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 12, 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 12, 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 dives into the related work on both a technical and an economical level. It is structured 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 is given. This includes both micropayments, subscription models and online advertising. Lastly, the possiblities that arise in the blockchain era are discussed.

2.1 Technical solutions

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

2.1.1 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% [2]. Meanwhile, the online advertising industry is very profitable and has evolved into the core business of the world wide web.

This section gives an overview of the current role model of the online advertisement industrsy and takes 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 is a publisher, such as *Der Spiegel* that provides online content, for example: 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 both demand and supply. Due to the wide range of different publishers and users that are reached by ad networks, it becomes really efficient to allocate ad space. Ad platforms can even be considered 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 is 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 [3].

The ad platform itself, consist of multiple components, that might also be run by different entities. Firstly, there is an ad network, which resells the ad space from a publisher to an advertiser. Secondly, another component on the ad platform is the ad exchange. These are auction based advertisement marketplaces where advertisers can bid on 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 to buy the ad space and thus offering a higher price [3].

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 [3].

2.1.1.1 Privad

The problem with the infrastructure mentioned above, is that everything can be controlled by one single entity. This single entity knows everything about all parties involved: advertisers, publishers and users. The behavior of a single user is tracked across multiple websites, which might be considered a privacy concern. Guha et al. [4] developed Privad, which they call a practical private online advertising system.

The model of Privad is slightly different from the original online advertising role model. The model also includes the user, publisher and advertiser. However, in this model there are also a dealer and a broker present. One key difference compared to the traditional online advertising model is that the profiling (building a profile of the user based on interests) is done on the users' computer and not by a central data aggregator. Secondly, the ad platform is split into two different entities: the broker and the dealer. The broker is comparable to the traditional ad platform and matches the profile with advertisements. The request, however does not come from the users' computer immediately: there is a dealer placed in between. The dealer anonymizes every request before it is sent to the broker and makes sure that click fraud is prevented. The dealer cannot eavesdrop on the request, because the request is sent in an encrypted form to the broker [4].

One concern with this approach is that a profile might be so detailed that the broker is able to find out an identity based on the profile. In orde to tackle this problem, Privad subscribes to a certain general profile that is shared with multiple other users. The user receives multiple advertisements and can pick locally which suits best.

Trust, however is still a key element in this approach. There is no way to find out if the dealer is trustworthy. If the dealer and the broker are secretly run by the same entity, it is possible to exchange data and learn more about the user.

2.1.2 Fee-based

The second business model as an alternative to online advertising is requiring payments in exchange for content. This section describes what different approaches there are in the field of

online payments, subscription models and third parties offering these services.

2.1.2.1 Paywalls

Even in the early days of the World Wide Web, the phenomenom of content that is only visible with a subscription existed. Such a mechanism, is called a paywall. For example, the *Wall Street Journal* implemented already a hard paywall in 1996, which is still in place today, with over 2 million subscribers as of February 2020 [5]. 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 [6].

Paywalls, however, are fairly easy to circumvent. This is especially the case 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, which enables the *New York Times* to prevent the free article from being served [7].

2.1.2.2 Micropayments

2.1.3 Donation-based

Other publishers are relying on the willingness of their users to compensate. These systems are either implemented on one particular website or offered as a service over multiple websites.

The Wikimedia Foundation is an example of a non-profit organiation that actively asks for donations on their Wikipedia website. However, this is not the main source of their revenue, because big companies like Amazon or Google are also donating to the foundation. Wikimedia raised a total of 91M USD in 2016-2017. [8]

For smaller websites, such a campaign is not very viable. Users are not likely to send a donation to each individual contentmaker they support. For these smaller publishers, Flatter[1] offers a user experience which is similar to the Facebook-like button. Although, the difference is that instead of just showing the interest in a certain page of website, the button also shows appreciation by making a small micropayment. Flatter offers a subscription with a minimum of 2 EUR per month. This monthly subscription fee is divided amongst all websites the user clicked the Flatter-button on [9].

2.2 Blockchain

Blockchain technology has been with us for more than a decade. Satoshi Nakamoto built the first practical application that used the so called blockchain as a decentralized ledger, where it is possible to transfer a digital currency without trusting a single party [10]. Since then, a lot has changed and all kinds of experiments using this technology are performed. For example,

[1] `https://flatter.com`

blockchain implementations are now capable of running scripts which are even Turing complete, which opens the door to programmable money [11] and all kinds of other assets that are stored on the blockchain.

2.2.1 Lightning Network

The general problem with blockchain technology is scalability and speed: the current average confirmation time of a Bitcoin transaction takes a couple of minutes [12]. With the current blocksize of 1MB, the amount of transactions is limited to seven per second. Therefore this technology is not suitable for micropayments, which are payments with a value less than a dollar [13].

In order to solve this problem, several researchers have experimented with alternative ways to circumvent these issues. The most promising system in this research field is the lightning network [14]. The goal of the lightning network is to send small payments immediately, without intervention of the blockchain ledger and with minimal fees. The lightning network features such a system by combining a smart idea with the capabilities of multi signature addresses. The system relies on two parties, for example Alice and Bob, opening a joint account (channel). Off-chain, there is an agreement on which part of the joint account belongs to whom. With this joint account, Alice can transfer money to Bob and vice versa by just updating the agreement about the joint account. However, it is still not very practical if Alice also needs to open a joint account with any other party, for example Charlie, that she wants to transfer money to. The lightning network solves this issue by finding a path from Alice to Charlie using multiple joint accounts. In this example, it might be the case that Charlie has a joint account with Bob. Using these two joint accounts, Alice can transfer money indirectly to Charlie via Bob. In practice, this system follows a hubs and spokes model, where a couple of big players are connected to a lot of individuals in order to create a reliable network. Every hop that is used by a single transaction can also receive a small fee, but these fees are insignificant compared to on-chain transactions [14].

2.3 Automated payments

In 2016, Brave Software launched a browser that blocks ads and trackers by default: the Brave Browser [2]. During the introduction, Brave Software also shared their plans for a Brave Publisher Ads program to pay publishers a fair share of their internet revenue. As of 2020, their service is called "Brave rewards program", in which any content creator can enroll in order to get paid for content.

2.3.1 Brave rewards

In order to achieve a system that makes it possible to reward content makers on the internet, Brave introduced the Basic Attention Token [15]. 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 this token can also be traded on a

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

cryptocurrency exchange. Brave Software is promoting this token to use it to reward internet users. What happens is that the brave browser is equipped with a standard ad blocker. The websites are filled with sponsored content by the brave browser. The difference with the original advertisements is that the user gets rewarded for viewing them in the BAT-token. The BAT-token can be traded for other crytocurrencies or even fiat currencies.

For this thesis, another application of the BAT-token is even more interesting. That is, the system also works the other way around: users can spend their BAT-tokens on websites of the publishers they support in an automated manner. The remainder of this section shows the inner working of this system. Furthermore it explains why even that is still suboptimal from a decentral perspective. First of all, the concept of the Brave Vault is explained. Secondly, the privacy and anonymity measures are analysed. Lastly, the monolopy of Brave in this ecosystem is discussed.

The Brave Vault is a private datastore where browsing information is stored [16]. The central part in this Brave Vault is the *persona*, that is used to identify and set your browsing behavior. The *persona* can also be synced with other browsers, so that one user still uses the same profile when switching devices. Another part of the Brave Vault is the *session*. The *session* is bound to the browser and does not have a predefined lifetime. Browser dependant information, like browsing history, is stored here [16].

In the *persona*, there is a setting that enables an ad-free browsing experience by paying a small contribution. The contribution amount is divided among the websites that the users visits. However, if these contributions are sent to the publishers directly, it would be very hard to guarantee privacy and anonymity. Based on the contributions, it is possible to reconstruct a profile that might be linked to an individual. To tackle these privcay concerns, Brave developed the Brave Ledger [17]. The Brave Ledger is a central system that processes micropayments for the contributions to the publishers. The system is designed on two core principles: anonymous and accountable. The former means that Brave should not be able to correlate publisher visits with contributions. The latter implies that Brave should only be able to have insights in the contributions on an aggregated basis.

In order to build a system on these principles, the Brave Ledger combines statistical voting with an anonymous voting scheme [17]. First of all, statistical voting means that if you only have one vote, but you would like to vote on multiple choices, you are picking a choice at random out of your preferred choices. If everyone follows this system, the result of the election would be roughly the same as if everyone had multiple votes. The benefit of combining such a system with making contributions to certain publishers, is that the user is not revealing his entire browsing history, but only the publisher he wants to reward. Secondly, the Brave Ledger makes use of an anonymous voting scheme called ANONIZE2 [18]. This system guarantees that every single user in a group of users is able to cast a maximum of one vote, while keeping the vote anonymous.

2.3.1.1 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, with a value of 35M USD at the time/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.

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] 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/

[3] 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.

[4] 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.

[5] 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/

[6] 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.

[7] T. M. Troupson, "Yes, it's illegal to cheat a paywall: Access rights and the dmca's anticircumvention provision," *NYUL Rev.*, vol. 90, p. 325, 2015.

[8] (2017) Wikimedia fundraising report 2016-2017.

[9] F. Loll, C. Mumme, and N. Pinkwart, "Flattr this! explorative evaluation von social (micro-) payments als alternatives bezahlmodell," 2010.

[10] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," Manubot, Tech. Rep., 2019.

[11] G. Wood *et al.*, "Ethereum: A secure decentralised generalised transaction ledger," *Ethereum project yellow paper*, vol. 151, no. 2014, pp. 1–32, 2014.

[12] T. Bamert, C. Decker, L. Elsen, R. Wattenhofer, and S. Welten, "Have a snack, pay with bitcoins," in *IEEE P2P 2013 Proceedings*. IEEE, 2013, pp. 1–5.

[13] J. Frankenfield. Micropayments. [Online]. Available: https://www.investopedia.com/terms/m/micropayment.asp

[14] J. Poon and T. Dryja, "The bitcoin lightning network: Scalable off-chain instant payments," 2016.

[15] B. A. Token, "Blockchain based digital advertising," *Whitepaper (13 March, 2018)< https://basicattentiontoken. org/BasicAttentionTokenWhitePaper-4. pdf> accessed*, vol. 22, 2018.

[16] Principles of using the brave vault. [Online]. Available: https://github.com/brave/vault/blob/master/documentation/Vault-Principles.md

[17] Principles of the brave ledger. [Online]. Available: https://github.com/brave-intl/bat-ledger/blob/master/documentation/Ledger-Principles.md

[18] S. Hohenberger, S. Myers, R. Pass *et al.*, "Anonize: A large-scale anonymous survey system," in *2014 IEEE Symposium on Security and Privacy*. IEEE, 2014, pp. 375–389.

Appendices

Appendix 1

```
1 for($i=1; $i<123; $i++)
2 {
3     echo "work harder! ;)";
4 }
```