

Unbundle-Rewrite-Rebundle: Runtime Detection and Rewriting of Privacy-Harming Code in JavaScript Bundles

Mir Masood Ali
University of Illinois Chicago
Chicago, USA

Chris Kanich
University of Illinois Chicago
Chicago, USA

Peter Snyder
Brave Software
San Francisco, USA

Hamed Haddadi
Imperial College London & Brave Software
London, UK

ABSTRACT

This work presents Unbundle-Rewrite-Rebundle (URR), a system for detecting privacy-harming portions of bundled JavaScript code and rewriting that code *at runtime* to remove the privacy-harming behavior without breaking the surrounding code or overall application. URR is a novel solution to the problem of JavaScript bundles, where websites pre-compile multiple code units into a single file, making it impossible for content filters and ad-blockers to differentiate between desired and unwanted resources. Where traditional content filtering tools rely on URLs, URR analyzes the code at the AST level, and replaces harmful AST sub-trees with privacy-and-functionality maintaining alternatives.

We present an open-sourced implementation of URR as a Firefox extension and evaluate it against JavaScript bundles generated by the most popular bundling system (Webpack) deployed on the Tranco 10k. We evaluate URR by precision (1.00), recall (0.95), and speed (0.43s per script) when detecting and rewriting three representative privacy-harming libraries often included in JavaScript bundles, and find URR to be an effective approach to a large-and-growing blind spot unaddressed by current privacy tools.

CCS CONCEPTS

• Security and privacy → Web application security.

KEYWORDS

Web privacy; Content blocking

ACM Reference Format:

Mir Masood Ali, Peter Snyder, Chris Kanich, and Hamed Haddadi. 2024. Unbundle-Rewrite-Rebundle: Runtime Detection and Rewriting of Privacy-Harming Code in JavaScript Bundles. In *Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security (CCS '24)*, October 14–18, 2024, Salt Lake City, UT, USA. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3658644.3690262>

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 the author(s) 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.

CCS '24, October 14–18, 2024, Salt Lake City, UT, USA

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0636-3/24/10
<https://doi.org/10.1145/3658644.3690262>

1 INTRODUCTION

An enormous body of research has established Web content filtering (e.g., blocking advertising, tracking, and other unwanted network requests on websites) as an important and effective technique for improving privacy[11, 20], security[16, 32], and performance[10, 24]. Most Web content filtering approaches rely on crowd-sourced lists of regular-expression-like rules that describe which URLs the browser should load and which should be blocked.

This approach—broadly, URL-based content filtering—works because URLs, in practice, provide useful and stable information about the resources they map to. In some cases, this is because of the text in the URL (e.g., browsers can make a reasonable guess about the purpose of JavaScript returned from a URL like (<https://advertising.example/tracker.js>), or because experts have manually evaluated the resource returned from a URL and found it to be similarly harmful to users.

However, modern Web development practices make URL-based content filtering increasingly difficult. Previously, Web applications were often delivered as a collection of discrete JavaScript files, each fetched independently from their own URL (e.g., `/script/library.js`, `/script/tracker.js`, `/script/app.js`), which allowed URL-based content filtering tools to easily block some parts of an application, but not others. Increasingly, though, developers integrate bundling tools as part of their build and deployment practices, compiling all of the libraries and application code into a single file unit, which is delivered to the browser from a single URL (e.g., `/script/bundle.js`).

These bundling approaches, inadvertently or otherwise, circumvent URL-based content filtering tools. When applications are delivered as a single bundled code unit, URL-based filtering tools can no longer block just parts of the application; blocking a site's JavaScript becomes an all-or-nothing proposition. And since blocking all JavaScript on a page breaks useful functionality on many sites, in practice, content filtering tools are reduced to blocking nothing, reintroducing the privacy, security, and performance issues the user wanted to avoid in the first place.

This work presents the design and implementation of Unbundle-Rewrite-Rebundle (URR), a system to enable content blocking in modern Web applications, even when Web applications are deployed as a compiled, single file JavaScript bundle. In other words, URR aims to enable browsers to avoid executing the code from

/script/tracker.js, while still executing the non-privacy harming code originally provided in /script/library.js and /script/app.js, *even when* all three libraries are bundled and delivered in /script/bundle.js.

URR is a novel, practical solution to a problem that a vein of related Web privacy research has explored. Works like [2] identify that bundled applications are widespread and pose a serious challenge to Web privacy, and [27] found that blocking these bundled JavaScript resources often broke the benign, desirable parts of websites. Systems like [28] showed that bundled applications could be automatically rewritten to prevent privacy harm, though with expensive precomputation, which rendered practical deployment prohibitive. URR is a first-in-class approach to solving the privacy and security harms caused by bundled JavaScript applications in a performant and practical method.

To do so, URR solves several non-trivial challenges:

First, the system must identify known privacy-harming code (e.g., the code delivered from /script/tracker.js) within the larger bundled application, without any information about the URL from which the code originally came. This identification must be robust even across the kinds of code modifications and transformations JavaScript bundlers make in their build processes (e.g., minification, dead-code elimination, tree-shaking).

Second, URR must remove known-privacy-harming code from the bundled application without breaking desirable functionality in the surrounding code. Just deleting unwanted libraries from a bundled application will, in practice, be counter-productive since the application will fail when trying to access now-deleted functions and classes defined by the deleted library. A useful solution must remove the unwanted, target libraries from the bundled application *without breaking surrounding code*.

Third, such a system must be performant and be able to Unbundle, analyze, modify, and reconstitute bundled Web applications at runtime, and quickly, in a way that maintains the usefulness of the Web application. If the performance overhead of a privacy-and-security preserving system is too costly, then the system is, in practice, unusable and so not meaningfully useful to benefit users.

URR is implemented in several parts: i. as a database of signatures of ASTs of real-world known-privacy-harming code, ii. a library of crowd-sourced privacy-preserving alternative implementations of privacy-harming code designed to remove privacy harming behaviors without impacting the surrounding application code, iii. a browser extension that, at runtime, decomposes a bundled JavaScript application into its constituent libraries, detects the sub-ASTs from the bundled application that come from known-privacy-harming libraries and rewrite the bundled application with the stub libraries in place of the privacy-harming versions.

More broadly, this work makes the following contributions to Web privacy and security.

- (1) A novel algorithm for efficiently generating fingerprints of bundled modules within JavaScript libraries. These fingerprints are robust to many of the modifications that bundling tools make during their compilation process (e.g., label minification, “tree-shaking”).
- (2) An open-sourced, empirically tuned system for:
 - (a) unbundling applications into their constituent libraries

- (b) generating fingerprints for each sub-library and checking them against a database of known unwanted JavaScript libraries
 - (c) replacing unwanted JavaScript libraries with compatibility-preserving “shim” implementations, which maintain the library’s API “shape”, while removing any privacy-or-security affecting behaviors
 - (d) reconstituting the resulting new application into a new bundle that can then be passed to the browser’s JavaScript engine for normal execution.
- (3) An empirical evaluation of the accuracy and performance of our system when applied to a representative crawl of the Web, finding that our system results in libraries being blocked on 7% of the top 10K sites in the Tranco list within practical performance bounds.
 - (4) An open-source implementation of our system as a Firefox extension, along with the complete dataset for all discussed figures and measurements, available at <https://github.com/masood/urr>.

2 BACKGROUND

This section first provides a primer on concepts relevant to bundling and their use in web development. It then presents a simple example that highlights the limitations of existing content-blocking approaches. The section concludes by outlining the properties of an effective solution.

2.1 Bundles and Relevant Concepts

As websites and web applications grow more complex, they require a plethora of functionality, often aided by numerous libraries and dependencies. Handling these dependencies can prove to be quite strenuous, especially when considering the range of platforms on which the code needs to execute correctly. JavaScript bundlers are essential tools used by web developers to streamline the handling of code and dependencies within complex web applications. At its core, a JavaScript bundler is a utility that gathers and wraps code from multiple JavaScript files. Bundles not only reduce the number of network requests required to load a web page but also optimize the handling of dependencies and various aspects of development and production environments. The simplicity of the example hides certain nuances necessary to understand the complexity of bundles. Below, we introduce a few fundamental concepts that can help better understand JavaScript bundles.

2.1.1 Modular Programming. Web developers design and create websites in different environments than browsers. These environments have their own caveats and follow different programming concepts and philosophies. Modular programming is a general programming concept where developers separate complex application code into independent pieces called *modules*. A module forms the atomic unit of a bundle and roughly corresponds to a code snippet relevant to a file or library that exports functionality that is consumed by other modules. Popular JavaScript development environments like Node.js¹ adopt a modular programming approach. Developers can create their own modules and use reuse modules available in the npm registry².

¹<https://nodejs.org/en/>

²<https://www.npmjs.com/>

2.1.2 JavaScript Module Systems. Like development environments and programming approaches, JavaScript module systems are also not a monolith. Depending on the context in which they are consumed and executed, JavaScript modules express functionality in different ways. The two most popular module systems for JavaScript are (1) the ECMAScript modules (ESM), which are consumed with `import` statements, and (2) CommonJS (CJS) modules, which are consumed with `require` statements. While Node.js supports both types of modules, web browsers only recognize `import` statements. Bundles, therefore, include wrappers around modules and provide workarounds for `require` statements within web browsers.

2.1.3 Inter-module Dependencies. Larger projects include multiple libraries and packages and, as a result, comprise numerous interdependent modules. When bundles gather and parse all the modules that need to be combined, they create a dependency graph that helps determine (1) the order and chain of dependencies between modules and (2) which code snippets can be combined within a module and which snippets need to be split across multiple modules.

2.1.4 Minification. Bundlers additionally perform a code transformation step that reduces the overall size of the code. Depending on various configuration and optimization options, minification returns an irreversible code output containing randomized, unidentifiable variable names and changes to the code syntax that only retains its underlying logic.

2.1.5 Source Maps. During development, bundles provide source maps as a key to reverse minification and debug parts of code. These files help map minified code back to their unminified counterparts and, hence, identify individual modules. However, source maps are unavailable by default in production, making it difficult to reverse engineer bundled code in the wild.

2.1.6 Popular Bundlers. Examples of popular JavaScript bundlers include Webpack³, Browserify⁴, Rollup⁵, and Parcel⁶, each of which offer unique features and advantages. In this work, we focus on Webpack because it is the most popular and mature JavaScript bundler as determined from GitHub stars⁷, NPM weekly downloads⁸, and prior work [25].

2.1.7 General Structure of Webpack Bundles. Listing 1 presents an example of a script comprising a webpack bundle. We describe the example below.

- Webpack wraps the bundled code in an **Immediately Invoked Function Expression (IIFE)**. As a result, the bundle is executed as soon as it is loaded, thereby making all necessary functions and variables available in the global scope.
- The bundle comprises a modular system where each module is represented as a function in an **array or object of modules**. Listing 1 contains a module array in L30–42.
- **Function wrappers around each module** handle dependencies on other modules and gather any variables and functions exported by the module.

³<https://webpack.js.org/>

⁴<https://browserify.org/>

⁵<https://rollupjs.org/>

⁶<https://parceljs.org/>

⁷<https://github.com/webpack/webpack>

⁸<https://www.npmjs.com/package/webpack>

```

1 (function (modules) {
2   // The module cache
3   var installedModules = {};
4
5   // The require function
6   function __webpack_require__(moduleId) {
7     // Check if module is in cache
8     if (installedModules[moduleId]) {
9       return installedModules[moduleId].exports;
10    }
11    // Create a new module (and put it into the
12    // cache)
13    var module = (installedModules[moduleId] = {
14      exports: {},
15    });
16
17    // Execute the module function
18    modules[moduleId].call(
19      module.exports,
20      module,
21      __webpack_require__
22    );
23
24    // Return the exports of the module
25    return module.exports;
26  }
27
28  // Load entry module and return exports
29  return __webpack_require__(0);
30 }) ([
31   /* 0 */
32   function (module, exports, __webpack_require__) {
33     const hello = __webpack_require__(1);
34     console.log(hello.sayHello("Webpack"));
35   },
36   /* 1 */
37   function (module, exports) {
38     exports.sayHello = function (name) {
39       return "Hello, " + name + "!";
40     };
41   },
42 ]);

```

Listing 1: A non-minified example of a webpack bundle.

- The `__webpack_require__` function (Listing 1, L6–26) loads and executes modules and **manages exports and dependencies**.
- The bundle executes an **entry point module** (module 0 in Listing 1). The entry point module uses `__webpack_require__` to load functionality from other modules (Listing 1, L33).

Overall, bundlers allow developers to organize their code into modules, manage dependencies, and apply various optimizations like minification and code splitting, resulting in more efficient, maintainable, and faster-loading web applications.

2.2 Motivating Example

In this section, we present an example that shows how typical content-blocking approaches work and why they are ineffective against bundles.

2.2.1 Generic Code Inclusion. Consider the toy example presented on the left half of Figure 1. Here, the website loads two scripts, each of which defines global functions. The first script, `setup.js`, loads from the website’s domain itself and contains benign code relevant to the website’s functionality. The second script, `track.js`, loads from a known tracking domain, `tracker.com`, and returns a script that creates a unique identifier to track the user. When the website includes the two scripts from two separate `<script>` tags,

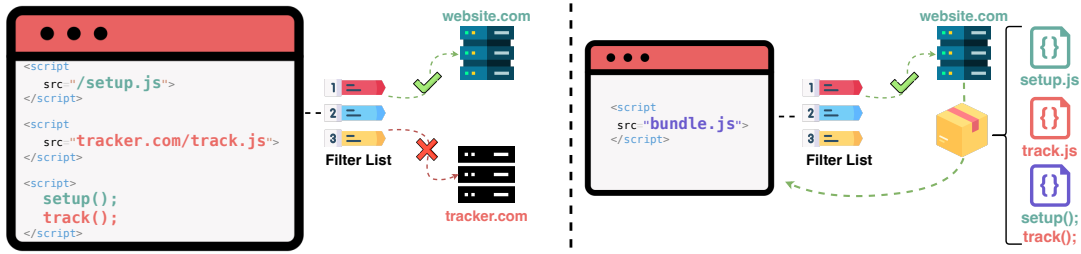


Figure 1: Motivating example.

it triggers two network requests, one to `website.com` and another to `tracker.com`.

2.2.2 Typical Content Blocking Approach. A typical approach to blocking privacy-harming scripts involves the use of a curated list of domains and regular expressions (for example, from EasyList, EasyPrivacy⁹). These lists are developed from manual contributions and include domains and paths to known privacy-harming resources. Content blocking tools (e.g., Adblock, uBlock) pull from these filter lists. The tools intercept outgoing network requests, compare them against entries in filter lists, and create interventions if they find a match. In the toy example, a filter list contributor might add the rule `||tracker.com/track.js`. Thereafter, when the website creates two network requests, the content-blocking tool permits a request to `script.js` but blocks the request to `tracker.com/track.js`. This way, the privacy-harming script is neither loaded nor executed in the user’s browsing session.

2.2.3 Limitations of Existing Content Blocking Approaches. Consider the scenario presented in the right half of Figure 1. Unlike the previous example, the website includes a single `<script>` tag that fetches code from its own server. The resulting network request does not have a corresponding entry in the filter list and is, therefore, not blocked by the content-blocking tool. The server responds with a bundled script that includes both benign code (`setup()`) and privacy-harming code (`track()`). Bundled resources expose multiple shortcomings of existing content-blocking approaches, which we briefly describe below.

- Privacy-harming code can be fetched from **multiple, variable resource paths**, including from first-party domains, making it impossible for contributors to manually detect and curate a comprehensive, non-exhaustive list of entries.
- Bundles include code from multiple resources. As a result, privacy-harming code is embedded along with necessary and benign functionality. Blocking the network request itself can break the website. Existing approaches must adopt an approach that targets **embedded resources** instead of the entire resource itself.
- Bundles also **mutate code included in modules** with wrappers and through various minification and obfuscation techniques. These mutations make it difficult to identify content with comparisons against regular expressions.

2.3 Properties of an Ideal Solution

We outline the properties of a general solution to identifying and replacing privacy-harming modules from bundled scripts.

First, a robust solution should be able to **identify and target specific modules**. Unlike typical content-blocking approaches, the solution cannot rely on domain-based blocking or attempt to replace the script entirely.

Second, as a corollary to the previous point, the solution should have **limited impact on benign functionality**, i.e., the solution should limit the effect of its intervention to privacy-harming code, leaving execution of other modules untouched. This also involves ensuring that any dependencies on the targeted module are handled in a way that limits side effects.

Third, the solution should be **generalizable across multiple dimensions**. It should apply to multiple privacy-harming libraries (and multiple versions of libraries) that interest content-blocking tools. Additionally, it should also be generalizable to different bundling strategies and robust against minification and obfuscation techniques.

Finally, the solution should have a **limited performance overhead**. While the solution executes in real-time, i.e., as scripts are loaded and executed in the browser, it needs to limit its effect on the usability of websites.

3 UNBUNDLE-REWRITE-REBUNDLE DESIGN

Unbundle-Rewrite-Rebundle (URR) adopts a static analysis approach that leverages the code structures of privacy-harming modules to identify and neutralize them when embedded in bundled scripts without disrupting the functionality of other application components. URR adopts a four-step process (see Figure 2):

- First, URR generates an Abstract Syntax Tree (AST) representation of a given script. It then analyzes the structure to identify if the script is a bundle that comprises one or more modules. If so, URR gathers the component modules for further analysis.
- Next, URR processes each module into an implementation agnostic representation, i.e., stripping variable names, function names, and object properties. It creates a bottom-up hash of the AST structure and uses the resulting representation for comparison.
- Next, URR compares each processed module against previously generated representations of privacy-harming modules. If URR finds a match, it marks the corresponding module for replacement.

⁹<https://easylist.to/>

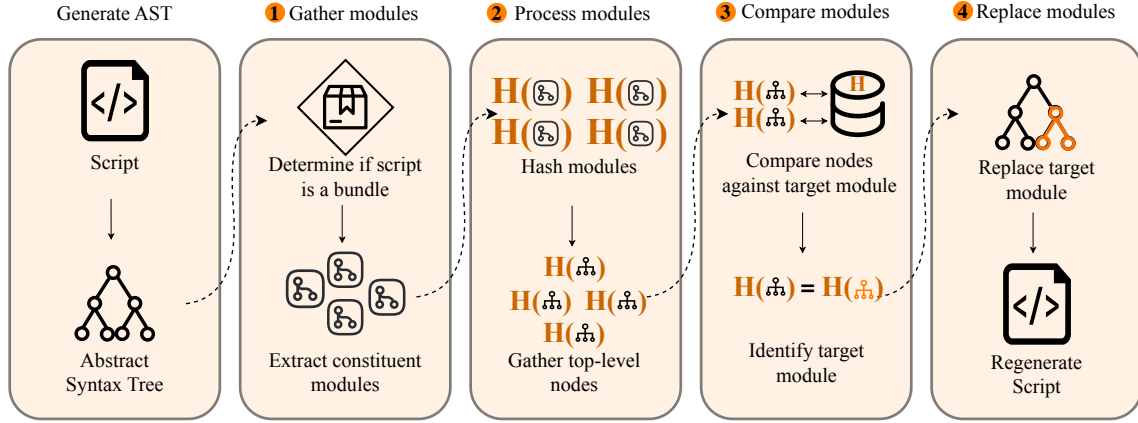


Figure 2: Overview of the framework.

► Finally, URR replaces each marked privacy-harming module with a corresponding benign replacement. In doing so, URR ensures that access and use of the replacement does not break other parts of website functionality. It then stitches the bundled script and moves it along for consumption and execution.

All the steps mentioned above are performed in real-time, i.e. when resources are loaded in the user's browser. However, processed representations of target modules and benign replacements are created and gathered offline.

3.1 Gathering Modules

URR first generates an Abstract Syntax Tree (AST) representation of a JavaScript resource. The AST representation helps gather the syntactic features of the script and provides insight into the structure of the code.

Thereafter, URR uses the AST representation to determine two aspects of the loaded script. First, URR determines if the script comprises a bundle with one or more modules. Second, URR gathers the sub-trees corresponding to each identified module.

Below, we describe the webpack-specific implementation.

3.1.1 Gathering JavaScript resources. To create a valid parser for bundles in the wild, we gathered examples of resources loaded on popular websites. We developed a puppeteer-based crawler that, upon visiting a page, intercepts network requests and stores a copy of observed script responses. We visited domains from the Tranco list [15] and gathered 30,930 scripts from 1,063 sites (1K crawl).

3.1.2 Generating ASTs. For each of the gathered scripts, URR uses acorn, a community-developed, open-source JavaScript parser to generate an AST¹⁰. The generated AST complies with the ESTree Spec to ensure a consistent, standardized representation that can be reproduced by alternative implementations¹¹.

3.1.3 Code Development and Refinement Methodology. We began with an understanding of the general structure of the output of a webpack bundle (see Listing 1). We developed a script that parses an

AST and looks for a module array or object, i.e., an array or object that comprises functions. We used this initial logic and adopted a code optimization methodology based on ground truth gathered through expert manual annotation. We describe the process below.

- (1) First, we randomly sampled 100 scripts (without replacement) from the JavaScript resources gathered during the 1K site crawl.
- (2) We manually evaluated the plaintext script and the AST of each sampled resource. In doing so, we checked for the presence of webpack-specific code patterns. We annotated each resource with a boolean value indicating whether we determined the script as a webpack bundle (**Annotation 1**). Note that these manual checks (e.g., keywords) were only used for annotation and not incorporated into the automated approach. The manual checks included:
 - (a) identification of code pattern similar to webpack's specific function that handles dependencies (see Listing 1, L6-26);
 - (b) identification of code patterns specific to webpack chunks, i.e., files comprising webpack modules separate from the entry point bundle;
 - (c) identification of objects and arrays specific containing functions with parameters, exports, and return statements similar to webpack's function wrappers;
 - (d) keyword searches that indicate the use of webpack in the use of webpack in the creation of the resource.
- (3) For any resource annotated as a webpack bundle, we also annotated the resource with the number of component modules we identified (**Annotation 2**).
- (4) We repeated Steps 1-3 until at least we had annotated 100 scripts as webpack bundles.
- (5) We then executed our code to automatically analyze and annotate each sampled resource in a similar manner, i.e., (a) **Annotation 1**: whether the script is a bundle, and (b) **Annotation 2**: the number of component modules identified in the script.
- (6) We compared the code-generated annotations against the manual annotations and gathered the set of differences. For each incorrectly labeled script, we analyzed the code outputs and refined its logic. Refinements included addressing edge cases, handling multiple IIFEs in a single resource, etc.

¹⁰<https://github.com/acornjs/acorn>

¹¹<https://github.com/estree/estree>

Table 1: Refining code to identify bundled webpack modules within scripts.

	Round 1 (n = 288)			Round 2 (n = 432)			Round 3 (n = 614)		Spot Check (n=300)	
	Manual	Code Initial	Code Refined	Manual	Code Initial	Code Refined	Manual	Code	Manual	Code
Annotation 1: Webpack Bundle										
# Webpack Bundles	104	153	111	156	162	163	207	217	99	102
Precision		0.68	0.94		0.94	0.96		0.95		0.97
Recall		1	1		1	1		1		1
Annotation 2: Component Modules										
# Modules	6,857	7,847	6,857	9,816	9,156	9,090	12,889	12,913	5,701	5,690
Precision		0.87	0.99		0.99	0.99		0.998		0.998
Recall		1	1		0.93	1		1		0.996

Algorithm 1 Processing modules.

Input: AST $\leftarrow$ module AST
Output: hashedAST $\leftarrow$ processed representation of the AST
procedure HASHMODULE(*node*)
 $childrenHash \leftarrow 0$
 for each *child* in *node.children* **do**
 $childrenHash \leftarrow childrenHash + HASHMODULE(child)$
 end for
 return $hash(node.type + childrenHash)$
end procedure

- (7) We used the refined version of our code to annotate the same sample and noted any improvements. This refined code version is the initial version of the subsequent phase against new samples.
- (8) For each subsequent round, we repeated Steps 1-7, appending the existing sample with 50 manually annotated webpack resources (Step 4). We stopped refining our logic when we observed negligible improvement in the precision, recall, and accuracy of code-generated annotations between consecutive phases.

We repeated the process for three rounds. Table 1 presents the numbers and metrics for both annotations and the associated metrics for each round of refinement. At the end of the third round, our code identified a given script as a webpack bundle with 95% precision and gathered modules with 99.8% precision.

We used the resulting logic to annotate all scripts gathered in our 1K crawl. We annotated a total of 11,995 scripts as webpack bundles. Finally, we performed a spot check of the generated annotation. We randomly sampled and manually annotated 300 previously unsampled scripts, of which we identified 99 scripts as webpack bundles. Upon comparing and verifying the code-assigned annotation, we observed a 97% precision rate in identifying webpack bundles and a 99.8% precision rate in gathering component modules.

3.2 Processing Modules

Each module gathered from the previous phase comprises a subtree, i.e., a partial AST of the larger AST representation of the script. The module's AST contains information about the structure of the

underlying code and the associated names of variables, functions, and properties included within the script. Script attributes like variable names are extremely volatile and have limited use in reliably identifying target modules. URR, therefore, only considers (1) the *structure* of the AST, i.e., the parent-child relationships between the nodes that comprise the AST, and (2) the *type* of each AST node, i.e., attributes which comply with the ESTree Spec, and are hence limited to a deterministic set of values.

To distill and represent only these specific attributes from a module's AST, we adopt a version of a cryptographic concept used in integrity verification: Merkle trees [19]. This approach presents a novel part of our methodology. For example, consider a list of data blocks that need to be verified (e.g., transactions in a blockchain). We represent this list of blocks in a tree structure. Consider a tree where each leaf node represents a piece of data, and each non-leaf node is a cryptographic hash of its child nodes. These hashes propagate upwards, converging into a singular root hash known as the Merkle root. Any alteration in the foundational data triggers a modification in the root hash, instantly signaling tampering or modifications. However, if the root hash is verified, all of its children are also verified.

URR adopts a similar methodology, briefly presented in Algorithm 1. Given a module's AST, it performs the following processing steps.

- (1) URR begins traversing the module AST from the root node.
- (2) For every *child node*, it recursively calls the function to gather the *child node*'s hash. The hash of each child, in turn, results from the hashes of its children.
- (3) It sums the hashes of all *child nodes* of the root node.
- (4) To generate a hashed representation of the root node, URR concatenates the *type* of the node to the sum of hashes of its children and hashes the resulting string.
- (5) It returns the processed representation of the AST.

This processed representation now comprises an alternative tree representation with equal depth to its original counterpart. However, tree-based comparisons are complex and add a large performance overhead. Recall that while processing an AST, top-level nodes contain information about underlying children. Therefore, comparing a limiting comparison to a few high-level nodes of the

Algorithm 2 Comparing modules against target libraries.

Input:
 $TOP_LEVEL_NODES \leftarrow$ list of top-level nodes of module
 $DATABASE \leftarrow$ hashes of known libraries
 $CANDIDATE_LIBRARIES \leftarrow$ empty set
Output:
 $MATCH \leftarrow$ weight of the associated match
for each $nodeHash$ in TOP_LEVEL_NODES **do**
 $librariesWithThisNode = DATABASE[nodeHash]$
 for each $[library, weight]$ in $librariesWithThisNode$ **do**
 $CANDIDATE_LIBRARIES[library] += weight$
 end for
end for
 $bestMatch = \max(CANDIDATE_LIBRARIES)$
if $bestMatch = targetLibrary$ **then return** $CANDIDATE_LIBRARIES[targetLibrary]$
end if

processed module can provide insight into the similarity of the module's AST. To this end, URR gathers a list of top-level nodes. In other words, noting that webpack wraps each module within a function, we consider the computed hashes of each statement within the function. Say, for example, when bundled within webpack, the module comprises three statements: one variable declaration statement, one function definition (which further contains multiple nested statements), and one export statement. URR will process this module and gather hashes for three top-level nodes – owing to the Merkle tree-based approach, each of the hashes will be computed from the top-level statements' children.

Additionally, URR associates each entry in this list with a weight that represents the number of child nodes it represents, i.e., it weighs a node representing a complex function higher than a node representing a simple variable declaration. Consider the same example as above. Here, the initial declaration statement and the export statement will represent fewer nodes than the function definition (which has children within nested statements). As a result, the function declaration statement will have a higher weight, corresponding to the ratio of nodes it represents to the total nodes in the module's overall AST.

3.3 Comparing Modules

URR uses the list of top-level nodes to compare against a database of top-level nodes of processed module representations corresponding to targeted libraries. While this database is generated offline, processed modules for these libraries are also gathered in a similar manner to the logic presented in Algorithm 1. We provide details regarding the creation of this database in Section 4. Here, we focus our discussion on the comparison and identification of processed modules.

Algorithm 2 presents an overview of the comparison strategy. We present a brief description below.

- (1) URR traverses the list of hashes of the top-level nodes representing the module observed in the wild.

- (2) It looks up the hash of each node in the database and identifies every library with the same top-level node. It considers any such library as a *candidate library*.
- (3) It appends the node's weight for each *candidate library* to previous matches, if any. Therefore, the weights associated with a *candidate library* increase each time it includes a match.
- (4) Once all top-level nodes have been traversed, URR identifies the *candidate library* with the highest associated weight.
- (5) If the highest match corresponds to the target library, URR returns the weight of this match, thereby marking the module as a candidate for replacement.

3.4 Replacing Modules

URR replaces each marked privacy-harming module with a corresponding, benign replacement. It can perform this action in one of two ways: (1) URR can replace the module's AST with an alternative, benign AST and then regenerate the script from the modified AST; (2) Alternatively, URR can identify the string indices for the module within the textual representation of the script, and place the benign replacement between these indices. Regardless of its approach, URR ensures that access and use of the replacement does not break other parts of website functionality.

Replacements for targeted modules are manually created. Given the nuances of each target library and its eventual mutated representation in webpack bundles, we describe the three categories of replacements to consider within each module.

- **Global Variables.** If libraries expose global variables on the window object, these variables must be made available. Additionally, the types of such variables need to be retained.
- **Exports.** The replacement must ensure that values previously exported from the target module remain available. When other modules consume the module, it must be made accessible, even if replaced with a benign version.
- **Webpack-based Function Wrappers.** When webpack wraps modules in functions, it passes specific parameters included in the function signature. Additionally, values exported by modules are appended as properties to webpack objects. Replacements need to ensure that these parameters and exports work seamlessly.

Finally, each replacement needs to ensure the consistency of variable types, i.e., functions be replaced with functions, constants be replaced with constants, and so on. We further provide examples of specific replacements in Section 4.

4 EVALUATION

In this section, we evaluate the effectiveness of URR in identifying specific libraries in bundled scripts captured in the wild. We select three libraries to target and describe the process of gathering their module representations for comparison. We then use URR to identify target libraries in scripts gathered from a crawl of popular websites. Next, we develop a deployment of URR as a Firefox extension and evaluate the performance overhead introduced by its interventions. We conclude by describing the creation and evaluation of benign replacements.

4.1 Evaluation Dataset

We first describe the selection of example libraries that we use to identify in the wild. We describe the offline process of gathering representations for these libraries, which URR can then use to identify libraries in the wild.

4.1.1 Target Library Selection. We select three libraries for our evaluation—FingerprintJS¹², Sentry¹³, and Prebid¹⁴. Each of these libraries is included in filter lists¹⁵ used by popular content blocking tools^{16,17}. Additionally, all three libraries provide the option for developers to use their npm packages with bundled code.

FingerprintJS¹⁸ is a popular browser fingerprinting library that tracks and identifies users. When embedded within a website that a user visits, the library performs various client-side operations that query browser attributes and stores a unique identifier for the user. Besides its use in user authentication [17] and ad fraud prevention¹⁹, the library de-anonymizes sensitive user activity across sites. Since the domains that the library uses are blocked by popular content-blocking tools, FingerprintJS recommends that developers use their npm package or self-host a copy of their scripts²⁰.

Sentry²¹ is an analytics tool that offers libraries for performance and error monitoring. The library helps developers gather details about user interaction, DOM events, console logs, and network calls. The library lets developers decide how they use the information gathered from users and offers multiple integrations, including multiple third-party analytics services²². Since the CDNs that host Sentry code are blocked by popular content blockers, the library recommends that developers get around this restriction by bundling Sentry's npm package into their app²³.

Prebid²⁴ is a popular advertising library that developers can use to add header bidding to their applications. The library integrates with numerous advertising, analytics, and user-tracking libraries, providing support for bidding on targeted advertisements. Similar to previously discussed libraries, Prebid is included within filter lists and content-blocking tools. The library is open source and is available to be added to bundles with an npm package. Prebid provides specific instructions on integration with webpack²⁵. The library must be compiled with Babel²⁶, and specific plugins must be used when loaded with webpack. It provides specific instructions for webpack configurations, which gives us insight into its use when integrated by bundles in the wild.

4.1.2 Bundler Configurations. Web applications are developed in a large variety of environments, frameworks, and configurations

¹²<https://github.com/fingerprintjs/fingerprintjs>

¹³<https://github.com/getsentry/sentry-javascript>

¹⁴<https://github.com/prebid/Prebid.js>

¹⁵<https://easylist.to/>

¹⁶<https://github.com/gorhill/uBlock>

¹⁷<https://adblockplus.org/>

¹⁸<https://fingerprint.com/>

¹⁹<https://fingerprint.com/case-studies/>

²⁰https://github.com/fingerprintjs/fingerprintjs/blob/master/docs/evade_ad_blockers.md

²¹<https://sentry.io/>

²²<https://develop.sentry.dev/development/analytics/>

²³<https://docs.sentry.io/platforms/javascript/troubleshooting/#dealing-with-ad-blockers>

²⁴<https://prebid.org/>

²⁵<https://github.com/prebid/Prebid.js>

²⁶<https://babeljs.io/>

Table 2: An overview of the build options used to gather representations for the target library.

Build Options	Possible Values
Module Systems:	
↪ Dependency Statement	[import, require]
Webpack Optimizations:	
↪ usedExports	[true, false]
↪ concatenatedModules	[true, false]
Minifier Options:	
↪ passes	[0, 1, 2]
↪ arrows	[true, false]
↪ dropConsole	[true, false]
↪ unsafeCompress	[true, false]
↪ unsafeMethods	[true, false]
↪ unsafeUndefined	[true, false]
↪ unsafeArrow	[true, false]
↪ unsafeCompare	[true, false]
↪ typeofs	[true, false]
Browser Support:	
↪ ie8	[true, false]
↪ safari10	[true, false]

before they are compiled and shipped to production. As a result, when the same npm package is bundled within different applications, its modules will generate a wide range of AST structures. These variations make it difficult to gather a definitive AST representation for the package that will return perfect matches when compared against scripts found in the wild. However, the AST of the module corresponding to each library is complex and, hence, unique. Our intuition is that the AST of a given library will be closer to differently configured ASTs of the same library than to ASTs of other libraries.

To gather multiple options for module representations and help us get the closest match to a target library in the wild, we gather a list of popular build options to be considered for each library. Table 2 These options cover four aspects that we describe below.

- (1) **Module Systems.** Depending on the target npm package and the development environment of the web application, the library may be consumed as a CommonJS module or an ECMAScript module. We gather configurations for both module types.
- (2) **Webpack Optimizations.** When bundling code from multiple files, webpack keeps track of imported and exported values with the help of a dependency graph. Depending on the module system, the output environment, and the developer's configurations, it provides options to automatically identify and discard unused parts of code and to reduce the size of the final output. The use of these options can change how the module appears in the output bundle.
- (3) **Minifier Options.** Once webpack creates a bundle output, it uses Terser²⁷ to minify the code. Developers can provide additional options to ensure that the output works with their intended production environment and that their applications

²⁷<https://terser.org/>

work in all supported environments. We consider a list of compression options that alter the syntax of the module in the bundled output.

- (4) **Browser Support.** Developers can additionally specify support for legacy browsers. These options override other minification and compression to ensure that the output bundle is also compatible with legacy browsers.

We gathered all combinations of the build options and created a total of 24,576 build configurations that we then used to bundle each target library.

4.1.3 Creation of Target Module Representation Database. We gather various hashes for these libraries with the help of a large set of versions and different webpack configuration options, resulting in multiple AST representations for each target library. We briefly describe the steps below.

- (1) **Library Versions.** We gather a list of past versions released by the library on the npm registry. We download and install each version in a separate barebones application.
- (2) **Barebones Application.** For each version, we create a barebones Node.js application with a single JavaScript file. The file consumes the target library with either an `import` or a `require` statement.
- (3) **Build Options.** For each library version, we build multiple webpack bundles, one for each combination of build options.
- (4) **Target Module.** When building a bundle, webpack provides the option to gather relevant info about individual modules²⁸. We use this information to identify the file, name, and position of the module corresponding to the target library. Thereafter, we gather an AST representation of the output bundle and extract the sub-tree corresponding to the target module.
- (5) **Processing AST.** We process the extracted AST in a similar manner to the steps described in Section 3.2. We then gather a list of hashes for the top-level nodes from each bundle along with their corresponding weights.

4.1.4 Additional Library Configurations. In addition to the three target libraries, we gathered a list of 10k popular npm packages based on their download counts²⁹. From this list, we collected a random sample of 10 versions of each of the 1k libraries. For each sampled library version, we generated 100 webpack bundles each with a different, random build option. We extracted and processed their modules and gathered a weighted list of hashes corresponding to the top-level nodes of their ASTs.

4.2 Framework Effectiveness

In this section, we describe our evaluation of URR’s effectiveness in identifying target libraries within scripts captured in the wild.

4.2.1 JavaScript Resource Dataset. We used a puppeteer-based crawler that visited sites from the Tranco list [15] and gathered JavaScript resources through network interception. We previously used this dataset to create a bundle parser, described in Section 3.1. The dataset comprises 30,930 scripts from 1,063 sites, of which 11,995 scripts are bundled resources.

²⁸<https://webpack.js.org/api/stats/>

²⁹<https://github.com/anvaka/npmrank>

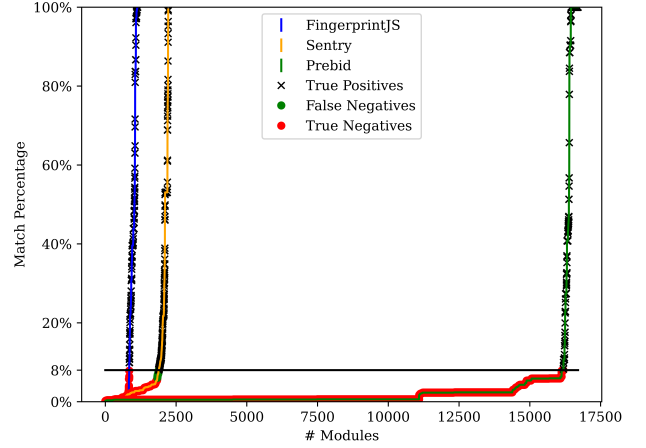


Figure 3: An overview of the target libraries identified from a crawl of the Tranco 10K. In addition to the 697 instances of true positive (x) matches above the 8% threshold, we observed >15K instances of true negative (•) matches of modules below the threshold.

4.2.2 Gathering Matches. For each bundle in the dataset, URR extracted and processed component modules (see Algorithm 1). It then gathered a weighted list of top-level nodes from the processed module. URR then compared each module against all libraries in the evaluation dataset (see Section 4.1). For each module, we noted the library in the evaluation dataset with the closest match. Finally, we gathered a list of modules for which the closest match was one of the three target libraries. We gathered a total of 73 matches for FingerprintJS, 324 matches for Sentry, and 3,532 matches for Prebid.

4.2.3 Manual Verification. For each match, we manually verified the code embedded between the corresponding indices in the script. We observed the code structure, the use of specific properties, and the number and types of exported variables, functions, and objects. We then manually annotated each match as a true positive or a false positive.

4.2.4 True Positive Threshold. We observed the tradeoff between precision and recall for the matches returned for all three target libraries. We looked for the lowest threshold match percentage for an AST for which all three libraries observed 100% precision, i.e., any AST that returned a positive match above this threshold was correctly annotated. We, therefore, arrived at a common match threshold value of 8%. Despite observing a small drop in recall for Sentry, we prioritized precision to ensure that URR never blocks benign modules even if it misses some privacy-harming scripts.

4.2.5 Evaluation on a larger crawl. Next, we performed a larger crawl of the web, gathering scripts from the top 10K sites in the Tranco list. We used URR to evaluate these scripts and employed the $\geq 8\%$ match threshold discussed earlier. We detected FingerprintJS on 205 sites, Sentry on 213 sites, and Prebid on 325 sites. Overall, URR identified at least one of the three target libraries on 7% ($n = 697$) of the top 10K sites. Figure 3 presents an overview of our observations. We note that the larger crawl has fewer matches than the Top-1K sites, which can possibly be explained by existing work,

Table 3: Time taken (ms) by the extension to process scripts using network interception.

Phase	# Scripts	Time (ms)		
		μ	σ	Median
Buffering script text	6,702	690.69	1,189.26	240
Gathering modules	6,699	46.17	98.25	9
Processing modules	2,135	326.31	635.58	114
Comparing modules	2,135	1.15	2.28	0

which states that (a) less popular websites are less complex [26], and (b) popular sites are more likely to engage in privacy-harming behaviors [23].

4.3 Performance

In this section, we provide a sample deployment of URR as a Firefox extension. We use this non-optimized deployment to evaluate an upper bound on the latency introduced by different phases of the pipeline and their effect on page load times.

4.3.1 Firefox Extension. Considering URR’s ability to recognize bundles given the entire contents of scripts, we determined that an approach that intercepts network requests and responses will be a useful initial deployment. Considering Chrome’s restrictions on content filtering from extensions³⁰, we developed a Firefox extension that intercepts and modifies the content of responses across all domains that the browser visits³¹.

4.3.2 Crawl. We evaluated the performance of a browser instance with and without the extension by visiting sites in the Tranco list [15]. For each web page, we performed the following steps.

- (1) First, we created and initialized a new browsing profile.
- (2) Next, we used Mozilla’s web-ext³² to load a Firefox instance with the web extension.
- (3) We used Puppeteer³³ to connect to the browser instance, visit the web page in a new tab and wait for 30 seconds.
- (4) The extension’s background script intercepts and evaluates all script-based network requests. It captures and stores the time taken by each step in the script evaluation pipeline.
- (5) The extension’s content script captures metrics relevant to web page performance.
- (6) We closed the browser instance and deleted the user profile directory and Firefox’s caches from the filesystem to ensure a fresh browsing state for subsequent visits.
- (7) We repeated steps 1-6 in a browsing profile without our extension installed and recorded relevant web page performance metrics for comparison.

4.3.3 Pipeline Performance. First, we discuss the evaluation of individual scripts. Table 3 shows the time taken by different phases

³⁰<https://developer.chrome.com/blog/improvements-to-content-filtering-in-manifest-v3/>

³¹<https://blog.mozilla.org/addons/2022/05/18/manifest-v3-in-firefox-recap-next-steps/>

³²<https://extensionworkshop.com/documentation/develop/getting-started-with-web-ext/>

³³<https://pptr.dev/>

Table 4: Time taken (ms) to load a page with and without the extension installed, from a crawl of $n = 963$ web pages.

Metric	w/o Extension			w/ Extension		
	μ	σ	Median	μ	σ	Median
First Contentful Paint	2,735.55	1,690.62	2,550	3,029.51	2,236.36	2,423
DOM Interactive	2,796.39	1,687.63	2,620	3,367.55	2,204.51	3,052
Page Load	4,790.39	3,156.72	4,130	7,107.45	4,759.42	6,078

Table 5: Memory space taken by the extension’s hash signatures to detect target libraries, comprising signatures of target libraries along with those of additional libraries (§4.1.4).

Signatures		#ASTs	#HashMap Entries	Size
Target	Additional			
All Hashes	All Hashes	258,788	265,962	3.55 GB
All Hashes	0.5 * (All Hashes)	176,963	142,173	2.77 GB
All Hashes	—	95,174	19,555	2.27 GB
Top-10K Matches	—	242	2,514	0.003 GB

of the pipeline. We observed the gathering modules and compared processed modules, which have the least impact, taking, on average, 46.17ms and 1.15ms, respectively. URR takes the longest time to process modules, i.e., recursively compute hashes for all nodes in an AST, which takes 326.31ms on average for each script.

However, we observed that the largest chunk of the time was consumed outside of URR’s evaluation and instead spent in buffering incoming response chunks to load the entire script. This is a limitation of the mode of deployment, i.e., Firefox extensions that modify the body of network responses bypass the browser’s optimized cache for scripts³⁴. The deployment waits for large network responses to complete before evaluating the script. In Section 5, we discuss alternative deployments that evaluate code at runtime and can bypass dependence on network load times.

4.3.4 Page Performance. Next, we discuss the effect of the sample deployment on page load. Recall that we used a content script to gather metrics from the page context for each visit with and without the extension. Table 4 presents a snapshot of our observations. We describe three collected measures below.

First Contentful Paint (FCP) is a timing measure that shows when the browser renders the first bit of content. The content could be any text, image, video, canvas, or non-empty SVG. The timing shows the first instance in which the user has an indication that the page is loading. We observed that the extension added 293.36ms to the average time a user would wait for such an indication.

DOM Interactive indicates the time taken for the Document Object Model (DOM) parser to finish its work on the main document, i.e., the time taken to construct the DOM. This time can be affected by the parser-blocking JavaScript. Note that the extension’s script evaluation runs outside the main thread. We observed that the extension added 571.66ms on average, i.e., a user would have to wait an additional 0.5s before the browser has parsed the DOM.

Page Load indicates the total time taken for all resources to load. It indicates that network requests for all scripts, images, and

³⁴<https://developer.mozilla.org/en-US/docs/Mozilla/Add-ons/WebExtensions/API/webRequest/filterResponseData>

other resources have been completed. This metric does not indicate actual end-user experience since it depends on device capabilities and various network conditions, but in our case, provides insight into the additional time added by the extension buffering responses and evaluating script contents. We observed that pages loaded with the extension took an additional 2,317.06ms on average before triggering the load event.

4.3.5 Signature Database Size. URR’s extension deployment uses a complete database of hash signatures, which includes those computed for target library configurations (based on Table 2) along with additional configurations of popular NPM libraries (see §4.1.4). To optimize the time taken to compare modules (§3.3), URR stores the signatures as a hashmap, with the key for each entry holding the hash of a *top-level* node and the value holding a dictionary comprising *root node* hashes along with corresponding weights (see Algorithm 2). We evaluated different database compositions to reduce its overall size, especially considering subsets of additional configs from popular libraries. Table 5 shows that even with the database solely comprising signatures of target libraries, it consumes 2.27GB of memory space (see Row 3 of Table 5). We observed that reducing the number of hashes included in the database reduced the extension’s use of memory space (from 3.55 GB to 2.27 GB), which can be beneficial to devices with limited resources.

We considered further reducing the database to only include target library signatures we found in our 10K crawl (see §4.2.5). While this subset greatly reduced the size of the database (3.54 MB), we noted that the majority of the ASTs (57%; $n = 139$) correspond to a single match. We comprised an extensive list of hash representations (based on Table 2) to make URR especially effective in detecting target libraries despite the websites adopting of a wide array of potential configurations. Using a match-based subset (as used in Row 4 of Table 5) will limit URR’s detection of libraries on websites that deploy target libraries within differently configured webpack bundles. We note that while greatly reducing target library hash representations may not be effective, developers can consider caching frequently encountered bundles. We further discuss caching in §5.

Finally, we repeated crawls of sites from the Tranco list (see §4.3.2), but this time, visited sites with each of the four database sizes deployed within URR. We observed that the different configurations had a negligible effect on page load, with the databases’ use in the *Comparing Modules* phase taking an average time of 1ms ($M = 0$ ms). Table 5 therefore focuses on the different configurations’ impact on memory usage.

4.4 Replacements

Prior work has extensively studied and determined the creation and use of non-breaking, benign replacements for privacy-harming code [28]. We follow similar principles and extend example replacements used by uBlock Origin³⁵.

4.4.1 Creating Benign Versions of Target Libraries. For each of the three libraries, we considered associated source code and documentation to create equivalent benign versions.

FingerprintJS (v3 and v4), when loaded as a library from its npm package, provides a function, `load()`, which returns a Promise that resolves to an object that can be used to `get()` a unique `visitorId` for the user. While the computation includes privacy-harming code, the benign replacement that we developed only returns a randomly generated `visitorId`. We used a similar approach to the existing replacement included within uBlock origin³⁶, but modified the same to a Node.js module. FingerprintJS v2 has developers `get()` a unique `visitorId` without first needing to `load()` the library. Our benign replacement exports values for all versions of FingerprintJS without needing to individually support 52 minor versions released across 3 major versions (v2, v3, and v4). Considering the small number of export replacements, manually creating a replacement for FingerprintJS only takes around 20 minutes.

Sentry provides a browser library (@sentry/browser). Below, we detail the development of replacements for the library.

- (1) *Enumerating Replacements.* Sentry’s browser library further imports related packages, which include utilities for tracing and profiling user action events and functionality to handle network requests, along with 24 functions (with void return values) from its core analytics library (@sentry/core). We used the official documentation as a guideline. Enumerating all imports, functions, and objects that needed replacements took us around 90 minutes of wall clock time.
- (2) *Functions.* We used a benign function that performed no operations (noop) as a stand-in for functions with no return values. We used the functions for the 24 functions imported from the core library, 3 functions imported to handle network requests, and one function each to support its additional “hub” extension and to initialize the library. The library additionally exposes a function, `close()` (used to flush pending events), that returns a Promise that resolves with a timeout. We retained this function as in the original library. Once listed, since all functions required straightforward replacements, this took around 20 minutes.
- (3) *Class Definition.* Sentry defines a `BrowserClient` class with a constructor that accepts any argument it is passed. Developing a replacement for this class only takes a couple of minutes.
- (4) *Global Variable.* Finally, Sentry exposes all its functionality by defining an object instance of its `BrowserClient` class and exposes all aforementioned functions as members of this class. The library makes the object available as a global variable (on the window object). Also, it exports the same functionality to the user’s code that imports it as an npm package. Defining the global variable and the exports takes under 5 minutes.

Updates between Sentry versions (v5-v8) reuse exports of the same name, ensuring that, unlike for FingerprintJS, our benign replacements did not need to accommodate additional changes. As a result, the benign replacement that we developed in around 2 hours by wall clock time replaces 335 minor versions of Sentry’s npm package.

Prebid provides a similar npm package to Sentry, i.e., its main functionality is included from an entry point module, while it imports additional functionality from related libraries. Our replacement focuses on the entry point module, i.e., `pbjs`, which, in turn,

³⁵https://github.com/gorhill/uBlock/tree/9123563895f0499849b4d85c4f95e1ed6ace2231/src/web_accessible_resources

³⁶https://github.com/gorhill/uBlock/blob/9123563895f0499849b4d85c4f95e1ed6ace2231/src/web_accessible_resources/fingerprint3.js

accesses other related functionality. The developer would need to refer to prebid's *publisher-facing APIs*³⁷ which exposes 45 exports that we individually replace. We developed 41 *function* replacements similar to those created for functions within Sentry (described earlier). Of these, 18 functions (`^get*()`) returned an object, which we replaced with an empty object response. The remaining 23 functions did not have a return value, which we replaced with an empty, noop function. We replaced the remaining 4 exports with empty arrays that developers can define and edit.

While Prebid's library itself has undergone numerous changes across versions (e.g., support for adapters, consent management, and user ID modules), its publisher-facing API has retained the same exports. We verified this against the library's release notes³⁸ and the library's source code. As a result, developing benign replacements for 41 functions and 4 arrays took only about an hour by wall clock time. Crucially, our replacement works against 186 minor library versions across 4 major Prebid version releases (v5-v9).

4.4.2 Limiting Breakage. We designed our replacements in a manner that would cause little breakage. Our replacements are expertly developed to maintain compatibility (uBO uses a similar approach for non-bundle replacements, e.g. FPJS³⁹). For each replacement module, we created a barebones application that accesses the library functionality. The barebones application helps evaluate each replacement, ensuring that the application remains functional. We package the resulting application and gather the webpack module corresponding to the benign replacement for use in the system's deployment. We evaluated our replacements by (i) ensuring that the JS execution of mock apps that bundle the library and the replacement were both successfully executed, (ii) our deployment returned metrics showing successful page load and DOM interaction events (Smith et al. [27] found network behavior highly predictive of breakage. Nisenoff et al. [22] found users consider page load and responsiveness the most prominent breakage category.) We ensured our replacements have the shape (scope and exports) we want and are effective (execution of mock apps, page load, DOM responsiveness in the real world) against extensive testing. We can add a manual breakage evaluation by sampling modified websites.

4.4.3 Manual Breakage Analysis. Core to understanding whether a website is *functional* or *broken* is the user's *subjective perspective* of the features that are needed within the context of the website. With this in mind, we adopted the methodology presented by Snyder et al. [29] to study the effect of URR's replacements on website functionality. We sampled 30 websites (10 sites per target library) that bundled target libraries and on which are extension deployed benign replacements (determined from the crawl in §4.2.5). We instructed three undergraduate workers to visit each website twice in a row. In the first visit, they were instructed to use the website in an unmodified Firefox browser. Considering that the average dwell time for a user on a website is under a minute (Chao et al. [18]), the workers were instructed to perform as many actions on the page as possible in one minute. On the second visit, they were instructed to visit and use the same website within a Firefox browser with

the extension enabled. We asked them to assign a score for each website: **1** if there was no perceptible difference between the two visits, **2** if the browsing experience was altered during the second visit, but they were still able to complete the same tasks as during the first visit, and **3** if they were unable to complete the same tasks as during the first visit.

We considered a website to be broken if the workers could not complete their intended task (i.e., the website was assigned a score of 3). We had the workers browse the same sites and record their scores independently and also had them include notes to provide context for their scores. We observed a high agreement ratio (96.67%) between the workers' scores, who reported that 29 of the 30 websites were functional with benign replacements. For the single site that the workers reported broken (<https://mlb.com>), we visited the site with a modified version of our extension that does not perform replacement. We observed that the site was broken even without the replacements, primarily owing to longer wait times in loading large scripts as fragments and the extensions' *blocking* network requests (see §4.3.3). Our observations were also in line with the workers' notes, which indicated that they considered the page broken because it takes "*too long to load*". We, therefore, concluded that the site was broken due to the extension's network interception and not the benign replacements.

We note that we did not collect or process private data or identifiable information from our evaluators, and the testing does not qualify as human subject research. Therefore, we did not seek an IRB review. This is consistent with prior work that used roughly the same evaluation setup [12, 13, 29].

4.4.4 Automated Breakage Analysis. We also performed an automated analysis by crawling websites from the Tranco Top-1K. We visited each site within a Firefox browser with and without deploying URR. We used a puppeteer script to capture all console errors and page errors that were observed during each visit. Visits with URR deployed observed an average of 5 errors ($\sigma = 15; M = 1$), while visits without URR observed a slightly higher but similar average of 5.66 errors ($\sigma = 14; M = 1$). We noted that visits to 84% sites ($n = 714/846$) with URR observed the same number of errors or fewer. Interestingly, we found more sites that produced fewer errors with deploying URR ($n = 152/18\%$) than sites that produced fewer errors without URR ($n = 132/15\%$). We further analyzed instances where deploying URR produced more errors and made two notable observations. First, a majority of these errors refer to (a) rejections in setting cookies, (b) CSP violations, and (c) Timeout errors (potentially resulting from waiting for network responses). Second, 63% ($n = 84/132$) of these instances involve a difference ≤ 2 errors between the two visits. We note that our analysis is limited and is only one of many insights into page breakage. With that caveat, considering that a large majority of websites respond similarly with and without deploying URR (overall average, #sites with $\leq$ errors, #sites with ≥ 2 errors), our observations show that URR retains web functionality and compatibility when deployed in the wild.

5 DISCUSSION

Applicability to Other Systems. In this work, we target and evaluate against JavaScript bundles generated by Webpack. We selected Webpack because it is the most popular bundling system

³⁷<https://docs.prebid.org/dev-docs/publisher-api-reference.html>

³⁸<https://docs.prebid.org/dev-docs/pb9-notes.html>

³⁹https://github.com/gorhill/uBlock/blob/9123563895f0499849b4d85c4f95e1ed6ace2231/src/web_accessible_resources/fingerprint3.js

on the Web. However, many other bundling approaches are used on the Web, some variations on the same approaches used in Webpack, some fundamentally different, and even working on different levels of the deployment process. We here briefly discuss these other bundling formats and how URR could be extended to apply to them (and with what difficulties).

Webpack is generally used to combine multiple different JavaScript libraries and code units and to process them into a single JavaScript file to make deployment easier and (in some cases) execution faster. While Webpack is the most common tool for this purpose, many others are used with different languages, build chains, testing frameworks, and, in some cases, bundling additional resource types beyond JavaScript files. Examples of these alternative bundlers include Browserify⁴⁰ and Gulp⁴¹, among many others.

URR could easily be extended to cover these other bundling tools, as at root they all operate in the same manner (i.e., consume multiple JavaScript files, preprocess them, and then generate code for the resulting combined AST). To do so would only require generating new signatures for ASTs of each target privacy-harming library generated by the bundler's preprocessing and rewriting phases and understanding the structure of each bundler combines the ASTs generated by each input library into the final, resulting code unit. This is work that would only need to be done once per bundler version and then could be shared across all URR clients.

Another approach to JavaScript bundling is directly combining each JavaScript code unit into a single archive and shipping the entire archive to the client alongside the website's initial HTML. This approach, exemplified by Google's WebBundles⁴² proposal, does not preprocess or otherwise modify the include JavaScript code units; included files are directly copied into the bundle archive. Cloudflare's Cloudflare's Managed Components⁴³ product can also be seen as a form of this kind of bundling, though instead of all combined into a single archive, they're instead delivered "on demand", as managed by an overriding "manager" application, making any URLs unpredictable. Extending URR to cover these kinds of bundled applications would be trivial since bundled JavaScript files are not modified in the bundling process. Identifying them within the bundled application is straightforward. Similarly, rewriting unwanted code is trivial since it only requires swapping the original file with the privacy-preserving alternative (either in the original archive or returned as a new subresource by the "manager" application).

Alternative Deployment Strategies. In this work, we implement URR as a Firefox extension. We choose a Firefox extension because only Firefox's extension API includes the ability for an extension to buffer and rewrite a fetched subresource⁴⁴ (like a JavaScript file), before the file is seen, parsed, and executed by the JavaScript engine.

However, the same approach we take in the extension could be identically deployed from other decision points, either to be more general across browsers (e.g., as part of a man-in-the-middle proxy

or more specific to particular browsers (e.g., as a modification to the V8 parsing pipeline). The majority of the resources needed for URR to work could apply equally across all of these intervention points (e.g., fingerprints of the ASTs of unwanted code, bundle "unbundling" logic, privacy-preserving replacement AST subtrees).

However, each of these intervention points would come with their own largely-predictable tradeoffs. Analyzing, unbundling, and rewriting JavaScript bundles as part of a man-in-the-middle proxy would work for any browser or network tool, though with a performance impact (since JavaScript engine optimizations like partial or lazy compilation would not be possible). Likewise, pushing URR logic into the JavaScript engine directly would likely allow for greater performance at the cost of development and maintenance cost, and require browser-specific implementations. That said, we reemphasize that most of the novel aspects of URR would be generic and shareable across all possible intervention points.

Diversity of Target Library Representation. URR requires a significant amount of precomputation to work effectively. Specifically, URR requires precomputing a signature for every AST for each library or code unit that should be rewritten at runtime. While this is a significant improvement over the existing state of the art ([28], for example, requires the precomputing each target *bundled application*, which will both be orders of magnitude larger in occurrence, but also enormously larger in terms of required disk space), it is still not trivial. The same target library can give different signatures depending on library version (e.g., FingerprintJS v2 vs v3), bundler version (e.g., Webpack v4 vs Webpack v3), bundler optimization strategies (e.g., "tree-shaking" or no), library integration method (e.g., CommonJS vs ECMAScript modules) among other dimensions.

As discussed in Section 3, URR uses several heuristics to effectively generalize signatures, to flatten the number of dimensions of signatures needed per target library (e.g., label stripping, AST simplification, etc.). Nevertheless, there is still a tradeoff between coverage—generating as many signatures as possible to correctly identify the same target code across a wide range of representations—and concision—minimizing the memory, matching time, and disk space used on each URR client at runtime.

Replacements. While other phases of URR are programmatically generated, we manually created benign replacements. Even with manual curation, we show in Section 4.4.1 that while benign replacements can take between 20 and 120 minutes to create, they can be scaled to replace a large number of versions of a target library (e.g., Sentry's replacement applies to >300 versions).

Regardless, this phase requires an expert understanding of both the specific bundler and the target library itself. While prior work has shown approaches that can automate this creation, we leave the adaption of a similar approach in the context of bundles as an avenue to explore in future work. Prior work like [28], which allows for the automatic creation of privacy-and-compatibility preserving versions of JavaScript libraries, could be leveraged to greatly expand the number of target libraries URR can identify and rewrite in bundled applications.

Caching Frequently Encountered Bundles. We designed URR to work against an extensive array of configuration options applied to bundles in the wild. Our deployment of URR makes an essential

⁴⁰<https://browserify.org/>

⁴¹<https://gulpjs.com/>

⁴²<https://webpack-wg.github.io/bundled-responses/draft-ietf-wpack-bundled-responses.html>

⁴³<https://managedcomponents.dev/>

⁴⁴<https://developer.mozilla.org/en-US/docs/Mozilla/Add-ons/WebExtensions/API/webRequest/filterResponseData>

tradeoff: its large database of target library signatures consumes notable memory space (2.27 GB - 3.55 GB). In Section 4.3.5, we discussed that reducing the database size can affect URR's detection of target libraries. However, developers can alternatively cache entire bundles that they frequently encounter. This way, while the browser extension would still block network requests, scripts do not need to be processed through URR on repeated encounters. For example, the service shopee.com has multiple related sites (shopee.com, shopee.co.id, shopee.co.th, shopee.tw), all of which load the same bundle (containing FPJS as one of the modules) from different URLs which would bypass filter lists. However, since the script's content is identical, a cache of this script can help perform detection and replacement without needing to process the script into component modules.

6 RELATED WORK

Filter Lists. This research is closely tied to a broad domain of research investigating the advantages, effectiveness, and responses to filter list-based content blocking. Note that existing filter lists (e.g., EasyList, EasyPrivacy) are developed manually and community-maintained. Merzdovnik et al. [20] performed a large-scale study that highlighted the effectiveness of filter list-based browser extensions and also indicated that these tools often lead to a *decrease* in overall CPU usage, even when considering their own overhead. Gervais et al. [11] quantified the privacy gain from ad-blockers and discovered that these tools can reduce interactions with third-party entities by up to 40% with default settings. However, other studies have pointed out significant inefficiencies in filter lists. Snyder et al. [30] highlighted the abundance of “dead-weight” rules that offer no discernible benefits in popular lists. Similarly, Alrizah et al. [1] found that popular lists contain a large number of false positives, which can take two or three months to be discovered.

Automated Content Blocking Approaches. Numerous studies have developed approaches to either help automatically generate filter list rules or develop alternative methods to block privacy-harming content. AdGraph [13] created a graph-based machine learning approach that used features like URL length and origin to differentiate between benign and privacy-harming resources. Bhagavatula et al. [4] also used URL-based features to train a machine-learning model for resource filtering. Chen et al. [5] used filter lists as ground truth and developed behavioral signatures based on the JavaScript event loop, while Sun et al. [31] classified JavaScript execution based on Web API calls. Le et al. [14] developed a reinforcement learning framework that generates filter list rules specific to a site of interest. They showed that their approach was comparable in visual breakage to manually-created filter lists.

JavaScript Analysis. Prior work that performs static or dynamic analysis on JavaScript to create “pre-filters” [7] for malicious scripts, detect malicious scripts that camouflage as benign scripts [6], and detect scripts that evade detection by adopting various obfuscation strategies [8]. In their analysis of the Top 10K sites in the Alexa list, Moog et al. [21] found that 90% sites contained a minified or obfuscated script. Fouquet et al. [9] presented a static analysis approach to webpack bundles but created signatures based on serializing their AST and hashing the result (compared

to the complex, Merkle tree-based approach) we present. Their work does not consider the effect of various bundling and minification options, which our work highlights as highly importance in ensuring the effectiveness of URR as a real-world deployment. Amjad et al. [3] employed stack traces (using dynamic analysis) to identify functions responsible for tracking within scripts. Their work addresses tracking in JavaScript but is not robust against bundles, which include nested functions and commonly introduce code transformations and stack information changes. Importantly, their proposal addresses “*tracking behavior*”, while URR identified specific libraries. Since our URR is not similarly limited in its ability to identify specific libraries, the benefits provided by the two approaches can be combined. URR can be incorporated as a first step to help narrow the scope of privacy-harming code to a small part of a large, bundled resource. Thereafter, [3]'s tool can identify specific functions within the library that perform privacy-harming actions, further increasing the granularity of detection and attribution.

More relevant to our work, Rack and Staicu [25] presented a method for detecting and partially reverse engineering bundles. Importantly, [25] overlaps with the *Gather Modules* phase (§ 3.1) for which we developed our own approach. To elaborate on the differences between the two approaches, we must divide the *Gathering Modules* phase into two steps. First, [25] determines if a script is a bundle by checking for specific text. Their use of keywords and consideration of a small set of configurations will have limited effect against bundle configurations in the wild. Our approach instead considers *code structure* and can handle code transformations that would frustrate [25]. Finally, they did not evaluate their approach's precision, recall, or accuracy against scripts gathered in the wild (see Table 1). Second, after identifying bundles, [25] relies on source maps and unmodified module names to parse bundles and break them down into component modules, restricting their approach to only work on 10% of real-world bundles. We instead present an automated approach (see § 3.2 and § 3.3) that can reverse engineer bundles regardless of the accidental availability of source maps. Our approach does not rely on either of these developer choices and works on a broad range of real-world sites.

7 CONCLUSION

Content blocking plays a crucial role in safeguarding privacy, enhancing performance, and preserving user autonomy online. However, the increasing use of bundlers presents a challenge – websites often intermingle tracking code with benign code within a single script, rendering traditional URL-based content blockers ineffective. We present a framework, Unbundle-Rewrite-Rebundle (URR), that detects bundles and reverse engineers them back to constituent modules. We develop an approach to identify the privacy-harming modules that we replace with benign alternatives. We demonstrate the effectiveness of our system in identifying Webpack bundles and further develop signatures for a fingerprinting library (FingerprintJS), an advertising library (Prebid), and an analytics library (Sentry). Our implementation can identify the bundled versions of these libraries in the wild via a similarity threshold that minimizes false negatives and prevents false positives within our training data. Leveraging our approach, we found the use of these libraries within bundled scripts on 697 sites of the Tranco 10K. Further,

we implemented a prototype deployment of URR as a browser extension and observed that the extension only adds 0.5s to the DOM Interaction Time on a page visit. URR can expand existing content-blocking approaches to combat the use of bundlers to hide privacy-harming code. It can be further adapted to combat alternative bundling strategies and can be further used to detect the use of other privacy-harming libraries to bolster web privacy protections.

ACKNOWLEDGEMENTS

We thank Victor Escudero, Marek Cwiek, and Zaheer Safi for performing the manual website analysis. This material is based upon work supported by the National Science Foundation (CNS-2247515).

REFERENCES

- [1] Mshabab Alrizah, Sencun Zhu, Xinyu Xing, and Gang Wang. 2019. Errors, Misunderstandings, and Attacks: Analyzing the Crowdsourcing Process of Ad-blocking Systems. In *Proceedings of the Internet Measurement Conference* (Amsterdam, Netherlands) (IMC '19). Association for Computing Machinery, New York, NY, USA, 230–244. <https://doi.org/10.1145/3355369.3355588>
- [2] Abdul Haddi Amjad, Dania Saleem, Muhammad Ali Gulzar, Zubair Shafiq, and Fareed Zaffar. 2021. TrackerSift: untangling mixed tracking and functional web resources. In *Proceedings of the 21st ACM Internet Measurement Conference* (Virtual Event). Association for Computing Machinery, New York, NY, USA, 569–576. <https://doi.org/10.1145/3487552.3487855>
- [3] Abdul Haddi Amjad, Zubair Shafiq, and Muhammad Ali Gulzar. 2023. Blocking JavaScript Without Breaking the Web: An Empirical Investigation. In *Proceedings on Privacy Enhancing Technologies*. <https://doi.org/10.56553/popets-2023-0087>
- [4] Sruti Bhagavatula, Christopher Dunn, Chris Kanich, Minaxi Gupta, and Brian Ziebart. 2014. Leveraging Machine Learning to Improve Unwanted Resource Filtering. Association for Computing Machinery, New York, NY, USA, 95–102. <https://doi.org/10.1145/2666652.2666662>
- [5] Quan Chen, Peter Snyder, Ben Livshits, and Alexandros Kapravelos. 2021. Detecting Filter List Evasion with Event-Loop-Turn Granularity JavaScript Signatures. In *IEEE Symposium on Security and Privacy (SP)*. 1715–1729. <https://doi.org/10.1109/SP40001.2021.00007>
- [6] Aurore Fass, Michael Backes, and Ben Stock. 2019. HideNoSeek: Camouflaging Malicious JavaScript in Benign ASTs. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security* (London, United Kingdom). Association for Computing Machinery, New York, NY, USA, 1899–1913. <https://doi.org/10.1145/3319535.3345656>
- [7] Aurore Fass, Michael Backes, and Ben Stock. 2019. JStap: a static pre-filter for malicious JavaScript detection (ACSAC '19). Association for Computing Machinery, New York, NY, USA, 257–269. <https://doi.org/10.1145/3359789.3359813>
- [8] Aurore Fass, Robert P. Krawczyk, Michael Backes, and Ben Stock. 2018. JASt: Fully Syntactic Detection of Malicious (Obfuscated) JavaScript. In *Proceedings of the International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*.
- [9] Romain Fouquet. 2023. *Improving Web User Privacy Through Content Blocking*. Theses. Université de Lille. <https://theses.hal.science/tel-04123409>
- [10] Kiran Garimella, Orestis Kostakis, and Michael Mathioudakis. 2017. Ad-blocking: A study on performance, privacy and counter-measures. In *Proceedings of the ACM on Web Science Conference*. 259–262.
- [11] Arthur Gervais, Alexandros Filios, Vincent Lenders, and Srdjan Capkun. 2017. Quantifying web adblocker privacy. In *European Symposium on Research in Computer Security*. Springer, 21–42.
- [12] Umar Iqbal, Steven Englehardt, and Zubair Shafiq. 2021. Fingerprinting the Fingerprinters: Learning to Detect Browser Fingerprinting Behaviors. In *2021 IEEE Symposium on Security and Privacy (SP)*. 1143–1161. <https://doi.org/10.1109/SP40001.2021.00017>
- [13] Umar Iqbal, Peter Snyder, Shitong Zhu, Benjamin Livshits, Zhiyun Qian, and Zubair Shafiq. 2020. AdGraph: A Graph-Based Approach to Ad and Tracker Blocking. In *2020 IEEE Symposium on Security and Privacy*. 763–776. <https://doi.org/10.1109/SP40000.2020.00005>
- [14] Hieu Le, Salma Elmalaki, Athina Markopoulou, and Zubair Shafiq. 2023. AutoFR: Automated Filter Rule Generation for Adblocking. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 7535–7552. <https://www.usenix.org/conference/usenixsecurity23/presentation/le>
- [15] Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Koczyński, and Wouter Joosen. 2019. Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. In *Proceedings of the 26th Annual Network and Distributed System Security Symposium (NDSS 2019)*. <https://doi.org/10.14722/ndss.2019.23386>
- [16] Zhou Li, Kehuan Zhang, Yinglian Xie, Fang Yu, and Xiaofeng Wang. 2012. Knowing your enemy: understanding and detecting malicious web advertising. In *Proceedings of the ACM conference on Computer and communications security*. 674–686.
- [17] Xu Lin, Panagiotis Ilia, Saumya Solanki, and Jason Polakis. 2022. Phish in Sheep's Clothing: Exploring the Authentication Pitfalls of Browser Fingerprinting. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, Boston, MA, 1651–1668. <https://www.usenix.org/conference/usenixsecurity22/presentation/lin-xu>
- [18] Chao Liu, Ryan W. White, and Susan Dumais. 2010. Understanding web browsing behaviors through Weibull analysis of dwell time. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (Geneva, Switzerland) (SIGIR '10). Association for Computing Machinery, New York, NY, USA, 379–386. <https://doi.org/10.1145/1835449.1835513>
- [19] Ralph C Merkle. 1987. A digital signature based on a conventional encryption function. In *Conference on the theory and application of cryptographic techniques*. Springer, 369–378.
- [20] Georg Merzdovnik, Markus Huber, Damjan Buhov, Nick Nikiforakis, Sebastian Neuner, Martin Schmiedecker, and Edgar Weippl. 2017. Block me if you can: A large-scale study of tracker-blocking tools. In *2017 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE, 319–333.
- [21] Marvin Moog, Markus Demmel, Michael Backes, and Aurore Fass. 2021. Statically Detecting JavaScript Obfuscation and Minification Techniques in the Wild. In *2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*. 569–580. <https://doi.org/10.1109/DSN48987.2021.00065>
- [22] Alexandra Nisenoff, Arthur Borem, Madison Pickering, Grant Nakanishi, Maya Thumpasery, and Blase Ur. 2023. Defining "Broken": User Experiences and Remediation Tactics When Ad-Blocking or Tracking-Protection Tools Break a Website's User Experience. In *32nd USENIX Security Symposium*. USENIX Association, Anaheim, CA, 3619–3636. <https://www.usenix.org/conference/usenixsecurity23/presentation/nisenoff-broken>
- [23] Emmanouil Papadogiannakis, Panagiotis Papadopoulos, Nicolas Kourtellis, and Evangelos P. Markatos. 2021. User Tracking in the Post-cookie Era: How Websites Bypass GDPR Consent to Track Users. In *Proceedings of the Web Conference (Ljubljana, Slovenia)*. Association for Computing Machinery, New York, NY, USA, 2130–2141. <https://doi.org/10.1145/3442381.3450056>
- [24] Enric Pujol, Oliver Hohlfeld, and Anja Feldmann. 2015. Annoyed users: Ads and ad-block usage in the wild. In *Proceedings of the Internet Measurement Conference*. 93–106.
- [25] Jeremy Rack and Cristian-Alexandru Staicu. 2023. Jack-in-the-box: An Empirical Study of JavaScript Bundling on the Web and its Security Implications. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security* (Copenhagen, Denmark). Association for Computing Machinery, New York, NY, USA, 3198–3212. <https://doi.org/10.1145/3576915.3623140>
- [26] Kimberly Ruth, Aurore Fass, Jonathan Azose, Mark Pearson, Emma Thomas, Caitlin Sadowski, and Zakir Durumeric. 2022. A world wide view of browsing the world wide web. In *Proceedings of the 22nd ACM Internet Measurement Conference* (Nice, France). Association for Computing Machinery, New York, NY, USA, 317–336. <https://doi.org/10.1145/3517745.3561418>
- [27] Michael Smith, Peter Snyder, Moritz Haller, Benjamin Livshits, Deian Stefan, and Hamed Haddadi. 2022. Blocked or broken? Automatically detecting when privacy interventions break websites. *arXiv preprint arXiv:2203.03528* (2022).
- [28] Michael Smith, Pete Snyder, Benjamin Livshits, and Deian Stefan. 2021. SugarCoat: Programmatically Generating Privacy-Preserving, Web-Compatible Resource Replacements for Content Blocking. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security* (Virtual Event, Republic of Korea). Association for Computing Machinery, New York, NY, USA, 2844–2857. <https://doi.org/10.1145/3460120.3484578>
- [29] Peter Snyder, Cynthia Taylor, and Chris Kanich. 2017. Most Websites Don't Need to Vibrate: A Cost-Benefit Approach to Improving Browser Security. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. Association for Computing Machinery, New York, NY, USA, 179–194. <https://doi.org/10.1145/3133956.3133966>
- [30] Peter Snyder, Antoine Vastel, and Ben Livshits. 2020. Who Filters the Filters: Understanding the Growth, Usefulness and Efficiency of Crowdsourced Ad Blocking. In *Abstracts of the 2020 SIGMETRICS/Performance Joint International Conference on Measurement and Modeling of Computer Systems* (Boston, MA, USA) (SIGMETRICS '20). Association for Computing Machinery, New York, NY, USA, 75–76. <https://doi.org/10.1145/3393691.3394228>
- [31] Jingxue Sun, Zhiqiu Huang, Ting Yang, Wengjie Wang, and Yuqing Zhang. 2021. A system for detecting third-party tracking through the combination of dynamic analysis and static analysis. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops*. 1–6. <https://doi.org/10.1109/INFOCOMWKSHPS51825.2021.9484564>
- [32] Apostolis Zarras, Alexandros Kapravelos, Gianluca Stringhini, Thorsten Holz, Christopher Kruegel, and Giovanni Vigna. 2014. The dark alleys of madison avenue: Understanding malicious advertisements. In *Proceedings of the Internet Measurement Conference*. 373–380.