, "Optimal resilience asynchronous approximate agreement," in *International Conference on Principles of Distributed Systems*. Springer, 2004, pp. 229–239.

[30] L. Lamport, R. Shostak, and M. Pease, "The byzantine generals problem," *ACM Transactions on Programming Languages and Systems*, vol. 4, no. 3, pp. 382–401, 1982.

[31] C. Dwork, N. Lynch, and L. Stockmeyer, "Consensus in the presence of partial synchrony," *Journal of the ACM (JACM)*, vol. 35, no. 2, pp. 288–323, 1988.

[32] Etherscan, "OCR2Aggregator," [Online]. Available: https://etherscan.io/address/0x7d4E742018fb52E48b08BE73d041C18B21de6Fb5#code, Accessed Jul. 2025.

[33] Infura, "JSON-RPC methods," [Online]. Available: https://docs.metamask.io/services/reference/ethereum/json-rpc-methods/, Accessed Jul. 2025.

[34] Chainlink, "Price feed contract addresses," [Online]. Available: https://docs.chain.link/data-feeds/price-feeds/addresses?network=ethereum&page=1, Accessed Jul. 2025.

[35] Dune, "Crypto's data platform," [Online]. Available: https://dune.com, Accessed Jul. 2025.

[36] Uniswap, "Uniswap Docs," [Online]. Available: https://docs.uniswap.org/, Accessed Nov. 2025.

[37] Sushiswap, "Sushi," [Online]. Available: https://docs.sushi.com/, Accessed Nov. 2025.

[38] C. F. Torres, R. Camino *et al.*, "Frontrunner jones and the raiders of the dark forest: An empirical study of frontrunning on the ethereum blockchain," in *USENIX Security Symposium (USENIX Security)*. USENIX Association, 2021, pp. 1343–1359.

[39] S. Wadhwa, L. Zanolini, A. Asgaonkar, F. D'Amato, C. Fang, F. Zhang, and K. Nayak, "Data independent order policy enforcement: Limitations and solutions," in *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2024, pp. 378–392.

[40] C. Jin, C. Yin, M. van Dijk, S. Duan, F. Massacci, M. K. Reiter, and H. Zhang, "PG: Byzantine fault-tolerant and privacy-preserving sensor fusion with guaranteed output delivery," in *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2024, pp. 3272–3286.

[41] Aave, "Aave Documentation," [Online]. Available: https://aave.com/docs, Accessed Jul. 2025.

[42] K. Marzullo, "Tolerating failures of continuous-valued sensors," *ACM Transactions on Computer Systems (TOCS)*, vol. 8, no. 4, pp. 284–304, 1990.

[43] T.-H. H. Chan, E. Shi, and D. Song, "Privacy-preserving stream aggregation with fault tolerance," in *International Conference on Financial Cryptography and Data Security (FC)*. Springer, 2012, pp. 200–214.

[44] H. Corrigan-Gibbs and D. Boneh, "Prio: Private, robust, and scalable computation of aggregate statistics," in *USENIX Symposium on Networked Systems Design and Implementation (NSDI)*. USENIX Association, 2017, pp. 259–282.

[45] S. Eskandari, M. Salehi, W. C. Gu, and J. Clark, "SoK: Oracles from the ground truth to market manipulation," in *ACM Conference on Advances in Financial Technologies (AFT)*, 2021, pp. 127–141.

[46] S. Werner, D. Perez, L. Gudgeon, A. Klages-Mundt, D. Harz, and W. Knottenbelt, "SoK: Decentralized finance (DeFi)," in *ACM Conference on Advances in Financial Technologies (AFT)*, 2022, pp. 30–46.

[47] I. Homoliak, S. Venugopalan, D. Reijsbergen, Q. Hum, R. Schumi, and P. Szalachowski, "The security reference architecture for blockchains: Toward a standardized model for studying vulnerabilities, threats, and defenses," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 341–390, 2020.

[48] T. Mackinga, T. Nadahalli, and R. Wattenhofer, "Twap oracle attacks: Easier done than said?" in *IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE, 2022, pp. 1–8.

[49] X. Deng, S. M. Beillahi, C. Minwalla, H. Du, A. Veneris, and F. Long, "Safeguarding defi smart contracts against oracle deviations," in *International Conference on Software Engineering (ICSE)*. IEEE/ACM, 2024, pp. 1–12.

[50] Y. Mo, J. Chen, Y. Wang, and Z. Zheng, "Toward automated detecting unanticipated price feed in smart contract," in *ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA)*, 2023, pp. 1257–1268.

[51] S. Park, O. Bastani, and T. Kim, "$Acon^2$: Adaptive conformal consensus for provable blockchain oracles," in *USENIX Security Symposium (USENIX Security)*. USENIX Association, 2023, pp. 3313–3330.

Appendix

A. Rationale for Estimating Maximum Natural Variation

Theorem 2. *Let ob_i denote the i-th value in an observations list L, the length of L is $3f+1$, such that $1 \leq i \leq 3f+1$. Let nv_{max} denotes the maximum natural variation between any two honest observation values in L. Then it follows that $nv_{max} \geq \min\{ob_{i+2f} - ob_i | 1 \leq i \leq f+1\}$.* ***In the main text, we use ξ to denote*** $\min_i\{ob_{i+2f} - ob_i | 1 \leq i \leq f+1\}$.

Proof: Let V_h denote the set of all honest observation values in L, ob_{h_min} denote the first observation value equal to $\min(V_h)$ in L, and ob_{h_max} denote the last observation value equal to $\max(V_h)$ in L. Then we have

$$nv_{max} = \max(V_h) - \min(V_h) = ob_{h\_max} - ob_{h\_min}. \quad (12)$$

Since L is sorted in non-decreasing order, it follows that $ob_i \geq ob_j$ for any $i > j$. Moreover, since there are at least $2f+1$ honest observation values in L, we have $h\_max \geq h\_min + 2f$. When $1 \leq i \leq f+1$, we have

$$nv_{max} \geq ob_{h\_min+2f} - ob_{h\_min} \geq \min_i\{ob_{i+2f} - ob_i\}. \quad (13)$$

B. Sensitivity Analysis of the Maximum Acceptable Natural Variation (Section IV-B2)

In Section IV-B2, we define the parameter ϵ in Theorem 1 to represent the maximum acceptable natural variation. We then apply the filtering conditions $ob_{3f+1} - ob_{2f+1} \leq \epsilon$ and $ob_{f+1} - ob_1 \leq \epsilon$ to extract honest observations lists. Here, we conduct a sensitivity analysis on the parameter ϵ.

Recall that we calculated ξ for each complete list. ξ denotes the minimum difference between the largest and smallest values in any $2f+1$-*subset* of a complete list. Since there are at least $2f+1$ honest observation values in a complete list, the parameter ϵ should satisfy $\epsilon \geq \xi$. **Let m denote the median of an observations list.** In historical ETH/USD price feed instances, the maximum value of ξ/m is 0.0851, and we set $\epsilon = 0.0851 \times m$ in Section IV-B2. For the sensitivity analysis of the parameter ϵ, we introduce an adjustable threshold t. The threshold t takes 10 uniformly spaced values in the interval [0, 0.0851], and we set $\epsilon = \max(\xi, t \times m)$ accordingly in ten runs of the experiment. **Note that the starting value of t, i.e., 0, represents that for every list the maximum acceptable natural variation takes its theoretical lower bound, which is an extreme condition that rarely occurs in practice.** When t takes the value 0.0851, we have $\max(\xi, 0.0851 \times m) = 0.0851 \times m$, which is equivalent to the setting $\epsilon = 0.0851 \times m$ used in the main text. We record, for each parameter setting, the maximum values reached by the width of the honest range, the metric *Byzantine-induced price deviation*, and the metric *maximum uncertain price deviation*. Considering ETH price volatility, **we normalize all corresponding experimental results relative to m.**

1) Width of the Honest Range: When t takes the values 0, 0.0095, 0.0189, 0.0284, 0.0378, 0.0473, 0.0567, 0.0662, 0.0756, and 0.0851 in order, the maximum width of the honest range is always 13.13%. This is because the lists with a wide honest range typically occur during periods of high price volatility, where the differences between observation values are large. Consequently, these lists inherently have sufficiently large ξ, and are preserved in our filtering regardless of the value of parameter t.

2) Byzantine-induced Price Deviation: As shown in Table XII, varying t does not change the maximum value of the metric *Byzantine-induced price deviation* in most cases. The difference caused by varying t from 0 to 0.0851 is at most 2.41%, which occurs only in the inflation case of Scenario II. Overall, the parameter setting used in the main text does not significantly inflate this metric.

3) Maximum Uncertain Price Deviation: Table XIII shows the maximum values of the metric *maximum uncertain price deviation* under each setting of t. Similar to the above, the parameter setting used in the main text does not significantly inflate this metric.

C. Supplementary Explanation of the Probabilistic Model

In Section V-C, our probabilistic analysis focuses on mapping the impacts of Byzantine behaviors to the underlying probabilistic rules governing index changes, thereby deriving a generic model. The only assumption of our model is that Byzantine oracle nodes' original observation values (i.e., before being falsified) occupy random positions in the ordering of a list. This assumption holds because each value's index in a sorted list is jointly determined by all values, and the f Byzantine oracle nodes cannot control the honest values. This generic model is index-oriented and independent of the concrete values contained in the list. Practical factors may affect the concrete values in an observations list. For instance, oracle nodes may query overlapping data sources, potentially making their observation values numerically close. However, the ordering of these values, which determines their indices, remains random. Thus, overlapping data sources do not affect our index-oriented probabilistic analysis, while their effects may be reflected in the experimental results in Section VI.

In our evaluation (Section VI), we align the values in each historical list according to their indices with our probabilistic model, thereby injecting concrete data into the generic model, and calculate the corresponding price deviations for each historical list. While the index changes caused by Byzantine behaviors are generic, the resulting price deviations are closely tied to the concrete values in each list, which may be affected by practical factors. For instance, overlapping data sources may produce minimal differences among observation values, resulting in smaller price deviations under the same index changes. Notably, we conduct our evaluation using real-world data, which accurately reflects all practical factors.

D. Empirical Study of Another Pair (BNB/USD)

To broaden our findings, we extend our empirical analysis to BNB/USD pair. Chainlink's DON for BNB/USD price feeds (on the Ethereum mainnet) consists of 16 oracle nodes (i.e., $n = 16$, $f = 5$), making its size roughly half that of the DON for ETH/USD pair. We collect the observations list of every BNB/USD price feed instance from block 11925663 (February 2021) to block 22647404 (June 2025), obtaining a total of 15,665 lists[8]. From these, we extract 14,530 complete lists (92.75%), each of which contains 16 observation values.

By Theorem 2, the maximum natural variation within an observations list is at least ξ. **Let m denote the median of an observations list.** Among the historical BNB/USD lists, the maximum of the ratio ξ/m is 0.0771. To determine the maximum acceptable natural variation for filtering honest lists, we adopt the same approach as in Appendix B and introduce an adjustable threshold t, then conduct a brief sensitivity analysis on t. Specifically, t takes 10 uniformly spaced values in the interval [0, 0.0771], and the maximum acceptable natural variation is $\max(\xi, t \times m)$. We apply the filtering step under different settings of t, and the resulting maximum width of the honest range is shown in Table XIV. **Considering BNB price volatility, all results below are normalized relative to m.** We find that when t is set to 0.06, decreasing it further does not change the maximum width of the honest range. To avoid

[8] The update interval for BNB/USD price feeds is 24 hours under normal circumstances, resulting in fewer observations lists than ETH/USD pair.

TABLE XII: The maximum of the metric *Byzantine-induced price deviation* under different t (ETH/USD)

	t	0	0.0095	0.0189	0.0284	0.0378	0.0473	0.0567	0.0662	0.0756	0.0851
Scenario I	inflation case	3.57%	3.57%	3.57%	3.57%	3.57%	3.57%	3.57%	3.57%	3.57%	3.57%
	deflation case	2.87%	2.87%	2.87%	2.87%	2.87%	2.87%	3.07%	3.07%	3.07%	3.07%
Scenario II	inflation case	6.06%	6.06%	6.06%	6.06%	6.06%	6.06%	6.48%	8.47%	8.47%	8.47%
	deflation case	6.00%	6.00%	6.00%	6.00%	6.00%	6.00%	6.12%	6.79%	6.79%	6.79%

TABLE XIII: The maximum of the metric *maximum uncertain price deviation* under different t (ETH/USD)

t	0	0.0095	0.0189	0.0284	0.0378	0.0473	0.0567	0.0662	0.0756	0.0851
Scenario I	7.56%	7.56%	7.56%	7.56%	7.56%	7.56%	7.56%	8.92%	8.92%	8.92%
Scenario II	13.13%	13.13%	13.13%	13.13%	13.13%	13.13%	13.13%	13.13%	13.13%	13.13%

inflating the results, we therefore set the maximum acceptable natural variation to $\max(\xi, 0.06 \times m)$. After filtering, 14,514 honest lists remain, and their widths of the honest range are presented in Figure 9.

TABLE XIV: Maximum width of the honest range under different t (BNB/USD)

t	0	0.0086	0.0171	0.0257	0.0343
max width	11.21%	11.21%	11.21%	11.21%	11.21%
t	0.0428	0.0514	0.0600	0.0685	0.0771
max width	11.21%	11.21%	11.21%	12.56%	14.20%

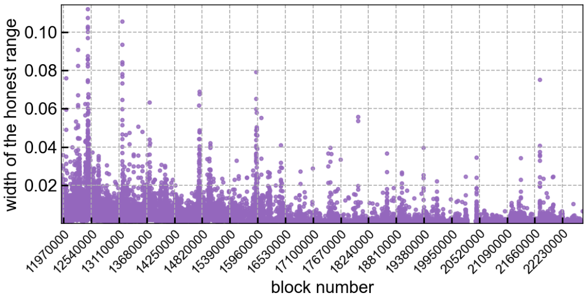


Fig. 9: Width of the honest range in historical lists (BNB/USD)

In terms of the overall distribution, the widths of the honest range of BNB/USD price feeds are comparable to those of ETH/USD price feeds. Using these honest lists, we evaluate the metric *Byzantine-induced price deviation* and the metric *maximum uncertain price deviation* for BNB/USD price feeds following the same approach detailed in Section VI. The experimental results are presented in Table XV and Table XVI, respectively. We note that the overall distributions of both metrics in BNB/USD price feeds do not differ significantly from those in ETH/USD price feeds. In general, wider honest ranges in price feeds lead to larger values of the two metrics. The honest range is closely associated with cryptocurrency price volatility, which is influenced by various factors. Therefore, assessing the impacts of Byzantine behaviors in the OCR protocol depends on the concrete price observation values.

TABLE XV: Values of *Byzantine-induced price deviation* (BNB/USD) corresponding to the top fractions of normalized results

Top fraction		0.007%	0.01%	0.1%	1%	10%
Scenario I	inflation	3.4%	3.26%	2.32%	1.02%	0.36%
	deflation	4.36%	3.97%	1.85%	0.79%	0.26%
Scenario II	inflation	6.17%	5.85%	4.37%	1.97%	0.68%
	deflation	8.36%	6.83%	4.24%	1.88%	0.70%

TABLE XVI: Values of *maximum uncertain price deviation* (BNB/USD) corresponding to the top fractions of normalized results

Top fraction	0.007%	0.01%	0.1%	1%	10%
Scenario I	8.04%	8.01%	4.54%	2.24%	0.82%
Scenario II	11.21%	10.75%	8.32%	3.67%	1.39%

E. Liquidations Analyzed in Section VI-C1

We analyzed 13 liquidations that executed on Aave within the block range [13179376, 13179389], the details of which are listed in Table XVII.

F. Application of Dolev's Approximation Function

In Section VII, we apply Dolev's approximation function to bound the metric *maximum uncertain price deviation*, and the theoretical basis is presented below. Notably, Lemma 1 was proven in [28], where it is referred to as Lemma 5.

Lemma 1. *Let V, W, and U be three multisets, where the multiset is a collection that may contain duplicate values*[9]*. The approximation function $d(\cdot)$ is defined as follows.*

$$d_{k,t}(V) = \text{mean}(\text{select}_k(\text{reduce}^t(V))). \tag{14}$$

Let $k > 0$, $t \geq 0$, and $m > 2t$ be integers. If the multisets V, W, and U satisfy the following conditions[10]*: $|V| = |W| = m$, $|V - U| \leq t$, $|W - U| \leq t$, and $|W - V| \leq k$, then we have*

$$|d_{k,t}(V) - d_{k,t}(W)| \leq \frac{\max(U) - \min(U)}{\lfloor \frac{m-2t-1}{k} \rfloor + 1}. \tag{15}$$

[9] The formal definition of the multiset is provided in [28].

[10] $V - U$ denotes the difference between two multisets, as defined in [28].

TABLE XVII: Details for each liquidation analyzed in Section VI-C1

Transaction Hash	Debt Asset	Debt Amount	Collateral Asset	Collateral Amount
0x3286......1c59	USDC	319561.55	WETH	102.50
0x202d......e093	DAI	152242.26	WETH	48.17
0x61fc......4b5f	USDC	469135.78	WETH	150.47
0xcdd2......2916	USDC	272464.56	WETH	87.39
0x788c......c0a4	DAI	907909.79	WETH	295.24
0x75b9......a326	DAI	188416.62	WETH	61.27
0x7260......7c09	USDC	342494.33	WETH	109.85
0xe1b3......0f25	USDT	34059.72	WETH	11.16
0xf326......bc3d	USDT	49149.50	WETH	16.10
0x5a37......3342	USDT	3567326.31	WETH	1168.90
0x0ebc......e97d	USDT	922844.10	WETH	302.39
0xbd99......f99e	USDT	81191.34	WETH	26.60
0xfd4f......ce85	DAI	23285.22	WETH	7.57

Theorem 3. *Let n denote the total number of oracle nodes in the DON, and let f denote the number of Byzantine oracle nodes among them, satisfying the condition $n \geq 4f + 1$. Assume that the f Byzantine oracle nodes exhibit falsifying behaviors, although the specific goals of their falsification remain unknown. Let L' and L'' denote the formed observations lists under any two possible cases caused by the falsifying behaviors. Let V_h denote the set consisting of observation values from all the honest oracle nodes, and let $d(\cdot)$ represent Dolev's approximation function. Thus, we have*

$$|d_{f,f}(L') - d_{f,f}(L'')| \leq \frac{\max(V_h) - \min(V_h)}{\lfloor \frac{n-3f-1}{f} \rfloor + 1}. \tag{16}$$

proof: Recall that the network is under the partial synchrony assumption [31]. With the grace period eliminated, the leader generates an observations list upon receiving observation values from $n - f$ distinct oracle nodes. Thus, $|L'| = |L''| = n - f$. Considering that some Byzantine oracle nodes may fall behind, up to f falsified values are included in the observations list. It follows that $|L' - V_h| \leq f$, $|L'' - V_h| \leq f$. Since the observation values from honest oracle nodes are identical in both L' and L'', it follows that $|L' - L''| \leq f$. L', L'' and V_h satisfy the hypotheses of Lemma 1. Thus, Eq. 16 holds, and we prove Theorem 3.

Artifact Appendix

A. Description & Requirements

1) How to access: The artifact is available on Zenodo with DOI: https://doi.org/10.5281/zenodo.17874648. The file AE_458.zip can be downloaded and extracted.

2) Hardware dependencies: A commodity desktop machine, e.g., an x86-64 CPU with 8 cores and 16 GB of RAM, running a recent Windows operating system.

3) Software dependencies: Python 3.13.9 is required. Below is a list of the required packages and their versions.

matplotlib == 3.9.2
numpy == 2.1.3
pandas == 2.2.3
web3 == 7.6.0
requests == 2.32.3
urllib3 == 2.2.3
Additionally, 7-Zip is required to extract a dataset used later.

4) Benchmarks: None.

B. Artifact Installation & Configuration

Our experiments were conducted on Windows 11 or Windows 10 system.

1) Enter https://www.python.org/downloads/windows/ in a browser, download and run the Windows installer (64-bit) for Python 3.13.9, and select the `Add Python to PATH` option during installation. In the following, path_to_python.exe denotes the path to the installed Python interpreter.
2) Enter https://www.7-zip.org/download.html in a browser, download and run the 64-bit Windows x64 installer.
3) After downloading, right-click the file (AE_458.zip) and select the option to extract it using 7-Zip. In the following, path_to_artifact denotes the path to the extracted folder.
4) Right-click on ens_case_data.7z.001(in the directory path_to_artifact/data_ae/), select `7-Zip` and `Extract Here`. The new file named ens_case_data_paper2.csv will be generated in path_to_artifact/data_ae/ directory.
5) Open Windows cmd, change directory to path_to_artifact, and enter the following commands.

```
path_to_python.exe -m venv env
env\Scripts\activate
pip install -r requirements.txt
```

C. Major Claims

- (C1): The ETH/USD price deviation range allowed by Chainlink's Off-Chain Reporting (OCR) protocol, i.e., the honest range of ETH/USD prices, is demonstrated by Experiment E1, whose results are reported in Figure 4 (Section IV-B).
- (C2): When f Byzantine oracle nodes (excluding the leader) in the OCR protocol attempt to inflate or deflate the ETH/USD price, the distributions of the resulting price deviation and the deviation ratio are demonstrated by Experiment E2. The results for the inflation and deflation cases are reported in Table IV and Table V, respectively (Section VI-A).
- (C3): When f Byzantine oracle nodes (including the leader) in the OCR protocol attempt to inflate or deflate the ETH/USD price, the distributions of the resulting price deviation and the deviation ratio are demonstrated by Experiment E3. The results for the inflation and deflation cases are reported in Table VI and Table VII, respectively (Section VI-A).
- (C4): When f oracle nodes (excluding the leader) in the OCR protocol exhibit Byzantine behaviors, the distributions of the maximum uncertain price deviation and the deviation ratio are demonstrated by Experiment E4, whose results are reported in Table VIII (Section VI-B).
- (C5): When f oracle nodes (including the leader) in the OCR protocol exhibit Byzantine behaviors, the distributions of the maximum uncertain price deviation and the deviation ratio are demonstrated by Experiment E5, whose results are reported in Table IX (Section VI-B).
- (C6): The cumulative impacts of Byzantine behaviors on ENS's revenue is demonstrated by Experiment E6, whose results are reported in Table XI (Section VI-C).

D. Evaluation

1) Experiment (E1): [1 minutes]: The file AE_458/data_ae/eth_usd_observations.csv records the historical ETH/USD price observation values in Chainlink's decentralized oracle network, i.e., the dataset described in Section IV-A. Experiment E1 uses this dataset to measure the honest range of ETH/USD prices and to reproduce the results shown in Figure 4 (Section IV-B).

[Execution] Execute the following command.

```
1) cd analysis_scripts_ae
2) python sec4_measure_eth_usd_honest_range.py
```

[Results] The experiment produces a figure identical to Figure 4 (Section IV-B). The figure is expected to appear during execution. If it does not appear, it may have been saved in the current directory as sec4_fig2.png.

2) Experiment (E2): [2 minutes]: Experiment E2 uses the dataset filtered in Section IV-B2 (i.e., eth_usd_honest_lists.csv) to evaluate the ETH/USD price deviation that f Byzantine oracle nodes (excluding the leader) can inflate or deflate.

[Execution] Execute the following command.

```
1) cd ../sec6_a_ae
2) python scenario1_inflation.py
3) python scenario1_deflation.py
```

[Results] The execution results of the second and third commands above are consistent with Table IV and Table V (Section VI-A), respectively. In the output, the data under *distribution of price deviations (USD)* correspond to the second row of the table, and the data under *distribution of price deviation ratios* correspond to the third row of the table.

3) Experiment (E3): [2 minutes]: Experiment E3 uses the dataset filtered in Section IV-B2 (i.e., eth_usd_honest_lists.csv) to evaluate the ETH/USD price deviation that f Byzantine oracle nodes (including the leader) can inflate or deflate.

[Execution] Execute the following command.

1) python scenario2_inflation.py
2) python scenario2_deflation.py

[Results] The execution results of the first and second commands above are consistent with Table VI and Table VII (Section VI-A), respectively. In the output, the data under *distribution of price deviations (USD)* correspond to the second row of the table, and the data under *distribution of price deviation ratios* correspond to the third row of the table.

4) Experiment (E4): [1 minutes]: Experiment E4 uses the dataset (i.e., eth_usd_honest_lists.csv) to evaluate the maximum uncertain price deviation that can be induced by f Byzantine oracle nodes, excluding the leader.

[Execution] Execute the following command.

1) cd ../sec6_b_ae
2) python scenario1_uncertainty.py

[Results] The execution results of the second command above are consistent with Table VIII (Section VI-B). In the output, the data under *distribution of price deviations (USD)* correspond to the second row of the table, and the data under *distribution of price deviation ratios* correspond to the third row of the table.

5) Experiment (E5): [1 minutes]: Experiment E5 evaluates the maximum uncertain price deviation that can be induced by f Byzantine oracle nodes, including the leader.

[Execution] Execute the following command.

1) python scenario2_uncertainty.py

[Results] Similar to E4, the execution results of the command above are consistent with Table IX (Section VI-B).

6) Experiment (E6): [80 minutes]: Using the data of 4,465,424 historical associated transactions (ens_case_data.7z.001), Experiment E6 evaluates the potential impacts of Byzantine behaviors in the OCR protocol on ENS's revenue.

[Preparation]

[Execution] Execute the following command.

1) cd ../sec6_c_ae
2) python ens_revenue_inflation.py
3) python ens_revenue_deflation.py

[Results] The execution results of the second command are consistent with the data in the *inflation case* column[11] of Table XI (Section VI-C), and the execution results of the third command are consistent with the data in the *deflation case* column[12] of Table XI.

E. Notes

The submitted artifact supports the main findings of the paper. The changes that we intend to carry out on the initially submitted paper are supplementary and will not affect the submitted artifact.

[11] In the camera-ready version, the column name *inflation case* is replaced with *Total loss*.

[12] In the camera-ready version, the column name *deflation case* is replaced with *Total gain*.