

# Demo: PhishChain: A Decentralized and Transparent System to Blacklist Phishing URLs

Shehan Edirimannage

University of Colombo School of Computing  
Sri Lanka  
shehan@scorelab.org

Charith Elvitigala

University of Colombo School of Computing - SCoRe Lab  
Sri Lanka  
charitha@scorelab.org

Mohamed Nabeel

Qatar Computing Research Institute  
Qatar  
mnabeel@hbku.edu.qa

Chamath Keppitiyagama

University of Colombo School of Computing  
Sri Lanka  
chamath@ucsc.cmb.ac.lk

## ABSTRACT

Blacklists are a widely-used Internet security mechanism to protect Internet users from financial scams, malicious web pages, and other cyber attacks based on blacklisted URLs. This demo introduces PhishChain, a transparent and decentralized system for blacklisting phishing URLs. At present, public/private domain blacklists, such as PhishTank, CryptoScamDB, and APWG, are maintained by a centralized authority, but operate in a crowd sourcing fashion to create a manually verified blacklist periodically. In addition to being a single point of failure, the blacklisting process utilized by such systems is not transparent. We utilize the blockchain technology to support transparency and decentralization, where no single authority is controlling the blacklist and all operations are recorded in an immutable distributed ledger. Further, we design a page rank based truth discovery algorithm to assign a phishing score to each URL based on crowd sourced assessment of URLs. As an incentive for voluntary participation, we assign skill points to each user based on their participation in URL verification.

## CCS CONCEPTS

• Information systems → Data exchange.

## KEYWORDS

Phishing URLs, Blacklisting, Blockchain, Smart Contracts, Crowd-sourcing, Truth Inference

### ACM Reference Format:

Shehan Edirimannage, Mohamed Nabeel, Charith Elvitigala, and Chamath Keppitiyagama. 2022. Demo: PhishChain: A Decentralized and Transparent System to Blacklist Phishing URLs. In *Companion Proceedings of the Web Conference 2022 (WWW '22 Companion)*, April 25–29, 2022, Virtual Event, Lyon, France. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3487553.3524235>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*WWW '22 Companion*, April 25–29, 2022, Virtual Event, Lyon, France

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9130-6/22/04...\$15.00

<https://doi.org/10.1145/3487553.3524235>

## 1 INTRODUCTION

Crowdsourced phishing blacklists such as PhishTank, CryptoScamDB, and APWG are quite useful in practice to defend against phishing attacks. Out of these services, PhishTank, has been the longest serving and most widely used crowd-sourced phishing blacklist. While it has been instrumental in maintaining a ledger of phishing URLs, it suffers from several limitations: (1) centralized architecture with a single point of failure, (2) lack of transparency, (3) lack of incentive to participate and (4) prone to manipulations. In fact, these are common issues for all existing centralized blacklisting systems. Focusing on PhishTank and for example, Figure 1 shows inconsistent assignment of final labels to URLs in PhishTank. The figure on the left shows a URL marked as phishing even though all 5 crowd sourced verifiers ascertain it is phishing. The figure on the right, on the other hand, shows that a URL is marked as not phishing even though no one has verified it. It is not clear how PhishTank arrived at the final decision and raises concerns on the transparency.

We propose PhishChain, a transparent and decentralized system to blacklist phishing URLs in a crowd-sourced manner while addressing the limitations mentioned above. We envision a blacklisting system where a consortium of mostly targeted organizations such as Paypal, Apple, Microsoft, and Facebook form an alliance to put our proposal to use. Our key design goals are as follows: (1) System should be decentralized where anyone can participate and no single authority should be controlling it, (2) All operations should be immutable and recorded, (3) URLs should be assigned a final phish score based on the participants' assessment, (4) Participants should be incentivized by assigning skill points instead of a cryptocurrency in order to encourage voluntary participation<sup>1</sup>, and (5) The system and data should be publicly available for anyone to consume.

There are a few notable efforts to build blockchain based decentralized solutions for phishing blacklisting or crowd sourcing in general. However, they suffer from one or two drawbacks. PhishLedger [6] is a private blockchain solution for blacklisting phishing URLs, and they do not support crowd sourcing. Further, the system is not available for public consumption and is managed in a closed environment. Hence, the verifications performed in this system are questionable. CrowdBC [4] and ZebraLancer [7] are blockchain

<sup>1</sup>PhishTank and APWG work through voluntary participation, but we observe that only a handful of participants contribute to the lists maintained.

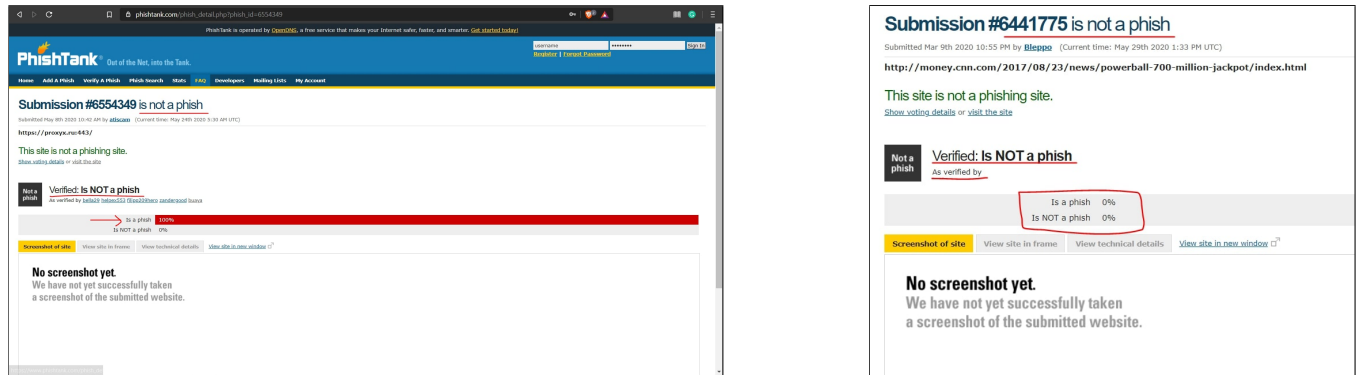


Figure 1: PhishTank inconsistent examples

based crowd sourcing systems, but they rely on a cryptocurrency to maintain the functionality.

**Demonstration Scenarios:** PhishChain is designed as the next generation decentralized phishing blacklisting system. We show the following three use cases. (1) URL verification, (2) URL detailed dashboard and transaction, and (3) URL verification graph and timeline along with the phish score and verifier skill points.

## 2 SYSTEM ARCHITECTURE

### 2.1 Overall System Design

Figure 2 shows the overall system architecture. The key components in the system are blockchain network, smart contracts, truth discovery algorithm, and API services modules. The system supports three main functions in addition to user and system management functionality: (1) URL submission, (2) URL verification and (3) URL lookup.

**Blockchain Network Module:** Blockchain network is the backbone of the decentralized infrastructure. Each node in our blockchain network is either a validator or a normal node. A validator node validates every transaction whereas a normal node is a read-only participant. Note that the validation node only participate in the blockchain block verification process similar to committing data to the blockchain ledger, whereas the verifier participates in the URL verification process itself. Istanbul Byzantine Fault Tolerance is utilized as the consensus algorithm since our threat model assumes that there can be malicious participants. In order to meet our system goals, a consortium type blockchain system called, Quorum [1], is employed as the blockchain network. Tessera [2] is used to manage transactions of the Quorum network.

**Smart Contract Module:** The crowd source functionality is implemented using smart contracts (Solidity version 0.5.0). We categorize the smart contracts into three main groups: (1) user management, (2) URL management and (3) URL verification.

**Truth Discovery Module:** In our crowd sourcing environment, URL verifiers are the workers and the task is to assess if a URL is phishing or non phishing. Due to the varying degree of expertise and experience of verifiers, one cannot simply take, for example, the majority answer as the final phish score of a URL. Instead, a truth discovery algorithm which takes potentially conflicting

status for a URL and verifier information, and arrives at the final phish score, should be employed. We observe that existing truth discovery algorithms developed by the database community [5] are performing poorly on our problem as these algorithms assume that the majority of the verifiers respond to each URL verification task, which is not true for verifying URLs. Our retrospective analysis of PhishTank URLs show that only a handful of verifiers verify a given URL even though there are thousands of them altogether. Hence, we design a new truth discovery algorithm by leveraging page ranking (PR) algorithm on *verifier-verifier graph*. We construct the verifier-verifier graph by using the “*follower*” relationship. We define verifier  $v_1$  follows  $v_2$  for a given URL if  $v_1$  verifies the URL before  $v_2$ . A directed edge from  $v_1$  to  $v_2$  is created in this case. The generated graph represent real world relationships among users which captures how users are following others and how they are being followed. Applying the page rank algorithm on top of this graph computes a rank, a real value number between 0 and 1, for each node (i.e. user), indicating how important it is with respect to the whole user population. It follows the intuition that, the more the number of followers are, the more recognition a user receive from its peers. We utilize the computed ranks to calculate the weighted voting score for each URL and assess if it is phishing or not.

Table 1 shows how our algorithm outperforms in comparison to popular truth discovery algorithms on a real-world dataset obtained from PhishTank. We collect URLs submitted and verified in PhishTank from January 1<sup>st</sup> 2020 to December 23<sup>rd</sup> 2020. The cleaned dataset contains 17,000 phishing URLs and 6,000 non phishing URLs<sup>2</sup>. We scrape the additional information of each of these URLs to identify the verifiers and their order of verification. We use a balanced dataset of 6,000 URLs from each class for the evaluation

We use our PR based algorithm to assign skill points to verifiers as well as final phish score to URLs. The skill points of verifiers are computed based on the verifier reputation score returned by the truth discovery algorithm and the proportion of correctly labeled URLs. The phish score of a URL is computed with the weighted score of phishing and non phishing verifiers.

Note that in PhishChain system, we enforce transparency and consistency on the data by following the consensus algorithm. In

<sup>2</sup>This dataset includes both consisted and non-consistent data

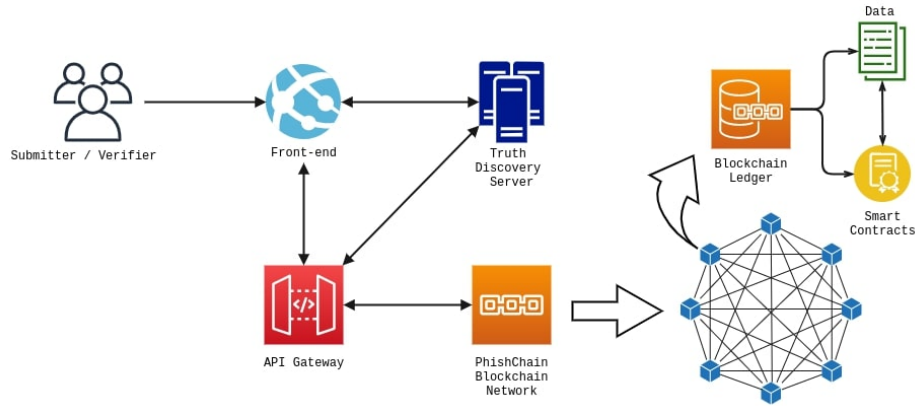


Figure 2: High Level System Design

other words, all the data related to URLs (URLs themselves, votes, status) on the blockchain is consistent and hardened from unauthorized modifications. While our system has a separate truth discovery server for URL verification due to the practical limitations in the Solidity framework utilized, as the algorithm and the data are publicly available and verifiable, any user can verify the results produced by the URL verification process is indeed correct.

Table 1: Truth Discovery Algorithm Comparison

Algorithm	Acc.	Prec.	Rec.
Our Approach	95.45%	96.74%	94.31%
EM	93.71%	91.01%	97.75%
GLAD	93.98%	91.72%	97.39%

**API Services Module:** This module facilitates the communication with the blockchain network. The demo web application we build is a consumer of this API.

### 3 DEMONSTRATION

In this section, we describe the main use cases of our system. We setup a consortium blockchain consisting of 7 validation nodes each possessing a copy of the blockchain ledger and participating in validating each transaction. In a real-world setup, these validation nodes are to be hosted by organizations mostly targeted by phishers (e.g. Paypal, Facebook, Microsoft) as such a system is in their best interest to defend against phishing attacks and a collective effort is likely to yield a significantly better blacklist compared to having individual systems.

#### 3.1 URL Verification

Any user can sign up and participate in the process of submission and verification of URLs. Figure 3 shows the UI for the verification of a given URL as phishing or not phishing. Once user submits their verdict, the validation nodes validate the transaction adhering to the consensus algorithm resulting the transaction shown in Figure 4. Once verified, if there are at least 3 votes for the URL, truth discovery model is invoked and assigns the URL's calculated score where a positive score indicates that the URL is a phishing

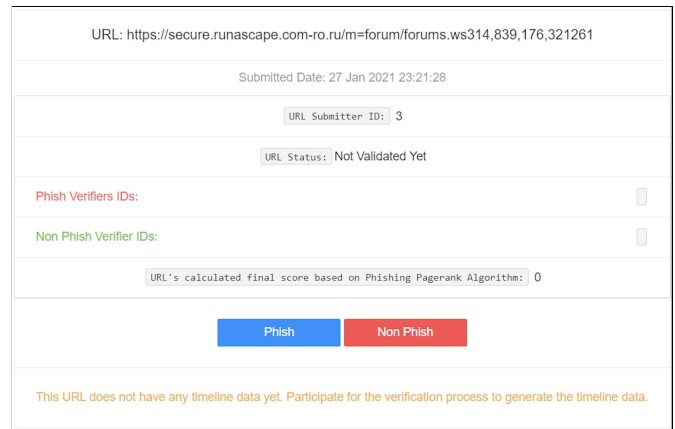


Figure 3: UI for URL Verification

URL. Note that the users in here are real-world users with varying degree of domain expertise who may either submit URLs or verify whether the URL is phishing or not. Further, in order to limit the abuse of the system, PhishChain enforces that each URL submitted to the system should be accompanied with a valid email referring to the URL<sup>3</sup>.

#### 3.2 URL Detailed Dashboard

Figure 5 shows the details of the URL submitted by user 1, and it is verified by users 3 and 4 as phishing and users 2 and 5 as non phishing. The system assigns an overall final phishing score of 0.3452 based on PR based truth discovery algorithm. Since it is a positive value, the system concludes that the URL is a phishing URL at that point in time.

#### 3.3 URL Verification Graph and Timeline

Figure 6 and Figure 7 show the verification graph and the timeline of verification respectively. Based on the participation in the system and the proportion of accurately labeled URLs, each user is assigned

<sup>3</sup>Similar to existing blocklists, the goal of PhishChain to assess suspicious URLs found in likely phishing emails.

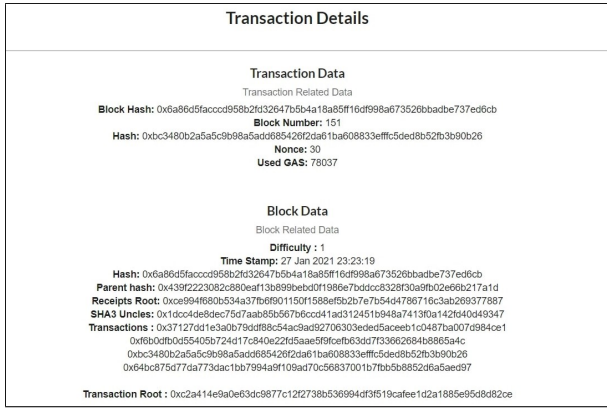


Figure 4: UI for URL Verification Transaction Details

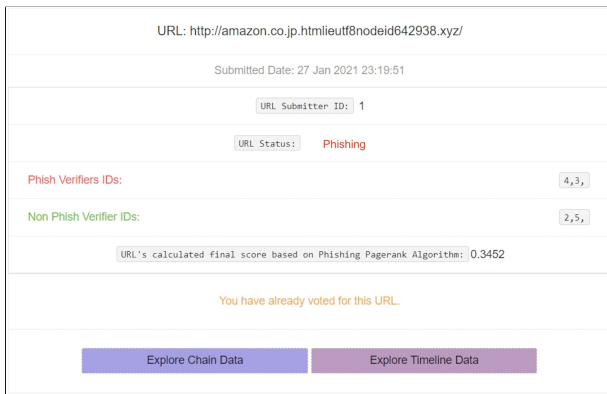


Figure 5: UI for URL Detailed Dashboard

skill points. For example, user 1 has the highest skill points of 153. Notice that after the 3<sup>rd</sup> vote, the URL is marked as phishing with a score of 0.64, but after the 4<sup>th</sup> vote, URL is still marked as phishing but with a lower phishing score of 0.35. The phishing score changes with every additional vote as truth discovery algorithm is invoked on affected URLs. Such a detailed timeline easily allows an analyst to not only assess the latest status but also understand how our system arrived at the decision.

#### 4 DISCUSSION

The motivation for introducing skill points is inspired from the question answering systems such as StackOverflow and StackExchange. Such systems show that even though participants do not receive any monetary benefits by contributing, these participants attain various “expertise” levels that have far more significance in their respective domains. A similar system for malicious URL verification could assist solve the problem of the limited participation in the volunteer based blacklisting systems currently available.

In this work, we assume that users do not maliciously try to manipulate the truth discovery based system. However, one may extend PhishChain to defend against such manipulations by introducing additional mechanisms that the reputation/recommender system community has already made available [3].

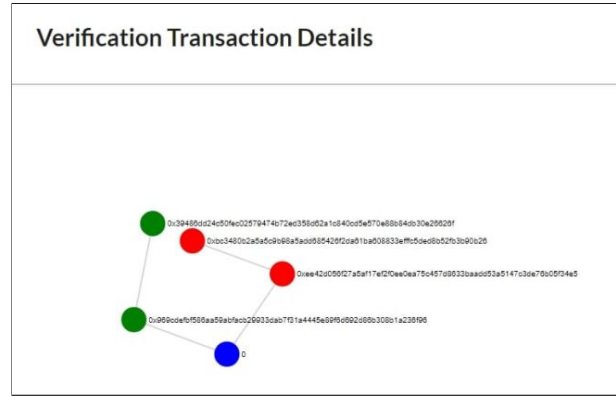


Figure 6: UI for URL Verification Graph

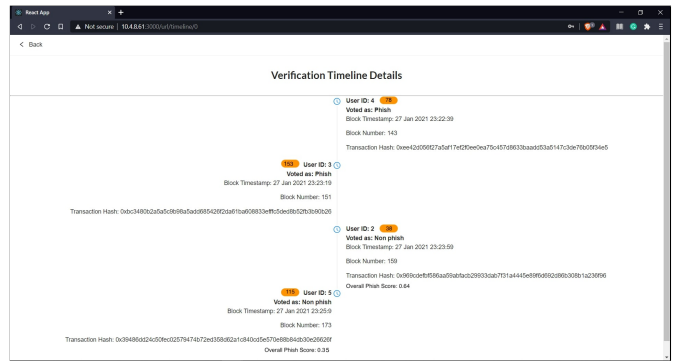


Figure 7: UI for URL Verification Timeline

#### 5 CONCLUSIONS

We propose, PhishChain, a truly transparent and decentralized phishing URL blacklisting system that is managed by a consortium instead of a centralized entity. Based on the proposed page rank based truth discovery algorithm, we compute a phish score for each URL as well as verifier skill points. Smart contracts are employed to implement the proposed functionality over the Quorum blockchain. We demonstrate three use cases: (1) URL verification, (2) URL dashboard and its transaction, and (3) URL verification graph and timeline. We believe our approach is generally applicable to any crowd sourced blacklisting system.

#### REFERENCES

- [1] ConsenSys. 2021. Quorum. Available at <https://consensys.net/quorum>.
- [2] ConsenSys. 2021. Tessera. Available at <https://github.com/ConsenSys/tessera>.
- [3] Eric Friedman, Paul Resnick, and Rahul Sami. 2007. Manipulation-resistant reputation systems. *Algorithmic Game Theory* 677 (2007).
- [4] M. Li, J. Weng, A. Yang, W. Lu, Y. Zhang, L. Hou, J. Liu, Y. Xiang, and R. Deng. 2019. CrowdBC: A blockchain-based decentralized framework for crowdsourcing. *IEEE Trans. on Para. and Dist. Sys.* 30, 6 (2019), 1251–1266.
- [5] Yaliang Li, Jing Gao, Chuishi Meng, Qi Li, Lu Su, Bo Zhao, Wei Fan, and Jiawei Han. 2016. A Survey on Truth Discovery. *ACM SIGKDD Explorations Newsletter* 17, 2 (2016), 1–16. <https://doi.org/10.1145/2897350.2897352> arXiv:1505.02463
- [6] D. Liu, W. Wang, Y. Wang, and Y. Tan. 2019. PhishLedger: A decentralized phishing data sharing mechanism. *ACM International Conference Proceeding Series* (2019), 84–89.
- [7] Y. Lu, Q. Tang, and G. Wang. 2018. ZebraLancer: Crowdsourced Knowledge atop Open Blockchain, Privately and Anonymously. July (2018), 1–17. arXiv:1803.01256