Automatic Universal In-Browser Payments

by

Daan Middendorp

Matriculation Number 397108

A thesis submitted to

Technische Universität Berlin
School IV - Electrical Engineering and Computer Science
Department of Telecommunication Systems
Service-centric Networking

Master's Thesis

May 5, 2020

Supervised by:
Prof. Dr. Axel Küpper

Assistant supervisor:
Philip Raschke, M.Sc.

Service-centric NETWORKING

Technische Universität Berlin

Statutory Declaration

I hereby declare that I have created this work completely on my own and used no other sources or tools than the ones listed.

Berlin, May 5, 2020 Daan Middendorp

Acknowledgments

I would like to thank my teddybear...

Abstract

In this thesis, we show that lorem ipsum dolor sit amet.

Contents

1 Introduction

The business of online advertising has evolved into a landscape which is not transparent anymore. A handful of large advertisement firms are controlling practically every online ad you see. Almost every movement during the visit of a regular website is sent in an obfuscated way to the advertisement broker, without any visible sign to the visitor. This makes the whole browsing experience obnoxious, especially now it turns out that entire societies are being influenced by the power of advertisement networks, as we have seen in the Cambridge Analytica scandal [**?**].

Several publishers have been experimenting with alternative ways of generating income. Currently, some of them are selling subscriptions, asking for donations or using the visitors' computer for cryptomining [1]. But these models do not seem to be a real substitute for advertisement networks.

In this master thesis, which is written at the Service-centric networking research group at the Technische Universität Berlin, the main focus lies at solving this so called unpaid content problem while assuring the privacy of the user and keeping the costs low. The increasing possibilties in the field of blockchain technology are of great use for such a solution and therefore also a key building block of the proof of concept.

The concept, as discribed in chapter 3, features a system that runs in the background while browsing the web. If the users visits a publisher that also supports the system, a message will be shown to the user indicating that it is possible to hide the advertisements and pay a small amount per pageview instead. When this permission is granted, the user will not see any advertisements on that particular website again, but contribute by sending small payments to the publisher instead.

As this research is made possible by public money, the entire process is kept as transparent as possible. This is achieved by publishing everything related to this thesis under a permissive free software license on Github [1].

1.1 Research statement

This reseach will investigate the possiblities of new technologies in order to solve the unpaid content problem. The following research question is defined:

How can the unpaid content problem be solved in a cheap, privacy preserving and transparent way?

This research question is split up in the following subquestions:

[1] `https://github.com/lightning-sprinkle`

- What current revenue models are used in order to solve the unpaid content problem?
- How is privacy preserved in the current models?
- What are the costs of the current models?
- What is the amount of transparency in the current models?
- What are the conditions, that an alternative model should adopt in order to be at least comparable to existing models?
- How to realize and implement a comparable revenue model that follows these conditions?

1.2 Methodology

In order to explain the different models and concept, a couple of roles will be used througout this thesis.

Unpaid content
: Content that is freely available on the internet (without a subscription or payment), such as news articles and video's.

Publisher
: The owner of the website that provides the unpaid content

User
: The visitor of the website that consumes the unpaid content

Ad broker
: A third party providing advertisements to the user in order to generate revenue for the publisher

2 Related Work

Revenue models that are applicable on the internet in order to monetize online content is a topic that has been actively researched and experimented with over the past few decades. This chapter will dive into the related work on both a technical and an economical level. It is structure in a way so that three different research areas will be discussed. Firtsly, the economic aspects of unpaid content are reviewed. Secondly, an overview of the technical solutions to the unpaid content will be given. This will include both micropayments, subscription models and online advertising. Lastly, the possiblities that arise in the blockchain era will be discussed.

2.1 Economic aspects

Even in the early days of the World Wide Web, the problem of unpaid content existed. These approaches include paywalls and web advertisements. For example, the *Wall Street Journal* implemented already a hard paywall in 1996, which is still in place as of today, with over 2 million subscribers as of February 2020 [2]. Alternatives to hard paywalls are soft paywalls. The difference between both types is that soft paywalls are trying to convince potential customers to subscribe by giving them a free sample of the content. For example, the *New York Times* has implemented a soft paywall with a limit of 5 free articles per month [3].

Paywalls, however are fairly easy to circumvent. This especially accounts for soft paywalls. Therefore, publishers are trying to implement counter measures in order to enforce a subscription. For example, the *New York Times* attacked one popular circumvention method, the use of an incognito window. With behavioral analysis, it is possible to find out that the user is using an incognito window, so that the *New York Times* is able to prevent the free article from being served.

2.2 Technical solutions

In order to generate revenue from online content, there are two technical solutions broadly adopted: (micro)payments and online advertising.

2.2.1 Online advertising

Advertising is a method to draw attention to a product, service or event in order to promote sales or attendance. Since the early days of the word wide web, this industry has also expanded to the internet. The first advertisement on the world wide web is possibly from 1994 on HotWired.com, which was bought by AT&T and had a click through rate of 44% [4]. Meanwhile, the online advertising industry is very profitable and has evolved into the core business

of the world wide web.

This section will gives an overview of the current role model of the online advertisement industrsy and take a closer look at the different approaches in online advertising and their privacy aspects. Lastly, the research field of privacy-friendly alternatives will be discussed.

Normally, there are sereval parties involved in the advertising ecosystem. On one side, there are publishers, such as *Der Spiegel* that provides online content, e.g. news articles. On the other side, there is an advertiser that provides the advertisement.

The most interesting part, however, is the ad platform. Ad platforms are entities that connect the publisher with the advertiser by providing them an interface to match demand and supply. Due to the wide range of different publishers and users that are reachable by ad networks, it becomes really efficient to allocate ad space. Ad platforms can even be concidered as the central hub in the online advertising industry. When a user visits the website of a publisher, the browser communicates with the webserver. The browser receives the content that is displayed to the user. Along with this content, additional scripts that are associated with an ad network are also delivered to the browser and executed. These scripts are triggering a connection to the ad exchange. The ad platform is able to serve extra commercial content (advertisements) over this connection, which will be embedded into the page by the script. This method makes it possible for ad exchanges to partner up with huge amounts of publishers and serve an amount of users that is several orders of magnitudes higher [5].

The ad platform itself, consist out of multiple components, that might also be run by different entities. Firstly, there is an ad network, which resells the ad space from to publisher to an advertiser. Secondly, another component on the ad platform is the ad exchange. These are auction based advertisements marketplaces where advertisers can bid on an ad space in realtime, which means that the auction takes place when the user visits the website of the publisher. Based on the profile of the user, certain advertisers might be more interested into buying the ad space and thus offering a higer price [5].

Thirdly, a data aggregator is an entity that which goal is to gather and aggregate data about the purchasing interest of the users. This data is used to provide insights to both the advertisers and the publishers to target their marketing decisions [5].

2.2.1.1 Privad

The problem with the infrastructure mentioned above, is that everything can be controlled by one single entity. This single entity knows anything about all parties involved: advertisers, publishers and users. The behavior of a single user is tracked across multiple websites, which might be concidered a privacy concern. Guha et al. [6] developed Privad, which they call a practical private online advertising system.

The model of Privad is slightly differs from the original online advertising role model. The model also includes the user, publisher and advertiser. However, in this model there is also a dealer and a broker present.

2.2.2 Payments

- Subscription
  - Paypal
  - Flattr
  - Blendle

2.3 Blockchain

2.4 Brave Browser

In 2016, Brave Software launched a browser that blocked ads and trackers by default, the Brave Browser [1]. During the introduction, Brave Software also shared their plans for a Brave Publisher Ads program to pay publishers back fair a share of internet revenue. As of 2020, their service is called "Brave rewards program", and any content creator can enroll in order to get paid for content.

2.4.1 Basic Attention Token

In order to achieve a system that makes it possible to reward content makers on the internet, Brave introduced the Basic Attention Token. This token, which works like any other cryptocurrency, represents user attention. Their goal with this token is to trade "attention" just like any other commodity, like oil and coffee. This means that the this token can also be traded on a cryptocurrency exchange.

2.4.2 Anonize protocol

2.4.2.1 ICO

Brave Software used an initial coin offering to introduce the new token to the market. The ICO happened in May 2017 and raised 156,000 ETH, which was worth around 35M USD at the time. The raised money is mostly used to pay for the development and other costs of the token. The development team exists out of 20 developers.

2.4.2.2 Brave ledger system

- Brave vault

[1] `https://www.brave.com`

2.4.3

2.5 Lightning Network

3 Concept and Design

The proposed solution is an implementation which is completely based on the existing infrastructure that is available on the web. This comes with the advance that it works across all different types of devices, from desktop computers to smartphones. Another feature of this approach is that it does not require any additional tools, which might need some effort to setup.

3.1 Architecture

There are basically two components in this system: there is a wallet, which takes care of the storage of encryption keys and is responsible for handling all communications with a cryptocurrency network. The second part is the publisher library, which can be embedded by any website who want to take part in the universal-pay ecosystem.

3.1.1 Wallet

The wallet will be, like all the other parts of the architecture, implemented in standard web technology, which means Javascript. For the convenience of the end user, the wallet will be hosted on a domain to make sure that for the system to work, still no additional configuration is needed. However, this requires trust. If the owner of the domain becomes malicious, the entire wallet might be stolen. Therefore, the user is free to host his own wallet on every desired location, even *localhost* is a possiblity.

3.1.2 Publisher library

Publishers can load an external library into their website, this library communicates with the wallet, as described in 3.1.1. When the user visits the page of the publisher, the loaded library will check if there is a wallet running on that local machine. If this is not the case, it will embed the hosted instance of the wallet in an iframe.

If the connection with the wallet is established, the publisher will ask the wallet for a payment. The wallet can accept this payment and create a transaction which is sent to the blockchain network.

3.1.3 Communication

One of the challenges with this architecture it the communication and how to make sure a connection is established with a publisher that actually is legit and provides content on the users' computer.

In order to make this possible, a structure with WebRTC is proposed. WebRTC is a technology which makes it possible for different websites (even accross different computers) to communicate with eachother. The technology was invented to make real time video and audio communication possible within the browser. However, the API makes it also possible to send data over the channel.

3.1.4 Micropayments

4 Implementation

Describe the details of the actual implementation here...

5 Evaluation

The evaluation of the thesis should be described in this chapter

6 Conclusion

Describe what you did here

List of Tables

List of Figures

Bibliography

[1] J. Ruth, T. Zimmermann, K. Wolsing, and O. Hohlfeld, "Digging into browser-based crypto mining," in *Proceedings of the Internet Measurement Conference 2018*. ACM, 2018, pp. 70–76.

[2] J. Benton. (2020) The wall street journal joins the new york times in the 2 million digital subscriber club. [Online]. Available: https://www.niemanlab.org/2020/02/the-wall-street-journal-joins-the-new-york-times-in-the-2-million-digital-subscriber-club/

[3] J. E. Cook and S. Z. Attari, "Paying for what was free: Lessons from the new york times paywall," *Cyberpsychology, behavior, and social networking*, vol. 15, no. 12, pp. 682–687, 2012.

[4] A. Lafrance. (2017) The first-ever banner ad on the web. [Online]. Available: https://www.theatlantic.com/technology/archive/2017/04/the-first-ever-banner-ad-on-the-web/523728/

[5] J. Estrada-Jiménez, J. Parra-Arnau, A. Rodríguez-Hoyos, and J. Forné, "Online advertising: Analysis of privacy threats and protection approaches," *Computer Communications*, vol. 100, pp. 32–51, 2017.

[6] S. Guha, B. Cheng, and P. Francis, "Privad: Practical privacy in online advertising," in *USENIX conference on Networked systems design and implementation*, 2011, pp. 169–182.

Appendices

Appendix 1

```
for($i=1; $i<123; $i++)
{
    echo "work harder! ;)";
}
```