

MyWebGuard: Toward a User-Oriented Tool for Security and Privacy Protection on the Web

Panchakshari N. Hiremath*, Jack Armentrout*, Son Vu, Tu N. Nguyen, Quang Tran Minh, and Phu H. Phung*

* Intelligent System Security Lab
Department of Computer Science, University of Dayton

<https://issec-lab-udayton.github.io/>

The paper is published in the 6th International Conference on Future Data and Security Engineering 2019 (FDSE'2019), Nha Trang City, Vietnam, November 27-29, 2019

Link: <https://bit.ly/2VsEAhM>

Privacy on the web- Tracker Example

The screenshot shows an Amazon search results page for 'gopro'. The search bar at the top contains 'gopro' and the page title is 'Stock up for college'. The search results are sorted by 'Featured' and show 1-16 of over 1,000 results. The top results are sponsored by GoPro and include:

- GoPro HERO 7 Black (4.5 stars, 850 reviews)
- GoPro HERO 7 Silver (4.5 stars, 851 reviews)
- GoPro HERO7 White - Waterproof Digital Action Camera with Touch Screen 1080p HD... (4.5 stars, 851 reviews)

Below these are two sponsored bundles:

- GoPro HERO7 Black Camera + Extra Rechargeable Battery + PNY Elite-X 32GB U3 microSDHC Card (Bundle) - Waterproof Digital Action Camera Touch Screen 4K HD Video 12MP Photos Live Streaming... (4.5 stars, 45 reviews, \$399.99)
- AKASO Brave 4 4K 20MP WIFI Action Camera Ultra HD with EIS 30m Underwater Waterproof Camera Remote Control 5X Zoom Underwater Camcorder with 2 Batteries and Helmet Accessories Kit (4.5 stars, 2,280 reviews, \$79.99)

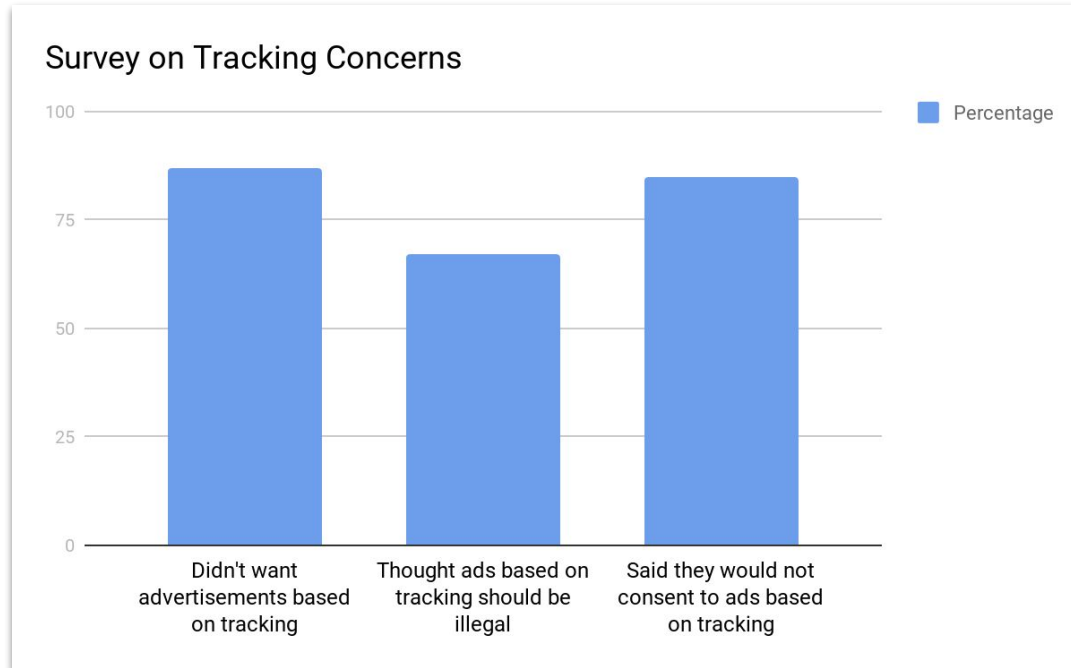
At the bottom, there is a result for GoPro HERO7 White — Waterproof Digital Action Camera with Touch Screen 1080p HD Video 10MP (4.5 stars, 851 reviews).

The screenshot shows a Facebook post from GoPro, dated 'about 3 months ago'. The post text reads: 'Got a HERO4? Trade it in for \$100 off a new HERO5 Black.' Below the text is a large image of a GoPro HERO4 camera with a 'TRADEUP' logo overlaid. The logo features the GoPro logo and the word 'TRADEUP' in large, bold, blue letters with an upward-pointing arrow. Below the image, the text reads: 'HERO5 Black', 'GoPro TradeUp', and 'SHOP.GOPRO.COM'. At the bottom of the post, there are engagement metrics: 121 likes, 50 comments, and 2 shares.

User Concerns

- Amount of malicious content on web?
 - Around 18.5 million websites are compromised at a given time each week
 - Average website attacked 44 times a day
- What are the options for a user?
 - Do Not Track- Websites do not have to honour
 - Blocker Extensions
 - Brave Browser

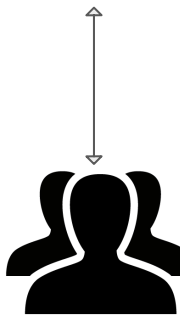
User Tracking Concerns



Existing solutions



Browsers



Web Users

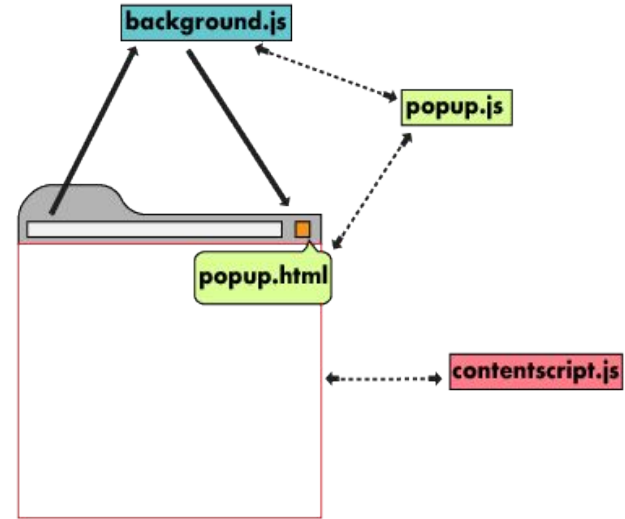
- There are more than 80 browsers in the industry, including popular ones such as Google Chrome, Apple Safari, Mozilla Firefox.
- Some of these are more committed to privacy than others such as Brave who disable third-party cookies to ensure user privacy, which in turn can limit usability.



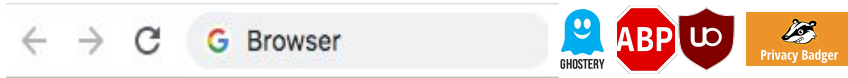
Is user safe from only browser?

Browser Extensions

- A browser extension is a small software module for customizing a web browser.
- Extensions are small software programs that customize the browsing experience. They enable users to tailor Chrome functionality and behavior to individual needs or preferences.
- They are built on web technologies such as HTML, JavaScript, and CSS.



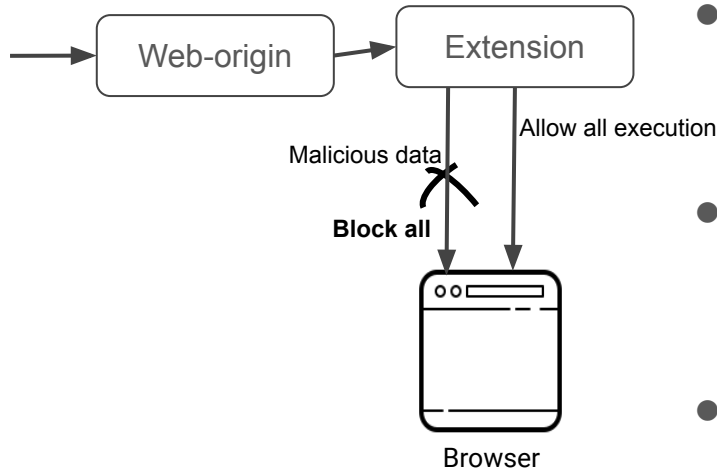
- There are many browser extensions for privacy and top extensions are, Ghostery, Ublock, Simple Blocker, and AdBlock.
- Still, can attackers gain access to user data?



Web Users

Is user safe from browser and extensions?

Limitations of existing solutions



Block all or do nothing.

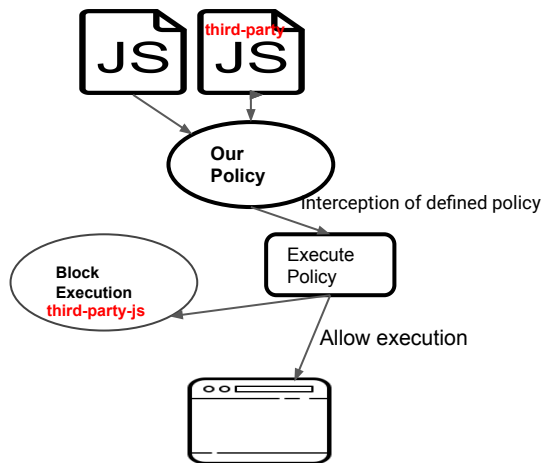
- Existing browser extensions only enforce block on all the attacks or allow there is no categorization.
- Extensions are predefined, so there are new trackers or other third-party content they are not blocked by these extensions.
- Popular security extensions like Ghostery and uBlock do not detect data leakage from sources and sinks.

Motivation

- Preserve user privacy on the web
 - Third-party trackers
 - Malicious injected webpage content
- Allow for an ethical middle ground for the collection of user data with the consent of the user
- User-centric
 - User defined control for third-party privacy accessibility
 - Allow user in real time to make privacy relevant decisions

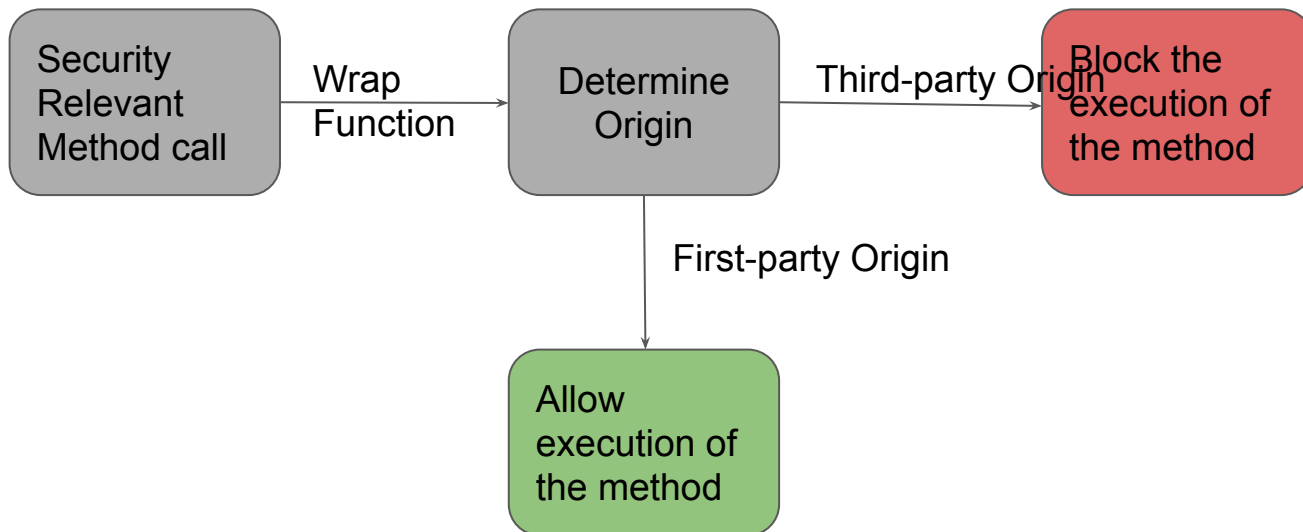
Our Approach

- Monitor JavaScript sources and sinks, distinguishing origin of code
 - Sources and Sinks: Where any JS API can execute
 - Utilize runtime stack to distinguish between first/third-party code
- Enforce policies on these channels to protect against privacy violations, based on code origin



How do we enforce a policy?

- When a monitored JavaScript API is called
 - Determine origin of code using runtime stack
 - Screen through relevant policy
 - Allow API to proceed or block the call in the case of a violation
 - Consult the user in the case of a privacy violation to allow for an override if requested



Policy Example: Read Cookie

If the determined origin is not allowed to read the cookie, then we block the request

```
Object.defineProperty(document, "cookie",
{get: function(){ //JavaScript attempts to read the cookie
  //Determine the origin of the code
    //if origin is allowed
      //return the original value
    //else
      //block the request

  },
  //...
});
```

Policy Example: Image Policies

Img sources are wrapped and passed through policies to protect against malicious content in img sources

```
var imgPolicy = {
  get: function(obj, prop) { /* policies for get */},
  set: function(obj, prop, value) { /* policies for set */}
};

//save the original image

class ImageWrapper {
  constructor(height, width) {
    //create image object from original image
    //pass image object through relevant policy
    //return sanitized image
  }
}

//replace the requested image with sanitized image
```

Interception of operation

The screenshot shows a web browser with the Forbes website loaded. The developer console is open, displaying a series of log messages from a script named 'mywebguard.js'. The messages include:

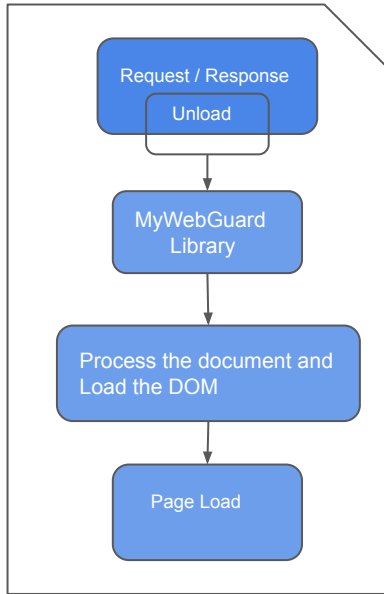
- log:starting mywebguard.js
- log:ImageObjectMonitor() is called
- log:WebSocketMonitor is called
- log:Cannot monitor chrome.webRequest: undefined!
- log:mywebguard.js is completely loaded
- log:document.createElement is monitored. elementType=script
- log:createElement is monitored. source=www.forbes.com/:15:1657
- log:document.createElement is monitored. elementType=script
- log:createElement is monitored. source=www.forbes.com/:15:1657
- log:document.createElement is monitored. elementType=img
- log:createElement is monitored. source=i.forbesimg.com/simple-site/dist/js/common-9c9b70b...js:10:5246)
- log:img type found
- log:document.createElement is monitored. elementType=fbs-ad
- log:createElement is monitored. source=i.forbesimg.com/simple-site/dist/js/homepage-26c1cab...js:1:43492)
- log:document.createElement is monitored. elementType=fbs-ad
- log:createElement is monitored. source=i.forbesimg.com/simple-site/dist/js/homepage-26c1cab...js:1:43492)
- log:document.createElement is monitored. elementType=fbs-ad
- log:createElement is monitored. source=i.forbesimg.com/simple-site/dist/js/homepage-26c1cab...js:1:43492)
- log:document.createElement is monitored. elementType=fbs-ad
- log:createElement is monitored. source=i.forbesimg.com/simple-site/dist/js/homepage-26c1cab...js:1:43492)
- log:document.createElement is monitored. elementType=fbs-ad
- log:createElement is monitored. source=i.forbesimg.com/simple-site/dist/js/homepage-26c1cab...js:1:43492)
- log:document.createElement is monitored. elementType=img
- log:createElement is monitored. source=i.forbesimg.com/simple-site/dist/js/homepage-26c1cab...js:1:43689)
- log:img type found
- log:getElementById is monitored. source=www.forbes.com/:84:379167)
- log:document.createElement is monitored. elementType=div
- log:createElement is monitored. source=www.forbes.com/:84:379212)

A
Billionaire's
Spouse
Donated
To Bernie
Sanders.
He's
Returning
The Check.

By Michela
Tindera Forbes Staff

"I don't
understand why
they would do
that," says Marta
Thoma Hall.
"That's ridiculous."

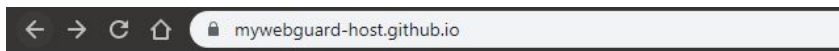
MyWebGuard Browser Extension Implementation



- We have developed a JavaScript library and deployed in the browser extension to self-protect the web users.
- Our interception library(CoreFlashJax) run first before a web page is loaded, we have implemented our JavaScript library code in innerHTML property so that when page loads it will be set as a first current page.
- We have implemented interception for data source access and data sink channels.

Evaluation Setup

- We set up on our host websites (first-party) with third-party JavaScript code implemented
 - We test our extension on the host website with simulated attacks from third party code



MyWebGuard evaluation

TODO steps:

1. Create some credentials, e.g., cookies
2. Use these credentials with JavaScript
3. Include some third-party code to perform some malicious actions

**Iframe Testing for
mywebguard**

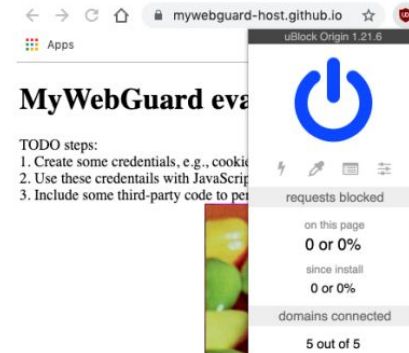
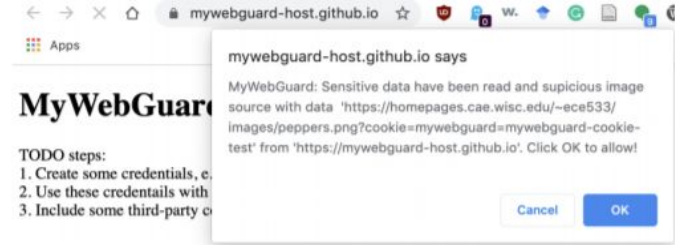
[MicroBenchmark](#)

```
<script  
src="https://mywebguard-thirdparty.github.io/script.js"  
></script>
```

Demonstration

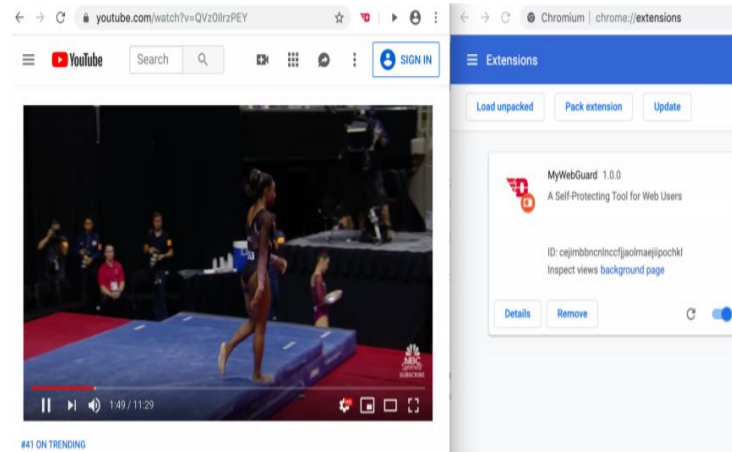
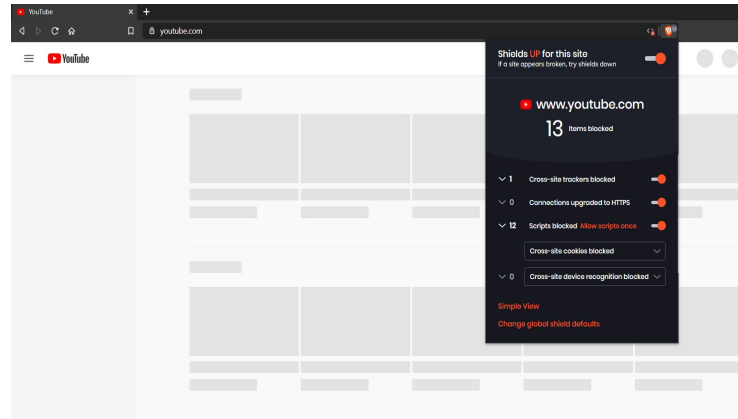
We notice that data leakage and tracking requests are caught by MyWebGuard and users would be notified

However, the simulated data leakage and tracking requests were ignored by uBlock, one of the popular browser privacy extensions



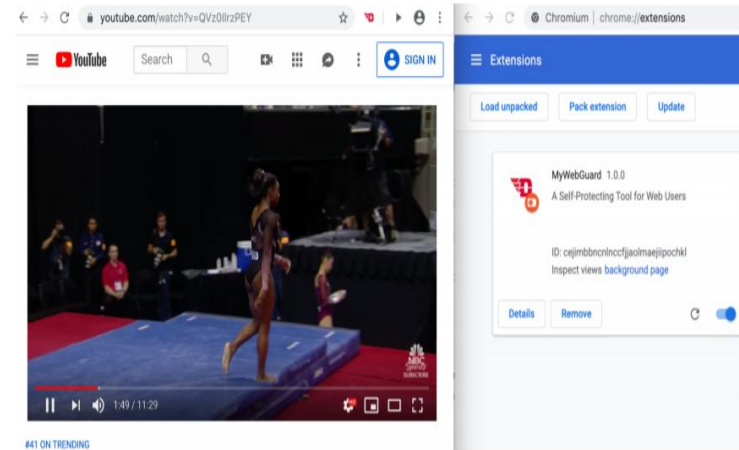
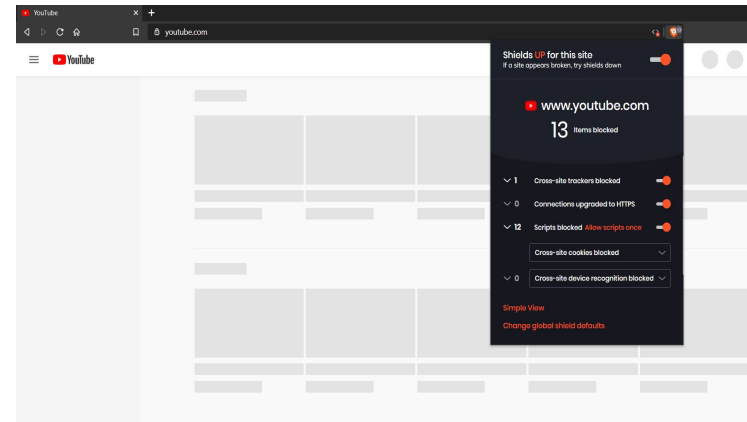
Functionality

- When examining Brave browser, which has a JavaScript-blocking mechanism, we test some websites
- We notice several breakages and loading issues (such as YouTube)



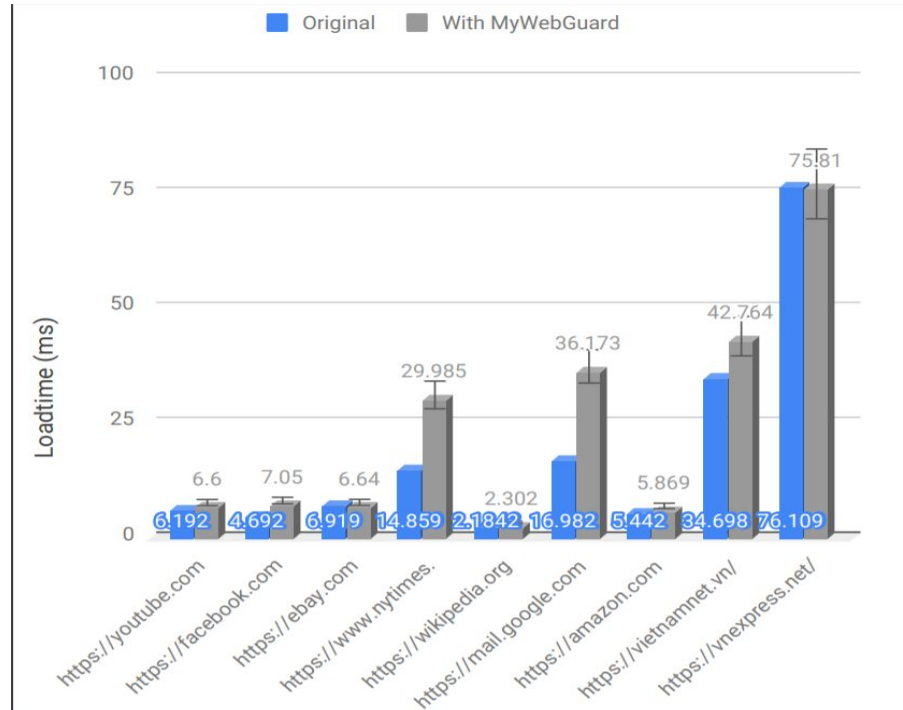
Functionality

- We do not notice any issues when loading those websites (like YouTube) with MyWebGuard



Runtime Evaluation

- We tested MyWebGuard with both Chromium and Brave browsers (on Ubuntu 18.04.2 LTS) on real websites
 - The overheads are not noticeable as shown in the graph



Contribution

- Browser-agnostic approach to preventing privacy leakage not monitored by contemporary solutions
- Present a compromise to “all-or-nothing” filter lists
- Advances conventional same-origin policy standard by enforcing different policies for each source of code
- Evaluation of approach overhead shows a lightweight yet effective implementation

Future Work

- Extend and refine policies and enforcement mechanism
 - Machine learning to produce practical policies
 - Effectiveness when built into a browser
- Allow for end-users to customize privacy preferences
- Perform large-scale evaluations of MyWebGuard
 - On top websites
 - Interference with co-existing browser extensions

Thank You

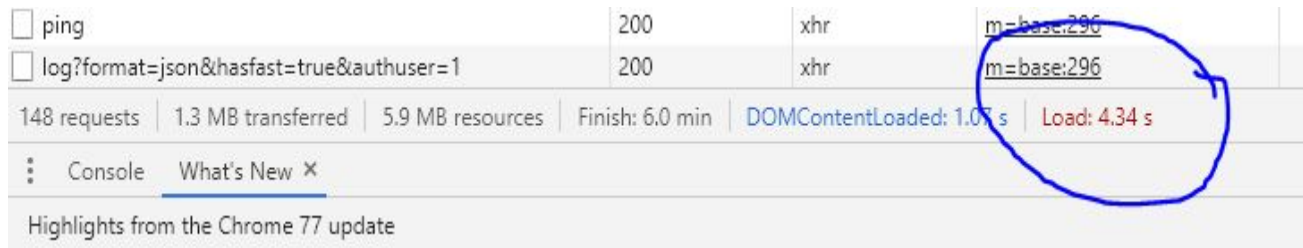
Intelligent System Security Lab
Department of Computer Science, University of Dayton

<https://issec-lab-udayton.github.io/>

Runtime Evaluation

We now test our extension on different sites (such as EBay, Amazon, FaceBook...)

To avoid anomalies, we test each site 10 times, recording the loading times (time for a site to finish loading its required contents)



<input type="checkbox"/> ping	200	xhr	m=base:296
<input type="checkbox"/> log?format=json&hasfast=true&authuser=1	200	xhr	m=base:296

148 requests | 1.3 MB transferred | 5.9 MB resources | Finish: 6.0 min | DOMContentLoaded: 1.07 s | Load: 4.34 s

⋮ Console What's New ✕

Highlights from the Chrome 77 update

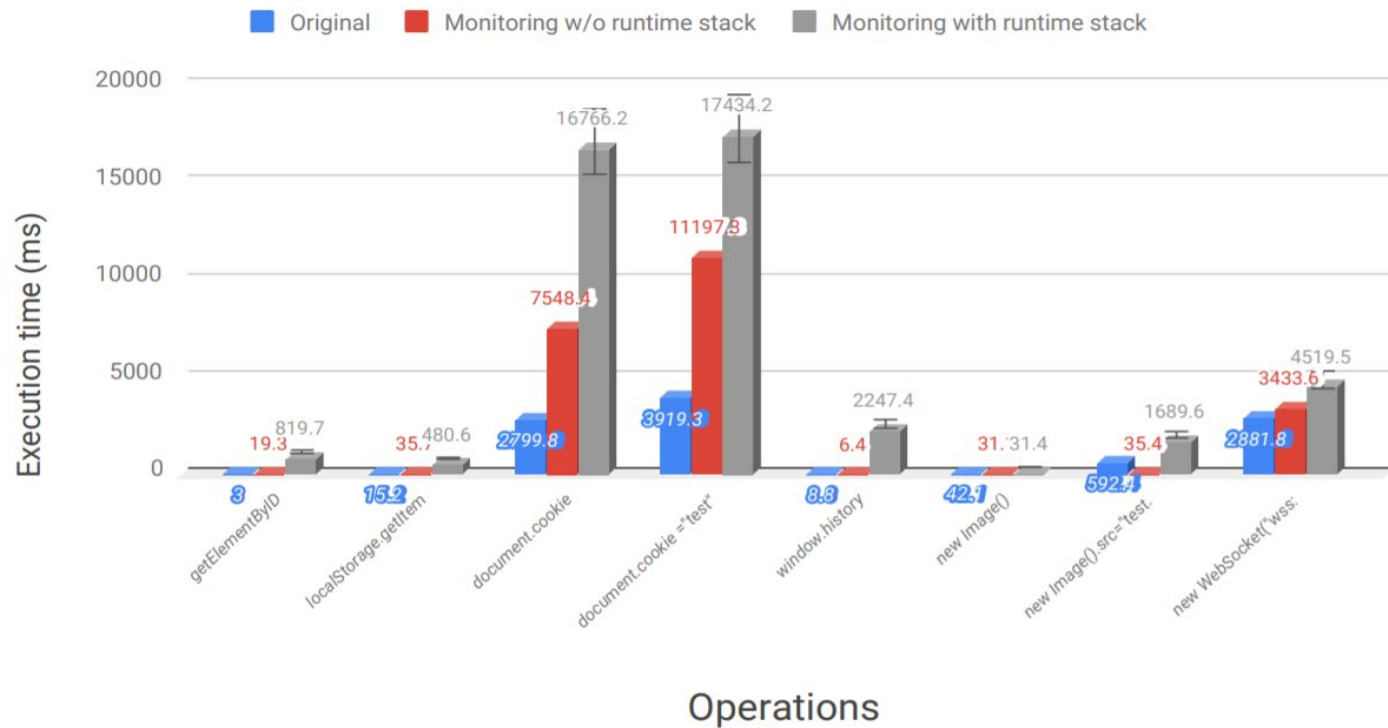
Runtime Evaluation

We tested MyWebGuard with both Chromium and Brave browsers (on Ubuntu 18.04.2 LTS) using our testing simulation

There exists a slowdown time on both browsers, which is due to the runtime stack that helps track the code origin



Performance overhead in Chromium



Performance overhead in Brave

