**Online Adblocker**

One Semester Individual Project

Final Report

CM3203

Author: **Oliver James Robert Storey-Young**

Student Number: **C1635943**

Supervisor: **Padraig Corcoran**

Moderator: **Federico Liberatore**

**Cardiff University Computer Science 2020**

# Abstract

When browsing the internet, it is all too common to be bombarded by adverts that will hinder and alter their browsing experience. Over the past few years, the number of adverts and unwanted traffic loaded when you visit a website has increased to a ridiculous amount. Adverts can often be unsafe, collect personal information and have even been used to influence political votes.

In this project I aim to create a reliable and effective adblocker that will be able to detect and circumvent adverts that may be displayed on any given website. Since there are already adblockers available I will spend time researching the techniques and algorithms that have previously been used to help identify adverts on a web page and begin to build my own method of detecting them.

The following report will outline and evaluate thoroughly the process I have gone through to complete this project.

# Acknowledgements

I would like to give a big thanks to all my friends and family who gave continued support and motivation throughout the duration of my project. I would also like to express my gratitude to Dr Padraig Corcoran for his continued guidance and help throughout the project.

Most of this project was completed during the lockdown brought on by the coronavirus outbreak of 2020. I would like to add further thanks to all the NHS staff, key and essential workers who have risked their lives to keep the country running during these times.

Contents

[Abstract 1](#_Toc40039784)

[Acknowledgements 2](#_Toc40039785)

[1 Introduction 5](#_Toc40039786)

[1.1 Project Overview 5](#_Toc40039787)

[1.2 Project Aims 6](#_Toc40039788)

[1.2.1 First Aim 6](#_Toc40039789)

[1.2.2 Second Aim 6](#_Toc40039790)

[1.2.3 Third Aim 6](#_Toc40039791)

[1.2.4 Fourth Aim 7](#_Toc40039792)

[1.3 Project Scope 7](#_Toc40039793)

[2 Background 8](#_Toc40039794)

[2.1 Existing solutions 8](#_Toc40039795)

[2.1.1 AdBlock — best ad blocker “Ex1” 8](#_Toc40039796)

[2.1.2 AdBlock Plus - free ad blocker “Ex2” 8](#_Toc40039797)

[2.1.3 uBlocker - Ad Block Tool for Chrome “Ex3” 9](#_Toc40039798)

[2.2 Evaluation of the current solutions 9](#_Toc40039799)

[2.2.1 The Balance of too many options and not enough options 9](#_Toc40039800)

[2.2.2 Default adblocking effectiveness 10](#_Toc40039801)

[2.3 Constraints and Limitations 11](#_Toc40039802)

[3 Approach 12](#_Toc40039803)

[3.1 Setup 12](#_Toc40039804)

[3.2 Research 12](#_Toc40039805)

[3.2.1 WebRequests 13](#_Toc40039806)

[3.2.2 Storage 13](#_Toc40039807)

[3.2.3 AdBlock Plus 13](#_Toc40039808)

[3.3 Development 13](#_Toc40039809)

[4 Specification and Design 15](#_Toc40039810)

[5 Implementation 16](#_Toc40039811)

[6 Results and Evaluation 17](#_Toc40039812)

[7 Future Work 18](#_Toc40039813)

[8 Conclusions 19](#_Toc40039814)

[9 Reflection on Learning 20](#_Toc40039815)

[10 References 21](#_Toc40039816)

# Introduction

## Project Overview

Websites have become littered with adverts. Adverts can come in many shapes and sizes, but all have one thing in common, to take something from you. As internet traffic has boomed over the past several years so have adverts. Adverts have evolved from simple pictures to embedded frames that can track you, collect your information and can even have malicious intent. Some websites are so heavily plastered with adverts that load times can be reduced and website functionality drastically lowered.

The aim of this project is to allow the user to easily manipulate their own browsing experience by culling adverts they do not want to or should not be seeing. Although ad-blockers already exist online they are often hard to customize for everyday users. I will be creating a chrome extension that can be installed into the chrome browser and then used to block adverts. Users will be able to manually select adverts to be blocked or opt to use predefined ad blocking lists that can be provided or added later to the blocker.

The application will offer the user simple yet powerful tools that allow for the manipulation of the websites they are visiting. Adverts that the user wishes to be blocked will be synced to their google account (this is needed to get the extension) and therefore will persist through any of their signed-in instances of google chrome. The application will offer two main streams of blocking that are element blocking, and domain blocking. Simply put, element blocking will allow for specific items on a web page to be hidden, while domain blocking will allow for adverts to be blocked before they are even loaded. Combining these two approaches will allow for a flexible and effective blocking extension.

The aim of this project is to develop a chrome extension that can be used to block adverts on websites inside of the chrome browser. The finished product will allow the user to manually select elements on the page to be blocked as well as offer a variety of address filtering methods for the user to block advert requests. The extension should be simple to use and easy to understand for the everyday web user. Although similar extensions already exist, the aim is to create a more user friendly and accessible version that will allow users to customise their filtering more simply.

My personal aims while completing this project are to increase my knowledge of JavaScript and to learn how to create and use the chrome extension API. This project will allow me to greatly expand my knowledge in both through use of documentation, trial and error and questioning. By the end I should have a clear understanding of the components that make up a chrome extension and the different functions and API endpoints I can use from chrome.

## Project Aims

The project being undertaken consists of a handful of complex parts that will allow for the overall ad-blocking experience. Therefore, to effectively plan and understand the road ahead I split my project into concise segments. Each segment holds a key role in the overall aim of this project therefore having detailed and well thought out acceptance criteria will help make sure they are met properly. Since I was working with the development process it was easy to make slight changes to my aims as needed whenever I encountered a major problem or discovered a new path somewhere else.

|  |
| --- |
| First Aim |
| The extension allows for the manual selection, storage and blocking of elements from any website the user wishes to block elements. |
| **Acceptance Criteria**   * It is obvious to the user which element they are selecting. * A reference to any item the user selects is stored (across sessions). * When a website loads, stored references are checked and applied to the loaded website to hide elements. * The user can edit the list of blocked elements they have selected. |

|  |
| --- |
| Second Aim |
| Adverts can be blocked via domain filtering selections made by the user. |
| **Acceptance Criteria**   * The user can select domains on an already loaded page to be blocked upon reload. * The user can manually type in domains which they wish to be blocked. * The user can edit the list of blocked elements they have selected. |

|  |
| --- |
| Third Aim |
| Adverts can be blocked via domain filtering lists that the user can load or present to the extension themselves. |
| **Acceptance Criteria**   * The user will be able to specify some sort of path to an already created filter list that they wish to use in their blocker. * The user can choose whether to enable this feature at any time. * The provided list works without conflict towards any other filtering. |

|  |
| --- |
| Fourth Aim |
| The user can easily and without confusion edit the settings of their application. |
| **Acceptance Criteria**   * All settings are loaded into the extension’s popup window, not in additional html files or background pages. * The settings have a good UX and will be understandable to users. * All settings are saved seamlessly upon change |

## Project Scope

The original scope for the project involved creating an extension for chrome, porting it over to other browsers like IE and even looking at blocking domains through the routers firewall. However, as the project took foot, I quickly discovered how hard it was going to be initially with learning another language and API. As the project has drawn to a close the scope has indeed shrunk greatly. The scope of the project ended up just being blocking adverts inside of the chrome browser. Initially I was too ambitious with what I wanted to do and reducing the scope meant; less pressure, better code, more time learning less languages and overall being able to think about how I could improve this smaller focus of mine. Overall, the ideas behind the adblocker have stayed the same, with the ideas for blocking elements being implemented, the domain filtering that I had previously thought about being done through the firewall or router being implemented straight into chrome.

# Background

## Existing solutions

When searching the chrome extension store its easy to see that there is not a shortage of adblockers out there for you to download for free. But what makes them different from each other, why is one better than another one. This short evaluation of existing solutions aims to look at some of the features and ideas behind a few of the adblockers. Gathering and evaluation these will help me make better choices when creating my own adblocker.

I will be looking at some of the options available with each of the blockers as well as carrying out some simple tests to see how well they block adverts across a couple of websites. Most of the adblockers seem to use the same underlying technologies so it will be interesting to see how they differ in action

|  |
| --- |
| AdBlock — best ad blocker “Ex1” <http://getadblock.com/> | 10,000,000+ Users |
| **Features:** Offers a free or premium ad-blocking experience, allow non-intrusive ads, show statistics of adds blocked, acceptable ads, custom filters, anti-circumvention measures, cryptocurrency mining protection. Open source and has a good support suite. + many more options |
| **Pros, cons, and thoughts:** Firstly, there are so many options for this adblocker, this is both good and bad. It boasts great customisation but can be a little overwhelming if you are not totally sure what you are doing. The popup is very simple to use and understand, also quite compact, however does not really have many options on it. The installation process on this one also takes a little longer than the others and does prompt you to donate once it has installed. Following on from the previous point the upgrade button is always displayed on the popup which can be a little annoying. With this adblocker when tested on *reddit.com* some ads were displayed, disappeared, and reappeared.  **Blocked:**  **R = {Claimed: 5, Visibility: Poor} S = {Claimed: 13, Visibility: Poor}** |

|  |
| --- |
| AdBlock Plus - free ad blocker “Ex2” [adblockplus.org](http://adblockplus.org/) | 10,000,000+ Users |
| **Features:** Whitelist website, whitelist page, basic statistics, different languages, tracking blocking, custom filters. Extensive help suite with lots of documentation. |
| **Pros, cons, and thoughts:** This adblocker have a good amount of options, not too many nor to few. The popup is very simple and therefore all the settings are hidden away. The UI for the adblocker is very clean and easy to understand. The adblocker installed fast and is totally free. When loading *reddit.com* no ads we are displayed to me. Overall, this adblocker is very simple to use and has a good amount of options that are easy to understand.  **Blocked:**  **R = {Claimed: 3, Visibility: Poor} S = {Claimed: 12, Visibility: Poor}** |

|  |
| --- |
| uBlocker - Ad Block Tool for Chrome “Ex3” 400,000+ Users |
| **Features:** Whitelist website, pause adblocker, statistics, allow non-intrusive add items through right click. |
| **Pros, cons, and thoughts:** This adblocker has a similar popup look to *AdBlock — best ad blocker.* It is very clean and shows you exactly what you need to see. Overall, the options available for this adblocker are very poor. However, I believe it provides a simple adblocking tool for the user who does not want any customisation etc. The extension has a much smaller user base. Once again, the options panel is hidden away in another page and is not very eye catching but cuts straight to the point.  **Blocked:**  **R = {Claimed: 4, Visibility: Excellent} S = {Claimed: 8, Visibility: Good}** |

## Evaluation of the current solutions

### The Balance of too many options and not enough options

Options are important when choosing the adblocker for you, the user needs to be in control of the content that is being allowed and blocked from the websites they visit. Finding a balance between needed and extra options is also something to think about as you do not want users to be overwhelmed or confused by the options that are available to them.

Firstly, *Ex1* had a huge range of options available which meant that the user was going to be able to fully customize their browsing and adblocking experience. However, I felt very overwhelmed when looking at the options as I was not sure what a lot of them meant or what they would do for me. Also having the premium and the options available with it constantly pushed in my face was annoying. Overall, I believe this adblocker had to many options available with it.

Secondly, *Ex2* had a good amount of options, it felt like the options that were available had been carefully curated to keep the user informed and not overwhelmed. Just like the previous the options were tucked away in the settings page. Overall, I believe that this adblocker had a brilliant balance between options and simplicity.

Finally, *Ex3* lacked a lot of options but seemed very simple and intuitive to use. This adblocker was missing a lot of features that I felt were needed such as being able to load in a custom filter list. This adblocker could really do with a few more options but if you are looking for something very simple and lightweight this may be a good option.

Overall, I believe that *Ex2* (AdBlock Plus) had the best balance between customization and simplicity with its careful balance of options available. I would have liked to see more options available on the popup though as all three had theirs hidden away in the settings page.

### Default adblocking effectiveness

An important factor to think about with these adblockers is how well they perform once they have just been installed, so no options being changed, the program in its default state.

I carried out a small test across two websites:

*R =* <https://www.reddit.com/>

S = <https://smallbusiness.chron.com/affiliate-marketing-tracking-tools-67773.html>

I recorded how many adverts are supposedly being blocked by the adblocker and then my own view on how well it had done to remove the adverts from the page. To do this I viewed both pages with and without the adblockers to see how well they did.

Firstly, *Ex1* Boasted the highest numbers being blocked but failed to block the ads which we are visible to me on the page. This happened to both websites in which I tested. I assume this means that adverts / tracking data in the background is being blocked but it has not recognised or filtered out the adverts that I am seeing.

Similarly, with the second, *Ex2*, both websites reported having a high number of ads blocked on the page, but all adverts were let though as usual. This is unfortunate for both bigger names adblockers to be letting through so many adverts I rated both websites as poor just like *Ex1*.

Finally, the smaller adblocker *Ex3* pulled through and did a really good job at blocking the adverts. This adblocker reported much lower number of adverts being blocked but then went ahead and blocked all but one of the visible adverts that I could see.

To conclude my adblocking tests I have discovered that the smaller of the blockers with much less options available had come out and blocked many more adds than it is bigger more popular competitors. There are many reasons this could be: Perhaps the other two need some setting up to work more effectively, Block may have a newer filter list.

## Constraints and Limitations

One major constraint I will face when undertaking this task is that I am not well informed with the technologies that I will be using to create my adblocker. To begin with I will be using JavaScript, this is a language I have used previously at university but have a very limited knowledge of. As this language is used often by many people I will preserver with the guidance and documentation that can be found online. Following this previous statement, I will also be learning and dabbling with jQuery, a shorthand for JavaScript that should help make coding simpler. Similarly, I will be working in a new environment to me – Chrome. I have never developed a chrome extension or worked with the API, so I do not know what is unattainable and what is realistic. Again, for all these documentation and further reading will help lessen this constraint.

One constraint that I have encountered that has affected this project is the coronavirus. At the time of writing this report university has been closed for quite some time. Universities have closed, society has shut down. The resources that are available when normally undertaking a university degree are not available. Computer labs, Libraries, and the ability to meet up with supervisors, peers and tutors have all been removed, potentially threatening the outcome of this project. As this project has progressed it has been noticeably difficult, being out of routine and stuck in a weird situation where no one is safe has impacted the ability to work properly will have had slight affects on this and many other projects.

The final constraint to this project is to do with filter lists that are used by adblockers. After doing research it is apparent that adblockers mainly use other people’s lists as filters rather than their own. Some of these lists are public and can be used but are written in schemas and formats that are not fully explained and may therefore be unusable. It is the aim of this project to at least allow for those filter lists to be loaded and used at least, as if they we’re processing the filters correctly.

# Approach

The approach that was taken towards the project was Setup, Research and Develop. This approach allowed for all the core pieces of application to be put into place before cycling back and forth through research and development. As each week came by and work began on the aims of the project, research was made into each individual part and used to develop that piece. This approach worked well as it worked well with the flexibility and granularity of the agile methodology. As the project move forward less and less research was needed, as familiarity grew, which left more time for development.

## Setup

The first part to the approach was the setup. During this time, the project had all its core components created and setup which would allow for everything else to flow nicely. To begin with was the task of setting up version management. GIT was chosen as it is easy to use and widely accessible. It was chosen that it would be hosted on a private repository on GitHub. Having the project hosted here meant it could be worked at from any computer on any device and the files would never be lost. The client that was chosen to mediate this was Gi Tortoise as it provides an easy and simple to use interface. GIT played a big role in this project and allowed for rapid development through trial and error.

Next on the setup phase was to create the initial chrome extension. I first had to learn what files would be needed and how they interact with each other. Over the course of a day the barebones application was created and an understanding of what each key individual file did was gained. Over the course of this project the skeleton evolved slightly with new discoveries and ideas coming into play. This played an important part as gaining a better understanding of how each component worked early allowed for progression with more pressing items.

The final Items that was part of the setup was a task board and notes page. Trello was used as an easy way to divide the project into tasks and subtasks that would allow for progress tracking. Trello is an online task board that works well with the agile methodology. Being online, like GIT it is useful because it could be accessed from any device and keep development up to date at any point. To keep notes a work in progress word document was created where anything from failed ideas to thoughts could be placed to help aid the writing of this report.

## Research

Using a language and API that was not so familiar meant a lot of research was needed continually throughout the development. Research was carried out and then put into practice in the development stage before that small cycle started again with the next component. There were three main areas of research that were carried out and are listed below.

### WebRequests

WebRequests is a crucial part of the chrome API that would be needed in the development of the extension. WebRequests allows for user intervention at several locations during the life cycle of an incoming or outgoing request. Using this would mean requests could be filtered and blocked to stop adverts from being loaded onto the webpage from external sources. On the API documentation page there are detailed and descriptive examples of how to use the code in a variety of scenarios. This and the overall chrome API documentation was a big help and made development a lot simpler.

### Storage

The Chrome API has an easy to use storage facility located in Storage.Sync. This storage facility allows for data to be stored inside of the browser, attached to your google account. Using this meant that users settings, choices and block preferences could all be saved easily and loaded to anywhere they log in. The storage API saved lots of time and effort as originally the plan was to save these things to documents. Storage.Sync has two simple functions, Get and Set. Get returns an object and set allows you to set an object. Combined these two powerful snippets helped create basic foundations of how the extension would solely function.

### AdBlock Plus

AdBlock plus is one of the already popular adblockers that is available as a chrome extension. They host lots of material on their website that talks about how their extension works, some of the algorithms and techniques behind the scenes. All their code is open source too. Combining these elements meant there was a good starting point and reference location to help throughout. Although greatly useful some documentation and examples were missing which meant I was not always relying on this and constantly came up with my own solutions.

## Development

The final part of the chosen approach was the actual development. Developing was straight forward and coupled with GIT allowed me to be flexible with my attempts as I had branches, I could always fall back on. Development would differ per task but would normally consist of a front end to back end style approach where possible. Since this extension is focussing on having concise and available options to the user, there was lots of front-end options that needed hooking up. The aim was to create one front end object first, connect it to storage and then hook-up the background logic after. Creating each option like this meant that it would be simple to trouble shoot and each commit was related to just a single item. Finally, as each bigger stage would be completed an overhaul including refactoring, reordering, and minifying could be applied to make the code better. Throughout the process there would be many problems and challenges but thanks to the approach I had selected and tools I had used to help preserve the projects integrity I would be able to work stress free and effectively code the project. Overall, the approach was a good fit for the extension I was creating and worked nicely with the agile framework.

# Specification and Design

# Implementation

# Results and Evaluation

# Future Work

# Conclusions

# Reflection on Learning

# References