What's in Your Wallet?

Privacy and Security Issues in Web 3.0

Philipp Winter[1], Anna Harbluk Lorimer[1]
Peter Snyder[1], and Benjamin Livshits[1,2]

[1] Brave Software
[2] Imperial College London

Abstract. Much of the recent excitement around decentralized finance (DeFi) comes from hopes that DeFi can be a secure, private, less centralized alternative to traditional finance systems but the accuracy of these hopes has to date been understudied; people moving to DeFi sites to improve their privacy and security may actually end up with less of both.

In this work, we improve the state of DeFi by conducting the first measurement of the privacy and security properties of popular DeFi applications. We find that DeFi applications suffer from the same kinds of privacy and security risks that frequent other parts of the Web. For example, we find that one common tracker has the ability to record Ethereum addresses on over 56% of websites analyzed. Further, we find that many trackers on DeFi sites can trivially link a user's Ethereum address with PII (e.g., name or demographic information) or phish users.

This work also proposes remedies to the vulnerabilities we identify, in the form of improvements to the most common cryptocurrency wallet. Our wallet modification replaces the user's real Ethereum address with site-specific addresses, making it harder for DeFi sites and third parties to (*i*) learn the user's real address and (*ii*) track them across sites.

1 Introduction

We are witnessing a movement in finance whose goal is to replace long-standing institutions with code. More than 65 billion dollars are currently locked in decentralized finance (DeFi). Guided by the principles of decentralization and self-custody, DeFi applications implement numerous instruments from traditional finance like insurance, loans, or exchanges which are now guided by algorithms rather than human intervention. DeFi applications differ significantly in their architecture from traditional websites: Instead of a complex Web front end that communicates with a database back end, DeFi applications expose a lightweight front end that interacts with a blockchain-based back end using wallet software that runs in the user's browser. This unorthodox design—coupled with the substantial and ever-increasing amount of money pouring into DeFi—raises several questions related to security and privacy: how well do DeFi sites protect the user's financial information? What is the attack surface that attackers could exploit to steal funds? What is the role of third-party trackers? This work sets

out to answer these questions. While smart contract security has been studied extensively over the last several years, the above questions have received little attention.

Security and privacy problems: We begin by compiling a list of 78 popular DeFi sites and proceed to study the problem through the lens of well-established Web security and privacy methods. In particular, we identify what third-party providers DeFi sites use, what user data those providers can obtain, and how they could misbehave. Our data shows that numerous sites use services like Google Analytics, which can leak sensitive information like the user's Ethereum address or financial activity. We also look at problems that are due to the architectural design of DeFi applications. If a DeFi site embeds scripts, those scripts are able to interact with the user's wallet API, which can facilitate phishing attacks.

Solutions: While well-established tools like script blockers go a long way toward protecting DeFi users from trackers, it is no panacea. We therefore further raise the bar by implementing a simple, yet effective privacy-enhancing patch for the MetaMask in-browser wallet. Our patch derives per-site "fake" Ethereum addresses while retaining usability.

Contributions: This work makes the following contributions:

- **Third parties result in privacy violations.** We show that script embedding—a well-accepted Web development technique—can be problematic in DeFi because it facilitates phishing and may allow third-party trackers to learn the user's Ethereum address.
- **Impact of third-party tracking on DeFi privacy.** To quantify the security and privacy risks, we determine how many DeFi sites rely on third parties and find that 66% of sites embed at least one script and 56% of DeFi sites embed at least one script from Google.
- **Wallet-based privacy mitigation.** To mitigate a user's exposure to third-party trackers, we implement a lightweight patch for the MetaMask in-browser wallet.

Paper organization: In the rest of this work, Section 2 provides background on DeFi, followed in Section 3 by an explanation of the privacy and security issues we discovered. We then measure the prevalence of those issues on popular DeFi sites in Section 4 and propose privacy-preserving countermeasures in Section 5. Section 6 discusses this work's limitations and makes recommendations to both users and DeFi developers. We conclude this work in Section 7 by contrasting it with past research and summarize the findings in Section 8.

2 Background

2.1 Web Tracking

Users are typically tracked on the Web by third parties that use either explicit or implicit information to track users across sites. Historically, third-party cookies

have been a popular way of tracking users but browsers like Firefox and Safari have recently started blocking third-party cookies by default, which makes them a less effective tracking vector. Numerous alternative tracking vectors exist though, like browser fingerprinting, whose idea is to uniquely identify users based on the nuances of their browser's configuration, e.g., screen resolution, available fonts, or installed extensions.

2.2 Ethereum

Ethereum's native currency is called Ether and users often manage (some of) their Ether in software wallets that are implemented as browser extensions. These extensions store the user's private key[3] and provide a UI for managing the user's funds. As of September 2021, the most popular in-browser wallet is MetaMask.[4]

Unlike Bitcoin, Ethereum implements an account-based model, similar to bank accounts. All transactions from and to a given Ethereum account are eternalized on the blockchain and can easily be linked to the user's account (but not necessarily their real-world identity), which means that at best, users enjoy pseudonymity.[5] While privacy layers have been proposed on top of Ethereum, their usage remains scarce [3, 17, 18].

Users are encouraged to keep their Ethereum address secret because knowing an address reveals the owner's funds (which may paint a target on one's back) and transaction history (which may reveal sensitive relationships). Regardless, many users freely share their Ethereum address online and, of course, the average user needs to reveal her address to some parties (e.g., Ethereum nodes) to use the network meaningfully. Given Ethereum's emerging financial ecosystem, we note that online advertisers have an incentive to obtain a user's Ethereum address because it may allow for targeted advertisement; this advertisement can include traditional e-commerce if wallet transactions are used for online marketplaces or, more likely, crypto-focused ads for NFTs or services such as crypto-friendly debit cards, for instance.

2.3 Decentralized Finance Software Stack

Centralized exchanges like Coinbase run counter to the philosophy of decentralization that is at the core of cryptocurrencies, which is why decentralized alternatives have emerged. Typically referred to as DeFi, these applications allow users to invest in liquidity pools, exchange tokens, send payments, or lend money. A particularly popular example is Uniswap, a site that allows users to swap ERC-20 tokens and invest in liquidity pools—tasks that have historically only been offered by centralized exchanges like Coinbase.

[3] Note that one can also store the private key in a hardware wallet and manage it via a software wallet.

[4] https://metamask.io

[5] It is however straightforward to create multiple Ethereum accounts to compartmentalize one's financial activity.

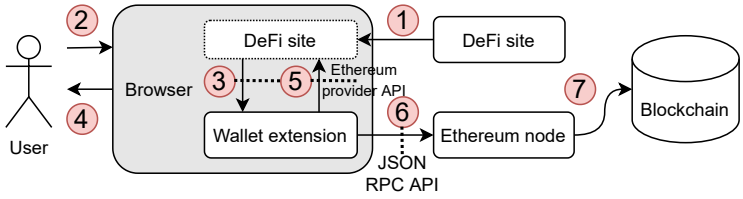

Fig. 1. The conceptual flow of DeFi sites: ❶ Users visit a DeFi site, ❷ click the "connect wallet" button, ❸ which prompts the DeFi site to ask the user's wallet for permission. ❹ The user then grants permission in the UI, ❺ after which the DeFi site can access the user's Ethereum account, and create transactions that make it ❻ via an Ethereum node ❼ to the blockchain.

As illustrated in Figure 1, DeFi applications essentially consist of a Web interface that bridges the gap between a user's Ethereum wallet (e.g., MetaMask) and the DeFi application's smart contract. DeFi sites therefore need to interact with the Ethereum blockchain *and* the user's wallet. Both types of interactions can (but don't need to) take place via the Ethereum provider API [9]—a JavaScript API that the MetaMask extension injects into the DeFi site's DOM. The DeFi site can then interact with the API via the `window.ethereum` JavaScript object. A subset of the Ethereum provider API is handled directly by MetaMask, e.g. the signing of transactions. The remaining API calls are *not* handled by MetaMask and forwarded to an Ethereum node, e.g. to an Ethereum-as-a-service provider like Infura.[6] Note that a DeFi site does not have to rely on MetaMask to interact with an Ethereum node; it could simply talk to an Ethereum node directly—many DeFi sites do—and limit the interaction with the Ethereum provider API to the signing of transactions. Unlike MetaMask, Ethereum nodes expose a JSON-based RPC interface. This interface consists of several dozen functions to call smart contracts, fetch gas prices, or obtain the number of the most recent block.

Note that MetaMask makes available some of its API functions only *after* the user gave permission—typically by manually clicking the "connect wallet" button[7] (to prevent unauthorized sites from accessing the user's Ethereum account information [5]), which prompts the DeFi site to ask MetaMask for permission, followed by a UI dialog asking the user to confirm the DeFi site's request. Once the user gave permission, the DeFi site is able to access the user's Ethereum address and balance, and create transactions (that still need to be signed off by the user).

For more information on DeFi, refer to Werner et al.'s 2021 arXiv report [21].

[6] https://infura.io

[7] E.g., via the following JavaScript call:
`window.ethereum.request({method: 'eth_requestAccounts'})`

3 Issues and Attacks in Web 3.0 Sites

In this section, we briefly outline the privacy and security issues we discovered among DeFi sites by using the example of 1inch.exchange (in short: 1inch). None of those issues are new—we could summarize them as "the past comes back to bite us again." What makes those issues relevant is the nature of DeFi sites, which differs from traditional websites in that (*i*) substantial amounts of money are involved and all that stands between a malicious DeFi site and the user's funds is often just a layer of JavaScript; (*ii*) a user's Ethereum address is effectively a unique, long-term identifier that is linked to the user's publicly visible financial history—ideal conditions for online tracking. A summary of issues and attacks covered below is shown in Figure 2.

Privacy: 1inch's landing page embeds scripts from several third parties, one of which is Google Analytics. This reveals arguably sensitive financial browsing activity to Google because 1inch encodes in its URLs what tokens the user is interested in exchanging, e.g. the URL `https://app.1inch.io/#/1/swap/COMP/USDC` (which makes its way to Google Analytics) reveals that the user selected the COMP and USDC token to swap. Furthermore, the use of Google Analytics also leaks the user's *Ethereum address* because 1inch happens to put the user's Ethereum address in an "event label" on Google Analytics. This allows Google to link the user's Ethereum address with the PII the company likely already has about the user. Worse, Google—and other analytics providers—also play a role in other DeFi sites, making it possible to track users *across sites*.

Analytics providers enjoy widespread use thanks to the convenience and insight that they provide but the amount of data they leak to third parties is highly sensitive in the context of DeFi.

Security: 1inch embeds a chat widget that allows users to contact 1inch support staff. The widget is provided by a third party and consists of JavaScript that is embedded in 1inch's first party context and therefore has full control over 1inch's DOM. That by itself is not surprising—virtually all major websites embed scripts but once the user connects her MetaMask wallet to 1inch, both 1inch *and* the embedded chat widget are able to interact with the Ethereum provider API that's injected by MetaMask. Crucially, the chat widget (or whoever compromised its distribution infrastructure) can modify 1inch's DOM to phish the user in an attempt to steal funds, or directly make a transaction via the user's wallet. While this transactions would have to be approved by the user, we believe that a well-crafted transaction at the right time would fool many users.

First-party script inclusion always brings with it a certain risk but we believe that this risk is highly elevated in the context of DeFi considering the amount of money that is at stake.

Issue type	Description
Privacy	• Ethereum address leaks to third parties. • Third-party trackers can track DeFi users across sites.
Security	• DeFi sites heavily rely on third-party scripts (and sometimes iframes) that could phish the user.

Fig. 2. A summary of the potential attacks and issues this paper presents.

4 Measurements

Having introduced potential privacy and security issues, we now turn to understanding how common those issues are. How many DeFi sites exhibit one or more of those issues? How common is the presence of Google Analytics?

We begin by discussing our measurement method (§4.1), followed by an analysis of how often a user's Ethereum address leaks to third parties (§4.2). We then take a step back and determine what third parties DeFi sites rely on and what this implies (§4.3).

4.1 Method

Collecting DeFi Sites We begin by compiling a list of DeFi sites to inspect from DeFi Pulse.[8] DeFi Pulse lists the top DeFi sites ranked by "total value locked", i.e. the amount of money that is currently "locked" inside the respective smart contracts—an apt proxy for the popularity of DeFi sites. We added the top 50 sites as of 2021-08-26 and augmented the list with 26 sites that we found of interest, resulting in a list of 78 DeFi sites, shown in full in Figure 8 in Appendix A. We then manually turned the list of domains into URLs so that when clicked, the browser lands directly on the page that asks users to connect their wallet, e.g. we turned tornado.cash into https://app.tornado.cash.

Recording Requests In the next step, we set out to visit each site and record the requests that it makes. To facilitate that, we built a Puppeteer-based crawler that spawns an instance of Google Chrome 92.0.4515.159 on Linux, visits all URLs in our DeFi list for 30 seconds, and records each request that the respective site makes. We used a fresh Chrome profile, added the MetaMask 10.0.3 extension, and completed MetaMask's onboarding process, resulting in a new Ethereum address. We are particularly interested in what requests a site makes *after* the user connects her wallet, which is why we manually click on the "connect wallet" button once our crawler visits a new site.

For each site, our crawler created a JSON file that contains metadata about the respective site and a list of requests, each consisting of (*i*) the request context (e.g., the site itself, or an iframe), (*ii*) the requested URL, and (*iii*) the type of request (e.g., a fetch, image, or script request).

[8] https://defipulse.com

4.2 Ethereum Address Leaks

Having recorded all requests of the most popular DeFi sites, we wondered if any of those requests leak the user's Ethereum address to third parties, e.g. did foo.finance send an XHR request containing our Ethereum address to bar.finance? To answer this question, we look for requests (*i*) whose destination has a different eTLD+1 than the origin[9] and (*ii*) whose URL contains our Ethereum address.

Table 3 illustrates the DeFi sites that leaked our Ethereum address, along with the number of leaks we found. Our script detected that 13 out of our 78 sites (17%) leaked our Ethereum address to third-party domains. Yearn issued four fetch requests to api.zapper.fi—to retrieve "Yearn Vaults" and token balances; DeFi Saver issued fetch requests to defiexplore.com and api.compound.finance; BiFi requested four images from heapanalytics.com whose URL contained our Ethereum address; and Zerion issued three fetch requests to ipfs.3box.io and maker.ifttt.com.

DeFi site	# of leaks
yearn.finance	4
defisaver.com	3
bifi.finance	3
zerion.io	3
loopring.io	2
bancor.network	2
dodoex.io	2
sablier.finance	1
reflexer.finance	1
impermax.finance	1
1inch.io	1
jelly.market	1
rarible.com	1

Fig. 3. DeFi sites that leaked our Ethereum address to third parties.

4.3 Cross-origin Dependencies

Recall from Section 3 that DeFi sites—like traditional banking sites—must be particularly careful about what external scripts they embed because those scripts can see wallet balances and potentially phish the user by modifying the page's DOM, which raises the question: how many DeFi sites embed scripts from third parties? We answer this question by extracting script requests from our DeFi list whose destination eTLD+1 differs from the site's origin.

Our data shows that 48 DeFi sites (66%) embed at least one script from a total of 34 third parties. Table 4 shows the top ten third parties that are embedded the most. Google's Tag Manager can be found on 28 DeFi sites, followed by Google Analytics on 21 sites. Intercom provides a chat support widget that can be found on 8 DeFi sites. Google's presence among DeFi sites is pervasive: Our data shows that 41 DeFi sites (56%) embed at least one script provided by Google.

Conversion analysis: Is Google able to monetize Alice's behavioral data as she navigates DeFi sites? For example, several DeFi sites leak their users' Ethereum

[9] We added a special case for the sites 1inch.exchange and balancer.fi because they also operate 1inch.io and balancer.finance, respectively—different eTLD+1 domains that are run by the same organization.

address directly to Google Analytics, giving the company the opportunity to link a user's real-world identity (which is PII) to her Ethereum address. What else can Google and other third party trackers learn about Alice?

As an advertisement company, Google has an incentive to track users through conversion funnels—an e-commerce term that refers to a user's journey from first hearing about a product, to considering a purchase, to finally purchasing it. Applied to the space of cryptocurrency, this could mean tracking Alice in each step as she (*i*) sees an "ad" for a currency or a token on a news site, (*ii*) inspects the asset's price chart on a price discovery site, and (*iii*) purchases the asset on a DeFi site. A concrete funnel could look as follows: Alice browses the popular news site CoinDesk where she learns about a new token that recently appreciated in value. To learn more about the token's price performance, she visits CoinGecko. Having gained faith in the token's performance, Alice decides to invest in it and visits Uniswap to make the purchase. An advertisement company tracking Alice through the funnel can enrich its profile about Alice, allowing for more effective targeted advertisement. Is Google able to do this? What about other entities?

Third party	# of sites
googletagmanager.com	28
google-analytics.com	21
intercomcdn.com	8
intercom.io	8
airswap.io	6
cloudflareinsights.com	5
facebook.net	3
crisp.chat	3
google.com	2
gstatic.com	2

Fig. 4. Top ten third party sites whose scripts were embedded the most.

To answer these questions, we need websites representing the top and middle part of the conversion funnel—a list of cryptocurrency news sites and of price discovery sites. Note that we already have a list of DeFi sites, which constitutes the bottom part of the funnel. We obtained a list of five news sites and six price discovery sites (see Figure 7) by searching Google for the top search results for keywords such as "cryptocurrency price" or "crypto news."

Equipped with three lists of sites, a subset of which Alice would traverse until she eventually conducts a transaction, we now turn to understanding what entities can track Alice.

For each of the sites in our three categories, we determine the third parties (identified by their eTLD+1) from which the sites embed scripts. We then determine what third parties can track users in all three categories—news site, price discovery, and purchase. Finally, we select the subset of entities that is able to track on at least 1% of sites in each of the three categories. In other words, we select entities that can track on at least 1% of news sites *and* on at least 1% of price discovery sites, *and* on at least 1% of DeFi sites. Figure 5 lists the five companies that survived our selection criteria, and the respective percentage of sites per category that they can track on.

What stands out is that Google is virtually omnipresent, observing *all* news and price discovery sites and *most* DeFi sites. Cloudflare, Facebook, Hotjar, and

Third party	% of news sites	% of price sites	% of DeFi sites
Google	100	100	56
Cloudflare	20	50	7
Facebook	60	33	4
Hotjar	40	33	3
Linkedin	20	17	1

Fig. 5. Companies whose scripts are embedded in at least 1% of sites of each of our three funnels parts—consisting of news sites, price discovery sites, and DeFi sites.

Linkedin are all only present on less than 10% of DeFi sites, resulting in less opportunity to track users through the conversion funnel.

Note that our set of news, price discovery, and DeFi sites is incomplete but we don't expect a larger set to change our results significantly because of the prevalence of Facebook, Cloudflare, and especially Google. We may however see other sites take the place of the less popular Hotjar and Linkedin.

5 Designing Countermeasures

This section discusses the design and implementation of a MetaMask patch that protects users against third-party address leaks. We begin with a conceptual overview of our approach (§ 5.1), followed by implementation details (§ 5.2), and finally evaluate the user experience implications of our patch (§ 5.3).

5.1 Conceptual Overview

The basic idea of our patch is to teach the wallet to not hand out the user's real Ethereum address to DeFi sites; instead, the patch hands out *deterministically-derived, site-specific* Ethereum addresses whose leakage to third parties is harmless.[10] Once the user wishes to make a transaction, we replace the derived address with the user's real address for the transaction to be valid. However, doing this alone would likely break both the DeFi site's UI and its functionality because DeFi sites often take into account the user's wallet balance to determine if a transaction is possible. To work around this issue, we intercept the DeFi site's RPC calls and replace fake with real wallet addresses in the request parameters, which effectively results in the DeFi site only ever learning the user's fake address but with its real balance.

5.2 Design and Implementation Details

We derive the (site-specific) Ethereum private key for example.com as follows:

$$K_{\text{example.com}} = \text{HMAC-SHA-256}(K, \text{“example.com”})$$

[10] Conceptually, this is similar to hierarchical deterministic wallets [22].

The key K to the HMAC is the private key of the user's real Ethereum account and its input is the DeFi site's eTLD+1. The HMAC's output is a 256-bit value, which serves as the private key of the derived Ethereum account. This key generation procedure guarantees that

1. repeat visits to example.com result in the DeFi site seeing the same Ethereum address, effectively creating an *address pseudonym*;[11]
2. accidental sending of funds to the user's fake rather than the real address *can be undone* because the user has control of the fake address's private key;
3. different sites see different addresses and third parties *cannot easily link* the user's addresses.

At this point, we can make example.com (and its third parties) believe that the user's Ethereum address is 0x123 instead of 0xabc but the site's UX is heavily impaired because it will not correctly reflect the user's balance[12]—worse, the DeFi site will not be usable because its UI won't allow the user to transact funds with an empty balance. We therefore need to make the site's UI believe that the fake address contains funds.

Many DeFi sites obtain the user's wallet balance via the `eth_getBalance` RPC call, so we modified MetaMask's internals to return the user's real balance if the DeFi site asks for the balance of the fake address. We noticed that DeFi sites use various other ways to obtain an address's balance; e.g., some invoke a smart contract via the `eth_call` RPC call. We therefore also intercept `eth_call` invocations and replace all occurrences of the user's fake with the real address. Finally, we noticed that some sites use Web requests to third party providers like Infura to obtain wallet balances. We used the tool Burp Suite [6] to replace addresses in those Web requests.

To recap, example.com now believes that the user's address is 0x123 instead of 0xabc *and* believes that 0x123's balance is that of 0xabc. One problem however remains: if the user opts to make a transaction on example.com, the resulting `eth_sendTransaction` RPC call will contain the fake address. Fortunately, fixing this problem is as simple as again replacing all occurrences of the fake with the real address. Figure 6 illustrates our approach.

Our proof-of-concept patch is based on MetaMask in version 10.0.3 and consists of approximately 100 lines of JavaScript code.[13]

Limitations: Note that we do not prevent DeFi sites from making an active effort to learn the user's real Ethereum address. The user's real address can easily be determined by taking the wallet balance and searching the blockchain for an address that has that exact balance. Our approach focuses on preventing

[11] One could make the keys effectively anonymous by adding a time parameter to the HMAC. Depending on the time granularity, subsequent visits to example.com will then use different Ethereum keys.

[12] The derived address is essentially a fresh Ethereum address and therefore contains no Ether.

[13] The code is available at `https://github.com/brave-experiments/defi-privacy-measurements`.

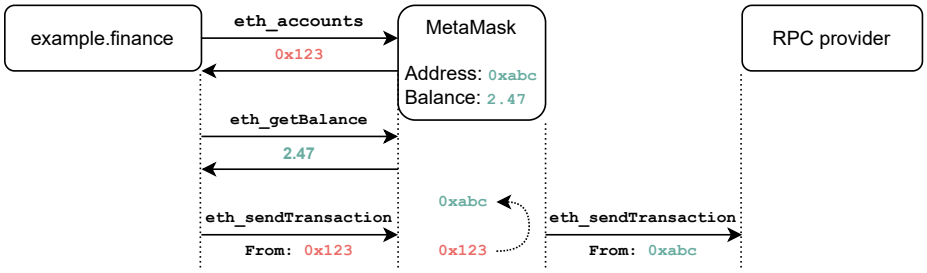

Fig. 6. When asked for the user's address, our wallet patch returns a derived address (0x123) and when the DeFi site asks for the fake address's balance, the patch returns the balance of the real address (0xabc). Finally, the patch replaces the fake address with the real address in transactions before forwarding them to the RPC provider.

leaks to third parties and making it more difficult for such third parties to track users across DeFi sites or link their Ethereum address (which is effectively a long-term, unique identifier) to their real-world identity.

We also note that some sites requests information about a user's Ethereum wallet without using MetaMask's Ethereum provider API, so we are unable to intercept those requests. In those cases, one has to use a separate tool like Burp Suite to intercept and rewrite requests.

5.3 Evaluation

To test our patch, we manually visited popular DeFi sites, made sure that our Ethereum address was replaced, wallet balances were adjusted, and it was possible to transact on those sites. While we believe that the core functionality of the sites we tested remained intact, we cannot rule out that there are niche features that we overlooked, which are broken by our patch. We deliberately designed our patch so that the worst case scenario—loss of funds—cannot happen because users control the private key of derived addresses and can therefore recover funds that were accidentally sent to fake addresses.

We believe that a sensible initial deployment plan for our MetaMask patch would include a whitelist that contains known-to-work sites. Adding sites to the whitelist involves ensuring that any instances of the fake address in RPC requests made by the site (including requests that do not pass through MetaMask) have their data field correctly modified to contain the user's real address. One could gradually expand that whitelist by fixing issues caused by our patch as they are encountered by users, or by repurposing our crawler from Section 4.1 to automate the process. Our whitelist currently consists of 23 DeFi sites including: Instadapp, Curve Finance, Uniswap, Sushiswap, and Tornado Cash. We note that the number of sites that work with our wallet patch is a *lower bound* as we were unable to test our patch with several of the uncommon tokens the DeFi sites in our list use.

6 Discussion

In this section we discuss our work's limitations (§ 6.1) and issue recommendations for both DeFi developers and users (§ 6.2).

6.1 Limitations

Our measurement method from Section 4.1 revealed some Ethereum address leaks to third parties but it is unable to reveal *deliberately disguised* address leaks. For example, unsophisticated address encoding schemes like Base64 could have evaded our detection method.

In this paper, we focused on 78 DeFi sites—most of which are among the most popular sites according to "total value locked." The total population of DeFi sites is much larger though and our results provide no insight into the privacy and security issues of the long tail of DeFi sites. Considering the little care and effort that goes into many DeFi sites, we expect the long tail to exhibit more problems than popular sites like Uniswap or Compound. Future work could take a more comprehensive and (perhaps longitudinal) approach to studying the problem.

Finally, our list of DeFi sites may exhibit *selection bias*, i.e. it may differ from the general population of DeFi sites in crucial aspects. For example, our list of popular DeFi sites may exhibit fewer security and privacy issues than the lesser-known long tail of DeFi sites.

6.2 Recommendations

What can DeFi developers do to improve security and privacy for their users, and guarantee the "De" in DeFi? How can users protect themselves? Below, we issue a set of recommendations for DeFi developers and users. For DeFi site users, we recommend the following:

Block analytics scripts: To prevent analytics providers from linking Ethereum addresses to real-world identities, we recommend browser extensions like Privacy Badger or browsers like Brave Browser, or Tor Browser.

Don't connect your wallet unless you have to: We recommend treating one's Ethereum address like credit card or bank account information, i.e. only revealing it selectively and when necessary. While our MetaMask patch from Section 5 mitigates this problem, it does not fix it.

For developers of DeFi sites, we recommend the following:

Use self-hosted analytics: We recommend the use of self-hosted analytics scripts over third-party services to minimize exposure to third parties. At the very least, if a DeFi site uses third party providers, it should ensure that their page URLs don't contain sensitive information that allow analytics providers to link Ethereum addresses to PII.

Consider address privacy a first-class design goal: Our work shows that many DeFi sites don't consider Ethereum addresses private.[14] We recommend that DeFi developers treat Ethereum addresses like credit card information.

Reconsider threat models: Our results show that well-established Web development methods like the use of innocuous support widgets can open the gates to phishing attacks.[15] We recommend that DeFi developers think carefully about what third parties should be allowed to modify their DOM. Not only does one have to trust that those third parties are *not malicious*, one also has to trust that they are *competent* and secure their infrastructure from compromise—a high bar that may be difficult to meet if substantial amounts of money are at stake.

7 Related Work

We conclude this paper by contrasting it with related work, which we divide into online privacy related to third-party tracking and cryptocurrencies, and finally user perception of cryptocurrencies.

Online Privacy and Third-Party Tracking: There exists a plethora of ways to track users online [1, 8, 7]. One of the most pervasive ways to do this is via third party trackers such as Google Analytics, which uses HTTP requests and first party cookies in addition to information about the user's browser and system to compile a profile of a user's online activity [19].

Online Privacy and Security of Cryptocurrencies:

Past work on Bitcoin privacy has mostly focused on the linkability of addresses [15] (along with some exploration of network-layer issues [2]), a PoPETs'18 paper [10] by Goldfeder et al. takes a first look at the intersection between Bitcoin and online privacy. An increasing number of online vendors now support payment by Bitcoin but the authors show that online trackers often collect sufficient sensitive information to link a purchase to its subsequent blockchain transaction. Worse, by taking into account auxiliary information, attackers could link together Bitcoin transactions that were anonymized via CoinJoin—a popular mixer at the time. We build on Goldfeder et al.'s first foray into the intersection of cryptocurrency and online privacy, showing that both third and first party scripts can facilitate security and privacy issues in DeFi sites.

A 2020 technical report by Béres et al. [4] takes a look at privacy in Ethereum, showing how attackers can profile and deanonymize users. The authors show that one can link several Ethereum addresses to the same owner by taking into account the time of day these addresses are typically used, the gas price, and

[14] For example, 1Inch Exchange's referral code contains the inviting user's Ethereum address, which is exposed to the invitee.

[15] Phishing constitutes a pervasive problem taking on numerous forms, ranging from scammers masquerading as support staff [23] or "helpful" social media users [24] to fake [11] or compromised wallet software [16].

unrelated addresses that are transacted with. Even mixers like the popular Tornado Cash are no panacea because they are frequently misused. For example, not understanding the nuances of mixers, some users use the same Ethereum address for deposit and withdrawal, effectively deanonymizing themselves. We expand on Béres et al.'s work by pointing out how common Web development methods lead to privacy and security issues in DeFi applications.

Li et al. reveal in their NDSS'21 paper [13] a DoS attack that makes it possible to disable RPC services that DApps rely on—e.g. to get the upper hand in an auction by preventing competing bidders from placing their bids. The attack exploits the fact that many RPC services don't impose a gas limit on the `eth_call` method, making it possible to make the RPC service engage in heavy computation, thus preventing it from serving other clients.

Refer to Werner et al.'s arXiv report [21] for a comprehensive overview of DeFi and the state of open research questions around DeFi.

User Perception:

In a FC'16 paper [12], Krombholz et al. surveyed 990 Bitcoin users about their understanding of Bitcoin security, privacy, and anonymity. Interestingly, nobody stored wallet backups on an air-gapped computer but several respondents used encrypted backups. 32% of respondents believe that Bitcoin is per-se anonymous but 80% think that it is possible to follow their transactions. The study's respondents were no strangers to loss of Bitcoin: 22% report that they have lost Bitcoins at least once—due to hardware or software failure, or because they lost access to their private keys. Finally, respondents saw vulnerabilities in hosted wallets as the top threat right after Bitcoin value fluctuation.

User misconceptions go beyond mixers. Drawing on data from twenty interviews with cryptocurrency users and non-users, Voskobojnikov et al. show in their FC'20 paper [20] that users express confusion about the concept of "gas prices," mistakenly believe that they own the private key for funds stored by the company Coinbase, or don't understand the idea behind public and private keys altogether. Faulty mental models can lead to critical mistakes like the loss of funds, highlighting the importance of safe defaults that protect users. This work contributes to our understanding of what privacy-preserving safe defaults look like.

Most recently, a SOUPS'20 paper by Mai et al. [14] qualitatively studied cryptocurrency users' (N=29) mental models and how these models are in conflict with security and privacy goals. Corroborating the findings of Voskobojnikov et al., the authors find that the idea of cryptographic keys is a frequent source of misunderstanding, prompting some participants to believe that miners or "the blockchain" create private keys. Other participants exhibit misunderstandings about the blockchain, believing that old transactions are eventually deleted, or that transactions are confidential and cannot be seen by third parties. This highlights that some users may have incorrect expectations of privacy.

8 Conclusions

This work is the first to examine the intersection between Web security and privacy, and DeFi.

This paper improves the state of DeFi by conducting the first measurement of the privacy and security properties of popular DeFi applications. We find that well-understood security and privacy risks are as widespread on DeFi applications as on other parts of the Web—but carry greater risk in DeFi given that money is involved. For example, we find that one common tracker has the ability to track DeFi users on 56% of websites analyzed. Further, we find that many trackers on DeFi sites can trivially link a user's Ethereum address with PII (e.g., name or demographic information) or phish users.

This work also proposes remedies to the vulnerabilities we identify, in the form of improvements to the most common cryptocurrency wallet. Our wallet patch replaces the user's real Ethereum address with site-specific pseudo addresses, making it harder for DeFi sites and third parties to (*i*) learn the user's real address and (*ii*) track them across sites. We test our improvement on popular DeFi sites and conclude that it protects the user's address from curious DeFi sites in the majority of cases.

Acknowledgements

We thank Will Scott, Roman Semenov, Arthur Gervais, Hamed Haddadi, and Gonçalo Pestana for their feedback on earlier versions of this paper.

References

[1] Gunes Acar et al. "The Web Never Forgets: Persistent Tracking Mechanisms in the Wild". In: *CCS*. ACM, 2014. URL: https://securehomes.esat.kuleuven.be/~gacar/persistent/the_web_never_forgets.pdf.

[2] Maria Apostolaki, Cedric Maire, and Laurent Vanbever. "Perimeter: A network-layer attack on the anonymity of cryptocurrencies". In: *FC*. Springer, 2021. URL: https://nsg.ee.ethz.ch/fileadmin/user_upload/publications/fc21final97.pdf.

[3] *Aztec*. URL: https://aztec.network (visited on 09/07/2021).

[4] Ferenc Béres, István A. Seres, András A Benczúr, and Mikerah Quintyne-Collins. *Blockchain is Watching You: Profiling and Deanonymizing Ethereum Users*. arXiv: 2005.14051v2 [cs.CR]. URL: https://arxiv.org/pdf/2005.14051.pdf.

[5] Paul Bouchon and Erik Marks. *Opt-in account exposure*. URL: https://github.com/ethereum/EIPs/blob/master/EIPS/eip-1102.md (visited on 09/07/2021).

[6] *Burp Suite – Application Security Testing Software*. URL: https://portswigger.net/burp (visited on 09/13/2021).

[7] Steven Englehardt and Arvind Narayanan. "Online Tracking: A 1-Million-Site Measurement and Analysis". In: *CCS*. ACM, 2016. URL: https://www.cs.princeton.edu/~arvindn/publications/OpenWPM_1_million_site_tracking_measurement.pdf.

[8] Steven Englehardt et al. "Cookies That Give You Away: The Surveillance Implications of Web Tracking". In: *WWW*. ACM, 2015. URL: https://senglehardt.com/papers/www15_cookie_surveil.pdf.

[9] *Ethereum Provider API*. URL: https://docs.metamask.io/guide/ethereum-provider.html#table-of-contents (visited on 09/13/2021).

[10] Steven Goldfeder, Harry Kalodner, Dillon Reisman, and Arvind Narayanan. "When the cookie meets the blockchain: Privacy risks of web payments via cryptocurrencies". In: *PoPETs* 2018.4 (2018). URL: https://www.petsymposium.org/2018/files/papers/issue4/popets-2018-0038.pdf.

[11] Yogita Khatri. "Fake MetaMask App on Google Play Store Hosted Crypto Malware". In: *CoinDesk* (Feb. 2019). URL: https://www.coindesk.com/markets/2019/02/11/fake-metamask-app-on-google-play-store-hosted-crypto-malware/ (visited on 08/31/2021).

[12] Katharina Krombholz, Aljosha Judmayer, Matthias Gusenbauer, and Edgar Weippl. "The Other Side of the Coin: User Experiences with Bitcoin Security and Privacy". In: *Financial Cryptography*. Springer, 2016. URL: https://fc16.ifca.ai/preproceedings/33_Krombholz.pdf.

[13] Kai Li et al. "As Strong As Its Weakest Link: How to Break Blockchain DApps at RPC Service". In: *NDSS*. The Internet Society, 2021. URL: https://www.ndss-symposium.org/wp-content/uploads/ndss2021_3C-1_23108_paper.pdf.

[14] Alexandra Mai et al. "User Mental Models of Cryptocurrency Systems - A Grounded Theory Approach". In: *SOUPS*. USENIX, 2020. URL: https://www.usenix.org/system/files/soups2020-mai.pdf.

[15] Sarah Meiklejohn et al. "A Fistful of Bitcoins: Characterizing Payments Among Men with No Names". In: *IMC*. ACM, 2013. URL: https://cseweb.ucsd.edu/~smeiklejohn/files/imc13.pdf.

[16] Andrey Shevchenko. "Founder of DeFi protocol Nexus Mutual gets hacked for $8M". In: *Cointelegraph* (Dec. 2020). URL: https://cointelegraph.com/news/founder-of-defi-protocol-nexus-mutual-gets-hacked-for-8m (visited on 08/31/2021).

[17] *StarkEx*. URL: https://starkware.co/product/starkex/ (visited on 09/07/2021).

[18] *Tornado Cash*. URL: https://tornado.cash (visited on 09/07/2021).

[19] *Tracking Code Overview*. June 2018. URL: https://developers.google.com/analytics/resources/concepts/gaConceptsTrackingOverview (visited on 08/31/2021).

[20] Artemij Voskobojnikov, Borke Obada-Obieh, Yue Huang, and Konstantin Beznosov. "Surviving the Cryptojungle: Perception and Management of Risk Among North American Cryptocurrency (Non)Users". In: *Finan-*

cial Cryptography. Springer, 2020. URL: https://fc20.ifca.ai/preproceedings/20.pdf.

[21] Sam M. Werner et al. *SoK: Decentralized Finance (DeFi)*. arXiv: 2101.08778v3 [cs.CR]. URL: https://arxiv.org/pdf/2101.08778.

[22] Pieter Wuille. *BIP 32: Hierarchical Deterministic Wallets*. 2012. URL: https://en.bitcoin.it/wiki/BIP_0032.

[23] wuzz1e. *$75,000 just disappeared from my Coinbase wallet*. May 2021. URL: https://www.reddit.com/r/CoinBase/comments/nhug9u/75000_just_disappeared_from_my_coinbase_wallet/ (visited on 08/31/2021).

[24] Martin Young. "MetaMask warns of new phishing bot". In: *Cointelegraph* (May 2021). URL: https://cointelegraph.com/news/metamask-warns-of-new-phishing-bot (visited on 08/31/2021).

A List of URLs

News sites	Price discovery sites
https://www.coindesk.com	https://www.coingecko.com/en
https://cointelegraph.com	https://coinmarketcap.com
https://decrypt.co	https://coinranking.com
https://cryptonews.com	https://coincodex.com
https://www.theblockcrypto.com	https://coincheckup.com
	https://www.cointracker.io/price

Fig. 7. List of cryptocurrency news and price discovery sites.

https://activate.codefi.network/staking/airswap/governance	https://app.1inch.io
https://app.aave.com/markets	https://app.alchemix.fi
https://app.badger.finance	https://app.balancer.fi
https://app.bancor.network/eth/data	https://app.barnbridge.com
https://app.bifi.finance	https://app.bifi.finance/lend
https://app.boringdao.com	https://app.compound.finance
https://app.coverprotocol.com	https://app.cream.finance
https://app.defisaver.com	https://app.dodoex.io
https://app.enzyme.finance/depositor/leaderboard	https://app.fei.money
https://app.flexa.network	https://app.impermax.finance
https://app.jelly.market	https://app.mai.finance
https://app.maple.finance	https://app.nexusmutual.io/swap
https://app.pickle.finance	https://app.rampdefi.com
https://app.rari.capital	https://app.rari.capital
https://app.reflexer.finance	https://app.reflexer.finance
https://app.ribbon.finance	https://app.rulerprotocol.com
https://app.sushi.com	https://app.swapswap.org/#/swap
https://app.tornado.cash	https://app.truefi.io
https://app.truefi.io/home	https://app.uniswap.org
https://app.vesper.finance	https://app.warp.finance
https://app.yield.is	https://app.zerion.io
https://beta.curve.fi	https://curve.fi
https://dashboard.keep.network/overview	https://debank.com
https://defi.instadapp.io	https://dmm.exchange/#/about
https://exchange.dfyn.network	https://exchange.loopring.io/swap
https://flash.wing.finance	https://for.tube/market/index
https://foundation.app	https://harvest.finance
https://homora-v2.alphafinance.io	https://idle.finance
https://inverse.finance/anchor	https://liquity.app
https://moonswap.fi/exchange/swap	https://notional.finance/portfolio
https://o3swap.com/swap	https://oasis.app/dashboard
https://opensea.io/assets	https://pancakeswap.finance
https://pay.sablier.finance	https://rarible.com
https://saddle.exchange	https://staking.synthetix.io
https://tinlake.centrifuge.io	https://trade.dydx.exchange
https://trader.airswap.io/	https://v2.opyn.co
https://wasabix.finance/#/app	https://www.akropolis.io/app/home
https://www.convexfinance.com	https://www.indexcoop.com
https://yearn.finance	https://zapper.fi/dashboard

Fig. 8. Our list of 78 DeFi sites.