A Proposed Methodology for Analysis of Cookie Tracking and Cookie Management Solution

by

Faryal Nosheen

NUST201261272MCEME35412F

MS-12(CSE)

Thesis Supervisor

Dr. Usman Qamar

Signature

In the name of Allah most beneficent most merciful

وَلَا يُحِيطُونَ بِشَيْءٍ مِّنْ عِلْمِهِۦٓ إِلَّا بِمَا شَآءَ

And they can’tencompass any thing from His knowledge, but to extend He wills [2:255]

This page is intentionally left blank

Acknowledgment

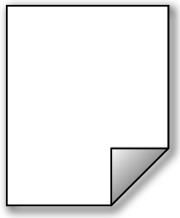
First and above all, I praise Allah, the almighty for providing me this opportunity and granting me the capability to proceed successfully. This thesis appears in its current form due to the assistance and guidance of several people. I want to offer my sincerest gratitude to my supervisor, Dr Usman Qamar, who has supported me throughout my thesis with his patience and knowledge whilst allowing me the room to work in my own way.

Sincere thanks to all my friends for their kindness and moral support during my study. Last but not least, my deepest gratitude goes to my beloved mother and husband for her endless love, prayers and encouragement. Therefore I like to offer my sincere thanks to all of them.

To my Mother, Husband and Supervisors.

Abstract

Advertisers place cookies in the browser to follow a user's steps across different websites, mostly without his knowledge or consent. This diploma thesis deals with the problems of privacy issues caused by the tracking activities of third parties on the web and purpose a Firefox extension as a solution. The add-on 'privacymanager' provides control to user to manage third party cookies by setting polices for each site. Relevance and impact of the extension are proven by the analysis of data, gained with automated visits of 5.000 websites. Furthermore, currently available Mozilla Firefox add-ons dealing with cookies and privacy are analyzed in detail.



Contents

1 INTRODUCTION **12**

1.1 BACKGROUND **13**
1.1.1 WHAT'S A COMPUTER COOKIE 13
1.1.2 USE OF COOKIES BY ADVERTISERS 13
1.2 WEB BROWSING BACKGROUND **15**
1.2.1 PAGE FETCHING 15
1.2.2 FIRST PARTY AND THIRD PARTY DOMAINS 15
1.2.3 HTTP COOKIE 16
1.2.4 FIRST AND THIRD PARTY COOKIES 16
1.2.5 USE OF COOKIES 17
1.2.6 LSO/FLASH COOKIES 17
1.2.7 TRACKING 18
1.3 CONCEPT OF THIRD PARTY WEB ADVERTISEMENT **18**
1.3.1 HOW DOES THIRD PARTY AD SERVING WORK 18
1.3.2 CATEGORIZES OF THIRD PARTY TRACKERS 18
1.4 THE MECHANISM OF ONLINE TRACKING **20**
1.4.1 FIRST PARTY COOKIES USED WITHIN SINGLE WEBSITE 21
1.4.2 THIRD PARTY COOKIES USED WITHIN SINGLE WEBSITE 21
1.4.3 THIRD PARTY COOKIES THAT TRACK USERS ACROSS THE INTERNET 21
1.5 TRACKERS BEHAVIOR **22**
1.6 CONTRIBUTION **22**
1.7 OUTLINE **23**

2 REVIEW OF RELATED WORK **24**

2.1 OVERVIEW **24**
2.1.1 TECHNOLOGIES/TOOLS TO CONTROL USER PRIVACY 25
2.1.1.1 OPT OUT COOKIES 25
2.1.1.2 DO NOT TRACK 25
2.1.1.3 BROWSER PROFILE CLEARING 26
2.1.1.4 PRIVATE BROWSING MODE 26
2.1.1.5 THIRD PARTY COOKIE BLOCKING 27
2.1.1.6 TARGETED COOKIE BLOCKING 27
2.1.1.7 EXECUTION BLOCKING 27
2.1.1.8 CONTENT BLOCKING 28
2.2 REVIEW OF RELATED WORK **28**

2.2.1 COOKIE MONSTER 28
2.2.2 COOKIE WHITELIST 29
2.2.3 COOKIE SAFE 30
2.2.4 EXTENDED COOKIE MANAGER 30
2.2.5 GHOSTERY 31
2.2.6 BEEF TACO 31
2.2.7 TARGETED ADVERTISING COOKIES OPT OUT 31
2.2.8 TRACKER BLOCK 32
2.2.9 BETTER PRIVACY 32
2.2.10 ADNOSTIC 33
2.2.11 PRIVAD 33
2.2.12 COLLUSION 33
2.2.13 FIREFOX'S NEW COOKIE POLICY 34
2.2.14 DOPPLEGANGER 34
2.2.15 SELF-DESTRUCTING COOKIES 35

3 THIRD PARY COOKIE MANAGEMENT POLICY AND IMPLEMENTATION 36

3.1 MOZILLA FIREFOX BROWSER AND ADD-ON DEVELOPMENT 36
3.1.1 PROGRAMMING AN ADD-ON 36
3.2 IDEA OF WORKING OF 'PRIVACOOKIE' 39
3.3 ENHANCEMENTS OVER 'PRIVACOOKIE' 40
3.3.1 USER INTERFACE 40
3.3.2 COOKIE STORAGE 41
3.3.3 PRIVACY/TRACEABILITY TRADE-OFF 42
3.3.4 SPATIAL AND TEMPORAL TRACKING POLICY 42
3.4 PROGRAMMING AN ADD-ON 44
3.4.1 XUL AND JAVASCRIPT FILES 44
3.4.2 DESIGN OF AN INTERFACE 50
3.4.2.1TOOLBAR-BUTTONS 50
3.4.2.2 CONTEXT MENU 51
3.5 FUNCTION TEST 54

4 ANALYSIS 55

4.1 WEB ANALYTICS 55
4.1.1 WEB MEASUREMENT TOOL-FOURTH PARTY 56
4.1.2 FOURTH PARTY TOOL DESIGN 56
4.1.3 ANALYSIS USING FOURTH PARTY TOOL 57
4.2 METHADOLOGY 57
4.2.1 INSTALLATION OF FOURTH PARTY TOOL 57
4.2.2 CRAWLER 58
4.3 AUTOMATION AND MEASUREMENT 59
4.3.1 IDENTIFICATION OF THIRD PARTIES 60
4.3.2 INSPECTING THIRD PARTIES 60
4.3.2.1 FREQUENCY OF USE BY FIRST PARTIES 61
4.3.2.2 NUMBER OF REQUESTS 61
4.3.3 LOOKING AT FIRST PARTIES 61

4.3.3.1 FIRST PARTIES AS THIRD PARTIES 62
4.3.3.2 FIRST PARTIES WITH TP CONTENT 63
4.4 EFFECTS OF 'PRIVACYMANAGER' ON BROWSING 65
4.5 DOM STORAGE, LSO & COOKIES 66
4.6 RATING MECHANISM OF THIRD PARTY DOMAINS 66
4.7 COMPARISON OF 'PRIVACOOKIE & 'PRIVACYMANAGER' 68
4.8 COMPARISON WITH OTHER APPROACHES 70
4.9 SUMMARY OF RESULTS 72

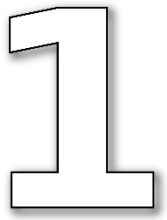
5 CONCLUSION AND FUTURE WORK 73

List of Figures

1.2.4 Figure 1.1 First Party And Third Party Cookies 16
2.2.2 Figure 2.1 Cookie WhiteList 30
2.2.12 Figure 2.2 Collusion 34
3.1.1 Figure 3.1 Sample file of 'install.RDF' 38
3.1.1 Figure 3.2 install.rdf 39
3.1.1 Figure 3.3 default.js 39
3.3 Figure 3.3 Add-on 'privacymanager' 41
3.3.1 Figure 3.5 Toolbar button along with context menu 41
3.3.1 Figure 3.6 'privacookie' user interface 42
3.3.2 Figure 3.7 43
3.3.3 Figure 3.8 'PrivaCookie' Privacy/Traceability Trade-Off 43
3.3.3 Figure 3.9 Privacy/Traceability Trade-Off 44
3.3.4 Figure 3.10 Spatial Tracking Policy 45
3.3.4 Figure 3.11 Temporal Tracking Policy 46
3.4.1 Figure 3.12 List of Files of Add-On 'privacymanager' 48
3.4.1 Figure 3.13 Duplication.js 49
3.4.1 Figure 3.14 ./chrome/content/logger.xul 50
3.4.1 figure 3.15 ./chrome/content/pref.xul 52
3.4.2.1 Figure 3.16 Toolbar-buttons 53
3.4.2.2 Figure 3.17 Third party list editor 54
3.4.2.2 Figure 3.18 Setting Dialogue 55
3.4.2.2 Figure 3.19 Log Console 56
4.2 Figure 4.1 Data Collection and Measurement Infrastructure 60
4.2.1 Figure 4.2 Fourth Party SQLite Database Schema 61
4.2.2 Figure 4.3 Web crawler Interface 62
4.3.2.1 Figure 4.4 frequency of use of third party cookies 68
4.3.3.2 Figure 4.5 Use of Third Party Cookies in First Parties 69
4.6 Figure 4.6 Rating by Privacy choice 70

List of Tables

3.1.1 TABLE 3.1 LIST OF MAIN FILES IN ROOT DIRECTORY 37
3.1.1 TABLE 3.2 LIST OF FILES IN DIRECTORIES 37
3.4.1 TABLE 3.3 PRIVACYMANAGER.SQLITE 44
4.3 TABLE 4.1 FIREFOX PROFILE FOLDER 60
4.3.2.2 TABLE 4.2 TOP 20 NUMBER OF THIRD PARTIES ON FIRST PARTY HOSTS 62
4.3.3.1 TABLE 4.3 TOP 20 FIRST PARTIES WITH FREQUENT USE OF THIRD PARTIES 63
4.5 TABLE 4.4 DOM STORAGE, LSO & FLASH COOKIES 66
4.6 TABLE 4.5 RATING BY PRIVACY CHOICE 68



The world is never going to be perfect, either on- or offline;
so let's not set impossibly high standards for online
ESTHER DYSON

Introduction

Internet marketing started in 1994, when the "Hot Wired", a stylish classification technology "Wired" magazine's electronic edition, started out a Website, about 12 vendors for the position of advertising ads throughout the website to pay costs. Thus, the development of a new commercial organization, which includes promoting marketing, innovative marketing and calculating how many individuals see the marketing companies.

With the appearance of the electronic economic system and the success of e-commerce, internet marketing becomes a multi-million money business. Social networking sites offer many individuals the foundation for international connections. These services provides at a cost of significant attack into the comfort of the customer, often without their knowledge and approval. Monitoring is also practical to Web customers who get customized details that coordinate their passions. However, this may result in the selection and the leak of users' private details which is really a risk to users' comfort. A data exchange is usually done by means of cookies.

In 1994, Lou Montulli, while working for Netscape, presented the idea of cookies in the perspective of a web internet browser [1]. The cookies mechanism allows a web server to store a bit of details on the computers of customers, which is then sent back to the web server upon request.

Many modern internet explorer now have a support for the rejection of third-party cookies and some even allow it by default but these browsers rejects all third party cookies. A browser's "Private Mode" is also available to help users to check out a set of websites without making records of their check out on their device, this method not only removes third party cookies but also first party cookies. By simply blocking many third-party cookies makes advertising a less relevant because it will be based solely on the present web site explored by the user (i. at the., context) instead of about what the customer have done in past times (behavioral tracking). The present mechanism to control third-party cookies does not allow advertisers to monitor its surfing behavior. It allow advertisers to track user activities either across all websites or none. A browser's "Private Mode" is also available to help users to check out a set of websites without making records of their check out on their device, this method not only removes third party cookies but also first party cookies

From this originates the need of additional tools, with which the internet user can minutely determine the level of disclosure of his details. Preferably, such tools provide comprehensive functionality for convenience of use so that they can also be used by non-technical individuals.

In the second chapter we survey current techniques to prevent private information from third parties (e.g., block third-party cookies) and then discuss a proposal in a third chapter, which keep balance between tracking by advertisers and user's control on his private information. We implement this proposal as an Add-on for Firefox 'privacymanager' and shows that users' can control an amount of information he wants to share on internet. The fourth chapter contains an analysis of the tracking by third party cookies, based on the visit of 5000 websites. Cookies and third-party relationships are hereby recognized and analyzed. Further, a comparison of privaCookie and privacymanager illustrates the implemented extensions.

1.1 Background

1.2.1 What's a computer cookie?

It's a small text file or pc ID created when your internet browser load a web page. It guarantees automated logins and verification when you want to access a web page again, and stores purchasing information needed for on the internet shopping trolleys to remember all the items you put in the trolley. So, cookies are useful and safe for the most part. But when used for advertising objectives, they get into your internet comfort. And when used to act as a form of malware, they become serious internet security threats

1.2.2 Use of Cookies by Advertisers

Advertisers can follow your motions from web page to web page and develop a data source of your interest on the internet with the help of the so-called "third-party cookies".

How do these monitoring cookies work exactly? Well, to begin with, you're always the first party, and the cookies you get when you check out a web page are the second-party cookies. Usually, sites let marketing systems place ads within their webpages. If you simply click an ad, another cookie is sent to your internet browser by the marketer – that's the third-party cookie. While this type of cookie doesn't cause an internet security risk, there's a comfort issue engaged. With every new web page you check out that's relevant to that particular marketer, the third-party cookie can be tracked. This way, the marketer understands about your interest on the internet routines and can develop up a customer user profile of you – that's behavior monitoring. Next step: it reveals you particular ads relevant your passions.

For example, assume that a user visits journey.com, whose home-page has a distant picture from monitoring.com. Therefore, as part of the process of making journey.com's home-page, the user's internet browser will demand the picture from

monitoring.com. The web server of monitoring.com delivers the picture along with an HTTP Set-Cookie headlines, establishing a cookie on the user's device, under the monitoring.com. Later, when the customer browses to other sites associated with monitoring.com, e.g., buy.com, the monitoring web page gets its formerly set cookie, identifies the customer, and makes a user profile of the user's surfing around internet. These third-party cookies, due to the negative results on a user's comfort and their immediate relationship with on the internet behavior marketing, taken the attention of both the research group [2], [3], [4] and the popular press sites [5] and, ever since, cause the public's pain [6], [7].

1.2 Web-Browsing Background

1.2.1 Page Fetching:

Every time a page needs to be fetched by the browser, an HTTP request was sent for a URL to the website for a URL in top-level effectiveness perspective for the web page (that fits with a user-visible screen with a web page title). The HTTP response may contain several type of sources '(HTML, scripts, pictures, fashion bed sheets and others)' that are processed and may makes further HTTP requests for sources. This process carries on recursively until pages loads totally.

1.3.2 First Party and Third Party Domain

A website may contains content from another domain. It may be in the form of an iframe, that included in the website but referred to the portion of the screen to the domain from which iframe belonged—this is considered the third-party site. The same-origin approach ensures that content in the two sites is actually isolated: any kind of script running within the iframe running from third party domain. A website may contains a script which belongs to another domain but that script run in the domain of the site which embeds it not from the script's source site.

1.3.3 HTTP Cookies

Web servers store and retrieve state of web customers by using “HTTP cookies” [10]. However, HTTP is a stateless connection, only cookies enable web applications to store persistent state over numerous HTTP requests. For instance, web shopping applications can utilize cookies to track which items a client adds to her shopping truck. At the point when a customer makes a HTTP request to a server, the server has the choice of including one or more Set-Cookie headers in its response. User will give back these cookies in HTTP requests utilizing the Cookie header. The Set-Cookie header has one obliged field, a name/worth pair of the structure NAME = VALUE. A web server utilizes this field to encode the state data it wishes to store on the client side. There are likewise four nonobligatory fields: “expires=Date, domain=DOMAIN, path=PATH PREFIX, and secure”.

An expire field shows for how long cookie is valid. Upon reaching expiry date, the customer's web program ought to erase the cookie. In case there is no expire field, then the cookie consider a “session cookie” and expired/deleted upon closing browser by user. Whereas persistent cookies have an expire field.

1.3.4 First And Third Party Cookies:

Web monitoring is based essentially on a website’s capability to store condition on the user’s device as do most features of today’s web. Client-side condition may take many types, most generally traditional browser cookie. Against a single cookie, three values (domain, key, value) store in the web browser across page visits, where domain is a website, and key and value are solid identifiers. Cookies that are set by the domain that the user visits directly (the domain shown in the browser’s address bar) are known as first-party cookies; cookies that are set by some other domain included in the top-level web page are third-party cookies.

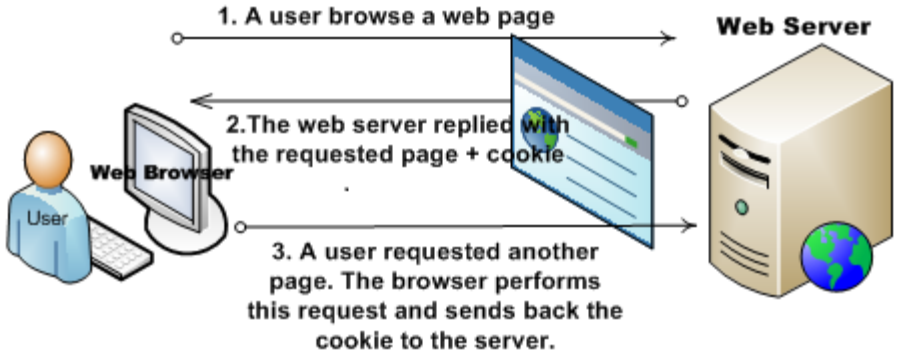


Figure 1.1 First And Third Party Cookies

1.3.5 Use of Cookies:

Cookies have many purposes: session state, customization, verification, and monitoring. Websites use cookies for customization to remember users' choices and configurations. For example, Google allows customers to change the structure of their search results and uses cookies to keep in mind these choices. Websites with user account also use cookies to authenticate users' sessions [12]: with a user login, a web site can set a session cookie on the user's machine to verify her requests. Websites can set persistent cookies to remember users and not require a subsequent visits. Finally, websites can use cookies to track customers and their activities.

Tracking cookies also make cookie management difficult. Many users might prefer not to accept tracking cookies due to the privacy risks. Most web mail services, e-commerce, and financial websites require users to accept and send cookies for authentication, and blocking cookies also declines customer's customization features. Blocking all cookies is consequently not suitable for most users

1.3.6 Local Storage Object (LSO)/Flash Cookies:

The concept of 'Local Storage Cookies/Objects' of 'Flash Cookies' introduced by HTML5 with new client-side storage systems for browsers. This is new concept and these cookies more persistent than browser based cookies which are easy to identify

and remove. These are also called “Local Shared Objects (LSO)/ flash-based cookies”. In particular, Local share objects save more permanently which sites can access with API calls, isolated by the same-origin policy. This concept mainly used in Banks where flash cookies are used to store customer specific information on users' computers to verify account owners and prevent fraud.

1.3.7 Tracking:

The majority of us don't understand how simple it is for our online conduct to be followed. ISPs, web sites, and advertising networks; every single have method for tracking your online activities. They track you to provide targeted advertising, classify you into a demographic group, and sell your information to advertisers. Tracking done by web sites by using cookies, IP addresses, browsing history, web Bugs and many different procedures. Here we mainly concentrate on tracking by cookies.

1.4 Concepts of Third Party Web Advertisements:

1.4.1 How does third-party ad serving work?

When a user visit a website like www.yahoo.com where the content of the site come from the domain web site but the ads running on the site, come from another domain/server. A browser gather information pertain to that site from different domains and all items appear on the same page. As a result, a third party cookie by third party website will be created in your browser

1.4.2 Categorizes of Third Party Trackers:

We can classify third party trackers/advertisers into six categories by [8]:

i. **“Advertising Companies:**

In starting 90's internet promoters directly deal with first party promoters but in late 90's and starting 2000s, growth in professional need and ad slot

provide (“inventory”) made it wrong for promoters and promoters to deal directly. Ad techniques offered a solution by allowing promoters to easily place ads with many promoters, and by allowing promoters to back up their content with many promoters. In the mid-2000s, Ad Dealings presented more pattern to this market by taking provides from many promoters via many marketing techniques (“real time bidding” or “RTB”) and a number of intermediary business models now are available in the exchange ecosystem i.e. Demand-side techniques, Supply-Side Systems and Information providers.

ii. **Analytics Services:**

Third-party statistics services provide resources for sites to better understand their guests, such as census, customer providers, and content opinions and communications. Some organizations (e.g. Adobe) provide statistics as professional help, while other organizations provide a free statistics service; they generate income from the information they gather by using it for ad focusing on (e.g. Quantcast), market knowing (e.g. Search engines Analytics)

iii. Social Networks:

Public integration uses one common business structure is social sharing gathering and enables sites to provide personalized content and single sign-on to online community customers. Solutions like AddThis, ShareThis, and Meebo provide free icons to sites that enable customers to share with dozens of social media sites. To generate income from their icons, the assistance collect tracking and usage data and sell it for ad targeting and market research.

iv. **Content providers:**

Content providers like ‘YouTube’ used to host video, maps, news, weather, stocks, and other media and embeds into websites. Some content providers, like Facebook, include third party widgets to website to indulge users and earn more revenue through in-widget advertising

v. **Frontend services:**

To decrease webpage load time, several third parties host Java Script libraries and APIs e.g. Google Libraries API and enable new page functionality

vi. **Hosting Platforms:**

For the assistance of publishers in distributing their own contents, few third parties provide services which may be blog platforms like wordpress.com and content distribution networks like Akamai.”

1.3 The Mechanism of Online Tracking:

It may be usually said that online "behavioral" marketing depends on systems that allow promoters to monitor a customer's surfing behavior. Among the main elements of this monitoring under current technological innovation is the "cookie." Cookies usually store a customer's preferences and other information about the user’s surfing behavior. Cookies are of different types.

1.5.1 First-Party Cookies Used Within a Single Website.

Students and experts on Internet business and privacy, can easily differentiate between "first-party cookies" and "third-party cookies." First-party cookies allow the web page to gather certain details from the user's internet browser and then

remember that details whenever that user revisits the web page. This can provide advantages, for example, because of first-party cookies, Amazon concisely knows what items a client has purchased or selected on a prior visit and can therefore create more focused and effective advertisement when the customer returns.

1.5.2 Third-Party Cookies Used Within A Single Website.

Sites may also use third-party cookies to offer "analytics" services for the web page. For example,Amazon.com utilizes a third-party company to offer its advertising "analytics." This third-party enterprise paths the client as he or she goes through different webpages within the web page. The third party examines and analyzes this behavior, makes certain presumptions about the client based upon it, and then suggests the web page owner on the best ways to get more traffic and make more sales to that client.

1.5.3 Third-Party Cookies That Track Users Across The Internet.

Of higher concern may be third-party cookies aid data brokers and aggregators. These types of cookies are not managed for the benefit of the particular website. Instead, they are used to monitor a web browser across several sites, and thus can attract many more implications about the particular customer. For example, so-called "ad networks" obviously place cookies on several of the most well-known sites and, as such, have the potential to monitor a single web browser as it trips the many sites within the "network." The ad system does not actually know the customer's personal identification, but more likely identifies the web browser or IP deal with.

1.6 Tracker's Behavior

Web monitoring/tracking includes several automated actions, beginning with the selection of Web information, the preservation of these information, and the use of

the information. By recombination, connection and de- contextualization, Web information can be used to create very specific predictive information of individual bevahior.

[9] defines classification framework' of web trackers on the basis of their behavior. But any tracker may exhibit more than one of these behaviors:

- **"Behavior A (Analytics):** In this behavior, a tracker acts as a third party analytics engine which can only track users within sites.
- **Behavior B (Vanilla):** In this type, a web tracker uses third-party storage and it can track users across websites. These trackers can get and set third party storage only from a third-party position.
- **Behavior C (Forced):** Such trackers forces users to visit domain as a first party position by redirecting, popup etc.
- **Behavior D (Referred):** this tracking behavior depends on B, C and E type trackers to get unique identifier of user, to track him across sites.
- **Behavior E (Personal):** The cross-site tracker is visited by the user directly in other contexts."

1.7 Contribution

We analyze the impact of tracking done by third party cookies, by custom designed web crawler along with web measurement tool and to deal with the problem of privacy while surfing around on internet but also to keep surfing user friendly, we suggests a fine-grained, per-site cookie control method. It makes the following contributions:

- A web crawler to automatically load 5000 web pages
- An add-on for cookie management
- Analysis of amount of tracking done by third party cookies by using web measurement tool and web crawler
- Evaluate the performance of cookie management proposal

1.8 Outline:

In the next chapter, we will review the related work and their methodologies. After that we will present implementation detail and in the end, will analyze the impact of propose approach

2

Privacy is well worth fighting for
Annicka Gunnarsson

Review of Related Work

In this chapter, we give a brief overview on the state-of-the-art of existing approaches of Third party cookies management. We mainly considered the work implemented as browser extensions. We briefly describe working of each approach and features.

2.1 Overview:

From few years, users get more concerned about privacy of their information on internet due to increase of online advertising, as reacted in coverage of the issues surrounding \behavioral tracking" in the popular press (e.g. the Wall Street Journal's \What They Know series [11])". Due to these privacy reservations, regulation has been passed in Europe [22] about the use of cookies.

Under these regulations, strict rules formed about the gathering, setting and utilization of cookies. For instance, storing cookies in a client's machine is permitted under the following conditions: (i) The client is defined about how this data is utilized; and (ii) the client is given the likelihood of refusing to store cookies. However, these regulations could not properly address and protect the privacy information of users on the web.

2.1.1 Technologies/Tools to Control User Privacy:

[27] has been reviewed the number of technologies/tools which offered consumers a control over third-party web tracking.

2.1.1.1 Opt Out Cookies:

From ten years, a minority of third-party trackers, most unmistakably those who are part of “self-regulatory Network (NAI) and Digital Advertising Alliance (DAA)”, have offered clients the capacity to set "opt-out"cookie.

In Opt-Out strategy of handling third party cookies, web site generates Opt-out cookies on client browser folder that empowers you to restrict that same site from creating future cookies. This mechanism advises the site not to install third party advertiser or other cookies on your browser. This keeps the third party ad server from following your page choices inside a site or among sites inside their system. The disadvantage of opting opt out cookies is that they are site specific. They can just piece treats from a particular server and won't square treats from different sites.

"opt-out" cookies mechanism do not block tracking, it only gives an option to users of not seeing ad-networks which track users—not protect from tracking.

2.1.1.2 DO NOT TRACK

Don't Track is an innovation and approach that empowers clients to “opt out” of tracking by sites they don't visit, including “analytics services”, “advertising networks”, and “social platforms”. Currently, third parties offer opt out strategy and other tools to strict tracking but these are not either user friendly not completely addressing the issue of privacy. Don't Track utilizes a HTTP header to indicate a client's choice to opt out of third party tracking but its main drawback is that it depends on third parties whether they honor users preferences which is not happening in reality [13].

2.1.1.3 Browser Profile Clearing

Users are encouraged to normally clear their “cookies”, “cache”, “browsing history”, and other profile settings of browser to avoid tracking by third parties. But due to several reasons, this approach does not address the issue in true sense:

i. In spite of many blocking technique, tracking methods still exists and work. Especially the tracking technique in which storing state in the browser not require, is totally unaffected. With respect to stateful tracking, the client must play “Whac-A-Mole” with third parties. To remove E-tag cookies, the client must clear the cache of browser. To evacuate “LSOs”, it is require to remove Flash plugin data. In short: the client need to scour wherever the browser or a plugin can store state.

ii. By clearing the browser profile, periodic protection is available only. When user clearing his browser settings, every tracking method works.

iii. Third, clearing the browser profile undermines the usefulness of saving state information. Huge numbers of the lost peculiarities bring about noteworthy inconveniences (e.g. stored logins and browsing history). Some even pose security vulnerabilities (e.g. stored authentication tokens and HTTP Strict Transport Security).

Therefore, clearing the browser profile is an unworkable solution

2.1.1.4 Private Browsing Mode

Implementation specification of each browser may vary but using an option of private browsing mode has a common objective: remove browsing history that stores on the client computer. Browse in a private mode function in a same way as clearing browser profile, except that a user declares session private which

automatically clear browser profile instead of clearing profile every time by user. The shortcomings of browsing in private mode is same as clearing the browser profile: it doesn't stop all tracking strategies, it gives just intermittent security (the client could be tracked inside a private scanning session), gainful web functionality breaks, and a user won't change a setting each time his browser starts.

2.1.1.5 Third-Party Cookie Blocking

All the significant web programs gives an option to user to block third parties from setting cookies. Since cookies are only one of numerous ways third parties track users, an option of blocking third party cookies gives limited protection. A user can protect privacy by setting browser to block third party cookies from being read, otherwise clicking a tracker's advertisement or surfing a tracker's site (e.g. Facebook or Google) once is sufficient to set an indefinite tracking cookie.

2.1.1.6 Targeted Cookie Blocking

Many browsers like Internet Explorer, Firefox, Chrome, etc. have the capacity to restrict cookies from specific domains from being read or set. Much the same as third party cookie blocking, this methodology does not mitigate non-cookie tracking advances. It additionally blocks many interactive functionality on websites that uses first party and third party trackers (e.g. Facebook or Google).

2.1.1.7 Execution Blocking

Other than third party cookies, JavaScript, Flash and other script contents can be used for tracking. Many tools are in market to prevent the execution of these scripts for privacy preserving. While there are numerous different reasons to use these tools (like security, speed, and power consumption), they only mitigate a subset of tracking methods.

2.1.1.8 Content Blocking

To ensure privacy, many privacy tools implements privacy policy with that much intensity that it prevents browser from even requesting certain third party content. No doubt content blocking can successfully prevent third party tracking but the effectiveness of tool is depends on its list of rules on what to block i.e. "blocklist". Most content blocking tools comprise of just an upgraded blocklist (or family of blocklists).

A Firefox extension, 'Request Policy', takes the inverse approach: it blocks all requests from third party domains, except those the user allows specifically. While 'Request Policy' offers complete protection from third party tracking but proper configuration of this extension requires patience and expertise as compare to average computer users.

Chrome, Safari, Mobile Safari and Android does not has a capacity to protect users against " third party content" but Firefox extensions provide this feature to its users and via installing 'block lists', Internet Explorer also provide facility to its users to block contents.

2.2 Review of Related Work

Here we review the closely related proposals with our proposed approach.

2.2.1 Cookie Monster [42]:

This add-on gives option to block all cookies or set temporary rules for various sites. It suits those users who DO NOT want to accept cookies by default, although it is not necessary. It offers choices to set general Firefox setting to block all cookies

or to block third Party Cookies. There is a panel which shows the current status of cookies for the current site and domain when we hover over the cookie status indicator icon on the status bar.

You can allow cookies from specific website by clicking on the 'Cookie Monster' icon on the toolbar or navigating to 'Cookie Monster' with right click on the page. You can choose to accept cookies from the website being referred or just temporarily allow cookies for the current session

It is preferable option for those users who wish to block cookies of individual websites quickly and particularly.

'Cookie Monster' does not address the issue of third party cookie management effectively, as rules one's set for particular domain, stored and that rules follow for that domain without considering the fact that whether domain is acting as First Party or Third Party

2.2.2 Cookie Whitelist

[29]'Cookie Whitelist' as name depicts, categorize the domains and create a list of websites 'Whitelist' from which user can chose to accept cookies permanently or temporarily. This extension gives an option to user to control which website can store cookies in your browser by clicking a 'Button.

By default, Firefox set to reject cookies from all websites. Those websites which user visits more frequently can add in 'Whitelist' by clicking the Whitelist button (+). User have an option to choose cookies permanently from whitelisted website or accept only from current session

User can turn the Cookie button (C) on, for those sites that do not visit very often. Upon turning off the Cookie button, closing the browser, cookies stored during this time will be removed.

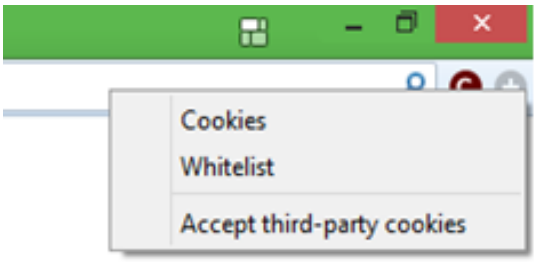


Figure 2.1 Cookie Whitelist

The add-on 'Cookie Whitelist' will be acknowledged by user good option for quick administration of a whitelist and is proposed when the user intends to visit a large number of websites on which he blocked cookies.

2.2.3 CookieSafe

'Cookie Safe' is Firefox extension, which allow you to easily control cookie permission by just clicking on 'Cookie Safe' icon [21]. With this add-on, user can allow, block or temporarily allow the current site to set cookies. Furthermore, user can view, edit or clear the cookies. It utilizes the standard browser cookie management so your settings will remain unchanged regardless of the possibility that you uninstall Cookiesafe.

2.2.3 Extended Cookie Manager

'Extended Cookie Manager' [17] is Firefox add-on but its current version is not compatible with Firefox anymore. Even on the previous versions of Firefox, complaints received from users on the Mozilla add-on reviews.

Users criticized this add-on and were of the view, that it provided inaccurate listings of blocked domains. In nutshell, 'Extended Cookie Manager' has a bad user rating

2.2.4 Ghostery

Ghostery add-on [32] works on Firefox browser and provides a large database of third party contents which updated frequently. When a third party content is find on website, it shows a notification to the user. User then can explicitly block or allow recognized third party, specifically. Ghostery not only control cookies but also control scripts, images and other elements. The main drawback of Ghostery is, it works in binary way i.e., the user can allow or disallow third-party contents, it does not provide fine-grained policy management. It depends on frequently updated database to identify third party contents.

2.2.6 Beef Taco – Targeted Advertising Cookie Opt-Out

Beef Taco [19] use the concept of opt out cookies. The concept of 'opt out' cookies actually works if the ad-networks honor them. Its main drawbacks are, it floods the browser with opt-out cookies and its interface is not configurable and user friendly.

2.2.7 Targeted Advertising Cookie Opt-Out (TACO)

TACO [30] works in the same way as DO NOT TRACK. It automatically blocks not only third party cookies but also many other technologies which can be used by advertisers to track users i.e. "web beacons, bugs.. TACO blocks links to many social websites i.e. 'Facebook Like', 'Google Plus', and others.

It also uses the concept of opt out cookies and permanently set opt out cookies to guard against behavior tracking. But actually, opt out cookies do not completely protect your privacy or prevent you from being tracked by advertisers. Opt out cookies can just request ad-networks to not track users.

So in terms of preserving privacy, this add-on is giving false impression to users that their privacy can be protected by opt out cookies concept which is actually not.

The improved versions of TACO are tracker-blocking and DO NOT TRACK ME.

2.2.8 TrackerBlock

[31] Trackerblock [31]is a Firefox browser add-on and it not only blocks specified tracking companies from setting and getting cookies from your computer but also delete Flash cookies that they may leave on your computer for future identification.

Tracker Block also has an 'opt out' option, if you choose this option, it will write opt out cookies from each tracking company to inform them that you do not want your behavioral tracking used to target ads. When you clear these cookies from your browser, Trackerblock re-write these cookies

Because of a not transparent methodology when blocking in hold, it is not suggested to use.

2.2.9 BetterPrivacy

BetterPrivacy add-on [12] protects tracking by dealing with Local Shared Objects (LSOs) and DOM storage but it does not handle third party cookies which is most frequently used method of tracking. It notify user when new LSOs create in browser and gives an option to user to modify them with an integrated editor. When browser starts or within specific time interval, a user can delete the LSO and DOM storage objects. In BetterPrivacy, only whitelist option is available to the user.

2.2.10 Adnostic

Adnostic [15] is closely related approach to ‘Do Not Track’ therefore its main problem is its reliance on third party ad-networks to respect user privacy . It handles the issue of online tracking with the help of in-browser system that uses ‘homomorphic encryption’ and ‘efficient zero-knowledge proofs’. Adnostic is constructed to prevent advertisers from learning about the user’s actions.

2.2.11 Privad

Privad [16] completely stops ad-networks to track users via third party cookies. It does not trust advertiser servers and inspect each bit of data sent by the user. This detail monitoring slows down its performance. Moreover it does not maintains a balance between the desire of ad-networks to track users and the user's desire to control the amount of information they want to share

2.2.12 Collusion/Lightbeam

Collusion [18] shows the connections between First Parties and Third Parties graphically and shows the tracking done by third-parties in the form of spider-web. The graph shows domains as nodes, data flows as connecting arrows, first parties are recognized by outgoing connections and third parties through numerous inbound connections of different domains. This graph is very complex (as shown in Figure 2.2). It does not fulfills any objective, it is just gives an information about first party and third party domains and tracking.

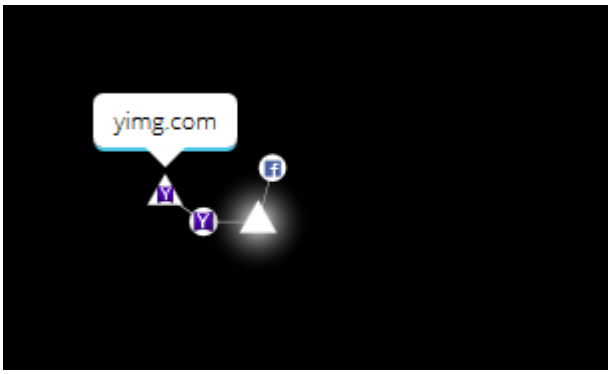


Figure 2.2 Collusion *(www.yahoo.com first party and third party sites are yimg.com & akamaihd.net linked to yahoo.com)*

2.2.13 Firefox's New Cookie Policy

To protect privacy of user on internet, a new third party cookie policy has been announced by [24,25,26], recently. According to [23]:

"1) if a domain is a first-party, it has an ordinary cookie permissions

2) if a domain is a third-party:

a) if the origin already has cookies, it has ordinary cookie permissions

b) otherwise, the origin gets no cookie permissions"

It depicts from this policy that if cookies already present in the browser then it remains unaffected. Moreover, it is instructed that [25], users should delete their cookies to take full advantage of this policy.

If as per policy, third party cookie has permission, no limit has been imposed by Firefox on amount of tracking.

2.2.14 Doppelganger

Doppelganger [20], implements fine-grained, privacy related cookies policies in browsers. But it has few drawback i.e. performance cost due to mirroring of every

action of the user and sending duplicate HTTP request back to the server. To define its policies, an extensive training require and it also depends on the comparison of main and mirrored window by the user. Moreover, this add-on simply blocks all third-party cookies.

2.2.15 Self-Destructing Cookies (for Android and Windows Plateform)

Self-Destructing Cookies [43] instantly removes cookies once they are no more used by current session of browser. This Add-on does not relay on 'blacklist' that also need to be maintained, to identify tracking cookies. It detect third party cookies by their behavior and upon recognition, removes promptly.

This specific add-on fits with blacklist-based solutions such as 'Adblock' as well as 'Ghostery' adequately. It is possible to whitelist web sites whose cookies as well as "LSOs" you desire to preserve without an active tab in the Firefox cookie exception list, that can also be quickly accessed from the add-on's preferences.

3

Privacy is bad for business
NETCOALITION

Third Party Cookie Management Strategy and its Implementation

In this chapter we discuss the detail of proposal of fine grained cookie management and its implementation as a firefox extension.

3.1 Mozilla Firefox Browser and Development of an Add-on

Mozilla Firefox provides the programmer an ability to design an extension based on JavaScript Framework in the form of available API1. After an introduction to the general procedure for the programming of add-ons, this chapter follows the detailed presentation of the implementation of proposal, an add-on, to manage third party cookies as per user choice.

This idea based on the proposal of Julien Freudiger at [1] and enhances its functional scope. We implement this idea as an extension of the Firefox web browser as a proof of concept

3.1.1 Programming an Add-on

There are two types of extensions for Firefox:

- Traditional or XUL extensions which are functionally powerful, but more complicated to build and require to restart Firefox to install
- Restartless extensions do not require to restart Firefox to install but their functionality is limited as compare to traditional extensions.

The core concepts of programming an add-on is ‘Javascript’ and ‘Extended Markup Language’, XUL. XUL creates Firefox interfaces like HTML creates web pages. Further Firefox extensions require a specific internal file structure, detail of which is as under:

The extension is supplied as a XPI, which is compressed to reduce download times. The XPI contains:

install.rdf	Information about the extension
chrome.manifest	Registration data for Firefox etc.
install.js	Installation script
Chrome	Directory containing the extension code
chrome/allcustom.jar	The extension jar
defaults/preferences	Directory containing a preferences file

Table 3.1 List of main files in root directory

Inside the ‘chrome’ directory, there are three directories:

content	XUL, JavaScript and other content that does not depend on the locale or theme
locale	Files for each locale
skin	Files for each theme

Table 3.2 list of files in directories

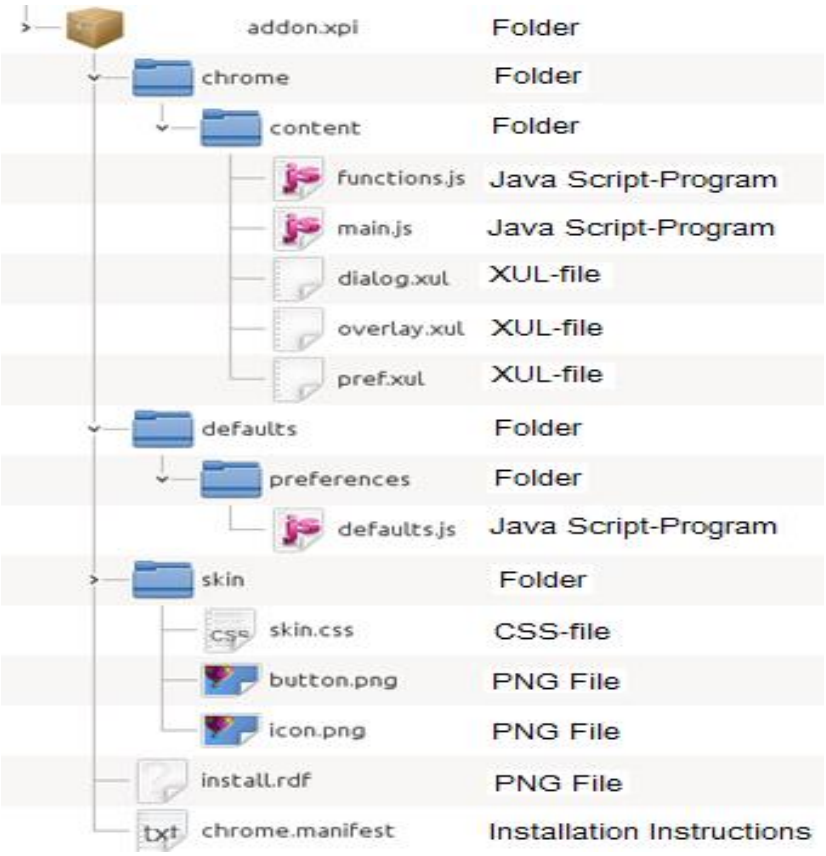


Figure 3.1 sample file of 'install.rdf'

i. 'install.rdf

The 'install.rdf' contains general information about your extension

ii. 'chrome.manifest'

The chrome of Firefox is everything around the content window. i.e. web browser toolbar, menus, status bar etc and all this detail is in 'chrome.manifest' file. This one, works in conjunction with install.rdf the key to how your extension will be added to Firefox, and how it will work.

iii. 'default.js'

The 'default' folder contains 'defaults.js' file, with default values that are used in the installation and can be changed later by the add-on or by the user.

```
<?xml version="1.0"?>

<RDF xmlns="..." xmlns:em="...">

<Description about="...:install-manifest"
             em:iconURL=".../icon.png">

  <!—Name of an Add-on -->
  <em:id>privacymanager@addon.com</em:id>
  <!-- Name, Version, Description, Author -->
  <em:name>privacymanager</em:name>
  <em:version>2.0</em:version>
  <em:description>Description</em:description>
  <em:creator>Faryal Nosheen</em:creator>
  <!-- Settings-Dialogs -->
  <em:optionsURL>.../pref.xul</em:optionsURL>

  <!—Details of the Target Application(Firefox)-->
  <em:targetApplication>
    ...
    <!—Identifier and Version Number -->
  </em:targetApplication>

</Description
```

Figure 3.2 install.rdf

```
pref("extensions.privacymanager.firststart", true);
pref("extensions.privacymanager.username", "none");
```

Figure 3.3 default.js

3.2 Idea of Working of privacookie [33]

Instead of completely blocking third-parties' cookies, the user can choose privaCookie's global configuration regarding how much private information he is willing to reveal. Options for setting the amount of third-party cookie usage are given by numeric values for their "maximum sendings" and their "maximum age". A blocked cookie is usually reset by the third-party with a new value. This enables the user to control amount of tracking done by advertisers.

Now, we give a short overview of privaCookie functionality. The add-on privaCookie monitors incoming and outgoing HTTP traffic for third-party requests. The domain name opened in browser's tab is refered as first-party and all communication from this tab with another domain is classified as third-party. This way of identifying third-party requests works in a reliable manner and does not depend on the initialising resources (scripts, frames and images etc).

In privaCookie, each third-party request is examined for cookies. Incoming cookies are stored and outgoing cookies are compared to a list of already known cookies. By storing the first-party along with its cookie contents, privaCookie ensures that no third-party will get the same cookie key-value pair from different first-parties. Add-on privaCookie additionaly checks the HTTP referrer and shortens it to remove possible tracking or identification strings.

The decision to use a given TP-cookie is based on a cost-benefit analysis that depends on the visited web site and the value of the TP-cookie. To valuate TP-cookies, it uses no spatial tracking across domains and limited temporal tracking policy. It first intercepts all TP-cookies existing/entering the system. If a TP-cookie should not be sent, then the cookie is removed from the existing HTTP request. The reply from the web server will then contain a new TP-cookie that the extension stores in its local table. This implementation does not deal with third party cookies set in Java script, these cookies simple blocked but such cookies are in minority [33].

In this add-on, the storage of cookie is not persistent, so that at each restart of the browser or add-on, the deletion of all currently existing cookies is required for the desired operation to be performed. For example, the login cookies of user is removed and he has to re-login after the start of an add-on on some websites. Also shopping cart items will also be deleted.

3.3 Proposed Solution for Cookie Management:

After the analysis of privacookie and other related add-ons ,this proposal enhances the functionality of 'privacookie' and gives more precise control to users for management of third party cookies. We proposes the use of three lists for managing third party cookies as per user choice i.e. whitelist, blacklist and trusted third party list.

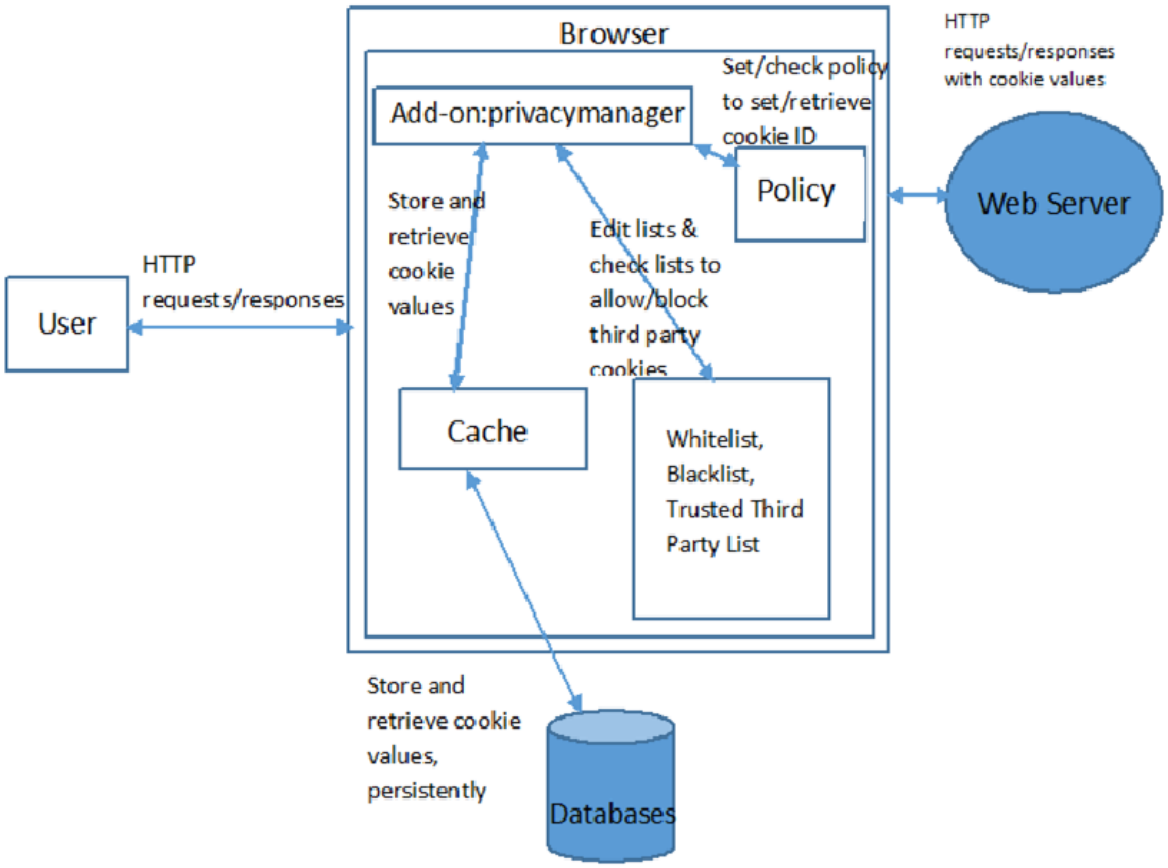


Figure 3.4 Add-on 'privacymanager'

3.3.1. User interface:

As contrast to privacookie, our add-on provides a user friendly interface . A user can add a toolbar button (see Figure. 3.5) to interact with add-on. It provides enabling/disabling option, state indicators and 'context menu' option in the form of drop down menu.

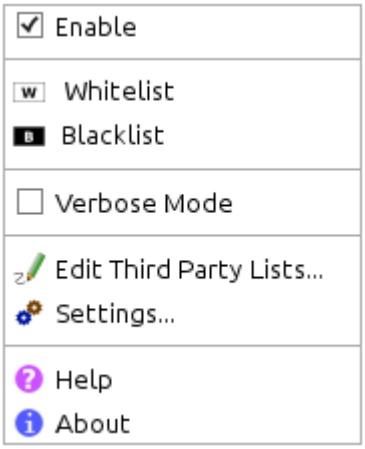


Figure. 3.5 Toolbar Button along with Context-Menu

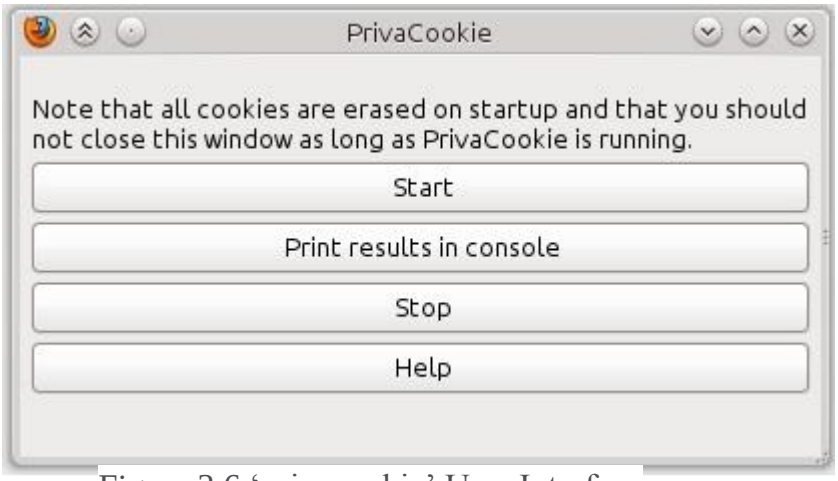


Figure 3.6 'privacookie' User Interface

The user interface of privacookie is shown in Fig 3.6:

3.3.2. Cookie Storage

In privaCookie, cookie storage is not persistent and each restart of the add-on requires a deletion of all available cookies to ensure its functionality. This will also delete cookies required by the user, e.g., login cookies or shopping baskets. Our add-on stores cookies persistently in a database. The add-on creates add-onnamehere.sqlite in the user's profile directory to save cookie information. One of the challenges we have faced is performance issue because of continuous querying the database during normal browsing. We overcame this challenge with the help of a cache layer which loads all entries once from the database and stores in memory for further operations. All changes are written back to the database when browser close.

3.3.3 Privacy /Traceablitiy Trade-off

The add-on privacookie manages all third- party cookies in a privacy-preserving manner.

According to privacookie:

"Our solution enables advertising to have differentiation capabilities without allowing for excessive tracking of users online."

The privacy/traceability trade-off as depicted in privaCookie can be seen in Figure 3.8.

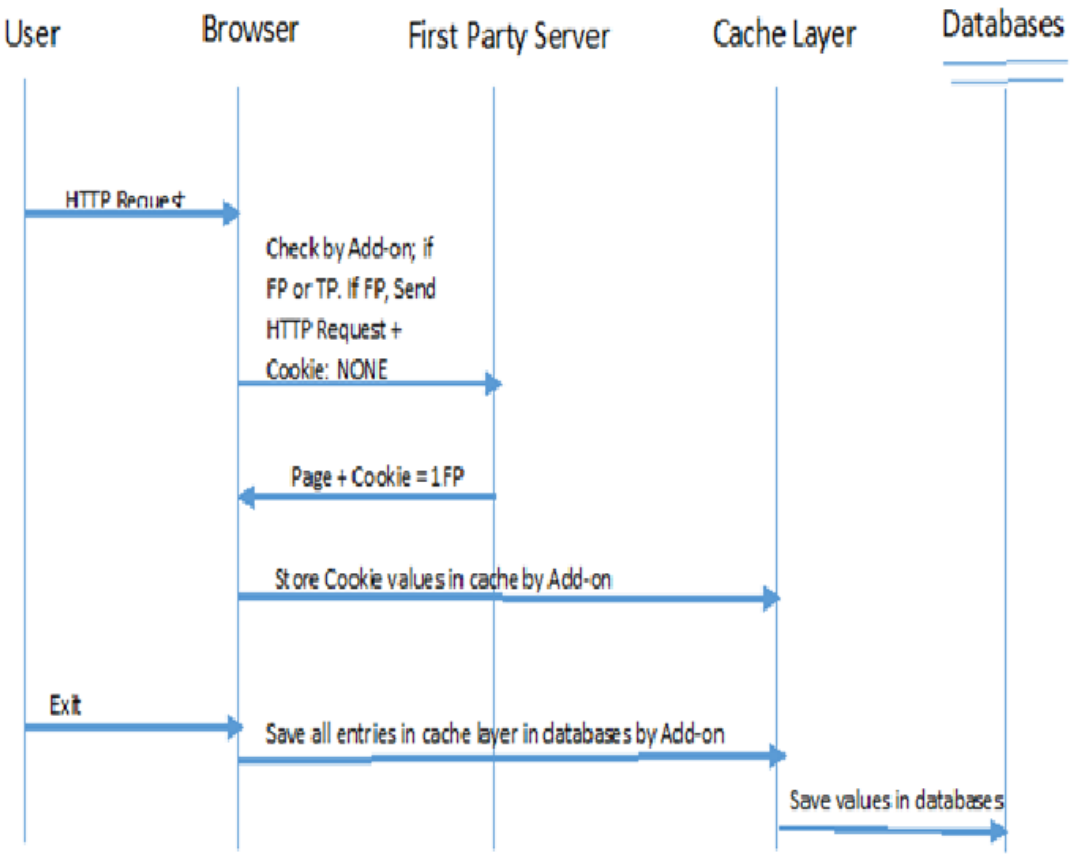


Figure 3.7 'Cookie Storage

At position (1,0), the third-party cookies management policy al- lows for the complete tracking of users online and at position (0,1), blocking all third-party cookies impedes online tracking by third parties. In our proposal, we have extended this trade-off by adding trusted third party lisrt functionality (see Figure. 3.7).

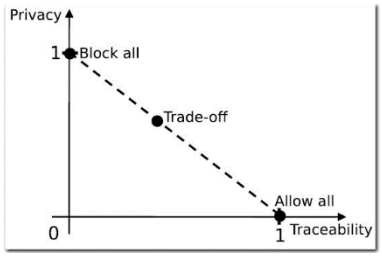


Figure 3.8. privaCookie privacy/traceability trade-off

3.3.4 Spatial And Temporal Tracking Policy

For consistency, we have chosen the same notation as described in privaCookie. It defines the set of first-parties as B with elements bi, while the different indices

represent distinct domains.

Third-parties are similarly described with D and dj. Cookies are contained in set C and identified by ci,j with the first index representing the first-party domain and the second index representing the third-party domain. Cookies and domain names are available in the browser history set i.e., H . There are two tracking policies defined in privaCookie [33]:

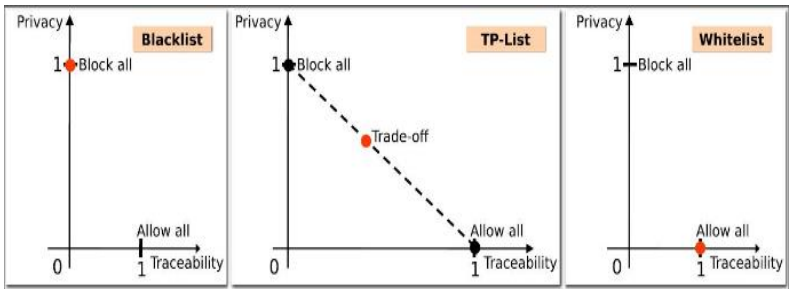


Figure 3.9. Privacy/traceability trade-off

- **Spatial Tracking Policy:**

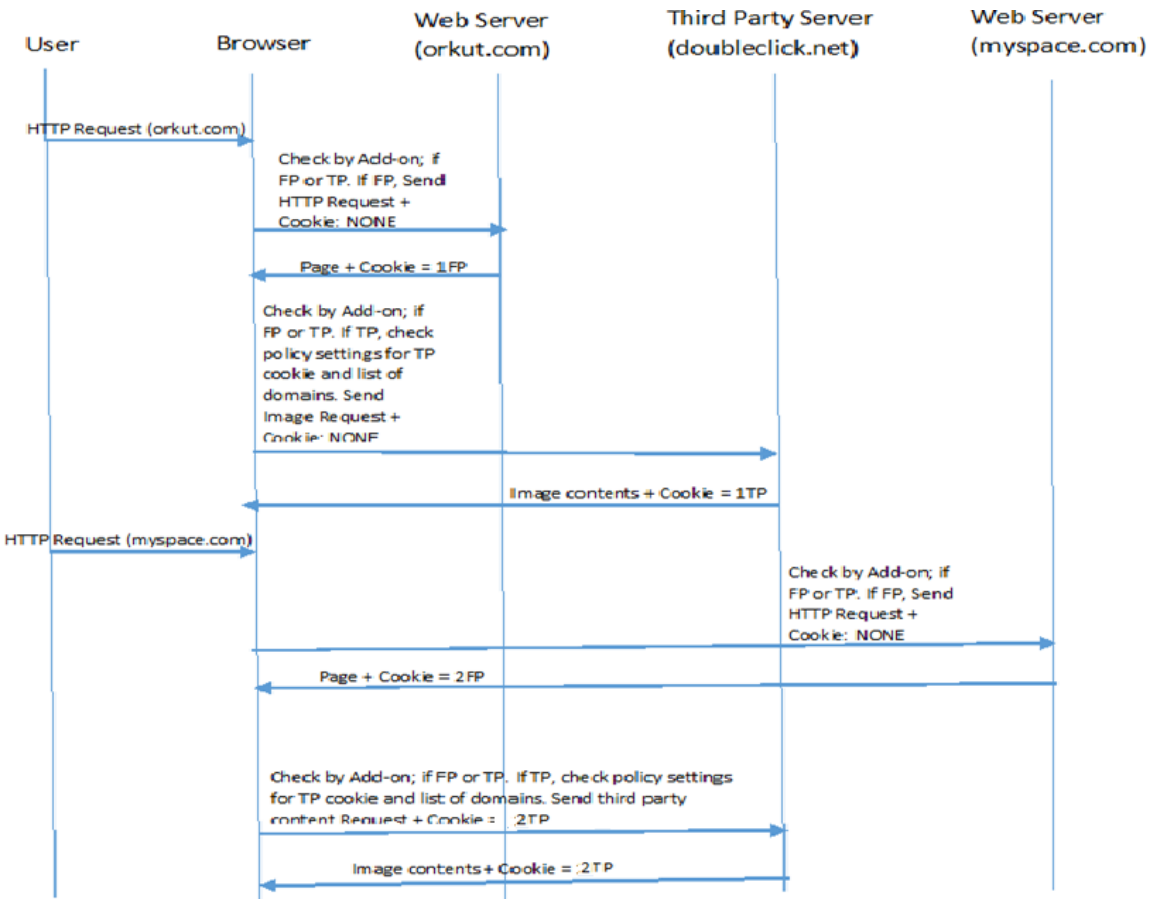
In order to protect user's privacy when a user visits multiple first-party domains that need to send cookies to the same third-party advertiser, new cookies will always be sent. The implementation of privaCookie only supports the hard coded value of LS i.e. LS = 1 where LS is a spatial tracking limit/threshold

$$\sum_{b_m \in H(B)} \beta(b_m, d_j, c_k) \leq 1$$

In case of our proposal, the value of LS depends upon user. User can choose the number of first-parties sharing the same third-party cookie.

$$\sum_{b_m \in H(B)} \beta(b_m, d_j, c_k) < L_S$$

User can set value of L_S



(Value of Ls =1)

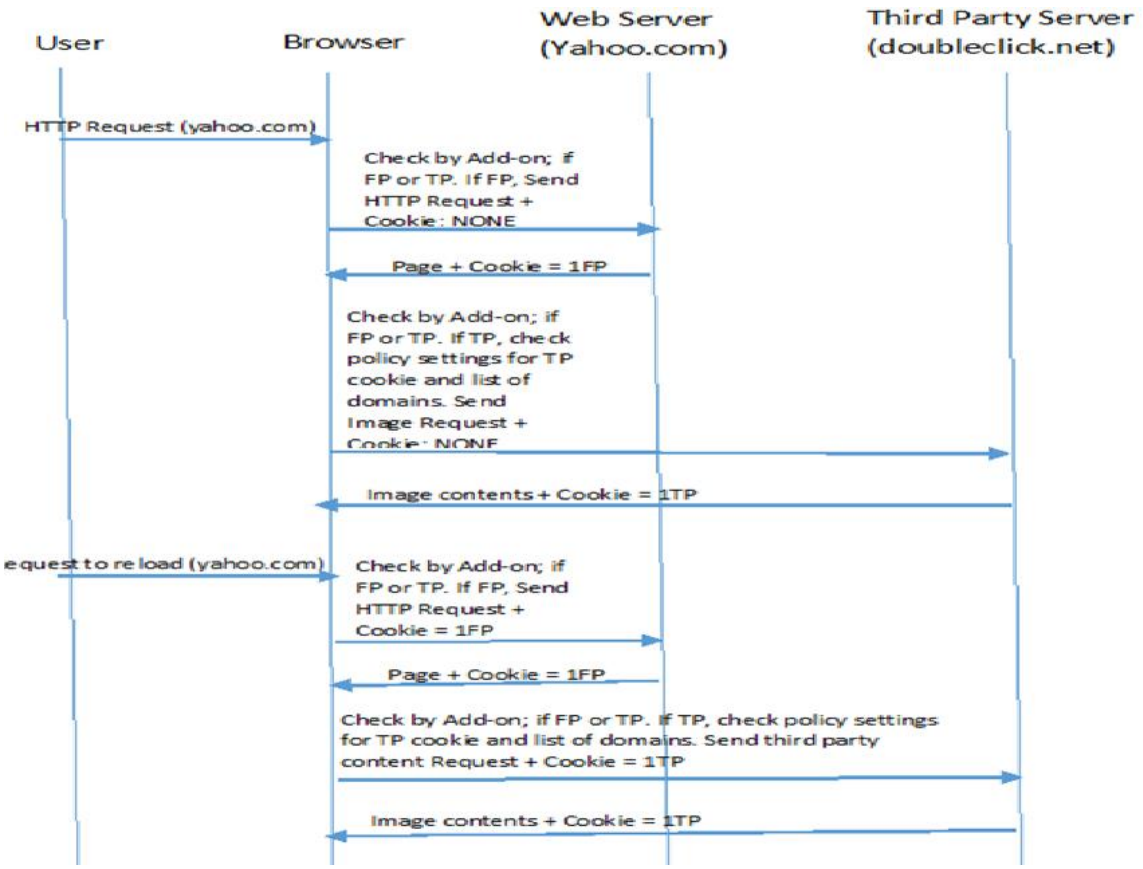
Figure 3.10 Spatial Tracking Policy

- **Temporal Tracking Policy:**

When a user visits a specific website, the same cookie to the same third-party advertiser is used only for a specified time period Lt or for a limited number of site's visits Lv

The above definition can be written in same form for LT. In our add-on, we have implemented this policy in the same spirit as given in privaCookie.

$$\sum_{b_m \in H(B)} v(b_m, d_j, c_k) < L_V \text{ Or } L_t$$



(Value of Lt=1)

Figure 3.11. Temporal Tracking Policy

3.4 Programming an Add-on 'privacymanager'

The distribution of add-ons is usually done in the form of an XPI file which can be exported in 'ZIP' format. The folder structure of the extension is shown in fig 3.8. In the main root directory, there are two text files i.e. 'install.rdf' and 'chrome.manifest' and sub folders 'chrome', 'default' and 'skin'.

3.4.1 XUL and Java Script Files

The actual code to run add-on, stored in Java Script format, in files functions.js and main.js. Firefox's user interface is written in XUL and JavaScript. XUL is an XML grammar that provides user interface widgets like buttons, menus, toolbars, trees, etc. User actions are bound to functionality using JavaScript. These files will be stored in 'chrome/content' folder within root folder.

i. Toolbar-Button (main.js/overlayDup.xul)

This file defines the central control function for the execution of add-on. It includes the initialization functions for starting and switching the operation status, update features for checkboxes and status indicators, monitoring functions for the response to change in settings and loading functions to dialog windows as well. The different XUL files provide for an integration of the add-ons with the Browser. For example, contains overlay.xul

ii. Core Functions (core.js)

The core component of privacymanager is realized as an object in core.js privacymanagerCoreObj, which is instantiated in main.js. It links the individual modules through centralized access to Third Party lists and log functions. Reading and saving text files as well as the application of statistical information is also managed by privacymanagerCoreObj.

iii. Cookie Storage (Storage.js)

From privacymanager detected cookies should be stored in association with the affected First Party, as well as information about the age and frequency of use, persistent. For a fast and flexible access to these data sets lists the add-on the SQLite database privacymanager.sqlite in the profile folder of the user (see Table 3.3)

id	Fp	tp	name	Value	date	sent	totalsent	changed

Table 3.3 privacymanager.sqlite

The CookiesEnvoyes () routine notifies the database interface each time you send a cookie by calling the privacymanagerCore.cookieDB.sendCookie function (fp, tp, name) and the numerator of the respective database entry be increased accordingly. But a direct and continuous access to the database led to the development phase of a severe performance problems.

These are dissolved in storage.js by an intermediate level as a cache for the data. Each entry is removed once from the database for reading and for the future use cached. All changes made during run time copied to the cache and would store in the database once the browser closed

Figure 3.12 list of files of Add-on ‘privacymanager’

```
var cookieDB = privacymanagerCore.cookieDB;
var tpLists = privacymanagerCore.tpLists;
var nvlEntete = "";

//Through Cookies of request
for (i in cookEnvoyes) {
  var cName = cookEnvoyes[i].name;
  var cValue = cookEnvoyes[i].value;
  var cookie = cookieDB.getCookie(fp,tp,cName);

  //Blacklist check
  if (typeof(tpLists.black.domains[tp])
      != 'undefined') {
    logConsole([...]); //Cookie blockiert
  //Whitelist check
  } else if (typeof(tpLists.white.domains[tp])
             != 'undefined') {
    nvlEntete += addCookie(tp,cName,fp,cValue);
    cookieDB.sendCookie(fp,tp,cName);
    logConsole([...]); //unmodified sent
  //Policy check
  } else {
    //Number of Sending
    if (cookie.sent >= envoisMax) {
      logConsole([...]); //sent too frequently
    //Age of Cookies
    } else if (cookieDB.getDate()
                > (cookie.date + tempsMax)) {
      logConsole([...]); //too old
    //Number of Common First Parties
    } else if (cookie.equal > share) {
      logConsole([...]); //
    } else { //No violation of Policy
      cookieDB.sendCookie(fp, tp, cName);
      nvlEntete += //Unmodified sent
        addCookie(tp, cName, fp, cookie.value);
    }
  }
}
//Modified Cookie Return
return nvlEntete;
```

Figure 3.13 duplication.js

```
<?xml version="1.0"?>
<?xml-stylesheet
  href="chrome://global/skin/"
  type="text/css"?>
<?xml-stylesheet
  href="chrome://projet/skin/skin.css"
  type="text/css"?>
<window title="privacymanager - Log"
        name = "privacymanagerLog"id="privacymanagerLog"
        xmlns="http://www.mozilla.org/..."
        onfocus = "autoScrollStop();"
        onblur = "autoScrollStart();"
        onload="startLogging();"
        onunload="cancelLogging();"
        height="300" width="600">
  <script src="helper.js"/>
  <script src="logger.js"/>
  <vbox flex = "1" id = "privacymanagerLogger">
  </vbox>
  <hbox>
    <button label = "Clear Console"
            image = "icon-console.png"
            id = "clearButton"
            oncommand = "clearLog();"/>
    <spacer flex="1"/>
    <button label = "Deactivate Verbose Mode"
            image = "icon-consoleclose.png"
            id = "closeButton"
            oncommand = "cancelLogging();"/>
  </hbox>
</window>
```

Figure 3.14: ./chrome/content/logger.xul

iv. Integration with 'privacookie'

The JavaScript implementation of PrivaCookie found in the file duplication.js, which, in modified form, as one of the components of core compo-privacymanager for further use.

All functions were remained unchanged, adjusted only the code related to the external visible interface to make it identical with privacymanager add-on (as shown in Fig 3.13).

vi. Third Party Editor (editor.js/editor.xul)

The window for the TP-editor (see Figure 3.17) is created by means of an XML file (editor.xul). To produce lists the element providing the functionality is used here listbox and the scripts helper.js and editor.js are involved.

When privacymanagerCookieCore.tp object initialized, editor reads the list box of whitelist and blacklist and stored temporarily.

There is an option to shift elements of one list to neighbor list by itemMove () to the correct alphabetical position in the target pasted list and deleted from the source list

vii. Settings (pref.js/pref.xul)

In pref.js, monitor the setting using addObserver () listener. In case of modification via setting dialog by the user or add-on itself, PrefsWatcher.observe () function transfers the amendments in the corresponding variables of privacymanager and where appropriate, calls other functions for processing the updates. For this purpose, uses the instance of nsIPrefService getter and setter functions for properties of the data types int, string and bool, which are also used by other components of the add-ons to access setting information.

A dialog for easy adjustment of settings is defined pref.xul in the XML document and opens when called from openPrefs () (in main.js) in a separate window.

vi. md5.js

In md5.js [37] of the MD5 hash algorithm is implemented in JavaScript and is used by privacymanager to generate a unique ID in a cookie stored in the database privacymanager.sqlite. For this purpose, the hash MD5 ("FP: TP: COOKIE NAME") formed.

[37] MD5-Implementierung in JavaScript http://www.webtoolkit.info/javascript-md5.html

3.4.2. Design of an Interface:

The aim of the development of ‘privacymanager’ is a practical and applicable add-on, which is why, in addition to the core technical components, the aspects of user interface design must be considered. Thus begins the design of the add-ons with the toolbar buttons, menus and dialogs.

```
<?xml version="1.0"?>
<prefwindow id="duplication-prefs"
     title="privacymanager - Settings"
     xmlns="http://...">
  <prefpane id="duplication-stock-pane">
    <!—Referencing the Settings -->
    <preferences>
      <preference id="pref_envois" type="int"
          name="extensions.privacymanager.envois"/>
      <preference id="pref_stats" type="bool"
          name="extensions.privacymanager.stats"/>
    </preferences>
    <vbox>
      <!—Adjustable Integer value -->
        <label control="envois" value="Maximum
        number of times a third-party cookie  can
        be sent:"/>
      <textbox preference="pref_envois"
              id="envois"
              type="number" min="1"
              decimalplaces="0" maxlength="4"/>
      <!—Reversible Boolean Value -->
      <checkbox preference="pref_stats" id="
        stats"/>
      <label control="stats" value="Collect
        Statistics"/>
    </vbox>
  </prefpane>
</prefwindow>
```

Figure 3.15 /chrome/content/pref.xul

3.4.2.1 Toolbar-Buttons:

It is the main controlling element and provides activating/deactivating options along with "state" indicators, "action" indicators and "context menu" option in the form of drop- down arrow. The add-on's state indicator is either "active (green color)" or "inactive (red color)". The "action" indicators start blinking when a certain action takes place (as shown in figure 3.16)

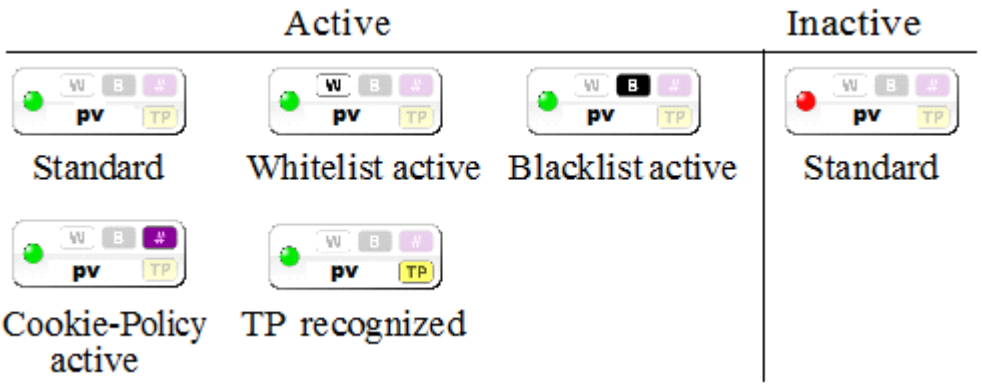


Figure 3.16 Toolbar-Buttons

The "action" indicators are:

• W: means whitelist third-party detected and that party can receive the cookie.
• B: means cookie has been blocked because the advertiser is part of user's blacklist. By default, whenever a new third-party detected, we added advertiser to the blacklist.
• TP: means trusted third-party list and the cookie has been sent if user-defined policy permits.
• #: When # sign starts blinking, it means that a cookie has been used too many times or is too old to be resent and therefore blocked.

The "context menu" provides options like "help", "verbose mode", "show whitelist", "show blacklist", "edit third-party list", "settings dialog" and "about" (as shown in Figure 3.4).

3.4.2.2. Context Menu

The context menu provides the user with access to advanced features as well as settings and information about the add-on.

i. Enable:

The checkbox behaves exactly the same direct click on the toolbar button and is needed by the user when 'privacymanager' not add to its toolbar and liked to use exclusively via the browser menu.

ii. Whitelist/Blacklist:

A message box with a list of the currently selected list is displayed to give the user a brief overview of settings

iii. Edit Third party list

We present a third-party list editor to the user (see Figure.3.17). The decision to

use the third-party cookie across different websites depends on lists (whitelist, blacklist and trusted third-party list). The user can use the arrow buttons to move entries in respective lists. The third-party editor dialog also provides an option to download predefined lists of advertisers. The downloaded entries are merged into the existing third-party lists. In order to develop, predefined lists, we have used Privacychoice (http://privacychoice.org/).

- Whitelist: The advertisers in this list get the third-party cookie unmodified.
- Blacklist: The advertisers in this list never get the third-party cookie.
- Trusted Third-party List: The advertisers in this list get the third-party cookie but according to the user's settings.

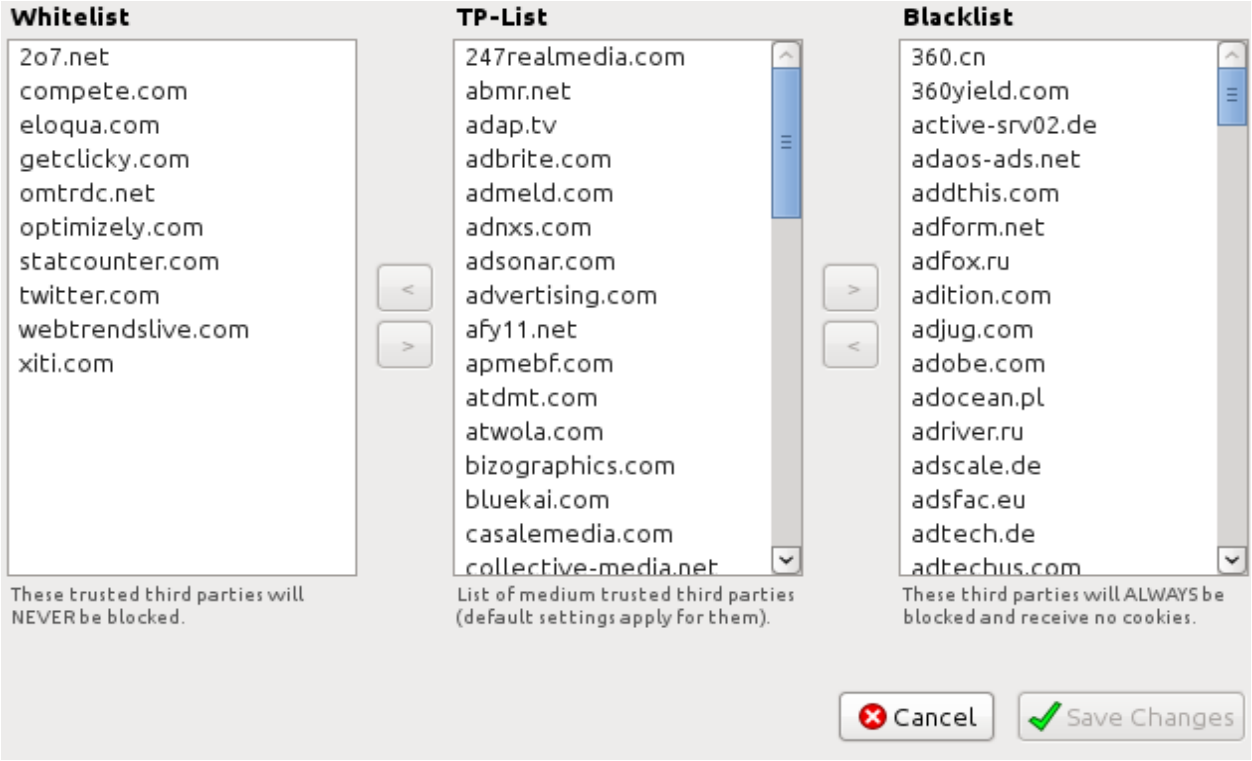


Figure 3.17 Third Party List Editor

iv. Settings

We also present settings dialog to the user so that the users can control the amount of tracking (see Figure. 3.18). More specifically, the following fine-grained policy management options are available:

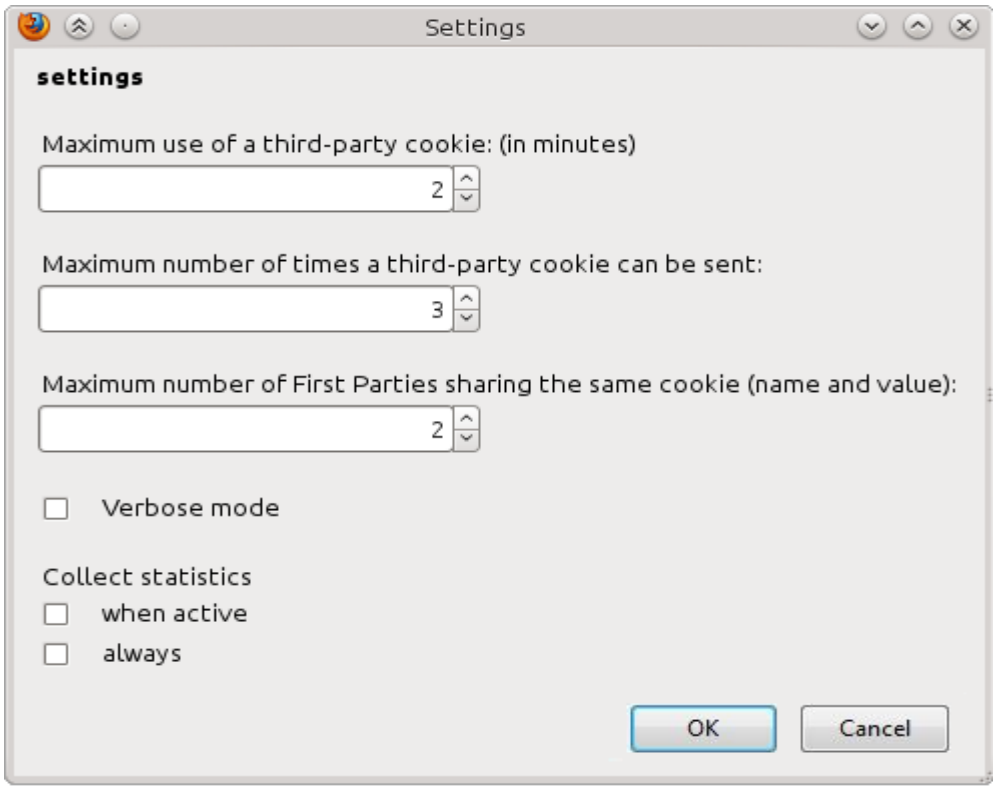


Figure 3.18 Settings Dialogue

• **Temporal Tracking**: In order to provide temporal tracking policy, the same cookie to the same third-party advertiser is used only for a limited number of times, when a user visits a specific website.

• **Spatial Tracking**: In order to provide spatial tracking policy, that is, to protect user's privacy when a user visit multiple sites that need to send cookies to the same third-party advertiser, new cookies will always be sent if the third-party advertiser is not in the whitelist and if policy permits.

v. Verbose Mode

Upon clicking 'verbose mode' via context menu the log console opens (as shown in Fig 3.19). The client may receive information to any manipulation made by 'privacymanager'. It shows detected third party cookies. All entries are chronologically sorted, grouped according to the time of occurrence and show the names of the parties involved, and optionally the name of the affected cookies. The button Clear Console "is for deletion of the entire previous detail.

The console log appears in a separate window. The window for the console log is created by means of an XML file (logger.xul, source code is shown in Fig 3.10) and its functionality encapsulates in logger.js file.

A call to privacymanagerCore.logMessage () returns status message in the form of an array, which is supervised by the log console output.

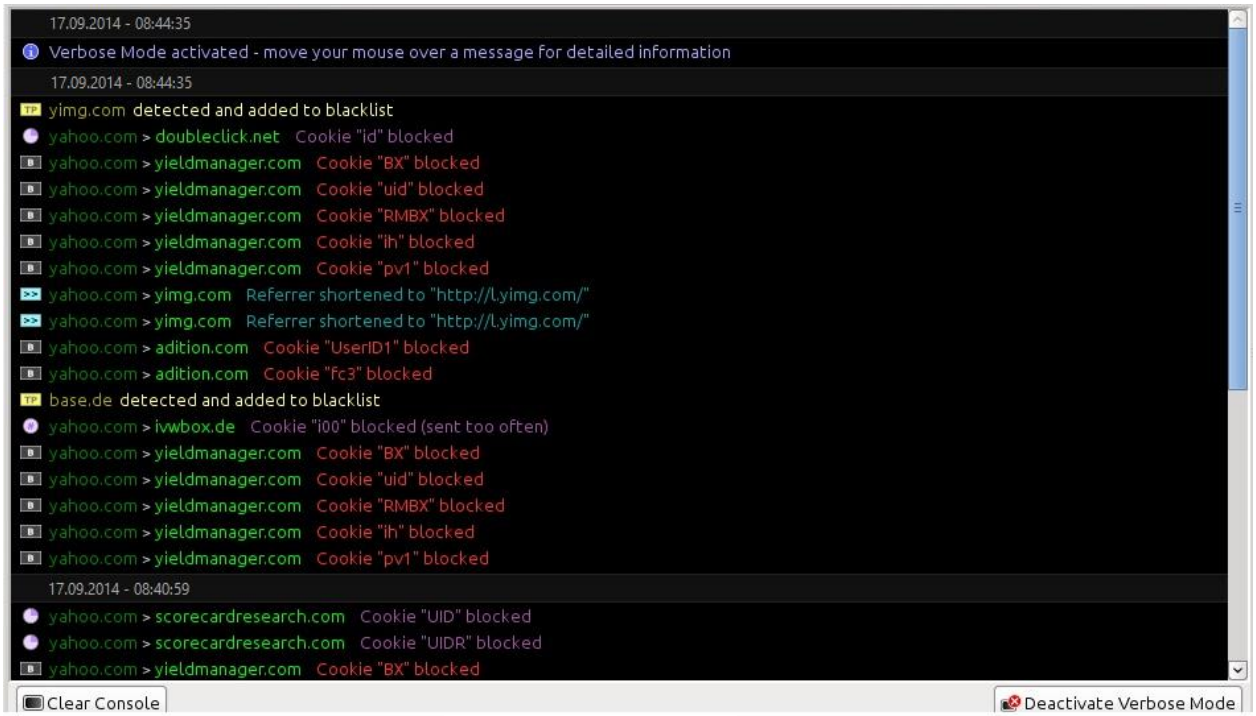


Figure 3.19 Log Console

vi. Help

A detail information about working of ‘privacymanager’ will open in separate browser

vii. About

A dialog box appear with concise information about author and other detail of ‘privacymanager

3.5 Function Test

During its development the add -on ‘privacymanager’ works in different versions of the Firefox browser (14.0.1 , 15.0.1 , 16.0.1 , 16.0.2) and under different operating systems (Windows 7 Home Premium , Linux 12.04 (KDE 4.8.5 / kernel 3.2.0))

and optimized. At the completion of the development the add -on caused in these environments no unexpected messages in the Firefox error console and is therefore classified as a stable version.

The correct observance of the user adjustable, policy in the form of black and white list values were verified by tests. For this purpose, test performed by using www.google.com. First it was placed in whitelist, it could be reloaded any number of times without changing the value of the cookie. The placement of www.google.com in blacklist caused no content for the cookie indicated and it was therefore never transmits to TP server.

To ensure that the rules of the default policy, put www.google.com in trusted third party list and set the following values:

- "Maximum age of third party cookie = 1 minute
- Maximum dispatch of third party cookies = 3 time
- Maximum number of First Parties sharing the same third party cookies = 2"

After loading the First Party-page, with a waiting period of 1 minutes, reload the page. There was no value for the third party cookie, hence policy interpreted successfully.

The following multiple reloading within a short period of time (< 1 minute) showed that the content of the cookie, after three ads, was not visible (in this case the cookie was blocked because it was sent too often) and when next load had a different value (the TP-server had set it new).

4

Privacy is suffering the death of a thousand wounds
SIMSON GARFINKEL

Analysis

Relevance and successful application of 'privacymanager' be verified by the analyzes in this chapter. There is the automated visit of 5,000 sites and the evaluation of the data obtained via third-party cookies as well as the relations between Third Parties and First Parties. In this chapter, applied methods are presented and the results illustrated by descriptions and graphical representations

In present scenario, an add-on for the management of TP-cookies is necessary and useful, when a considerable of amount of privacy information sharing by websites with advertisers.

To prove that a large part of the First Parties uses TP cookies, as well as to identify the most relevant Third Parties, the Top 5,000 Domains (worldwide) of Alexa Rankings automatically called in this work and by using FourthParty tool [34] recorded in an SQLite database.

For evaluation of collected data in different formats, it is transferred to SQL database tables. Through SQL queries in PHP scripts we do analysis and comparison, using this database

4.1 Web Analytics

Web analytics is not just a tool for measuring web traffic but can be used as a tool for business and market research, and to assess and improve the effectiveness of a web site.

[35] have found several advantages to placing web measurement at the center of our methodology.

i. Web measurement provides objective, reliable evidence that both furthers public understanding and establishes a sound basis for policymaking.
ii. Web measurement is fast.
iii. Web measurement facilitates longitudinal study. Often the very same hardware and software can be reused to collect and analyze data even years apart.
iv. Web measurement can often be automated. Once a generic measurement tool has been built, it can be trivially applied to millions of websites.

4.1.1 Web Measurement Tool-Fourth Party

Since Firefox is a modern browser used by many users on the web and also provides a strong plugin infrastructure. This allows us to use FourthParty, a web measurement tool built by the Stanford Security Lab, which serves as our primary method for logging browser data

FourthParty is a Firefox extension which monitors all dynamic web content on the browser, including HTTP requests and responses, cookie interactions and JavaScript calls. In particular, we hooked into Fourth- Party's HTTP request and response tables to determine which cookies are transmitted under each page load, and the cookie table to detect and monitor unique identifiers. FourthParty's SQLite database enables quick aggregation and searching of cookie values through standard SQL. FourthParty offers usability benefits over a pure HTTP

We use FourthParty tool to conduct analysis along with custom designed crawler.

4.1.2 Fourth Party Tool-Design

It was developed around three design principles [35]:

1)General-purpose instrumentation:
By implementing comprehensive instrumentation and logging only once, FourthParty avoids the need for many purpose-built tools, decreases duplication of effort, and trims development time.
2) Production web browser:
Building on a production browser allows reuse of existing add-ons, including for automation, and closely emulates real-world browsing.
3) Standardized log format:
A standardized, easy-to- manipulate log format facilitates data sharing and cuts back on redundant data gathering

4.1.3 Analysis Using Fourth Party Tool

Analyzing FourthParty data is fast. All of the FourthParty results generated for analysis of proposed approach in this thesis, most of which took seconds to execute with databases including visits to thousands of popular websites. Analyzing FourthParty data is also easy for researchers who are already familiar with SQL syntax.

4.2 Methodology:

We drive our data collection and measurement infrastructure (see Figure 4.1) using a web crawler and fourth party tool

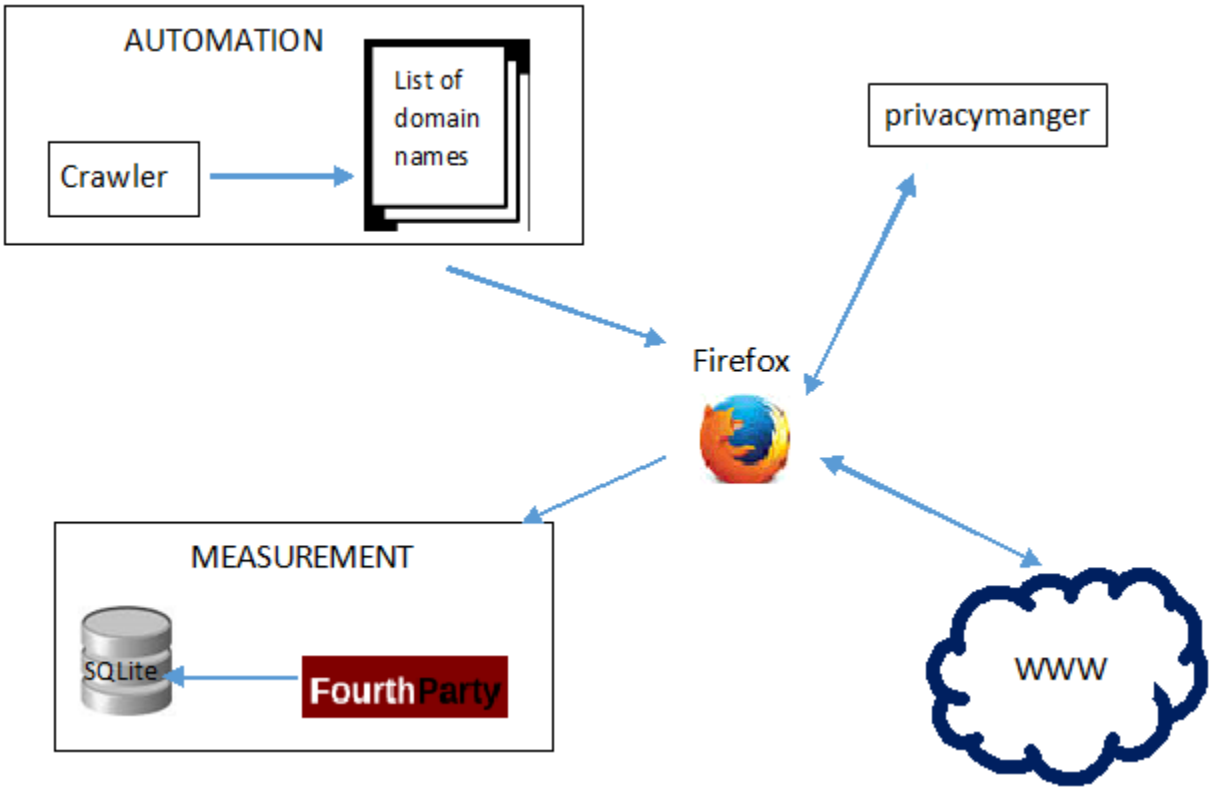


Figure 4.1 Data Collection and Measurement Infrastructure

4.2.1 Installation of Fourth Party tool

Download latest fourth party tool from http://fourthparty.info/get_started in 'zip' format and unzip the files in a folder named 'fourthparty'. To activate fourth party and start with firefox, run the following commands:

- ☆ cd <fourthparty-folder>
- ☆ cd addon-sdk/addon-sdk-1.0b5-fourthparty
- ☆ source bin/activate
- ☆ cd ../../../extension
- ☆ cfx run

Now when we do browsing on firefox, FourthParty automatically generates a logging database, fourthparty.sqlite, in the current profile. *Before* closing Firefox copy the FourthParty database to another directory. Closing Firefox will delete the temporary profile, including the FourthParty database.

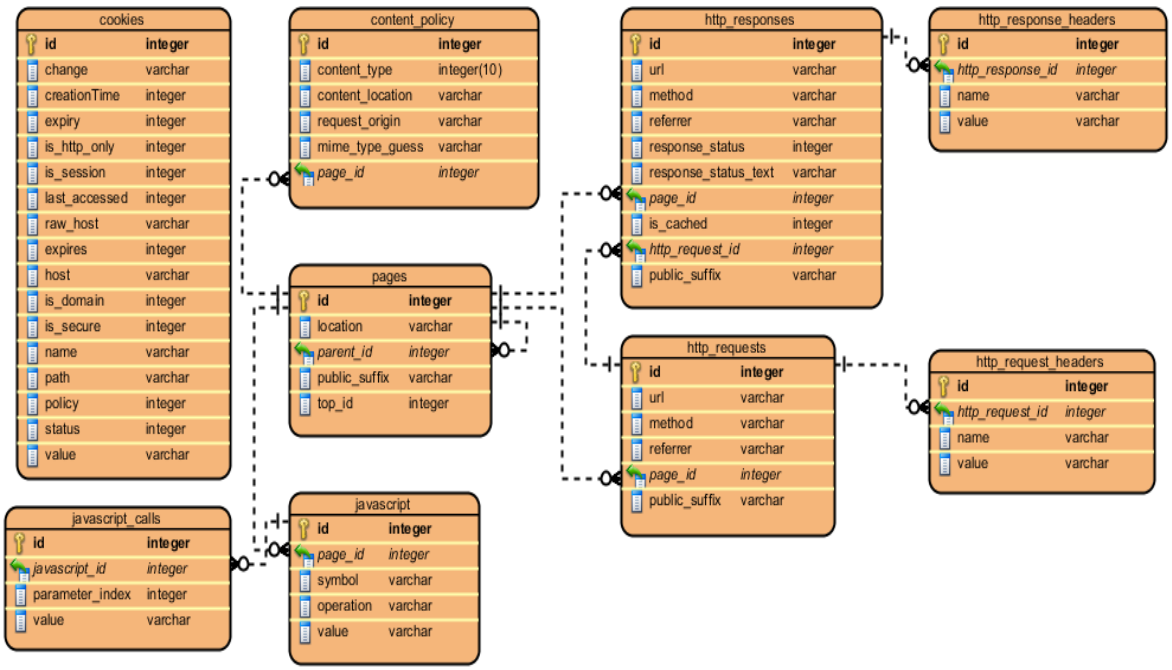


Figure 4.2 Fourth Party SQLite Database Schema

4.2.2 Crawler:

To automatically visit top 5000 web sites for empirical study, a crawler is require. There is a list of crawler suggested on fourth party website for automation of visit of websites like MozMill, JSSh, Selenium, or Watir but all these crawlers are complex to integrate. Therefore we design a small addon for firefox browser which automatically visits a given list of 5000 web sites and fourth party maintains automatic databases at background. A

detail of design and implementation of this crawler is omitted here (related to previously presented design and implementation detail of ‘privacymanager’ add-on). What is relevant is that the user by means of a toolbar button, provided by the add-on, can open a box to select a file name.

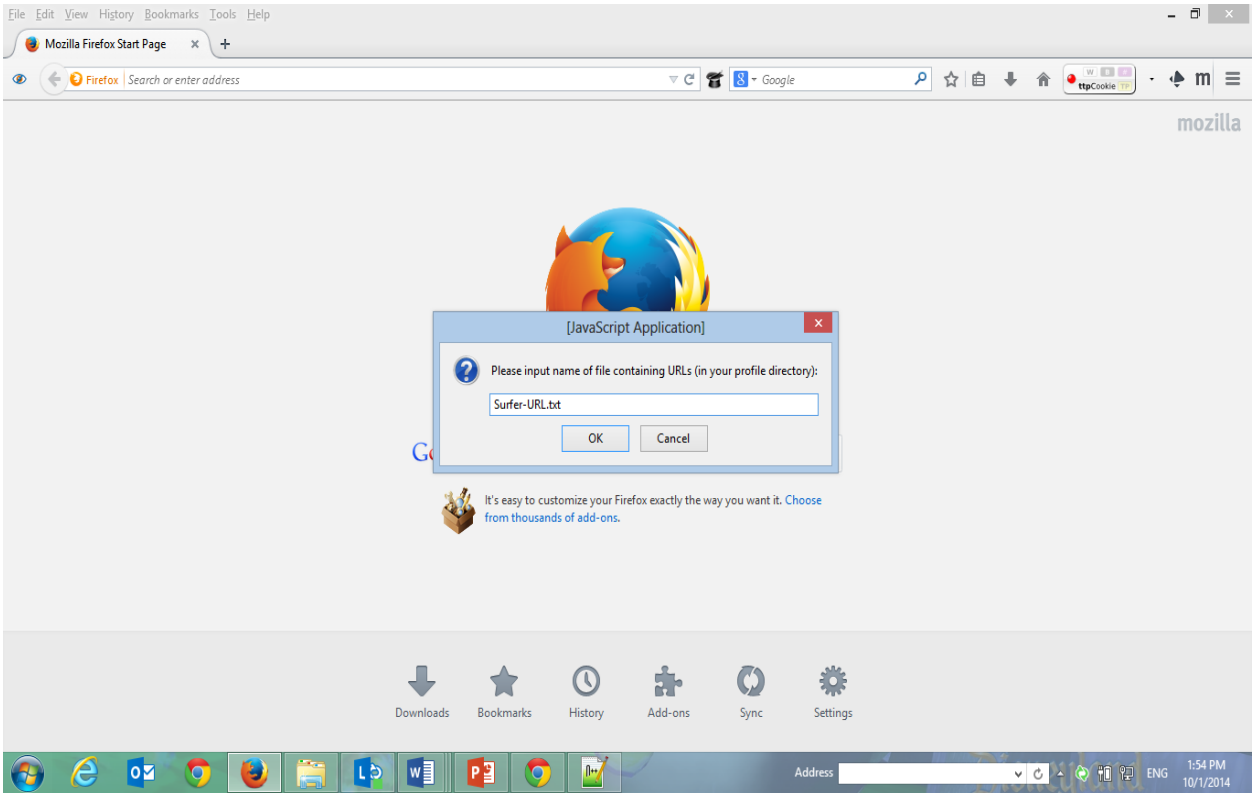


Figure 4.3 Web Crawler Interface

A user to select a file name of ‘.txt’ file, which contains list of URLs. The file is read from crawler and each row contained interpreted as a URL that should be invoked in the current tab after maximum 10 seconds. After complete loading a page (or after a maximum of 10 seconds), it navigates to the next entry. During the stepwise processing of the list items in the currently loaded site, a progress bar is displayed to the user about the number of visited sites e.g. ‘333 of 5000’

4.3 Automation and Measurement

After launching Firefox using Fourth-party tool (for installation instructions see section 4.2.1) and installing crawler (drag and drop .XPI file in the browser) & add-on ‘privacymanager’, we need to restart a browser.

Enable add-on ‘privacymanager’ and set the following values to test run:

i. Maximum age: 2 minutes
ii. Maximum dispatch of third party cookies: 3 times
iii. Maximum number of first parties sharing third party cookies: 2

Downloaded a list of one million most visited websites by Alexa, provided by [44]; extracted top 5000 entries and separated as a text fil, save this text file in the profile folder of current Firefox instance and used by crawler as a URL list to call the sites automatically.

All collected data is retained from the temporary profile directory of FourthParty tool SQLite databases and transferred into SQL database tables to facilitate further analysis. A list of source files used for data analysis is mentioned in table 4.1.

We have opened SQLite databases using Sqliteman [36] program and export in SQL command and create tables using phpmyadmin [37] for data analysis. The fourth party.sqlite database contains several tables whose names were maintained and keep with the prefix fp e.g. fp_content_policy, fp_cookies,fp_http_requests, fp_http_request_headers, fp_http_responses, fp_http_response_headers, fp_javascript and fp_javascript_calls

Firefox-Profile Folder	
cookies.sqlite	Firefox-cookies
fourthparty.sqlite	HTTP-Requests/Java Script
privacymanager.sqlite	privacymanager-cookies
webappsstore.sqlite	DOM storage
privacymanager-Statistics-YYYY-MM.txt	log file of privacymanager

Table 4.1 Firefox Profile Folder

For further use of statistics of ‘privacymanager’ there was a conversion of the structured text format into SQL commands (using a PHP script) and their direct execution for the transfer of entries in the database tables.

4.3.1 Identification of Third Parties

Using the SQL Table, by means of an SQL query (1), all TP-domains that were contacted during the test run by First Parties, be selected in alphabetical order. The number of third parties is determined by an additional query (2)

(1) SQL> SELECT DISTINCT tp FROM statistics ORDER BY tp;
(2) SQL> SELECT COUNT(DISTINCT tp) FROM statistics;

The result is a list of 1,200 third-party domains. The comparative analysis of the SQL table 'cookies' (i.e. the internal memory of Cookie Firefox) brings in a query (3) for comparison with the SQL table 'domains'. During comparison, we have ignored already known 5000 first-parties. As a result, we found 909 distinct domains acting as third-party only. This turns out that approximately 25% of third-parties are also acting in the role of a first-party and that's why they belong to the list of 5000 visited domains.

(3)SQL> SELECT DISTINCT baseDomain FROM cookies WHERE baseDomain NOT IN (SELECT domain FROM domains);

4.3.2 Inspecting Third Parties

We ordered our list of identified third-parties by their absolute usage count in distinct first-party websites (see Table 4.2 and by the relative proportion of the total number of requests

4.3.2.1 Frequency of Use by First Parties

To analyze the frequency of use in relation to the total number of First Parties and Third Parties, all different TP domains from the SQL table 'statistics' are first selected (4) and then in a PHP loop (5) for each entry, the number of different FP domains determined.

(4) SQL> SELECT DISTINCT tp FROM statistics;

(5) PHP> foreach (☆tpList as ☆tp) {
☆q = mysql query("SELECT COUNT(DISTINCT fp) FROM
statistics WHERE tp = '☆tp';")
}

We found out that all identified third-parties are used from 20 – 1130 distinct first-parties

(see Figure 4.4). The list leader is doubleclick.net which covers ≈ 23% of our website visits, during automated visits of 5000 websites.

4.3.2.2 Number of Requests

By analysis of the SQL table 'fp_http_requests', we recorded a total number of 49.054 third-party requests an average count of 10 (= d49.054/5.000e) requests per first-party domain.

4.3.3 Looking at First Parties

We consider first parties in terms of distinct third parties usage. Furthermore, we analyze first parties acting as third party and those who do not use cookies but integrate third party contents.

When the number involved the TP-FP-domains per page varies the value for the 5,000 visited sites is 0-57 (Fig 4.5). The website 'knowyourmeme.com' came out as the heaviest user of third-parties with 57 requests to distinct third-party domains. Figure 4.4 shows that the majority of visited websites make use of third-parties and approximately 50% of them utilize more than one third-party domain.

Pos.	Domain	# FPs	% FPs
1.	doubleclick.net	1,130	22.60
2.	facebook.com	977	19.54
3.	google.com	970	19.40
4.	scorecardresearch.com	579	11.58
5.	twitter.com	512	10.24
6.	quantserve.com	391	7.82
7.	imrworldwide.com	210	4.20
8.	adnxs.com	180	3.60
9.	yieldmanager.com	179	3.58
10.	2o7.net	163	3.26
11.	yadro.ru	146	2.92
12.	baidu.com	144	2.88
13.	revsci.net	139	2.78
14.	serving-sys.com	139	2.78
15.	addthis.com	131	2.62

16.	yandex.ru	128	2.56
17.	gemius.pl	127	2.54
18.	atdmt.com	116	2.32
19.	ivwbox.de	113	2.26
20.	criteo.com	111	2.22

Table 4.2: Top 20 Number of Third Parties on of FP Hosts

4.3.3.1First Parties as Third Parties

7.22% of the First Parties act simultaneously as third parties for other domains. This has been found by means of a comparison query (6) in the SQL table 'statistics'.

(6) SQL> SELECT COUNT(DISTINCT tp) FROM statistics WHERE tp IN (SELECT domain FROM domains);

291 first-parties (5.82%) are acting as third-parties for other visited first-party domains and therefore categorized as "hybrid-parties". These are hard to identify for a user as they cannot be detected by simply watching the cookies stored in Firefox browser. A general blocking of cookies for "hybrid-parties" would stop their third-party activities but it can affect browsing experience when visiting as first-party.

Pos.	Domain	TPs
1.	knowyourmeme.com	57
2.	hongkiat.com	44
3.	digitalspy.co.uk	42
4.	sport1.de	41
5.	cuantarazon.com	36
6.	dailycaller.com	29
7.	premierleague.com	29
8.	thisissouthwales.co.uk	28
9.	mysearchproperties.com	28
10.	anime44.com	28
11.	wetter.com	28
12.	radaronline.com	27
13.	bostonherald.com	26
14.	allkpop.com	25
15.	gamestar.de	25

16.	socialmediaexaminer.com	25
17.	boston.com	25
18.	mediatakeout.com	24
19.	freekaamaal.com	23
20.	breitbart.com	23

Table 4.3: Top 20 First Parties with frequent use of third parties

4.3.3.2 First Parties with TP content

The analysis of the SQL table 'fp_http_responses' in conjunction with 'fp_http_headers' reactive response shows that 25.12% (1,256 of 5,000) of the accessed web sites set no FP cookies. This does not mean that also no TP-Cookies created by this First Party.

A comparative SQL command (7) returns the domains from which no FP cookies are saved, and those who have themselves generated TP-cookies by embedded TP content

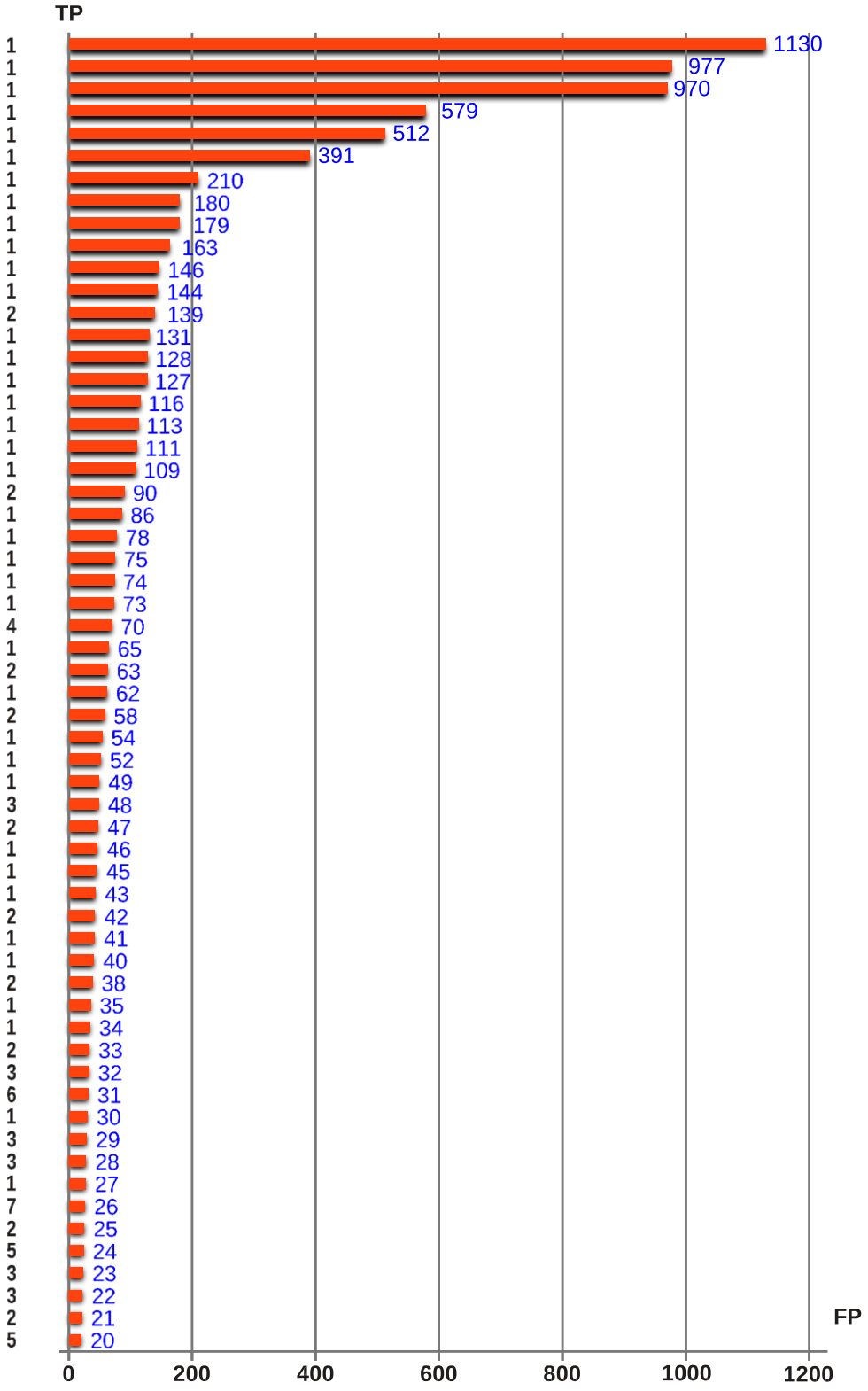
(7)
```
SQL> SELECT COUNT (DISTINCT fp) FROM statistics
        WHERE NOT EXISTS (
            SELECT DISTINCT baseDomain FROM cookies WHERE
            cookies.baseDomain = statistics.fp
        );
```

During analysis, we counted the domains which have no first-party cookies but indirectly created cookies by including third-party content. The result is 244 First Parties, ie 4.88% of the 5000 FP domains and 19.43% of the 1,256 sites without FP cookies.

4.4 Effects of 'privacymanager' on Browsing

In a reference-evaluation of the 5000 Overall visits 364,867 HTTP requests were recorded in FourthParty, without the use of 'privacymanager'. After installation and activation of 'privacymanager' 359,326 HTTP requests received and stored. It shows deviation is only 1.52% and is therefore considered negligible. 'privacymanager' is not blocking HTTP requests. For this reason, no significant difference in the observed requests with and without 'privacymanager' be seen but only removes selected cookies from requests if they are violating the user- defined policy. We found 49,054 (out of 359,326) third-party requests during automated visit of 5000 websites and of those 83.29% (40,857) were modified by the add-on (at least one third-party cookie was blocked) because of the given policy violation. We also found out that the third-parties

do not initiate further tracking actions if they detect a missing cookie – they simply assign a new value in the following regular connection.



F i g 4.4: Frequency of Use of Third Party Cookie

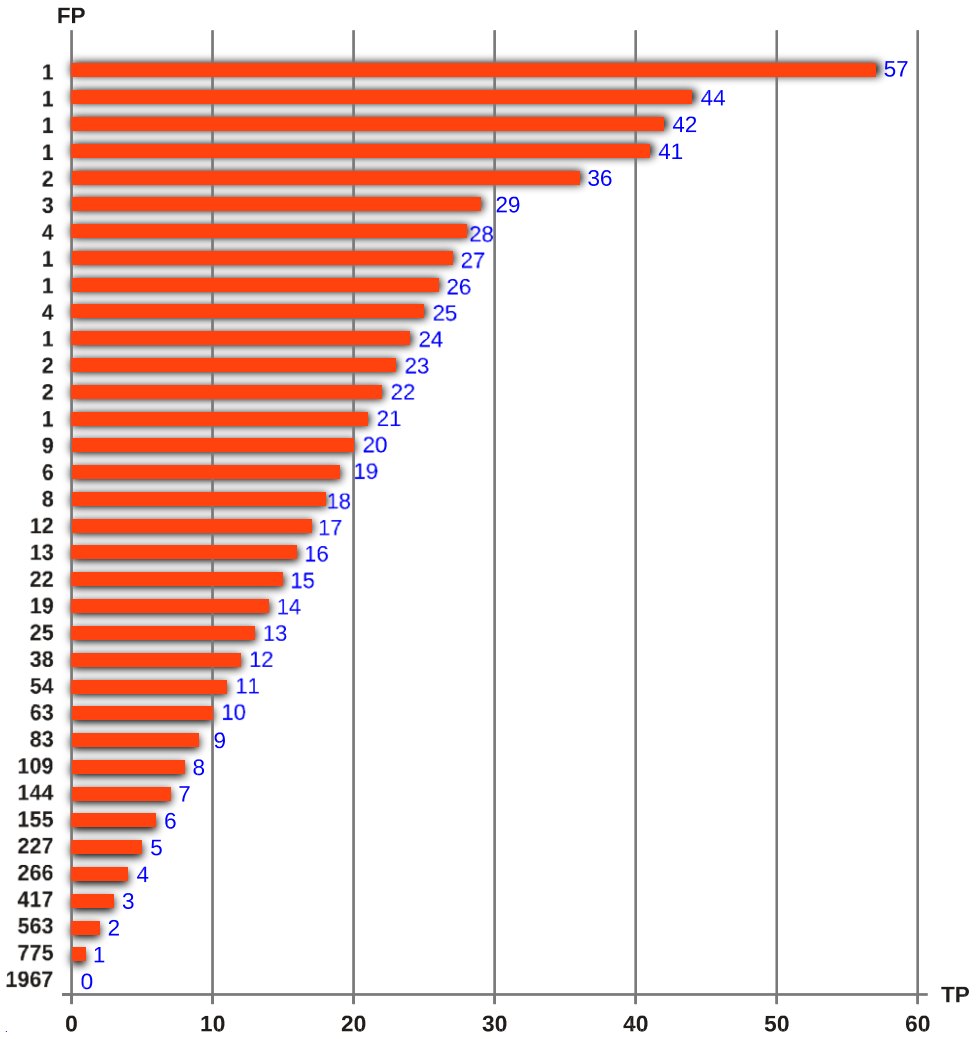


Figure 4.5: Use of Third Parties in First Parties

4.5 DOM Storage, Local Shared Objects (LSO) and Cookies

After 5000 websites automated visits, Firefox contained a total number of 20654 cookies from 4560 different domains in its internal storage. We also found 392 elements in the DOM Storage created by 253 distinct domains and at the same time there were 90 distinct domains which created Local Shared Objects (LSOs). Only a small subset of third-parties are making use of DOM storage and LSO as compared to the use of cookies(see Table 4.4).

Type	Count	Domains	Third Parties
Cookie	20.645	4.560	909
DOM Storage	392	253	11
LSO	90	90	49

Table 4.4 DOM Storage, LSO and Flash Cookies

4.6 Rating Mechanism of Third Party Domains

Third Party editor of 'privacymanager' offers users predefined whitelist, blacklist and trusted third party list. Users able to move domain name from one list to another and also can add name of their choice in any list. For this purpose, the reliable assessment of third parties require. We also want to identify the relations between third-parties and categories to detect their trust level. e.g., a third-party that is only used on adult websites could be an entry in the blacklist for some users.

Figure 4.6. Rating by privacychoice

There exists different online categorization platforms but they do not provide enough information to finalize the categorization process. An automated crawl of alexa.com and dmoz.org assign categories for 2,877 out of the requested 5,000 domains. But 2,123 of the investigated sites (42.46%) remain without categorization and on this basis, no reliable classification of Third Parties take place, so we need a method to

fill the list

The alternative solution is a direct classification of third-party domains provided by the tracker and advertiser database on privacychoice [38,39]. Privacychoice offers detailed information for a large number of third-parties, combined with a rating system. Free automated access is not possible and we have to check each domain manually. Therefore, the list of 1200 identified third- parties is reduced to 172 entries by choosing only those which are included from minimum 10 different first-party domains.

Privacychoice's rating system assigns classification values between zero & 50 and marks them with supportive colors (see Figure 4.6). The rating "No Priva- cyscore" is the most positive and shows that a third-party is not taking user data for tracking or does not even collect it at all – therefore we place these providers on whitelist. Third-parties rated with "comfort" are moved into trusted third- party list and each other domain (classified as untrusted by "caution", "concern", or unknown state) goes to the blacklist. Table 4.5 shows the results of our rating of 172 third-party domains. We refer to [20,21] for the detailed description of the rating system.

Rating (privacychoice)	Number	Target List
"No Privacyscore"	10	Whitelist
"comfort" (45-50)	43	Trusted third-party List
"caution" (40-44)	7	Blacklist
"concern" (0-39)	71	Blacklist
unknown	41	Blacklist

Table 4.5 Rating by Privacychoice

4.7 Comparison of privacookie and privacymanager:

General notation used, defined as i.e. B with elements b_i, while the different indices represent distinct domains. Third-parties are similarly described with D and d_j. Cookies are contained in set C and identified by $c_{i,j}$ with the first index representing the first-party do- main and the second index representing the third-party domain. Cookies and domain names are available in the browser history set i.e., H

i. Tracking Polices:

Spatial Tracking Policy

Spatial tracking policy implemented by 'privacookie' as:

$$\omega 1: \sum_{b_m \in H(B)} \beta(b_m, d_j, c_k) \leq 1$$

('1' is hardcoded)

And implemented by privacymanager as

$$\omega 1: \sum_{b_m \in H(B)} \beta(b_m, d_j, c_k) < L_S$$

User can set value of L_S

Temporal Tracking Policy:

This policy implemented by 'privacymanager' in the same spirit as implemented in 'privacookie'

$$\omega 2: \sum_{b_m \in H(B)} v(b_m, d_j, c_k) < L_V$$

ii. Website/URL Categorization:

'privacookie' does not categorize domains where as 'privacymanager categorizes the domains in three lists and do further working on the basis of these lists. If domain lies in whitelist, third party cookie will be set, if domain name is in blacklist, no third party cookie will be set and if in trusted third party list, then set cookie if user policy permits.

iii. Cookie Storage

Cookie storage in ‘privacookie’ is not persistent and when add-on restart, all previously saved cookies will be deleted. ‘privacymanager’ stores cookies persistently in a database in the user’s profile directory but also maintains a cache layer to handle performance issue.

‘privacymanager’ stores cookies persistently in a database in the user’s profile directory but also maintains a cache layer to handle performance issue.

iv. User Interface

Interface provided by ‘privacookie’ is not user friendly as shown in Figure 3.5. In contrast, ‘privacymanager’ provides user friendly interface with toolbar button that gives control to user to set cookie blocking policy on per site level

4.8 Comparison with Other Approaches:

i. Cookie Monster

As discussed in section 2.2.1, this add-on does not precisely address the issue of third party cookies. When rule sets for particular domain, that rule will be followed regardless of that fact that whether domain acting as first party or third party. When choosing between ‘privacymanager and Cookie Monster, ‘privacymanager’ should be preferred because it provides the user an effective configuration and settings for transparent dealings with third parties.

Cookie Monster cannot be used together with ‘privacymanager’, as both add-ons actively intervene in the cookie management of the browser.

ii. Cookie Whitelist

Cookie Whitelist also uses the idea of whitelist of domains, a list from which user wants to accept cookies. Together with ‘privacymanager’, it offers the better control of First and Third Party cookie management.

iii. Cookie Safe

The add-on should not be used together with 'privacymanager' because can cause problems in the simultaneous manipulation of the cookie data. A disadvantage of the use of Cookie Safe with 'privacymanager' is the lack of distinction between First Parties and Third Parties, it can only block all/allow all without going in fine detail.

iv. Ghostery

Ghostery depends on a frequently updated database. In our add-on, a user can set the fine-grained cookie policy and it does not depend on frequently updated database. We also give the user more control over his/her private information.

v. Better Privacy [12]

In BetterPrivacy, only whitelist option is available to the user and at the same time it does not deal with third- party cookies which is by far the most frequently used tracking method. Our add-on gives better control regarding the commonly used tracking technique and provides fine-grained cookie policy management.

vi. Do Not Track

The main problem with DNT is its reliance on third-parties to honor users' preference which is not happening in reality. In our solution, the control is in user's hands and we do not assume any sort of regulation.

The common use of 'privacymanager' and DNT is recommended, as it overcomes the drawback of DNT

vii. Adnostic

Similar to DNT, the main problem of Adnostic is its reliance on third-parties to honor the users' settings. In our approach, we give control to the user so that he can decide. Our proposal enables tracking with minimal privacy risks.

viii. Privad

Privad anonymizes every piece of information sent by the client, this anonymization impacts on performance. Our solution is better than Privad in a sense that user can trust on some ad-networks with the help of policy

ix. Collusion

Collusion does not have opt-out tracking option and will depends on a global database of web trackers. Our proposal allows users to set policies and gives more control over their private information

x. Beef Taco

As discussed in section 2.2.6, Beef Taco crowded the browser with opt-out cookies and also does not provide option for configuration. Our proposal gives more control to use to set policy about third party cookies

xi. Firefox New Cookie Policy

To take full advantage of this, user need to clear all cookies and if user allows third party cookie, there is no limit on the amount of third party tracking. In 'privacymanager', user can set limits by setting values for 'spatial tracking and temporal tracking'.

xii. Doppelganger

The drawbacks of Doppelganger are: performance cost due to 'mirroring mechanism' and blocking all third-party cookies. In 'privacymanager', we give user a functionality to make editable lists (whitelist, blacklist and trusted third-paryt list) of service providers to set policies about third party cookies.

4.9 Summary of Results

The large number of blocked cookies shows that the majority of TP requests is used for tracking user activities. In contrast, DOM Storage and LSOs are rarely used and are largerly the storage of FP content that does not violate privacy.

Despite the small number of LSOs (90), result shows the high proportion of TP domains (54.44%) with Flash cookies that should be considered critical and should be controlled by the user. The majority of websites has no restrictions on the use of Flash cookies. Therefore, the installation of a Flash-blocker add-ons [40], [41] and the global deactivation of Flash is recommended in the Firefox browser along with privacymanager

The implementation of presented add-on privacymanager contributes successfully to protect the user privacy and reliably filers all cookies as per user defined policy. By setting the values of following three parameters in setting window, it blocks substantial

portion of third party cookies per site level.

- “Maximum age of third party cookie = 1 minute
- Maximum dispatch of third party cookies = 3 time
- Maximum number of First Parties sharing the same third party cookies = 2”

This illustrates the relevance of the proposal with monitoring and controling of third-party activities. The comparison with PrivaCookie shows that the enhanced add-on in the form of ‘privacymanager’ offers the user an even more precise possibility for setting the third-party filtering.

5

Conclusion and Future Work

With the completion of this diploma thesis, the Firefox add-on ‘privacymanager’ is testesd. The realisation of functional requirements, as an extension of privacookie, was presented in detail and proved to be working.

Analysing 5.000 website visits pointed out the intensive use of cookies by third parties and led to the necessity of an effective management system. The comparison with already existing extensions showed that ‘privacymanager’ forms a new model of cookie management and that the combination of different add-ons is able to provide an extensive and controlled protection of privacy.

For future extensions of ‘privacymanager’ there should be a usability study with users from different groups of age and knowledge to detect possible difficulties and to solve them during the developement process. Useful might be the integration of automatic ratings for third parties by accessing the database of privacychoice [40] or with another similar method for automatic black- and whitelist management. We are also planning to test this solution for Andriod platform and implement it.

References

[1] The New York Times - John Schwartz, “Giving the Web a Memory Cost Its Users Privacy,” http://www.nytimes.com/ 2001/09/04/technology/04COOK.html.

[2] B. Krishnamurthy, “Privacy leakage on the Internet,” presented at IETF 77, March 2010.

[3] B. Krishnamurthy and C. E. Wills, “Generating a privacy footprint on the Internet,” in Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement, ser. IMC ’06, New York, NY, USA, 2006, pp. 65–70.

[4] F. Roesner, T. Kohno, and D. Wetherall, “Detecting and defending against third-party tracking on the web,” in NSDI’12: Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation. Berkeley, CA, USA: USENIX Association, 2012, pp. 12–12.

[5] The Wall Street Journal, “What They Know,” http://blogs.wsj. com/wtk/

[6] J. Turow, J. King, C. J. Hoofnagle, A. Bleakley, and M. Hennessy, “Americans Reject Tailored Advertising and Three Activities that Enable It,” 2009.

[7] B. Ur, P. G. Leon, L. F. Cranor, R. Shay, and Y. Wang, “Smart, useful, scary, creepy: perceptions of online behavioral advertising,” in Proceedings of the Eighth Symposium on Usable Privacy and Security, ser. SOUPS ’12. New York, NY, USA: ACM, 2012, pp. 4:1–4:15.

[8] J. R. Mayer, J. C. Mitchell. Third-Party Web Tracking: Policy and Technology. 2012

[9] F. Roesner, T. Kohno, D. Wetherall. Detecting and Defending Against Third-Party Tracking on the Web. 2012

[10] Persistent client state: HTTP cookies, Preliminary specification. http://wp.netscape.com/newsref/std/cookie_spec.html.

[11] J. Valentino-DeVries. What they know about you. The Wall Street Journal, July 31 2010.

[12] BetterPrivacy Add-on.: https://addons.mozilla.org/de/firefox/addon/betterprivacy/

[13] Do Not Track. http://donottrack.us/

[14] Tracking the Trackers.: http://cyberlaw.stanford.edu/node/6694

[15] V. Toubiana, H. Nissenbaum, A. Narayanan, S. Barocas, and D. Boneh.: Adnostic: Privacy preserving targeted advertising. In NDSS 2010.

[16] S. Guha, B. Cheng and P. Francis.: Privad: Practical Privacy in Online Advertising in NSDI 2011

[17] Extended Cookie Manager Add-on.: https://addons.mozilla.org/en-us/firefox/addon/extended-cookie-manager/

[18] Collusion Add-on.: http://www.mozilla.org/en-US/collusion/

[19] Beef Taco (Targeted Advertising Cookie Opt-Out) Add-on.: https://addons.mozilla.org/en-us/firefox/addon/beef-taco-targeted-advertising/

[20] U. Shankar and C. Karlof.: Doppelganer: Better browser privacy without the bother. In CCS, 2006.

[21] Add-on Cookie Safe https://addons.mozilla.org/en-US/firefox/addon/cookiesafe-ff-4-compatible/

[22] http://ec.europa.eu/ipg/basics/legal/cookies/index_en.htm

[23] Block cookies from sites I haven't visited.: https://bugzilla.mozilla.org/show_bug.cgi?id=818340

[24] Firefox getting smarter about third-party cookies.: https://blog.mozilla.org/

privacy/2013/02/25/firefox-getting-smarter-about-third-party-cookies/

[25] The New Firefox Cookie Policy.: http://webpolicy.org/2013/02/22/

the-new-firefox-cookie-policy/

[26] C is for Cookie.: https://brendaneich.com/2013/05/c-is-for-cookie/

[27] Jonathan R. Mayer. Tracking the trackers: Selfhelp tools. https://cyberlaw.stanford.edu/blog/2011/09/tracking-trackers-self-help-tools

[28] https://addons.mozilla.org/en-US/firefox/addon/cookiesafe-ff-4-compatible/

[29] https://addons.mozilla.org/en-US/firefox/addon/cookie-whitelist-with-buttons/

[30] Add-on Targeted Advertising Cookie Opt-Out (TACO), https://addons.mozilla.org/en-US/firefox/addon/targeted-advertising-cookie-op/

[31] Add-on TrackerBlock, Version 2.2 https://addons.mozilla.org/de/firefox/addon/trackerblock

[32] Add-on Ghostery https://www.ghostery.com/en/

[33] http://infoscience.epfl.ch/record/135690/files/Freudiger09VH.pdf

[34] FourthParty Web Measurement Platform, Version n/a http://fourthparty.info/

[35] http://cyberlaw.stanford.edu/files/publication/files/trackingsurvey12.pdf

[36] Desktop-Application Sqliteman, Version 1.2.2 http://www.sqliteman.com/

[37] Web-Application phpmyadmin, Version 3.4.10.1 http://www.phpmyadmin.net/

[38] Privacy Choice.: http://privacychoice.org/checkprivacyscores

[39] Privacy Score.: http://privacyscore.com/#

[40]Add-on FlashDisable, Version 0.8.4 https://addons.mozilla.org/de/firefox/addon/flashdisable

[41] Add-on Flash OnOff, Version 0.3 https://addons.mozilla.org/de/firefox/addon/flash-onoff

[42] Add-on Cookie Monster https://addons.mozilla.org/en-US/firefox/addon/cookie-monster/

[43] Add-on Self Destructing Cookies https://addons.mozilla.org/en-US/firefox/addon/self-destructing-cookies/

[44] Alexa Top 1.000.000 Sites http://s3.amazonaws.com/alexa-static/top-1m.csv.zip