A First Look into DeFi Oracles

Bowen Liu
Singapore University of Technology and Design
Singapore
bowen_liu@mymail.sutd.edu.sg

Pawel Szalachowski
Singapore University of Technology and Design
Singapore
pawel@sutd.edu.sg

ABSTRACT

Recently emerging Decentralized Finance (DeFi) takes the promise of cryptocurrencies a step further, leveraging their decentralized networks to transform traditional financial products into trustless and transparent protocols that run without intermediaries. However, these protocols often require critical external information, like currency or commodity exchange rates, and in this respect they rely on special *oracle* nodes. In this paper, we present the first study of DeFi oracles deployed in practice. First, we investigate designs of mainstream DeFi platforms that rely on data from oracles. We find that these designs, surprisingly, position oracles as trusted parties with no or low accountability. Then, we present results of large-scale measurements of deployed oracles. We find and report that prices reported by oracles regularly deviate from current exchange rates, oracles are not free from operational issues, and their reports include anomalies. Finally, we compare the oracle designs and propose potential improvements.

KEYWORDS

Blockchain; DeFi Platforms; Price Oracles; Decentralization

1 INTRODUCTION

One promise of open cryptocurrencies is to make money and payments universally accessible without needing trusted parties. Decentralized Finance (DeFi) aims at extending this promise, proposing novel and traditional financial tools built on the top of a blockchain-based smart contract platform. DeFi offers multiple advantages over traditional finance. First, it inherits the blockchain properties, like decentralization, openness, accessibility, and censorship-resistance. Second, DeFi is highly flexible, allowing for rapid innovation and experiments by combining, stacking, or connecting different financial instruments. Finally, DeFi provides interoperable services. Generally, new DeFi projects can be built or composed by combining other DeFi platforms.

An increasing trend within the DeFi ecosystem is hybrid protocols which try to offer all advantages of DeFi, but eliminating the high volatility of cryptoassets [24], which hinders broader DeFi adoption. They do so by connecting their cryptoassets to conventional financial instruments. A prominent example is decentralized lending protocols that have achieved more recent attention than any other categories of DeFi. MakerDAO [11], a collateral-backed *stablecoin* whose value is stable relative to USD, enables anyone to leverage their collateral assets to generate new tokens [12] through a dynamic system of collateralized debts. Once new assets are generated, they can be used in the same manner as any other cryptocurrency. After paying down the debt and stability fee, the users can withdraw collateral and close their loans. Following the MakerDAO success, other DeFi lending platforms, like Compound [7], were launched. By offloading the traditional credit checking and reducing the costs with automation, Compound markets are actually a pool of assets that algorithmically derive interest-rates based on the supply and demand for the particular asset. Lenders and borrowers of these assets interact with the protocol directly in order to earn and pay respectively a floating interest rate, without having to negotiate any kind of terms such as maturity or interest rate. As of March 2020, DeFi Pulse reports that active outstanding loans from four open lending protocols – Fulcrum [23], dYdX [8], MakerDAO and Compound are above $200 million [16].

Another example of a project that aims at value stability is AmpleForth [1], the first DeFi protocol with supply elasticity. In response to the changes in demand, the platform always seeks a price-supply equilibrium based on the states of market and CPI index by universally expanding to, or contracting from holders, aiming at making it robust to economic shocks and runaway deflations. Synthetix [14], another recent DeFi project, allows the creation of "synthetic assets" – *Synths* whose prices can track currencies, cryptocurrencies, and commodities. The holders firstly lock their Synthetix native SNX tokens as collateral to mint the Synths which are tokens intended to track the value of the target asset (e.g., USD or gold).

All these systems require real-time information about the market price of the assets used as collateral and redemption. It is necessary for their security as the value (expressed in fiat) of cryptoasset collaterals is volatile. To implement this functionality, DeFi protocols introduce oracles, third parties reporting the price of assets from real-world (off-chain) sources. An oracle acts as a source of data that is being fed

to a smart contract. Although oracles play a critical role in the DeFi ecosystem, the underlying mechanics of oracles are vague and unexplored. Firstly, their deployment practices, including how frequent the price updates, how to aggregate the price value from multiple nodes, etc., are not transparent nor accountable, leaving room for various misbehaving. Secondly, the level of trust placed in oracles is unclear and most likely unknown to many participants of the ecosystem. Finally, the impact of a potentially malicious oracle (or a group of oracles) on the DeFi ecosystem is not investigated.

In this work, we shed light on these issues, presenting the first (up to our best knowledge) study on DeFi oracles. First, we explain oracle designs deployed in practice. Second, we systemically investigate the deployment of oracles for four popular open DeFi platforms - MakerDAO, Compound, AmpleForth and Synthetix- which rely on external oracles for price feeds. We conduct detailed measurements on the price deviation that comes from the differences between the information provided by external oracles and real-world prices. Moreover, we measure the robustness and deployment practices of oracles by transaction graph analysis. Finally, we compare the deployed platforms, and we give insights on potential improvements.

2 BACKGROUND

Many DeFi protocols aim at providing low volatility of their cryptoassets by using crypto collateral with its price pegged to some real-world assets.[1] Unlike in the real world, in DeFi protocols communicating exchange prices is not trivial, since these protocols are implemented as smart contracts deployed on the blockchain, without having access to any external resources (like current asset prices). Therefore, in such a design, price oracle is a fundamental component bridging cryptoassets with the external information about their intended value. In this section, we give a background on prominent DeFi protocols and their oracle designs. All of these platforms, as well as the vast majority of all DeFi platforms, are built upon Ethereum [20].

2.1 MakerDAO

MakerDAO is the most popular decentralized lending protocol where each its native token SAI is pegged to USD and is backed by collateral in the form of cryptoassets. Since dealing with cryptocurrency volatility is a problem, MakerDAO offers the programmability of crypto without the downside of volatility that you see with traditional cryptocurrencies like Bitcoin or Ethereum. By leveraging their cryptoassets into a Collateralized Debt Positions (CDP) contract as collateral, users are able to generate multi-collateral SAI tokens that can be traded in the same manner as any other cryptocurrencies. In return, the CDP accrues the debt known as overcollateralized loans determined by collateralization ratio (i.e., C-Ratio), locking them out of access until the outstanding debt is paid. The C-Ratio currently is set to 150% that helps the platform manage the risk of the borrowers by overcollateralizing the underlying assets. When the users want to retrieve their collateral, they have to pay down the debt in the CDP, plus the stability fee that continuously accrues on the debt over time, which can only be paid in MakerDAO's native tokens (MKR). In addition to payment of the stability fee, the MKR token also makes it possible to vote on the evolution of the platform and plays an important role in the governance of MakerDAO, in proportion to the number of MKR each owner has. The combination of SAI as a stablecoin and MKR as a governance token ensures the equilibrium of the system. Holders of MKR benefit directly from the use of SAI and the usage of the SAI is managed by the holders who are able to protect the system.

MakerDAO introduces an oracle module to obtain the real-time price of assets. The accuracy of this information is critical as it determines whether a CDP has enough collateral assets locked up and when to trigger liquidations. The oracle module is composed of a number of whitelisted oracle addresses and an *aggregator* contract. Oracles send periodic price updates to the aggregator which aggregates them, computes the median price as the reference price, and updates the platform by this reference price. Each asset type has an independent aggregator contract to collect information from authorized oracles. We give a high-level overview of this architecture in Figure 1. The aggregator contract implements access control logic allowing addition and removal of price oracle addresses. This operation is determined by the governors – MKR holders, who vote and update the changes on oracle addresses. Moreover, the logic allows governors to set other parameters that control the aggregator's behaviors, e.g., the minimum number of oracles necessary to accept a

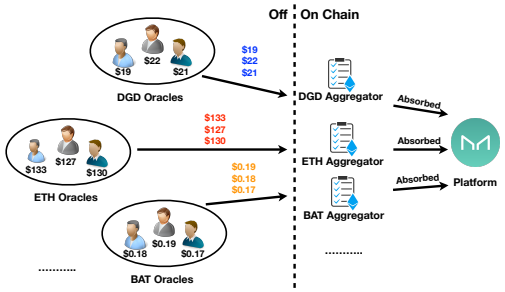


Figure 1: Oracle hierarchy of MakerDAO platform.

[1] We note, that there are other designs that do not require pegs or collaterals but these systems are out of scope this submission and we refer the reader to recent surveys [26, 32].

new median value. Consequently, in such a decentralized governance mechanism, the oracles could be manipulable by MKR holders. Similar in style to 51% attacks, a coalition can profitably manipulate governance to "steal" system collateral [9].

2.2 Compound

Compound is a blockchain-based borrowing and lending platform where participants can lend their cryptoassets out and earn interest on it. Participants deposit their cryptoassets to the Compound smart contract as collateral, and borrow against it. This contract automatically matches borrowers and lenders, and adjusts interest rates dynamically based on supply and demand. Similar to MakerDAO, Compound employs oracles for price feed which are managed and controlled by *administrators* – holders of Compound's native COMP token. Compound platform is governed and upgraded by COMP holders who propose, vote and implement any changes through the administrative functionalities. Proposals can include changes like adjusting an interest rate model or collateral factor, managing the aggregator contract, and choosing the source of the oracle.

The logic of price updates in Compound is depicted in Figure 2 and as shown, at the beginning, the administrator deploys an anchor contract and then creates an aggregator contract with *min*, *anchor* and *tolerance* set, where *min* is the minimum number of reports necessary to calculate a new median value which is set to one by default, *anchor* indicates the address of anchor contract and *tolerance* rate is set to 10%. The oracle system in Compound allows multiple authorized sources, known as reporters, to report price data to the aggregator contract. Reporters can be exchanges, other DeFi projects, applications, over-the-counter (OTC) trading desks, etc. The aggregator receives the reference prices from reporters, verifies them and calculates the median value from them, and then stores it in order to be accessed by the Compound market. The mechanism of updating the reference price of the asset is based on an anchor price (reported by an anchor address), together with upper and lower bounds that the median prices calculated by the aggregator are checked against. Should the ratio between a new median price and the anchor price be out-of-bounds, the official reference price of the asset would not be updated.

2.3 AmpleForth

Traditional commodities like gold or even cryptocurrencies like Bitcoin (produced based on a fixed supply schedule) cannot efficiently respond to changes in demand, making them vulnerable to destabilizing economic shocks and runaway deflation. To address this shortcoming, AmpleForth creates AMPL tokens to automatically propagate asset price-information into supply. By universally and proportionally expanding or contracting the quantity of tokens from each holder based on the price exchange rate between the AMPL and USD, this automatic price-supply equilibrium is countercyclical and non-dilutive. AMPL is initially pegged to USD, however not into perpetuity, since the platform takes the Consumer Price Index into account to balance future USD inflation. Thus, AmpleForth aims for purchasing power stability by altering the supply based on the demand for AMPL tokens. More specifically, whenever there is more demand than supply, the platform automatically increases the total

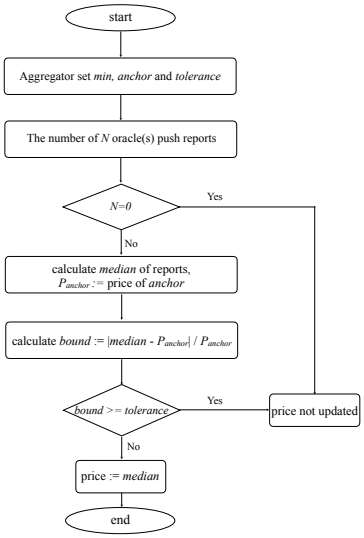


Figure 2: Oracle logics in Compound.

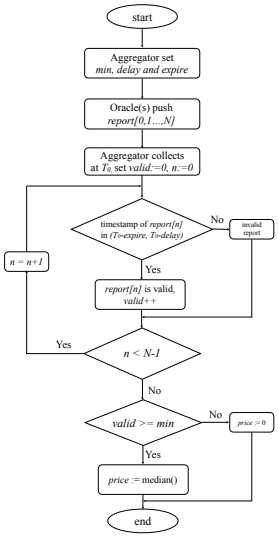


Figure 3: Oracle logics in AmpleForth.

quantity of AMPL to prevent the relative price of goods from rising. Similarly, when the total demand is less than supply, it decreases the quantity accordingly.

It is critical for such protocol aiming at a price-supply equilibrium to have a reliable and accurate source of market price information. This core functionality of AmpleForth is depicted in Figure 3. The platform administrator sets the *min*, *delay*, and *expire* parameters of the aggregator contract during its initialization, where *min* (one by default), indicating the minimum number of providers with valid reports to consider the aggregated report valid, *delay* is the number of seconds since reporting that has to pass before a report is usable (set to one hour, by default), and *expire* represents the number of seconds after which the report is deemed expired. In AmpleForth, this value is 12 hours by default. A valid report must exist on-chain publicly for at least 1 hour before it can be used by the supply policy and will expire on-chain if a new report is not provided before 12 hours elapses. That

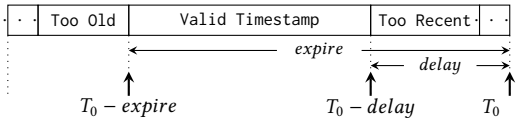


Figure 4: The valid range of a report in AmpleForth.

means that only the reports submitted within the *valid timestamp* are considered as valid reports. We depict this logic in Figure 4 and let us assume that the aggregator retrieves price information at T_0. The correct AMPL/USD price rate is median calculated by aggregator based on the reports that were submitted in

$$T_0 - expire \leq valid\ timestamp \leq T_0 - delay$$

from authorized oracles.

Chainlink. Smart contract platforms, like Ethereum, lack the ability of connecting smart contracts with off-chain resources (like web) natively. Chainlink [4] aims to resolve this issue by acting as a decentralized oracle network bridging on-chain smart contract with the off-chain environment. (In Section 5, we discuss competing designs to Chainlink.) It implements this by giving smart contracts APIs allowing to request off-chain resources such as market data, bank payments, retail payments, back-end systems, events data, or web content. Chainlink consists of a network of multiple decentralized, and independent oracles and aggregators that collect and process off-chain data and deliver it (processed) to smart contracts on request. AmpleForth is an example of a platform that is integrated with Chainlink [4].

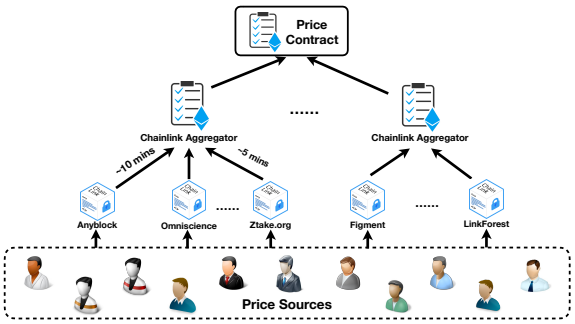


Figure 5: Oracle hierarchy of Synthetix platform.

2.4 Synthetix

Synthetix [15] is a platform that allows users creating and exchanging synthetic versions of assets such as gold, silver, cryptocurrencies, and traditional currencies. The purpose of Synthetix is to allow the creation of "synthetic assets" whose prices can track currencies, cryptocurrencies and commodities. Synthetix involves two distinct types of token. Users first purchase and lock up Synthetix's native token SNX into the Synthetix contract that gets staked as collateral to back other Synths tokens. [2] These Synths are synthetic assets created through Synthetix platform. Please note that Synthetix platforms always values sUSD, one of synthetic assets, at one USD. The price of Synths is determined through oracles that report external real-world price of asset to the aggregator contract which then proceeds the median calculation. As shown in Figure 5, the current oracles and aggregators are provided through Chainlink. Each asset type provides an independent Chainlink's aggregator which maintains a number of oracle sources. To ensure accurate data feeds, the oracles update the on-chain price in a short period, like 5 or 10 minutes.

3 MEASUREMENTS

In this section, we present details and results or our measurement studies. We focus on the AmpleForth, Synthetix, MakerDAO, and Compound platforms, and we measure and report on the following: a) market price volatility of the platforms' assets (Section 3.1), b) deviations between the market prices and prices reported by oracles (Section 3.2), c) anomalies that may indicate oracle malfunctions or misbehaving (Section 3.3), d) transaction graphs of oracles showing their interactions with the ecosystem (Section 3.4).

3.1 Price Volatility

In this section, we demonstrate the price volatility of DeFi assets that aim at removing volatility. We summarize daily

[2]The C-Ratio of Synthetix is 800% as for now.

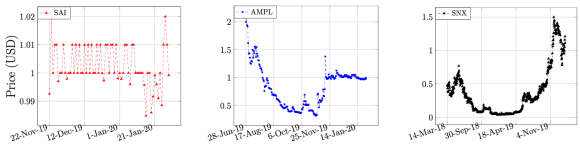
Table 1: Price volatility of DeFi assets.

Assets	Obs.	Daily change			
		±1%	±5%	±10%	±20%
DAI	69	47 (68%)	22 (32%)	0	0
SAI	764	502 (66%)	253 (33%)	8 (0.9%)	1 (0.1%)
AMPL	216	65 (30%)	92 (42%)	32 (15%)	27 (13%)
SNX	687	78 (11%)	272 (39%)	213 (31%)	124 (19%)

changes of the market prices (in USD, as reported by https://coinmarketcap.com) for each discussed platform in Table 1, where `Obs.` is the number of observations (i.e., days that a platform is in operation). It can be observed that despite of aiming a stability, all platforms experienced 1% or 5% price changes within one single day. Moreover, in around 30% of the investigated days, the market price of AMPL has more than 10% daily price change. The price volatility over time is depicted in Figure 6, where rapid prices changes are mainly caused by changing volume, external events (like banning cryptocurrencies by countries), or speculation. All the results presented indicate that these DeFi protocols and protocols relying on their assets require real-time accurate reference price data to hedge the risk of high volatility.

3.2 Price Deviation

In this section, we measure the deviation between real-time market prices and prices reported by the oracles of four major DeFi platforms. We also investigate the possible reasons for "outliers" – oracle reports with unusually higher deviation than other reports. To conduct the study, we select active oracles, with most frequent reporting, from MakerDAO, Compound, and Synthetix, reporting ETH/USD rates. For AmpleForth, we investigate its official market oracle which is supposed to report an AMPL/USD rate every 12 hours. We use Ethereum's BigQuery database [22] to get data about oracle interactions with their DeFi platforms. For each oracle, we analyze its all transactions by extracting their data and metadata, parsing the data to a readable price format, and comparing it with the real-world price sources that the oracle is supposed to be following (oracles may use different price sources). In our experiments, we consider an oracle's claimed source as its baseline for the given asset price. Moreover, as the methodologies of price reporting by oracle are not strictly specified, for each baseline source, we also show its real-time "raw" price and median values over 1, 5, 10, and 60 minutes.



(a) SAI **(b) AMPL** **(c) SNX**

Figure 6: Historical prices of selected DeFi assets.

Results. We first investigate the Synthetix oracle,[3] by analyzing its 3,308 price posting and comparing them with the data from different exchanges. As Synthetix is integrated with Chainlink, we find that the oracle's claimed ETH price sources [13] are Coinmarketcap, Bitfinex [2] and Bittrex [3]. Therefore, in Figure 7 we show the ETH/USD price deviations between the oracle reports and its price sources. As we can see, the number of deviations is substantial with most deviations standing within ±2% range.

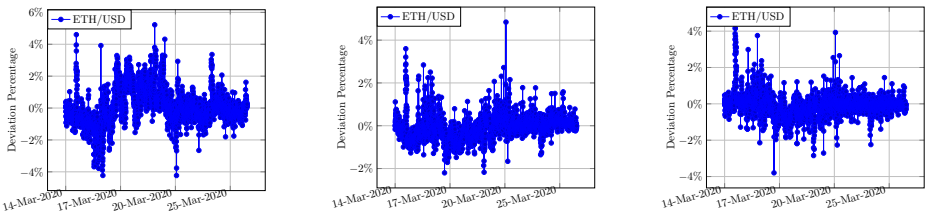
Similarly, in Figure 8, we illustrate deviations of the MakerDAO ETH/USD oracle[4]. As the oracle does not specify its sources, we use the same baselines as in Synthetix for evaluation, except for Bittrex which provides the information of ETH/USD rate only since June 2018 [18] (we measure the oracle interactions starting from Jan 2018). As we can see, there is a substantial number of deviations, with most of them being within the 5% range, indicating that the MakerDAO oracle is less effective than the previous Synthetix oracle. Moreover, there are a few outlier reports, deviating more than 10% (we investigate them further in Table 3).

In AmpleForth, the declared sources of oracle is Anylockanalytics [17, 19], however, it does not provide a public API to retrieve real-time prices for individuals. Therefore, we consider the same baselines as in Synthetix, except for Bittrex which does not track AMPL/USD rates. We analyze 980 transactions for the AmpleForth oracle[5] and check the results against Coinmarketcap and Bitfinex. As shown in Figure 9, the majority of deviations are within the 5% range, similarly as for the Synthetix oracle. Interestingly, there is a single outlier report with an extremely large deviation value (i.e., 273.7%), which we discuss further in this section.

The oracle in Compound use Kraken [10] and CoinbasePro [6] as its ETH/USD sources [5]. We conduct the evaluation of the Compound oracle[6], analyzing its 2,144 transactions in total, and presenting the obtained results in Figure 10. As we can see, most of the deviations are below the range of 4%, and there are only a few deviations above 5%.

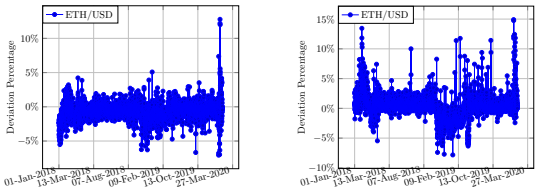
Comparison. In Figure 11, we show the average deviation of each source for all platforms. Please note, that the average deviation is calculated as $D = \sum_{n=1}^{N}(|D_n|)/N$, where D_n is the percentage of each point (i.e., transaction) and N is the sum of points (transactions). In most cases, the average deviation is below 2%, which given the volatility of cryptoassets, can

[3] Address: 0xac1ed4fabbd5204e02950d68b6fc8c446ac95362
[4] Address: 0xfbaf3a7eb4ec2962bd1847687e56aaee855f5d00
[5] Address: 0x8844dfd05ac492d1924ad27ddd5e690b8e72d694
[6] Address: 0x3c6809319201b978d821190ba03fa19a3523bd96



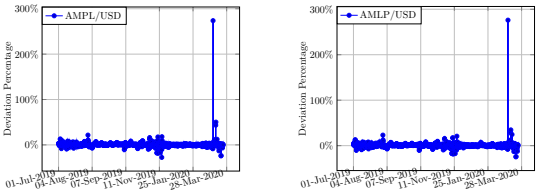
(a) Coinmarketcap (b) Bitfinex (c) Bittrex

Figure 7: Price deviations of the Synthetix oracle from its price sources.



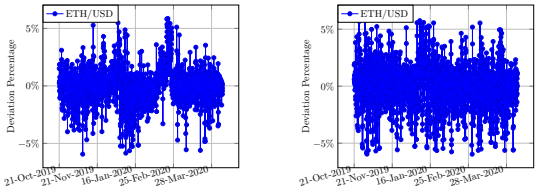
(a) Coinmarketcap (b) Bitfinex

Figure 8: Price deviations of the MakerDAO oracle from its price sources.



(a) Coinmarketcap (b) Bitfinex

Figure 9: Price deviations of the AmpleForth oracle from its price sources.



(a) CoinbasePro (b) Kraken

Figure 10: Price deviations of the Compound oracle from its price sources.

be seen as relatively precise information. From Figure 11a to Figure 11c, we can see the average deviation increasing from raw data to 60-minute median. In Synthetix, Bittrex is more precise source compared to Coinmarketcap and Bitfinex. In MakerDAO, Bitfinex is more accurate for the real-time price, 1-minute, and 5-minute median, while Coinmarketcap is

Table 2: Deviation comparison between different platforms.

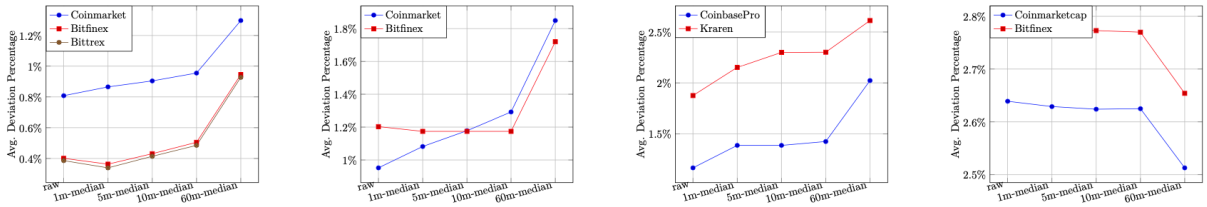
Oracle	Obs.	Source	Rule	# of dev. in range (%)				Avg dev (%)
				≤1	≤5	≤10	>10	
Synthextic (ETH/USD)	3308	CMC[1]	raw	2388	969	1	0	0.808
			1m	2245	1061	2	0	0.865
			10m	2115	1183	10	0	0.955
			1h	1733	1502	72	1	1.297
		BF[2]	raw	3088	220	0	0	0.400
			1m	3159	149	0	0	0.363
			10m	2914	393	1	0	0.505
			1h	2227	1046	35	0	0.945
		Bittrex	raw	3048	260	0	0	0.386
			1m	3121	186	1	0	0.339
			10m	2925	381	2	0	0.485
			1h	2276	997	35	0	0.927
MakerDAO (ETH/USD)	4707	CMC	raw	2974	1708	22	3	0.952
			1m	2733	1928	43	3	1.082
			10m	2404	2226	73	4	1.292
			1h	1799	2649	231	28	1.849
		BF	raw	2903	1679	108	17	1.203
			1m	2813	1791	93	10	1.174
			10m	2813	1791	93	10	1.174
			1h	1991	2472	224	20	1.720
Compound (ETH/USD)	2144	CBP[3]	raw	1309	816	19	0	1.168
			1m	1207	859	78	0	1.388
			10m	1121	933	89	1	1.426
			1h	871	1084	153	36	2.023
		Kraken	raw	1059	1030	55	0	1.876
			1m	1020	1055	69	0	2.152
			10m	1001	992	151	0	2.302
			1h	979	989	143	33	2.612
AmpleForth (AMPL/USD)	980	CMC	raw	418	449	86	27	2.639
			1m	411	458	84	27	2.629
			10m	413	451	83	28	2.625
			1h	448	423	83	26	2.513
		BF	raw	387	455	112	26	2.792
			1m	388	455	110	27	2.777
			10m	397	444	112	27	2.770
			1h	417	430	107	26	2.654

[1] Coinmarketcap. [2] Bitfinex. [3] CoinbasePro.

better for the 10- and 60-minute median. For Compound, it is observed that CoinbasePro is much more accurate than

Table 3: Outliers analysis.

Platform	Tx hash	Time	Oracle (/USD)	Baseline (/USD)	Dev. (%)	Possible explanations
AmpleForth	0x9c61..	2020-03-05 00:31	6.23	1.66	273.7	Oracle wrong input (most likely)
	0x67fb..	2020-03-13 00:39	0.855	0.569	50.2	Exchange price sudden drop
Compound	0x33ba..	2020-02-17 16:02	271	256	5.85	Exchange price sudden drop
	0xde2f..	2019-11-08 18:12	181	191	-5.9	Exchange price sudden increase
MakerDAO	0x698a..	2020-03-13 02:44	121.49	107.73	12.8	Exchange price sudden drop
	0x38b0..	2020-03-13 02:36	110.43	98.43	12.2	Exchange price sudden drop
Synthetix	0xd9f0..	2020-03-19 16:29	138.66	131.78	5.2	Exchange price sudden drop
	0x3860..	2020-03-15 00:30	127.82	122.19	4.6	US stock market volatility



(a) Synthetix **(b) MakerDAO** **(c) Compound** **(d) AmpleForth**

Figure 11: Deviation comparison (per platform) among different sources.

Kraken. Furthermore, the results of the AmpleForth oracle indicate an opposite trend to the other three platforms. That is most likely caused by this oracle processing the price averaged over a longer period of time before reporting to an aggregator.

To illustrate the differences between the oracles better, we also give the specific numbers in Table 2. We can observe that the deviations of most posting behaviors are $\leq$1% and $\leq$5% except for AmpleForth, where its oracle introduces relatively higher deviation. A possible reason may be that the baselines measured by us are different from used by them (as mentioned before, AmpleForth oracles do not reveal their price sources). However, the average deviation in our measurement is around 2.5%, which compared with other platforms seems to be tolerable.

Outliers. As every oracle may face some inevitable outliers due to sudden changes in real-time price or mistakes by oracles itself, in Table 3 we present selected reports with large deviation values observed in Section 3.1. In AmpleForth, the market oracle had one evident posting error on 5 March 2020[7], when the oracle submitted the price with the hex format of `0x5667f2bb31e073c7` which introduced 273.7% deviation from the current exchange price. We have not found any reason for this anomaly and we suspect an input error.[8] Another interesting anomaly report had 50.2% deviation.[9], This inconsistent input was most likely due to the sudden drop of the exchange rate, hitting the lowest price in the past four months. The similar situation happened to the oracle of MakerDAO submitting two reports, deviating 12.8% and 12.2% respectively, due to the sudden drop of cryptoasset exchange rates in the past three months[10]. In Synthetix and Compound, the percentage of top two outlier reports is much smaller than for the previous two platforms, with just around 5%. The largest outlier in Synthetix also comes from the sudden drops of real price[11] while the second largest is most likely due to the recent US stock market volatility.

3.3 Failures

In this section, we investigate oracle failures. For MakerDAO, Compound and AmpleForth, we check all transactions which were submitted by their oracles but which got processed by the Ethereum network unsuccessfully (either rejected by the network or reverted by the oracle itself). For Synthetix, due to the integration with Chainlink, we inspect the oracle nodes of all supported assets, find out the real sources they collect from, and then measure those oracles.

[7]Transaction info: https://bit.ly/2KHiTFE

[8]When changing the first digital number of the transaction payload, the deviation lowers to 2.9%, which is a standard range for this oracle.

[9]Transaction info: https://bit.ly/2K5kSDF

[10]Transaction info: https://bit.ly/3ep74BO, https://bit.ly/2K3NcGb

[11]Transaction info: https://bit.ly/34Av0xo

Table 4: Oracle failures.

Platform	Address	Type (/USD)	Total	Failures	Remark
MakerDAO	0x000d..	ETH	7042	54 (0.77%)	out of gas
	0x005b..	ETH	7545	164 (2.17%)	out of gas
	0x0032..	ETH	11032	154 (1.39%)	out of gas
	0xa8eb..	ETH	12633	67 (0.53%)	out of gas
AmpleForth	0xd035..	AMPL	953	30 (3.15%)	reverted
	0xcaef..	AMPL	295	34 (11.53%)	reverted

MakerDAO. The reference price ETH/USD in MakerDAO is updated by an aggregator, which collates price data from a number of external sources. As shown in Table 4, there are 54 failed transactions occurred in the total 7,042 due to *out-of-gas* error[12] in one ETH oracle[13], with 0.77% failure rate. The similar issues can be found for the another three oracles[14] with 2.17%, 1.39% and 0.53% failure rates, respectively.

AmpleForth. Next, we investigate all transactions originated from AmpleForth's oracles. In Table 4, we show our findings that 30[15] and 34[16] reverted transactions happened in two market oracles of AmpleForth by April 2020.

Synthetix. Synthetix is fully integrated with Chainlink for feed services. Each asset type supported in Synthetix has a collection of corresponding Chainlink's nodes to complete oracle-like jobs. We further review the nodes for all exchange pairs – ETH/USD, BTC/USD, AUD/USD, EUR/USD, CHF/USD, GBP/USD, JPY/USD, XAG/USD and XUG/USD, to find out potential failures.

In Table 5, we give the details on oracles, their respective sources, and the issues encountered. We found that Omniscience, Ztake.org, Anyblock, and Simply VC submitted transactions which subsequently have been reverted by the Ethereum network. By 14 February 2020, Alpha Vantage received 17 requests from Chainlink but ignored them, submitting no successful responses. LinkPool employs two external sources, i.e., CryptoCompare and Alpha Vantage, for ETH, BTC, AUD, and XAG rates. However, Alpha Vantage is unreliable which renders LinkPool to be unreliable too. Fiews, Cosmostation, Validation, etc,. are stable nodes with no anomalies found, while the sources of stake.fish and Chainlayer are unknown to the public, thus cannot be publicly audited.

[12]This error happens when a transaction to be completed requires more computing resources than provided by its sender.
[13]Address: 0x000df128eb60a88913f6135a2b83143c452c494e
[14]Addresses: 0x005b903dadfd96229cba5eb0e5aa75c578e8f968, 0x0032ad8fae086f87ff54699954650354bb51e050, 0xa8eb82456ed9bae55841529888cde9152468635a
[15]Address: 0xd0352aad6763f12d0a529d9590ea2f30421667a6
[16]Address: 0xcaefaf2130f0751520d5a6a62f3b9c9eaa4739f4

Table 5: The price oracles of Synthetix.

Oracle	Price (/USD)	Issues
Omniscience	ETH, GBP, XAU	Reverted transaction(s)
Anyblock	BTC	Reverted transaction(s)
Ztake.org	AUD, BTC, XAG	Reverted transaction(s)
Simply VC	BTC, EUR, CHF, GBP, XAG	Reverted transaction(s)
Alpha	AUD, EUR, CHF, GBP, JPY	17 requests from ChainLink but 0 successfull responses
LinkPool	ETH, BTC, AUD, XAG	Two sources: CryptoCompare and Alpha Vantage.
stake.fish	ETH, AUD, JPY	Oracles are unknown to the public
Chainlayer	ETH, AUD, JPY	Oracles are unknown to the public
Fiews	EUR, CHF, JPY, XAG	N.A.
Cosmostation	ETH	N.A.
Validation	ETH, CHF	N.A.
SDL	BTC, EUR	N.A.
LinkFprest	CHF	N.A.

3.4 Transaction Activity Analysis

An oracle address may interact with a large number of Ethereum addresses that could be an ERC-20 token contract, an on-chain service, an entity from other protocol or an external account address, etc. In this section, we focus on the transaction activities analysis for oracles of the DeFi platforms. We crawl the entire transaction history of oracles by using BigQuery, then we build transaction graphs, and find what are most common addresses oracles interact with, what entities or external accounts they communicate with, and what interesting activities they are involved in.

AmpleForth. We collect 132,119 transactions from the market oracle of AmpleForth, and find out that there are 47 different addresses interacting with the oracle. As depicted in Figure 12a, the large proportion of entire transactions are interactions with Chainlink's aggregators, colored in green. The red node indicates that the market oracle has 161 transactions with `UpgradeProxy` contract of AmpleForth to set or update certain parameters. The blue and orange ones are external account addresses with 48 and 1 transaction involved. The oracle sent 48 transaction without input data to the blue node[17] in succession on 13 March 2020 (most likely, these are testing activities).

MakerDAO. Similarly, we use the ETH/USD MakerDAO oracle as our measurement target, extract 4,914 transactions,

[17]Address: 0x43eb83a6b54a98b2d051c933b8e4a900d6bacbee

Table 6: Comparison of the oracle designs.

Platform	Freq. (/hour)	Design	Hierarchy
MakerDAO	1, 6	Centralized aggregator	Centralized
Compound	2, 5	Centralized aggregator	Centralized
AmpleForth	12	Aggregator + Chainlink	Semi-decentralized
Synthetix	1/12	Chainlink nodes	Decentralized

analyze them, and present results in Figure 12b. As depicted, the oracle interacts with four types of entities from seven different addresses. Most transactions are about price posting behaviors, however, there are two failed one due to out-of-gas error. It has four proxy activities (red nodes) and only three token transfers (blue nodes) of SAI and DAI. Moreover, there is one migration activity when the platform decided to make SAI and DAI conversion.

Compound. We select an ETH/USD oracle in Compound and analyze its all 11,458 transactions. All transactions are about the price reporting actions interacted with three on-chain aggregators in total. In contrast to the oracles of other platforms, the transaction history does not contain interactions with other actors or services.

Synthetix. The 142,422 transaction graph of the Synthetix oracle is depicted in Figure 12d. Similar to other platforms, this ETH/USD oracle mainly interacts with the active aggregator contracts marked in green color. An exception is one of the aggregators (orange node)[18] which is self-destructed without a clear reason. Besides that, the oracle is also involved in 2,056 transactions with the Synthetix network contract, most are about getting the value of parameters from the platform. One interesting activity (in red) is 667 transactions sent in total to itself without input data (most likely for testing purposes).

4 DISCUSSION

4.1 Decentralization

In the background section, we discuss different designs of DeFi oracles. Some of them rely on centralized aggregators to retrieve reference price while others develop the partnership with Chainlink's feed providers. In this section, we investigate how oracles systems are implemented in practice and how that can influence the aimed decentralization of the platforms.

Table 6 describes the selected properties of oracles that influence their decentralization. MakerDAO and Compound have similar architectures – they employ one single aggregator to periodically retrieve price information from external whitelisted oracle nodes. This design introduces inherent centralization drawbacks even though other components of these systems are deployed on decentralized smart contract platforms. AmpleForth employs Chainlink to provide oracle functionalities, which also mitigates (due to the design of Chainlink) the centralized risks of a single aggregator. However, it still relies on an aggregator contract to collect data from four oracles. In a near future, AmpleForth is planned to be fully integrating with Chainlink for data feeds, thus, we classify it as a semi-decentralized design as of now. Synthetix has announced that Synthetix-Chainlink integration is now operational on Ethereum [25], providing fully decentralized price feeds. The data feed will be offloaded to the decentralized oracle network of Chainlink and the reference price rates are transferred on-chain by a number of independent nodes backed by economic incentives rather than any central parties. Therefore, as for now, its design is the closest to being decentralized.

4.2 Recommendations

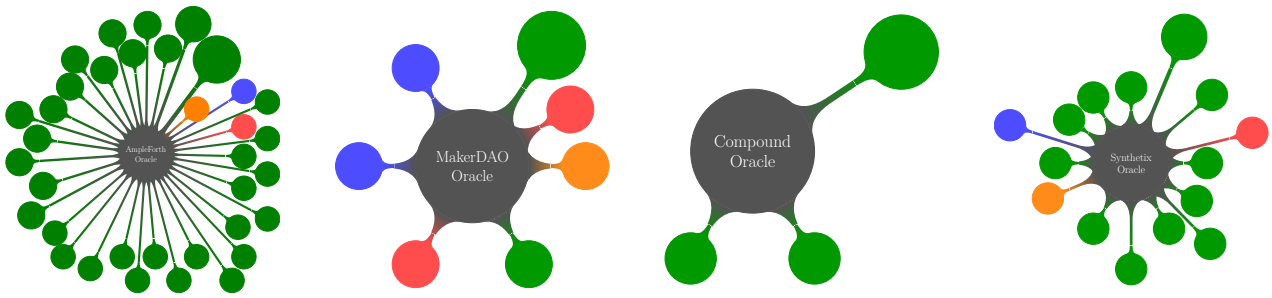
Our study, which can be seen as initial, indicates that the oracle ecosystem is immature. Therefore, in this section, we try to learn from our observations and give insights on the potential improvements of future oracle platforms.

Transparency. As discussed in Section 3.2, the methodologies of price processing by oracles are not clearly stated. Even the sources that the oracle retrieves from are ambiguous or unknown to the platform users. This results in a lack of transparency and potentially undetectable misbehavior of oracle platforms as currently no entity is able to provably examine the precision of price reported by oracles. Our first recommendation for future oracle designs is to require oracles to explicitly declare their manifest. Such a manifest would contain oracle metadata (like oracle contact information), deployed data sources, intended frequency of oracle updates, and precise description of the price derivation. Due to its properties, we see the underlying blockchain platform as a natural place of publishing such manifests.

Accountability. We believe that oracles, becoming critical trusted parties, should be held accountable for their actions. In the blockchain ecosystem, we can envision that a feasible way of implementing accountability is crypto-incentives. Therefore, to incentivize oracles to report accurate prices in promised frequency, we can imagine that the platforms implement mechanisms that would punish an oracle violating its manifest or platform policies, e.g., events like late or missing reports, or provable misbehavior like a high price deviation. Such a mechanism could be partially implemented by a smart contract, but it would require oracles to deposit some substantial amount of cryptoassets.

Operational Robustness. We found it surprising that despite relatively simple oracle interactions, they are not free

[18] Address: 0x5c545ca7f9d34857664fdce6adc22edcf1d5061f



(a) AmpleForth (b) MakerDAO (c) Compound (d) Synthetix

Figure 12: Transaction graphs of the DeFi oracles

from basic operational issues (e.g., causing out-of-gas errors). Since oracle reports play a crucial role in the DeFi ecosystem, we encourage operators to overprovision them by increasing their gas and gas price. The former has to guarantee that there is enough gas to be consumed by the entire execution of the transaction. The latter parameter is important for the latency of reports, which may be especially important when the Ethereum network is congested [21]. A high gas price would allow oracle reports to be appended to the blockchain much faster, since the blockchain network prioritizes more expensive transactions to be added first.

5 RELATED WORK

We are not aware of any work measuring or analyzing DeFi oracles, however, below we discuss work on the most related topics.

Oracle Designs. Town Crier (TC) [37] is an authenticated data feed system for smart contracts. The entity TC acts as a bridge between smart contracts and existing web sites, which are already commonly trusted for non-blockchain applications. It combines a blockchain front end with trusted hardware (i.e., the Intel SGX technology [27]) back end to scrape HTTPS-enabled websites and serve source-authenticated data to relying upon smart contracts. Due to the integration of the SGX enclave, it is possible in TC to conduct a remote attestation that the correct code was executed. TC establishes a secure TLS connection with a website and parses its content, which then can be used as an input to smart contracts. However, one potential limitation of TC is it positions Intel as a trusted party required to execute a remote attestation.

TLS-N [35] is a general TLS extension that provides secure non-repudiation to the TLS protocol. TLS-N modifies the TLS stack such that TLS records sent by a server are authenticated (in batches). Therefore, TLS-N clients can present received TLS-N records to the third parties which can verify it, just trusting the server (without any other third trusted parties). In general, TLS-N generates non-interactive proofs about the content of a TLS session that can be efficiently verified by third parties and blockchain-based smart contracts. As such, TLS-N increases the accountability for the content provided on the web and enables a practical and decentralized blockchain oracle for web content. However, the main drawback is its deployability. It requires significant changes to the TLS protocol and adoption processes are pretty slow.

Practical Data Feed Service (PDFS) [29] is a system that extends content providers by including new features for data transparency and consistency validations. It allows content providers to link their web entities with their blockchain entities. In PDFS, data is authenticated over blockchain but without breaking TLS trust chains or modifying TLS stacks. Moreover, content providers can specify data formats they would like to use freely, thus data can be easily parsed and customized for smart contracts. PDFS keeps content providers auditable and mitigates their malicious activities (e.g., data modification or censorship) and allows them to create a new business model. The shortcoming is the validation logic placed in smart contracts is not too lightweight or efficient and the design with shorter proofs could be a potential improvement for PDFS.

DeFi Surveys. [26] provides an understandable survey of on mainstream DeFi protocols, mainly focus on designs of stablecoins. [32] and [33] systematically discuss the general designs of existing DeFi platforms. By decomposing the designs into various components, i.e., peg assets, collateral amount, price information and governance mechanism, such surveys aim to explore the strengths and drawbacks of DeFi platforms to identify future directions.

Attacking DeFi. [34] presents a detailed exploration of the flash loan mechanism on the Ethereum network for the DeFi ecosystem. It analyzes two existing attacks vectors with ROIs beyond 500k%, then formulates finding flash loan-based attack parameters as an optimization problem. It also shows how two previously executed attacks can be "boosted" to result in a profit of 829.5k USD and 1.1M USD, respectively, which is a boost of 2.37× and 1.73×, respectively. Lewis et

al. [30] explore how the weaknesses of design could lead to DeFi crisis. In their paper, over-collateralization and governance attack on MakerDAO are discussed and a new financial contagion is proposed.

For the pump-and-dump scheme in cryptocurrency problem, [36] investigates 412 pump-and-dump activities and discovers patterns in crypto-markets associated with pump-and-dump schemes by building a model that predicts the pump likelihood of all assets listed in a crypto-exchange prior to a pump. Josh et al. [31] examines existing information on pump-and-dump schemes from classical economic literature, synthesizes this with cryptocurrencies, and proposes criteria that can be used to define a cryptocurrency pump-and-dump. The patterns can exhibit anomalous behavior, are utilized to locate points of anomalous trading activity in order to flag potential pump-and-dump activity.

Philip et al.[28] present the arbitrage strategies on decentralized exchanges that are executed by bots who pay high transaction fees and optimize their network latency to front-run ordinary users' trades. By empirically studying the bot's profit-making and blockchain-specific strategies, [28] formally models the behavior of the bots competing against each other for miner-supplied transaction priority in priority gas auction, demonstrating that in many cases, the revenue of bots from pure revenue arbitrage far exceeds the Ethereum block reward and transaction fees.

6 CONCLUSIONS

In this paper, we present the first study of DeFi oracles. We first dispel the mist of oracle designs of mainstream DeFi platforms. By conducting large-scale measurements of deployed oracles for four prominent open DeFi platforms – MakerDAO, Compound, AmpleForth and Synthetix, we investigate the details of price deviation that comes from the differences between the price information provided by real-time price and oracle nodes. We compare the price deviations of the deployed platforms, conduct the detailed measurement on the stability, accountability, and deployment patterns of oracles. We find that deviations from the claimed sources as well as operational failures are quite frequent. Finally, we give the discussion on the potential security vulnerabilities that such platforms may suffer from and give recommendations that could improve some of the found drawbacks.

REFERENCES

[1] 2019. Ampleforth. https://www.ampleforth.org.
[2] 2019. Bitfinex. https://www.bitfinex.com.
[3] 2019. Bittrex. https://global.bittrex.com.
[4] 2019. ChainLink. https://chain.link/.
[5] 2019. The claimed sources of Compound's oracles. https://bit.ly/3bNiuh5.
[6] 2019. Coinbasepro. https://pro.coinbase.com.
[7] 2019. Compound. https://compound.finance.
[8] 2019. dYdY. https://dydx.exchange.
[9] 2019. How to turn $20M into $340M in 15 seconds. https://bit.ly/2VNilEM.
[10] 2019. Kraken. https://www.kraken.com.
[11] 2019. MakerDAO. https://makerdao.com.
[12] 2019. SAI. https://sai.makerdao.com.
[13] 2019. The sources of Chainlink ETH/USD. https://feeds.chain.link/eth-usd.
[14] 2019. Synthetix. https://www.synthetix.io.
[15] 2019. Synthetix whitepaper. https://www.synthetix.io/uploads/synthetix_litepaper.pdf.
[16] 2020. The active outstanding loans from open lending protocols reported by DeFi Pulse. https://defipulse.com/defi-lending.
[17] 2020. Anylockanalytics. https://www.anyblockanalytics.com.
[18] 2020. Bittrex ETH/USD rate. https://bit.ly/2VQHTRv.
[19] 2020. The claimed sources of AmpleForth's oracles. https://bit.ly/3aMaqfc.
[20] 2020. Ethereum DeFi ecosystem. https://defiprime.com/ethereum.
[21] 2020. Ethereum gas price shot up. https://bit.ly/2y15AgR.
[22] 2020. Ethereum in BigQuery: a Public Dataset for smart contract analytics. https://bit.ly/3aNU8lU.
[23] 2020. Fulcrum. https://fulcrum.trade.
[24] 2020. Most volatile cryptos. https://yhoo.it/2YhINbg.
[25] 2020. Synthetix integrates with Chainlink. https://bit.ly/3bJfLFc.
[26] Jeremy Clark, Didem Demirag, and Seyedehmahsa Moosavi. 2019. SoK: Demystifying Stablecoins. *Available at SSRN 3466371* (2019).
[27] Victor Costan and Srinivas Devadas. 2016. Intel SGX Explained. *IACR Cryptology ePrint Archive* (2016).
[28] Philip Daian, Steven Goldfeder, Tyler Kell, Yunqi Li, Xueyuan Zhao, Iddo Bentov, Lorenz Breidenbach, and Ari Juels. 2020. Flash Boys 2.0: Frontrunning in Decentralized Exchanges, Miner Extractable Value, and Consensus Instability. In *41st IEEE Symposium on Security and Privacy*.
[29] Juan Guarnizo and Pawel Szalachowski. 2019. PDFS: practical data feed service for smart contracts. In *24th European Symposium on Research in Computer Security*.
[30] Lewis Gudgeon, Daniel Perez, Dominik Harz, Arthur Gervais, and Benjamin Livshits. 2020. The Decentralized Financial Crisis: Attacking DeFi. *arXiv preprint: 2002.08099*.
[31] Josh Kamps and Bennett Kleinberg. 2018. To the moon: defining and detecting cryptocurrency pump-and-dumps. *Crime Science* (2018).
[32] Amani Moin, Emin Gün Sirer, and Kevin Sekniqi. 2019. A Classification Framework for Stablecoin Designs. In *24th Financial Cryptography and Data Security*.
[33] Ingolf Gunnar Anton Pernice, Sebastian Henningsen, Roman Proskalovich, Martin Florian, Hermann Elendner, and Björn Scheuermann. 2019. Monetary Stabilization in Cryptocurrencies-Design Approaches and Open Questions. In *2nd IEEE Crypto Valley Conference on Blockchain Technology*.
[34] Kaihua Qin, Liyi Zhou, Benjamin Livshits, and Arthur Gervais. 2020. Attacking the DeFi Ecosystem with Flash Loans for Fun and Profit. In *arXiv preprint: 2003.03810*.
[35] Hubert Ritzdorf, Karl Wüst, Arthur Gervais, Guillaume Felley, and Srdjan Capkun. 2017. TLS-N: Non-repudiation over TLS Enabling-Ubiquitous Content Signing for Disintermediation. *IACR Cryptology ePrint Archive* (2017).
[36] Jiahua Xu and Benjamin Livshits. 2019. The anatomy of a cryptocurrency pump-and-dump scheme. In *28th USENIX Security Symposium*.
[37] Fan Zhang, Ethan Cecchetti, Kyle Croman, Ari Juels, and Elaine Shi. 2016. Town crier: An authenticated data feed for smart contracts. In *23rd ACM SIGSAC conference on computer and communications security*.