

卒業論文 2017年度(平成29年)

ユーザー属性埋込み型 HTTP 通信を利用した Web 広告手法の提案

慶應義塾大学 総合政策学部

桑原 誠尚

徳田・村井・楠本・中村・高汐・バンミーター・植原・三次・中澤・武田
合同研究プロジェクト

2017年1月23日

ユーザー属性埋込み型 HTTP 通信を利用した Web 広告配信手法の提案

近年 Web は人々の生活に欠かせないものとなった。それに伴い、企業広告が Web 上に掲載される機会も増加している。

しかし、既存の Web 広告配信手法、主にターゲティング広告の手法では、各 Web サービス・広告配信ネットワーク（以後、アドネットワーク）がユーザーの行動を追跡し、そこから推測した属性情報を元に広告を配信している。この追跡のデータは、ユーザーに開示されていない。また、アドネットワーク側が開示しない限り、ユーザーはどのようなユーザーとしてシステムに認知されているか知ることが出来ない。これはユーザーにとって好ましくない状況であり、アドネットワークに対して不信感を抱く要因である。アドネットワーク側にとっても、現在の手法では一定の割合で属性の推測に誤りが生じてしまう課題を抱えている。

そこで本研究では、ユーザが自身の好みにより配信される広告を決定できるシステムを提案する。ユーザーの好きな属性と嫌いな属性をアドネットワークに発信することで、ユーザーが間接的に配信される広告を決定することができる。アドネットワークにユーザーの属性を発信する手段として、属性情報を HTTP 通信に埋め込んで通信する。そして、アドネットワークが Web サイトを経由して受信した属性情報を元に広告配信する。この提案手法によって、ユーザーが自分の属性情報を管理でき、不適切な広告配信を防ぐことができるため、Web 広告におけるユーザーエクスペリエンスを向上させることができると考えた。

本手法を検証するため、自ら属性情報を選択し HTTP リクエストヘッダにユーザー属性を埋め込んで通信、発信された属性を元に広告を配信するシステムを実装した。また、達成すべき要件を明確化するため、Web 広告に関する印象と属性情報の設定の実態について調査を行った。この調査から、属性情報をユーザーの実世界の状況に応じて切り替える必要性が明らかになった。

評価として、Web 広告におけるユーザーエクスペリエンスが向上されたことを確かめるため、提案手法が満たすべき要件である、透明性、ホワイトリスト属性の発信、ブラックリスト属性の発信、属性の切り替え、の4つが達成されているか検証、達成された。

本研究の成果により、属性情報の管理をユーザー自身が行うことができ、ユーザーのアドネットワークに対する不信感の解消が期待される。また、従来は実現できなかった、ユーザーの実世界で置かれている状況に合わせた Web 広告の配信が可能になり、Web 広告に関するユーザーエクスペリエンスの向上が見込まれる。

キーワード:

1. Web 広告, 2. AdNetwork, 3. privacy, 4. HTTP, 5. Web tracking, 6. Web Intelligence

慶應義塾大学 総合政策学部
桑原 誠尚

Web advertisement system with User Attributes in HTTP headers

Recent Web has earned a vital role in people's lives.

On the contrary, current Web advertising methods, as represented by targeting advertisements, each Web services / Web advertisement distribution networks (hereinafter referred to as "ad network") tracks and stores the user's behavior. Advertisements are determined and distributed by the user attributes that are taken from these behavioral data. In this system, the tracking data are not disclosed to the users, making them ignorant of what kind of data that they are being tracked on. In addition, unless the ad network are open of its user attribution information, users are not aware of what attribution they are related to. This is an unfavorable situation for the users, which leads to a sense of distrust. There are problems even for the ad networks. In the current method, an error occurs in estimation of attributes at a fixed rate.

This paper proposes a method where users can choose their liking and dis-liking attributes, which is embedded in HTTP transaction. Ad Network delivers advertisements based on attribute sent by users. From this method, the uses themselves are able to manage their own attribution, prevent inappropriate ad delivery, which could lead to encourage user experience improvements.

As for the evaluation, to verify our assumptions and achieve our goals, there are four requirements we need to consider, which include transparency, the ability to transit whitelist attribution, the ability to transit blacklist attribution, and the ability of attribution modification.

As a result of this research, users are able to hold privileges to their attributes, which could lead to the mitigation of distrust to user's ad networks. Furthermore, it can be expected to realize web advertisement replacement according to the actual situation of the user, which will lead to UX improvements.

Keywords :

1. Web Advertisement, 2. Ad Network, 3. Web tracking, 4. privacy,
5. HTTP, 6. Web Intelligence

Keio University, Bachelor of Arts in Policy Management
Sena Kuwabara

目次

第1章	序論	1
1.1	本研究の背景	1
1.2	本研究の着目する課題	3
1.3	本研究の手法	3
1.4	本論文の構成	4
第2章	Web 広告	5
2.1	既存の Web 広告手法	5
2.1.1	サイトターゲティング	5
2.1.2	コンテンツターゲティング	6
2.1.3	検索連動型広告	7
2.1.4	行動ターゲティング	8
2.1.5	リターゲティング	10
2.1.6	デモグラフィックターゲティング	11
2.1.7	ジオグラフィックターゲティング	12
2.2	関連技術	14
2.2.1	Cookie を利用したトラッキング技術	14
2.2.2	Browser Fingerprint	14
2.2.3	Canvas Fingerprint	16
第3章	提案手法	18
3.1	既存手法の問題点	18
3.2	課題に対する調査	18
3.2.1	調査概要	19
3.2.2	調査手法	19
3.2.3	調査結果	20
3.2.4	調査結果についての考察	22
3.3	技術概要	24
3.3.1	HTTP 通信の仕組み	24
3.4	課題解決のための要件	28
3.4.1	透明性	28
3.4.2	ホワイトリスト属性の発信	28
3.4.3	ブラックリスト属性の発信	28

3.4.4	属性の切り替え	28
3.5	提案手法	29
第4章	システムの設計と実装	30
4.1	システム概要	30
4.2	本手法における属性値の規格	31
4.3	実装環境	32
4.3.1	サーバーサイドの実装	32
4.3.2	クライアントサイドの実装	33
4.3.3	広告配信プログラムの実装	34
第5章	評価および考察	36
5.1	動作検証	36
5.2	定性評価	36
5.2.1	要件1：透明性	36
5.2.2	要件2：ホワイトリスト属性の発信	37
5.2.3	要件3：ブラックリスト属性の発信	37
5.2.4	要件4：属性の切り替え	37
第6章	結論	38
6.1	本研究のまとめ	38
6.2	本研究の課題と展望	39
6.2.1	ユーザーの興味における潜在性	39
6.2.2	属性の切り替えの高度化	39
6.2.3	レコメンドシステムへの転用	40
	謝辞	41
	付録A 付録	44

目 次

1.1	2016年日本の広告費 出典：株式会社 電通 [1]	1
2.1	サイトターゲティング概略図 [14][15][16]	6
2.2	コンテンツターゲティング概略図 [14][15][16]	7
2.3	検索連動型広告概略図 [14][15][16]	8
2.4	行動ターゲティング概略図 [14][15][16]	10
2.5	リターゲティング概略図 [14][15][16]	11
2.6	デモグラフィックターゲティング概略図 [14][15][16]	12
2.7	ジオグラフィックターゲティング概略図 [14][15][16]	13
2.8	OSごとのフォントの描画結果一覧 [24]	17
3.1	実験の様子1	19
3.2	実験の様子2	19
3.3	Google, 広告カスタマイズページ (好みのトピック)[26]	20
3.4	Google, 広告カスタマイズページ (好みではないトピック)[26]	20
3.5	被験者53人のGoogle広告設定分布	21
3.6	HTTPの仕組み [28]	25
3.7	HTTPの仕組み [29]	26
3.8	HTTPの仕組み [29]	27
3.9	提案手法概略図	29
4.1	システム構成図	31
4.2	シチュエーション1	34
4.3	シチュエーション2	34
4.4	シチュエーション3	34
4.5	シチュエーション4	34
4.6	シチュエーションの切り替え	35
A.1	調査2・調査3の回答	45
A.2	調査2・調査3の回答	46
A.3	調査2・調査3の回答	47
A.4	調査1・調査4・調査5・調査6の回答	48
A.5	調査1・調査4・調査5・調査6の回答	49
A.6	調査1・調査4・調査5・調査6の回答	50

表 目 次

2.1	Result of Browser Fingerprint[20][21][22]	16
3.1	調査概要	19
3.2	質問項目	19
4.1	Open User Attribute の規格	31
4.2	使用ソフトウェアおよびハードウェアのバージョン	32
4.3	本実装において想定したユーザーのシチュエーションと属性情報	33
5.1	本研究の課題解決のために満たすべき要件	36

第1章 序論

本章では、本研究における背景・課題及び手法を提示し、本研究の概要を述べる。

1.1 本研究の背景

World Wide Web は 1990 年の登場以後急速な成長を遂げた。そして情報通信技術の進歩，スマートデバイスの登場も相まって，現在では人々の生活に欠かせない情報インフラとしての地位を確立した。利用者が増えたことで，広告を掲載する媒体として利用されることも増え，Web 広告市場は順調な成長を続けている。更に図 1.1 が示す通り，新聞・雑誌・ラジオなどの，既存媒体からの移行も順調に進んでおり，現在の広告市場において，最も成長著しい媒体であると言える。

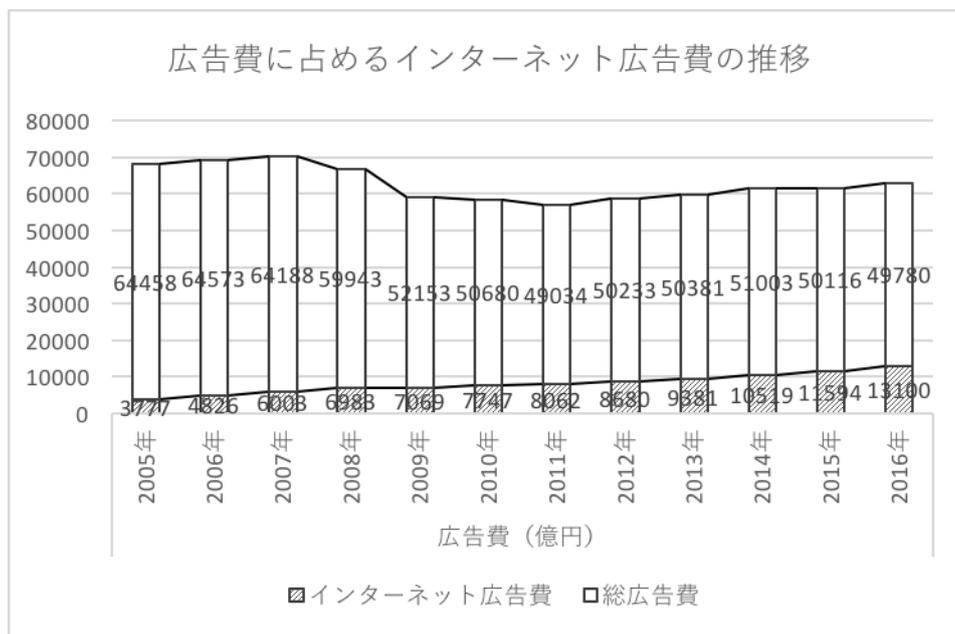


図 1.1: 2016 年 日本の広告費 出典：株式会社 電通 [1]

このように注目を集めている Web 広告であるが，その配信システムは非常に多様だ。サイトターゲティングやコンテンツターゲティングのような，サイトの内容に合わせて配信するもの，検索ワードに連動して配信する検索連動型広告，Web 上のユーザーの行動に連動して配信する行動ターゲティングなど，アプローチも様々である。しかし，抽象的な

仕組みの話は明らかにされていても、実際に広告を配信する過程のアルゴリズムが、公開されることはない。これは、そのアルゴリズムこそが、競合と差別化を図る重要なポイントの 1 つであり、企業が最も力を入れているポイントだからである。また同様の理由で、各アドネットワークがどの程度ユーザーの行動を追跡できているか、どのような属性を持つユーザーとして認識しているか、共有されることもない。結果として、多くのアドネットワークのシステムにおいて、ユーザー自身がトラッキングされた内容や自らの属性を確認することができないのである。これによりユーザーが、アドネットワークにインターネット上での行動や自身のプライバシーをどの程度追跡されているか、実態に気付くことを難しくしている。そしてこれは、ユーザーがアドネットワークそして Web 広告に対して不信感を抱く要因にもなる。なお、本研究では、Web 広告の配信の過程で、各ユーザーに適した広告を配信するために、アドネットワークが割り出した各ユーザーごとの興味のあるトピックの情報を、「属性情報」と定義する。

このインターネット上におけるプライバシーについての議論は、現在世界中で行われており、2013 年 6 月の Edward Snowden 氏による告発¹ によって、国家レベルで一般市民に対して個人情報収集を行っていることが明らかになった。このような社会状況から、市民レベルでもインターネット上でのプライバシーに対して、強い関心が向けられている。そのような社会的な需要もあり、企業がユーザーのプライバシーをどの程度把握しているのか明かされること、そしてそれがユーザーが管理可能であることの重要性は大きい。

一方で、現在の Web 広告のシステムは、アドネットワーク側にとってもデメリットを抱えている。

まず抑えなければならない点として、そもそも現在の Web 広告のシステムが、アドネットワークの思惑だけで設計されたものではないということである。巻末の参考文献を見てもらえれば分かる通り、ユーザーの行動を追跡する Web トラッキングに関する論文は、1 つは 1976 年 [3] のものであり、その他も、1986 年 [4]、1989 年 [5]、1990 年 [6] と、20 世紀終盤のものである。これらは、World Wide Web がインターネット上で利用可能なサービスとして公開された、1991 年 8 月 6 日以前のものであり、Web トラッキングの技術は、World Wide Web 黎明期から存在した。そして、Web の広告利用が叫ばれ、トラッキングの技術のこの分野での利用が議論され始めたのが、1990 年代の後半である。[7][8][9][10][11] この時代、コンピューターのスペックは今に比べて低く、Web でできることやブラウザの機能も簡素であった。この環境において、ユーザーの好みを詳細に反映することは難しく、サイトの内容や、ユーザーが起こした行動に対して広告を配信することしかできなかったのである。そして現在においてもユーザーの行動から、アドネットワークがユーザーの好みを推測して、広告を配信する構造は変わっていない。しかしながらこの間、インターネット環境の整備、スマートフォンの普及と SNS の登場によって、自らの私生活や好きなもの、プライバシーを公開しあう文化が生まれた。プライバシーの保護に気を使

¹当時 NSA(アメリカ国家安全保障局) 及び CIA(中央情報局) の局員で、システム分析官として情報収集活動に関わっていた Edward Snowden 氏が、NSA による全世界規模の個人情報収集活動について告発したもの。これには米国人一般市民を対象とした盗聴や同盟国内での通信傍受、そして「PRISM」と呼ばれる通信監視作戦の内容が含まれていた。この「PRISM」は、Google や Facebook、Microsoft などの米大手 IT 企業が協力し、各社のサービスに NSA が通信傍受をしやすいようバックドアを設置、NSA が容易に個人情報にアクセスできる体制を築くものだった。[2]

う人々がいる一方、プライバシーを発信する土壌が育まれていたのである。[12][13] このような流れの中において、Web の通信のレイヤーから個人を発信できる設計についても、議論されるべきである。

アドネットワークがユーザーの属性を推測する構図の最大の問題点は、推測の誤りが発生してしまう点である。技術の進歩により、ユーザーの好みや興味を極めて詳細に推測することが可能になったが、ユーザーの行動にも誤った行動が含まれており、その行動も属性の推測に利用されると、誤った属性が付与されてしまう。また、人間の興味は恒久的でなく、常に変化する。現在の仕組みだと、ユーザーのが様々な事情で変化した場合、アドネットワークはその変化に対応しきることができない。このような問題によって、ユーザーに対して不適切な広告が配信されることは、アドネットワークにとって大きな機会損失であり、収益を下げるものである。また、アドネットワークはユーザーの好みの属性は推測できても、無関心な属性と嫌いな属性を識別することは難しい。これらはアドネットワークにとってのデメリットであり、ユーザーの興味や好きなもの・嫌いなものの情報は彼らが欲している情報である..

1.2 本研究の着目する課題

前節で述べたように、現在の Web 広告システムにおいて、広告を配信するアドネットワーク側が持つ属性情報を、ユーザー側で管理することができない。これはユーザーにとって好ましくない状況であり、ユーザーが自分の追跡されている内容、システムに認識されている属性の内容を簡単に知ることができないため、アドネットワークに対して不信感を抱く要因になっている。また、ユーザーのインターネット上での行動が全て広告配信に利用されるため、間違って訪れたサイトの情報も、広告配信に利用される。他にも、興味のある商品について、その商品の購入後も関連の広告が配信され続けるなど、ユーザーエクスペリエンスを低下させている例は多い。そしてこれはアドネットワーク側のデメリットでもある。更に、広告が配信されるタイミングによって、ユーザーの実世界での状況は変化している。その際、ユーザーの置かれた状況に対して適切な広告の内容が変わるが、ユーザーの状況の変化をアドネットワークは検知できないため、適切ではない広告が配信されてしまうことが予想される。

1.3 本研究の手法

本研究では、既存の Web 広告の手法における問題を解決し、ユーザーが自らの属性を管理し、Web サイトを開く際の自らの状況によって属性を使い分けられる事を目的とする。上記の目的を達成する手段として、ユーザーが自分で属性を選び、自らアドネットワークに発信する手法をとる。HTTP リクエストヘッダーにユーザー属性を埋め込み、Web の通信をすることで、Web サイトを開くと同時にサイトにリンクされているアドネットワークに属性を発信する。ユーザーが自らの属性をユーザー自身で入力することで、属性の内容を把握でき、公開をしたくない情報が公開されてしまうことを防ぐことができる。こ

れによりユーザーのアドネットワークに対する信頼性が高まると考えられる。また、ユーザーの好みの属性のリスト（以後、ホワイトリスト）と嫌いな属性のリスト（以後、ブラックリスト）の内容を状況によって変更することができるため、ユーザーの実社会での状況に適した Web 広告を配信することが可能になる。

1.4 本論文の構成

本論文における以降の構成は次の通りである。

2 章では、既存の Web 広告の手法と関連技術について議論し、本研究の背景を明確化する。3 章では、本研究における問題の定義と、その問題に対する調査結果、解決するための要件と手法について説明する。4 章では、3 章で提案した、ユーザー属性埋込み型 HTTP 通信を利用した Web 広告システムを、設計・実装したことについて解説する。5 章では、3 章で求められた課題に対しての評価を行い、考察する。6 章では、本研究のまとめと今後の展望についてまとめる。

第2章 Web 広告

本章では，既存の Web 広告全般について説明する．

2.1 既存の Web 広告手法

本節では，現在利用されている Web 広告の手法について，各手法の構造と違いを明確にし，特徴を解説する．

2.1.1 サイトターゲティング

特定の Web サイトにのみに広告を配信する手法．そのサイトを訪れるであろうユーザーの属性を分析し，適した広告を配信している．図 2.1 のように，アドネットワーク側で絞り込んだ属性を配信に利用しており，属性から割り出せるユーザー像を 1 つに絞ることができる場合，有効な手法である．

■ サイトターゲティング

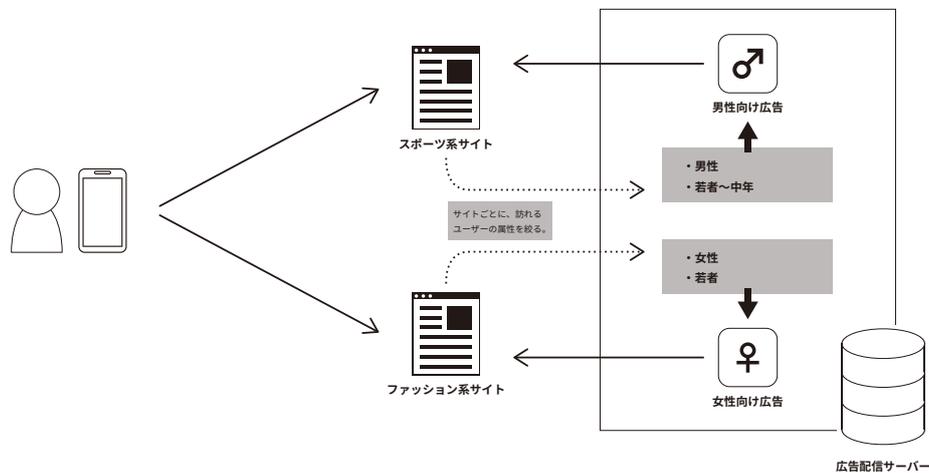


図 2.1: サイトターゲティング概略図 [14][15][16]

2.1.2 コンテンツターゲティング

登録したキーワードに対し、関連が高い Web サイトに広告を配信する手法。HTML ヘッダのメタキーワードなども利用されている。2.2 のように、ブログ・ニュースサイトにおいて、記事ごとに広告を配信することができるため、記事の内容によって同じニュースサイト内でも配信広告を切り替えることが可能。サイト単位ではなく、コンテンツ単位で訪問ユーザーの属性が変わる場合に有効な手法である。

■ コンテンツターゲティング

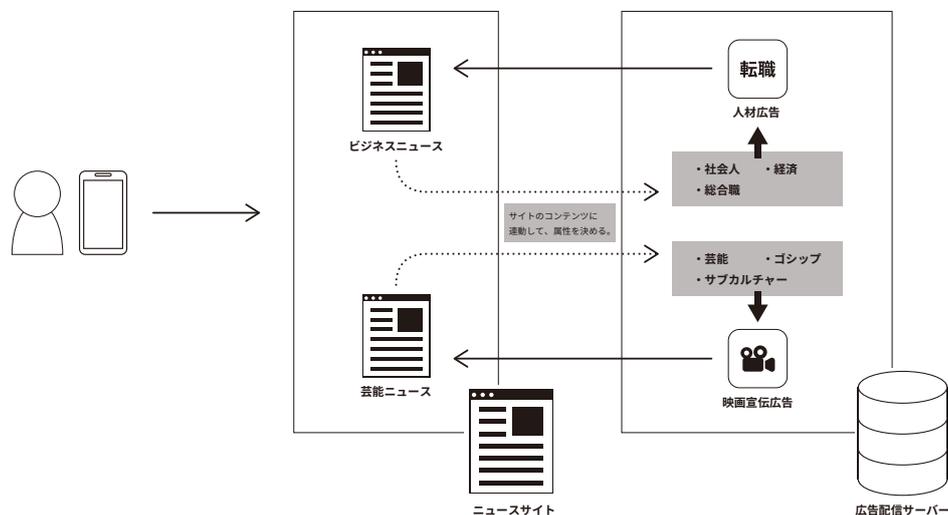


図 2.2: コンテンツターゲティング概略図 [14][15][16]

2.1.3 検索連動型広告

リスティング広告とも呼ばれる手法。図 2.3. 検索サービスでユーザーが検索を行った際に、検索結果と共に配信される広告。検索サービスを運営している会社自体が広告の配信事業を行っている。ユーザーがその時に興味を持ったワードを基に広告が配信されるため、他の手法に比べ、その広告にニーズのあるユーザーに対してのみ広告を配信し、ニーズを持たないあるいは少ないユーザーに対して広告が配信されることを防ぐことができる。

しかし、検索したユーザー以外に配信することができないため、配信可能なユーザーの範囲が限定的である。また費用に合わせて広告を配信する検索ワードをこまめに確認する必要があり、アドネットワークに広告を出稿期間中はコンバージョンやクリック数などを基に調整することで、広告効果が高まる。

■ 検索連動型広告

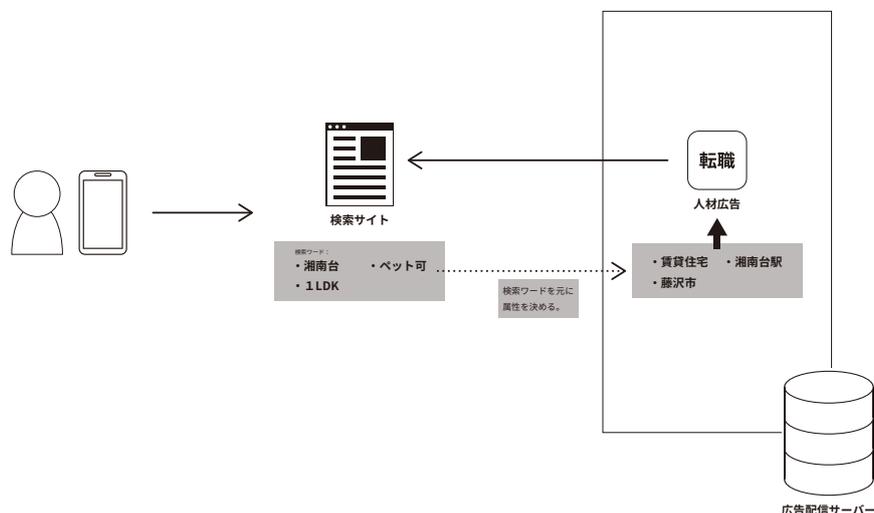


図 2.3: 検索連動型広告概略図 [14][15][16]

2.1.4 行動ターゲティング

Web ブラウザの Cookie から取得したユーザーの行動履歴をもとに、ユーザーの関心・興味を推測し、適切な広告を配信する手法。図 2.4。また、広告へのクリック履歴や、e コマースサイト（以後、EC サイト）での購入履歴を元に、広告を配信するものもある。特定の事象に対して確実に関心を持つユーザーに、広告を配信できる。しかし、ターゲットとなるユーザーの母数は少ないため、配信可能な数は限られている。このような詳細な行動履歴を利用するには、ユーザーの承認を得ることが必要¹。

この手法において行動の追跡には、Cookie を始め、以下の手法がある。

Cookie

Cookie(クッキー)とは、Web ブラウザ内に蓄積される来歴情報で、Web サイト側が Web ブラウザを通じ、ユーザのコンピュータに一時的にデータを書き込んで保存させる仕組みを持つ。Web サイト側が訪問者を識別するための仕組みであり、ログイン情報の自動入力などに使われている。ユーザの意思とは

¹ オプトイン・オプトアウトの設定

無関係に、自動的にユーザに関連する情報、そのサイトの訪問回数が記録される。住所や名前を毎回入力する手間を省くことができる一方、保存される情報は一切暗号化されておらず、セキュリティ関連の情報については慎重になる必要がある。この Cookie を用いたユーザーの追跡技術については、2.2.1 で詳しく述べる。

携帯電話の契約者 ID

携帯電話で Web サイトを閲覧した時に、サーバーに送信される識別子。これは各携帯電話事業者が独自に付与するもので、ID の形式は事業者によって異なる。

ディープパケットインスペクション

コンピュータネットワークのパケットフィルタリングの一種。インスペクションポイントをパケットが通過する際に、パケットのデータ部と、場合によってはヘッダ部を検査すること。インターネットサービス・プロバイダ（以後、ISP）にこの装置を設置し、全ての通信を分析することで、その ISP の契約者の属性を把握することができる。

■ 行動ターゲティング

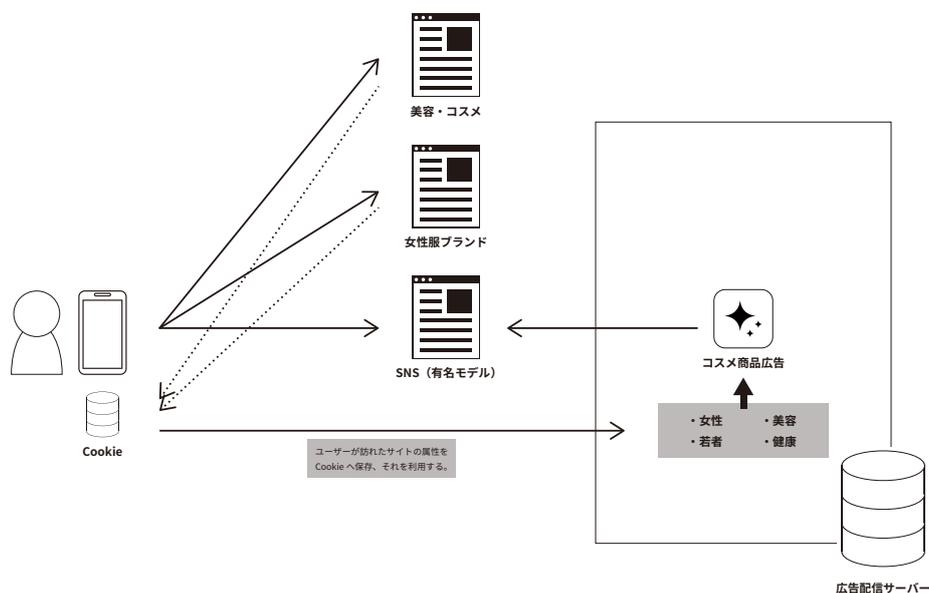


図 2.4: 行動ターゲティング概略図 [14][15][16]

2.1.5 リターゲティング

行動ターゲティングの一種。図 2.5. 広告主の Web サイトを訪れたユーザーに対し、別のサイトへの来訪時に、広告主のサイト自身もしくは関連の広告を配信する手法。あるサイトで一度配信された広告と同じものを、別のサイトでも配信することもできる。すでにサイトを訪問している見込みの高いユーザーに対して、広告を配信できる一方、一般的な行動ターゲティングよりもターゲットは狭い。

また、同じ広告が何度も表示されることに対し、ユーザーからの抵抗を招く恐れがあり、配慮が必要である。

■ リターゲティング

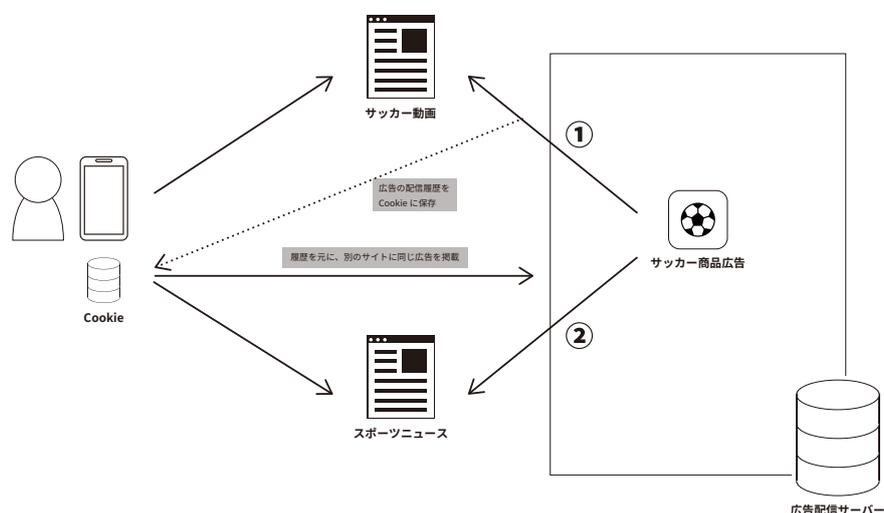


図 2.5: リターゲティング概略図 [14][15][16]

2.1.6 デモグラフィックターゲティング

行動ターゲティングの一種。ユーザーごとに個々のアカウントを持つ、アカウント紐付け型のサービスで行われる手法。会員情報はもちろん、アカウント情報に紐付いたユーザーの行動履歴も Cookie と合わせて利用されている。図 2.4 は、会員情報を利用して広告配信をした例である。

この手法においてサイトの履歴の取得は、Cookie の情報を用いる以外にも、ブラウザにアカウントがログインされていれば、そのアカウントに紐付いて履歴を収集することができる。また、行動履歴はサイトの履歴の利用に留まらず、サービスによって異なるが、サイト内でカーソルが動いた地点と軌跡、各サイトを閲覧した時間や、EC サイトでの購入履歴、そのサービスのアカウントでログイン中であれば、GPS やマップアプリでの行動履歴から、生活圏を割り出すことも可能である。GPS を利用した生活圏の推定は、次の小節 2.1.7 で触れる、ジオグラフィックターゲティングの手法でもある。このようにデモグラフィックターゲティングでは、非常に多岐にわたる情報が収集されている。

■ デモグラフィックターゲティング

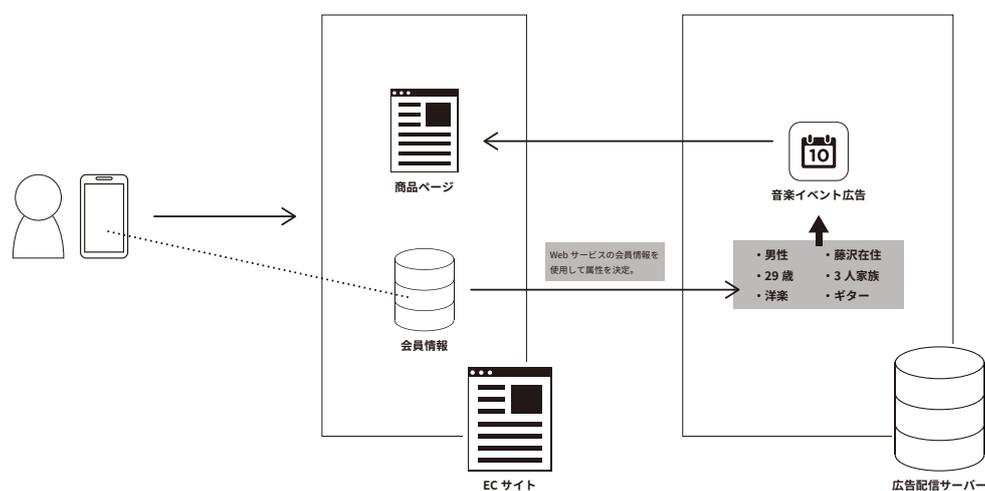


図 2.6: デモグラフィックターゲティング概略図 [14][15][16]

2.1.7 ジオグラフィックターゲティング

ユーザーの位置情報を利用して、広告を配信する手法。位置情報の取得には、Wi-Fi や GPS の情報、IP アドレスなどが利用され、これらの情報から取得した位置情報からユーザーの現在地を推定し、近隣の店舗の広告や、店舗のクーポンなどを配信する事が可能。また、位置情報を保持しておくことで、ユーザーの生活圏の特定も可能である。オンラインとオフラインを繋げたプロモーションを行う際に有効な手法である。

■ ジオグラフィックターゲティング

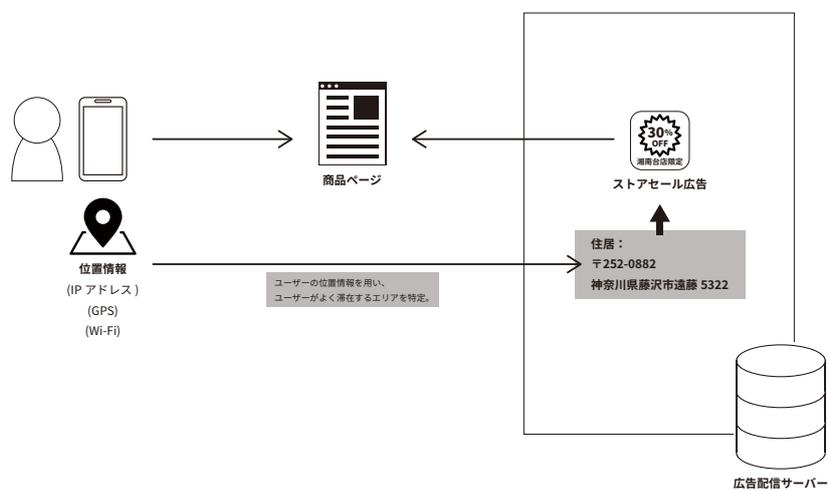


図 2.7: ジオグラフィックターゲティング概略図 [14][15][16]

2.2 関連技術

本節では、2.1 で紹介した Web 広告手法において、ユーザーを特定し、その行動を追跡・蓄積するためのトラッキング技術について解説する。

2.2.1 Cookie を利用したトラッキング技術

現在最も多く利用されているトラッキング技術が、ブラウザの Cookie を利用するものである。これは、サイトに埋め込まれた Web 広告を配信しているアドネットワークが、トラッキング Cookie (もしくはサードパーティクッキー) と呼ばれるものをユーザーの端末に送り込む手法である。このトラッキング ID にはユーザーの識別子が埋め込まれており、アドネットワークからユーザーを識別することが可能になる。具体的な仕組みを述べると、まずユーザーが訪れたとある EC サイトに A 社の Web 広告が配信されている。ここで A 社がトラッキング Cookie を発行しユーザーの端末に保存されるとともに、そのサイトの訪問履歴が A 社側に残る。次にユーザーが A 社が広告を配信しているニュースサイトを訪れると、保存されていた A 社のトラッキング Cookie が送信され、以前訪れた EC サイトで閲覧していた商品の広告が配信される。またこの時ニュースサイトの閲覧記録が A 社側に保存される。これを繰り返していくことで A 社にはユーザーの行動履歴が蓄積されていき、ユーザーの属性情報の推測の精度が向上していくのである。[17][18]

以下は実際に Google からトラッキング Cookie が発行された際のレスポンスヘッダの内容である。

```
1 HTTP/1.1 200 OK
2 Date: Mon, 22 Jan 2018 06:59:12 GMT
3 Expires: -1
4 Cache-Control: private, max-age=0
5 Content-Type: text/html; charset=Shift_JIS
6 P3P: CP="This is not a P3P policy! See g.co/p3phelp for more info."
7 Server: gws
8 X-XSS-Protection: 1; mode=block
9 X-Frame-Options: SAMEORIGIN
10 Set-Cookie: 1P_JAR=2018-01-22-06; expires=Wed,
11     21-Feb-2018 06:59:12 GMT; path=/; domain=.google.co.jp
12 Set-Cookie: NID=122=HwUL6sh1_***OPdwoCd;
13     expires=Tue, 24-Jul-2018 06:59:12 GMT;
14     path=/; domain=.google.co.jp; HttpOnly
15 ***
16 Accept-Ranges: none
17 Vary: Accept-Encoding
18 Transfer-Encoding: chunked
```

このレスポンスヘッダ内の Set-Cookie が発行された Cookie の内容である。

2.2.2 Browser Fingerprint

2.2.1 の Cookie を利用したトラッキング技術が一般的になり、プライバシー保護の意識の高いユーザーは、Web 利用時の匿名性を維持するため、日常的に Cookie を削除したり、

ブラウザのシークレットモードやプライベートブラウジングの機能を使って Cookie を自動的にブロックする，他にも VPN などの手法で IP アドレスを隠蔽したりなど，さまざまな解決策を実施するようになった。

このような環境の中で広まったのが，ブラウザ・フィンガープリントである．これは Web サイトを閲覧しているコンピューターに対して，サイト側が把握できる情報を収集し組み合わせる．そうして集めた情報のコンピューターごとの違いから，どのコンピューターからのアクセスか，つまりどのユーザーなのかを特定するというものである．ブラウザが収集可能な情報としては，ブラウザの種類とバージョン，OS の種類とバージョン，画面の解像度，利用可能なフォント，プラグイン，時間帯，言語とフォントの設定，など多岐に渡る．他にもタッチパネルが付いているかなどハードウェア構成も把握することができ，これらの情報は 1 つだけでは，アクセスしているコンピューターを特定することは難しい．しかし，他の情報と組み合わせることで，数百万人に 1 人程度の精度まで，絞り込むことができる．

そこで実際にどのような情報がブラウザから取得可能なかを検証するため，ブラウザフィンガープリントのデモサイト [19] を用いて，筆者のブラウザから取得可能な情報を調査した．以下がその結果である．なお，個人特定を防ぐために幾つかの結果を”****”を用いて伏せ字にした．

表 2.1: Result of Browser Fingerprint[20][21][22]

Browser Characteristic	bits of identifying information	one in xbrowsers have this value	value
Limited supercookie test	0.38	1.31	DOM localStorage: Yes, DOM sessionStorage: Yes, IE userData: No
Hash of canvas fingerprint	***	***	***
Screen Size and Color Depth	***	19.53	1440x900x24
Browser Plugin Details	7.13	***	***
Time Zone	7.21	148.27	-540
DNT Header Enabled?	1.23	2.34	FALSE
HTTP_ACCEPT Headers	***	***	text/html, /*; q=0.01 gzip, deflate, br ja, en-US;q=0.9,en;q=0.8
Hash of WebGL fingerprint	***	***	***
Language	8.51	363.76	ja
System Fonts	6.5	90.82	Andale Mono, Arial, Arial Unicode MS, Book Antiqua, ***
Platform	3.05	8.31	MacIntel
User Agent	***	***	Mozilla/5.0 AppleWebKit/537.36 Chrome/63 ***
Touch Support	0.58	1.49	Max touchpoints: 0; TouchEvent supported: false; onTouchStart supported: false;
Are Cookies Enabled?	0.2	1.15	Yes

上記の結果から分かる通り、多岐にわたる項目において、それぞれ細かい数値やパラメータの値が取得できていることが分かる。

2.2.3 Canvas Fingerprint

Canvas Fingerprint は 2.2.2 で述べた、Browser Fingerprint の手法の 1 つである。これは HTML5 で導入された `canvas` 要素を用いる手法である。`canvas` 要素は JavaScript によって図形や文字を描画することができる。Web サイト内で、グラフを描く、写真を合成する、アニメーションを作成する、動画のリアルタイム加工やレンダリングを行うために利用される。[23]

この canvas による描画結果は、OS やブラウザの種類とバージョン、端末の物理構成によってピクセル単位の誤差が生じる。具体的には以下のように誤差が生じる。[24]

- ・2d モードでフォントをレンダリングする際、ブラウザのバージョンやインストールされているライブラリ・フォントにより、ピクセル単位でレンダリング結果が異なる。

- ・WebGL を使ってレンダリングする際、マシンの GPU によってレンダリング結果が異なる。

Canvas Fingerprint ではこの誤差を用いる。`jcanvas`要素を用いて複数の文字や図の描画させ、結果のピクセルデータをハッシュ化することによって扱いやすくしている。ユーザーごとのピクセルデータはそれぞれ固有の描画結果を持つので、ユーザーの識別に使うことができる。Canvas Fingerprint の結果は、「パフォーマンスの向上を目的に、ブラウザが OS の機能やハードウェアと密接に結び付いている場合、Fingerprint は固有のものになる」[24]とされており、識別の精度は9割ほどである。[25]

以下は OS ごとの描画結果の違いの例である。

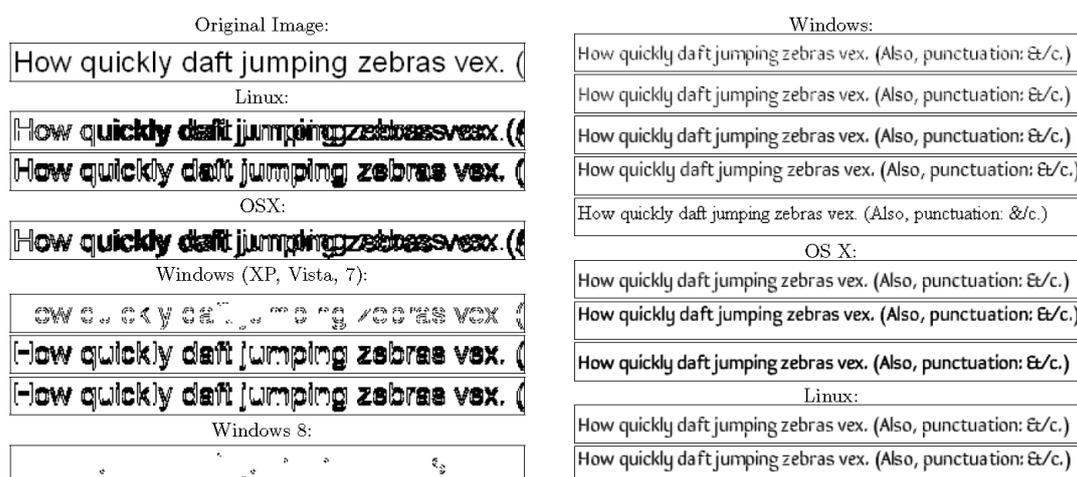


図 2.8: OS ごとのフォントの描画結果一覧 [24]

第3章 提案手法

本章では、1章で述べた背景に対して、詳細を明らかにするための調査、そこから考察される課題解決のための要件について議論し、提案システムを概説することで本研究で用いるアプローチについて述べる。また、提案手法を実現する上での技術概要についても説明する。

3.1 既存手法の問題点

本節では、2章で紹介した既存手法の問題点について述べる。2.1で紹介した7つの手法を見ると、既存のWeb広告の配信手法は、全てアドネットワークがユーザーの好みを推測する、というアドネットワークからユーザーへの一方的な関係であることが分かる。サイトのコンテンツの内容、検索ワード、トラッキングしたデータなど手段は様々であるが、ユーザーの好みをアドネットワークが一方的に探る構図は変わらない。アドネットワークやそのトラッキングに対し嫌悪感を抱くユーザーは、アドネットワークによるCookie利用を拒否する、Cookieを定期的に消去する、アドブロックと呼ばれるWeb広告を非表示にするソフトウェアを用いる、などの対策をしている。しかし、2.2で述べたBrowser Fingerprintのように、Cookieを用いずとも個人を特定しトラッキングできる技術も開発されている。このような一方的な関係しか選択肢がない状況は、ユーザーにとって大きな問題であり、アドネットワークに不信感を抱く原因である。

一方でアドネットワーク側にとっても問題を抱えている。既存のWebトラッキング手法において、ユーザーの行動に含まれる誤った行動を判別することが出来ない。ユーザーが間違えて訪問したWebサイトの情報も推測に利用されるため、誤った属性が付与されてしまう。また、ユーザーの行動から好みのトピックを推測することはできるが、その他多くの関心のないトピックと嫌いなトピックを判別することはできない。これらの問題から、Web広告をその広告の属性が嫌いなユーザーに配信してしまった場合、広告効果が見込めないどころかその企業のイメージを損ねてしまう恐れがある。

3.2 課題に対する調査

3.1で述べた問題点について、現状を明らかにし課題を明確にするため、インタビュー調査を行った。

3.2.1 調査概要

以下が調査の概要である。

表 3.1: 調査概要

日時	2018 年 1 月 15 日 (月), ~ 2018 年 1 月 19 日 (金)
調査対象	慶應義塾大学 湘南藤沢キャンパスの学生 53 人 (総合政策学部・環境情報学部・看護医療学部) 男子：28 名女子：25 名
調査内容	Google が自身に設定した属性に対する印象調査。 属性の参照元：Google 社の広告カスタマイズページ [26]

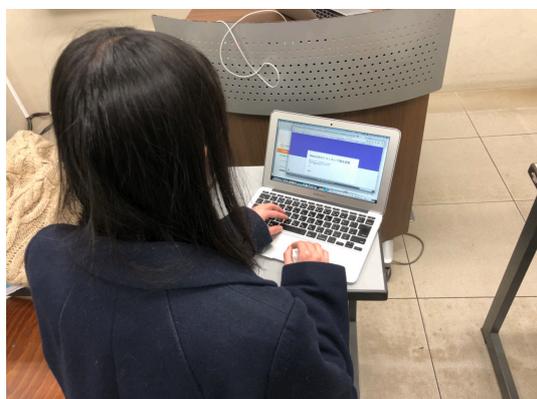


図 3.1: 実験の様子 1

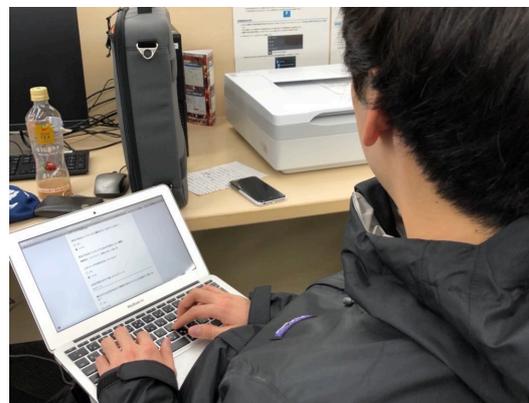


図 3.2: 実験の様子 2

3.2.2 調査手法

本調査は全て、筆者と被験者の一対一による対面のインタビュー調査で行った。調査にはアドネットワークの中でも、ユーザーに属性情報を開示している Google の広告システムにおける、属性情報を利用した。また、質問の項目は以下である。

表 3.2: 質問項目

調査 1	Google の広告のカスタマイズページを知っていたか。
調査 2	現在、自身に設定されているトピック（属性）（図 3.5）の数
調査 3	自身に設定されている好みのトピック（ホワイトリスト属性）（図 3.5）の中で、興味のないもの（不適切な属性）
調査 4	好みではないトピック（ブラックリスト属性）（図 3.4）に自分が追加したいもの
調査 5	Web 広告の内容で困ったエピソード
調査 6	自分の属性を自分で管理する必要があるか。 (Google が属性情報を保持することで実現している便利機能を合わせて説明)



図 3.3: Google, 広告カスタマイズページ (好みのトピック)[26]

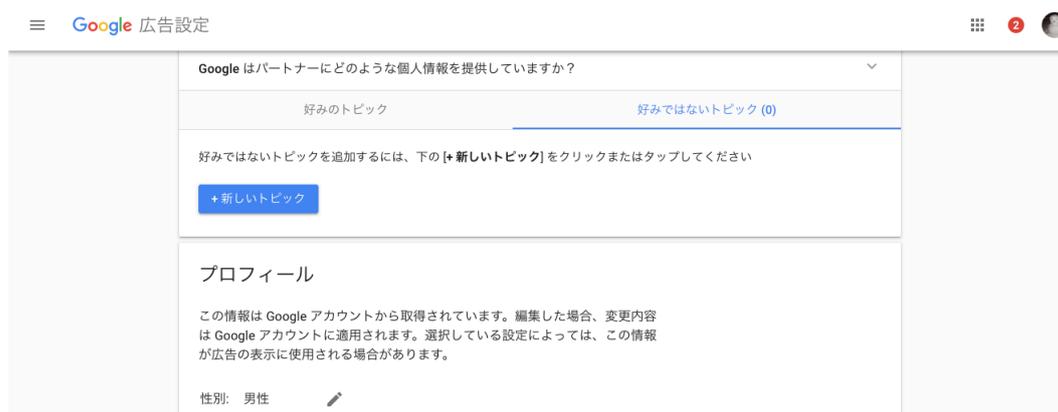


図 3.4: Google, 広告カスタマイズページ (好みではないトピック)[26]

3.2.3 調査結果

本調査の結果についてまとめる。巻末の付録 A に付録として本調査の回答一覧を添付する。

Google 広告設定の認知度

Google の広告設定 [26] では、ユーザーの行動情報を用いた広告のパーソナライズを拒否することができる。(初期設定は許可。) 今回の調査の被験者においては、Google のアカ

アカウントを所持していなかったのが 5 人，初期設定のままで広告のパーソナライズを許可していたのが 43 名，拒否していたのが 5 名だった。また，この広告設定管理画面の存在を知っていたのは，パーソナライズを拒否していた 5 人の内の 2 人である。それ以外の 3 名は Google のアカウント作成の際に，パーソナライズ設定を拒否にしていたものと思われる。

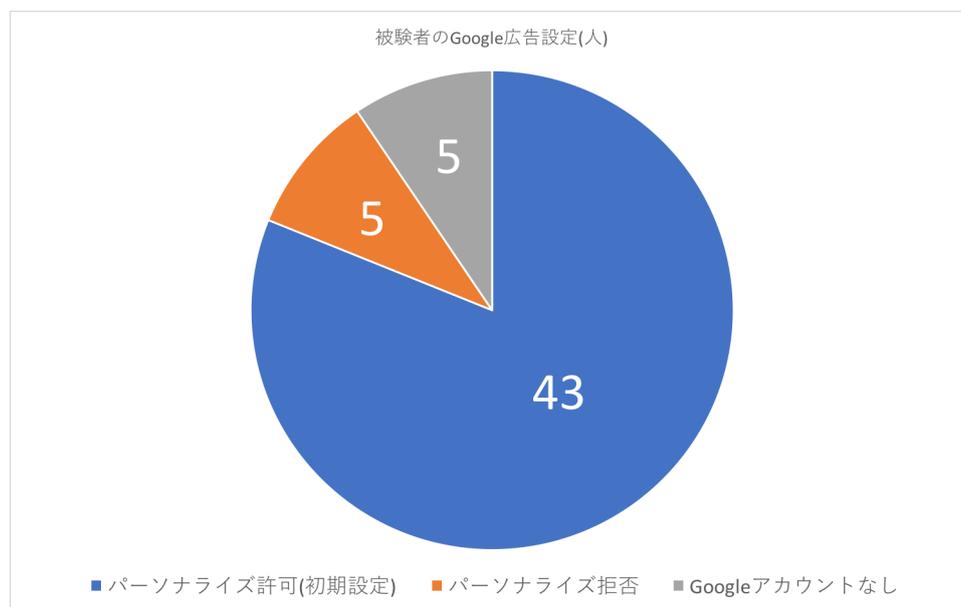


図 3.5: 被験者 53 人の Google 広告設定分布

属性推測の精度

パーソナライズ設定を許可していた 43 名それぞれに設定されているトピック（属性）の数は，51.88 個である。またトピック数が最も多かった例は 89 個で，最も少なかった例は 15 個であった。

次に，設定されているトピックの中で，興味のないトピックを挙げてもらった。全トピック数における興味のないトピックの割合は 14.43% であった。興味のないトピックが最も多かった例は 40.00% (15 個中 6 個) で，最も少なかった例は 2.27% (44 個中 1 個) であった。

ブラックリスト属性

被験者 53 人全員に対して，ブラックリスト属性に追加し配信されないようにしたい広告の属性を調査した。その結果，最も表が集まったのが男性向けアダルトコンテンツで 15 票，その次が出会い系サービス・アプリの広告とゲームの広告で 13 票の同数だった。また，仮想通貨に関する広告が 9 票を集めるなど，社会のトレンドも反映された結果となった。

広告内容に対する不満

調査5で、表示された広告の内容に関してユーザーが困ったエピソードを調査した。最も多く表を集めたのがアダルト系の広告で16票を集めた。理由としては、女性を中心にそもそもの広告の内容や写真・イラスト自体に不快感を感じる、という意見がある一方で、他人に見られてしまう環境でアダルト系の広告が出てしまった際に、恥ずかしい・誤解を生む、という意見が多く見られた。他、代表的な意見としては、プレゼントを贈ろうと探していた時期に、送る相手と一緒にWebサイトを見ていたら送る予定の商品がWeb広告としてでてしまい、プレゼントがバレてしまう、という意見。すでに購入済みで興味を失った広告が出続ける、という意見、Wi-Fi環境ではないにも関わらずアプリのインストールを促す広告が配信され、仮にインストールしたくてもできないという意見、などがみられた。なお、広告の内容以外のものに対する不満は除外した。

属性情報を自ら管理する必要性

ユーザーに付与されているトピックがトラッキングを元にしたものである事実と、一方でそれによりGoogleのサービス内で実現されている便利機能について解説した上で、属性情報をアドネットワークではなくユーザー自らが保持・管理する必要があるか、質問した。結果として53人中44人が必要であると答え、9人が必要ないと答えた。必要ないと答えた被験者の意見としては、「便利な機能を提供してくれるのであればそれでいい」「この程度のプライバシー情報であれば問題ない」というものだった。

3.2.4 調査結果についての考察

本調査によって、Googleによる属性情報推測の精度が明らかになった。結果としてこれは平均で14.85%であった。この事実からも分かる通り、Webトラッキングを使った属性情報の推測には、一定数の誤った属性が含まれてしまう。この要因としては、3.1で触れた、誤った行動情報が含まれていることに加え、本調査でも見られたユーザーの興味が失われてしまった属性が残っている、という要因が想定される。時間経過による興味の消失以外にも、すでに購入した商品についての属性が残っているなどの例が考えられる。

3.2.3の結果から、ある特定の広告にユーザーの不満が集まっていることが分かる。これは2.1.4で示した通り、行動ターゲティング広告ではターゲットとなるユーザーの母数が少ないため、配信可能な数は限られてしまうことに起因する。つまりクライアントから広告の配信依頼を受けたユーザーの人数に対し、依頼された広告属性情報に適したユーザーの数が足りていないのである。そして合致する属性を持たないユーザーにも広告が配信されるが、その中にはその広告を嫌いな属性としていたユーザーも含まれる。するとこのようにユーザーが不快感を持つ広告として印象に残ってしまうと考えられる。

また、3.2.3の結果から約8割のユーザーが、属性情報をユーザー自らが保持・管理する必要性を感じていた。3.2.3においても1割ほどのユーザーが広告のパーソナライズを拒否に設定しており、プライバシー保護に高い意識を持つユーザーの存在も明らかになっ

た。本調査を通して、全てのユーザーが調査項目のどれか1つにはアドネットワークに対して否定的な回答をしており、ユーザーが不満の方向性は様々であるが、なんらかの不満を抱えていることは明らかである。

状況に応じて属性を切り替えることの必要性

3.2.3 で多く見られた意見として、広告そのものが悪いのではなく、ある状況下においてだけ問題となるケースだ。もっとも多く意見が集まったのが他人と画面を共有しているというシチュエーションで、友人に画面を見せる、プレゼンで自分の画面を全員が見る、などの状況では、最適な属性が通常時とは大きく異なることが分かった。具体的な例としては、アダルト系が16票を集めた。アダルト系の広告が配信された画面を友人に見せるのが恥ずかしい、友人から自分の趣味を疑われる、という意見だった。他にも物件の広告で自分の所在地がバレてしまう、商品広告からプライバシーが明かされてしまうなど、自分が見るか他人が見るかでユーザーの広告に対する価値観は大きく変わるのである。また、他のシチュエーションとしては、Wi-Fi環境ではないのにアプリやゲームのインストールをさせる広告が配信され、仮に興味を持ってダウンロードできない、というものがあつた。

この結果から Web 広告はユーザーの実世界の状況に応じて、最適な内容に切り替える必要があると考える。スマートデバイスの登場によって Web サイトが閲覧される場所は、自宅のデスクトップの前からユーザーの生活のあらゆる場所が変わった。しかし Web 広告配信の根本的な構図は変わっておらず、ユーザーの生活スタイルの変化が反映されていない。また昨今のデバイス性能の向上によって、ユーザーの状況を認識しそれを Web 広告の内容に反映することは性能的に実現可能な範囲である。本研究ではこの機能を実現すべき重要な要件であると考えられる。

3.3 技術概要

本研究では、提案手法を実現する上で、データ送信の方法として、HTTP を用いる。本節では、HTTP の規格や仕様について説明する。

3.3.1 HTTP 通信の仕組み

[27][28][29] HTTP とは Hypertext Transfer Protocol の略で、Web サーバとクライアントが HTML などのコンテンツの送受信する際に用いられる、アプリケーション層のプロトコルである。トランスポート・プロトコルは通常の場合 TCP を使用する。HTTP では、データを要求する HTTP リクエストと、それに応答して要求されたデータを送り返す HTTP レスポンス、この 2 つのやりとりを繰り返して Web ページを表示している。仕組みとしては、まずユーザーがブラウザに URL 入力すると HTTP の処理が開始される。通信には通常 80 番のポート番号が使用される。次にクライアントからサーバへ、ポート 80 番で TCP コネクションの確立が行われる。ここで確立された TCP の通信路を利用してコマンドやデータの送受信が行われていく。

レスポンスの内容である HTML, JPEG といったデータを成形し描画するのは、Web ブラウザの機能である。描画結果はブラウザの仕様によって変わるため、全ての Web ブラウザで見た目が同じになることが保証されているわけではない。また、ブラウザの中にはテキストしか表示できないテキストブラウザというものも存在する。

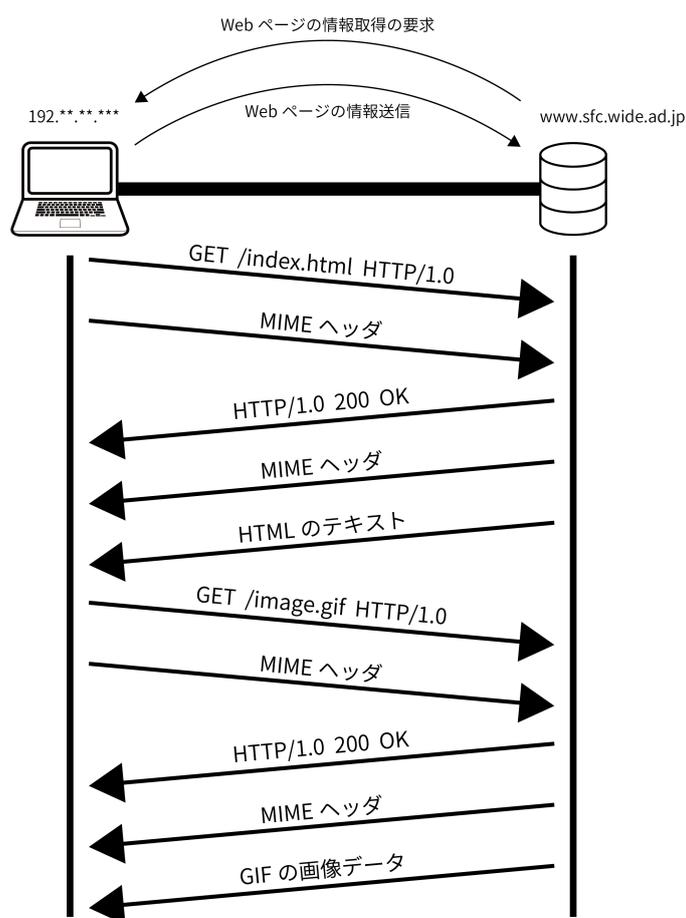


図 3.6: HTTP の仕組み [28]

リクエストメッセージ

[27][28][29] HTTP リクエストのメッセージ内容について述べる。HTTP リクエストのメッセージはリクエスト行、メッセージヘッダ、空白行、メッセージボディで構成されている。

リクエスト行には、メソッド、URI、HTTP バージョンの情報が含まれる。メソッドとは Web サーバに要求を示すコマンドのことである。URI とはリクエストの対象となるデータを指す情報のことである。HTTP バージョンとは、Web ブラウザがサポートする HTTP プロトコルのバージョンのことである。メッセージヘッダには、Web ブラウザ側でサポートするデータのタイプ、OS の種類、データの圧縮方法、ブラウザの種類、Cookie などの情報が含まれる。なお HTTP リクエストのメッセージヘッダを以後リクエストヘッダと定義する。空白行は、Web サーバに対してメッセージヘッダの終了を伝えるために使用される。メッセージボディは Web サーバにデータを送るために使用する。具体例としては Web ページ上で入力欄がある場合、入力したテキスト情報を Web サーバに送るために使用される。入力情報がなければ空白のまま送信される。[29]

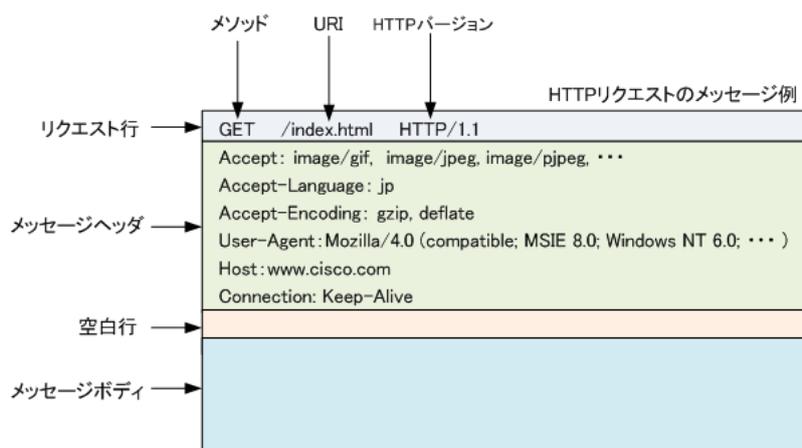


図 3.7: HTTP の仕組み [29]

レスポンスメッセージ

[27][28][29] HTTP レスポンスのメッセージ内容について述べる。HTTP レスポンスのメッセージはステータス行、メッセージヘッダ、空白行、メッセージボディで構成されている。

ステータス行には、HTTP のバージョン、ステータスコード、説明文などの情報が含まれる。Web ブラウザに、Web サーバでの処理結果を伝えるために使用される。メッセージヘッダには、サーバの種類、返信するデータのタイプ、データの圧縮方法などの情報が含まれる。2.2.1 で述べた Web トラッキング用の Cookie を設定する Set-Cookie もこの部分に含まれるパラメータである。Web ブラウザに Web サーバの情報を示すために使用される。空白行は、Web サーバに対してメッセージヘッダの終了を伝えるために使用される。メッセージボディは、HTML 文書、画像ファイル、動画ファイルなどのデータを格納するために使用される。

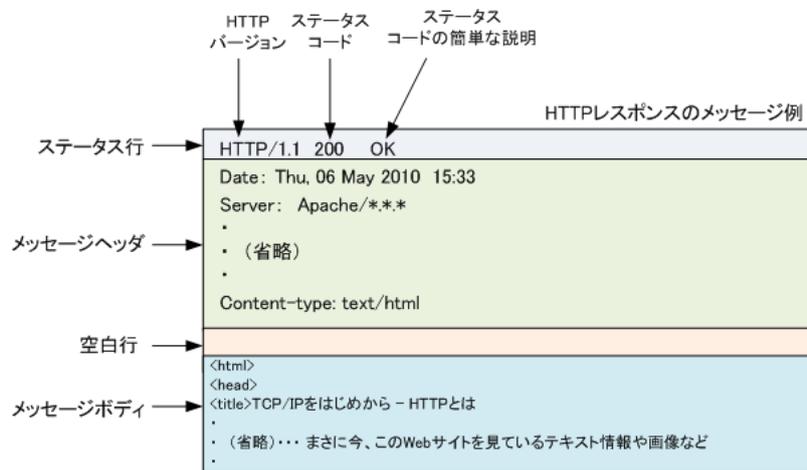


図 3.8: HTTP の仕組み [29]

3.4 課題解決のための要件

本節では、3.1 節で述べた既存手法の問題点と、3.2 節の調査の考察を踏まえ、Web 広告のシステムに対するユーザーの不信感を解消し、アドネットワーク側のデメリットを解消するために、必要な要件を述べる。

3.4.1 透明性

ユーザーが Web 広告のシステムに対して不信感を頂く要因の一つが、アドネットワークがユーザーの行動を追跡し、属性を勝手に推測し、広告を配信する、という配信過程における透明性の欠如である。事実、3.2 の調査の結果を見ても、属性情報をユーザー自ら管理する必要性を、被験者の%が感じていた。そこで、この透明性の実現のため、属性情報がユーザーの管理下にあることが求められる。

3.4.2 ホワイトリスト属性の発信

3.1 で述べた通り、現在の Web 広告の配信システムは、アドネットワークからユーザーへの一方向的な関係が中心で、ユーザーが自分の好みを発信する構造が乏しいため、アドネットワークはユーザーの閲覧履歴を始めあらゆる行動から、ユーザーの好みを探っている。しかし、1.1 で触れた通り、ユーザーが自分の好みを発信することに抵抗を感じないユーザーが増えている。そこで、Web 広告の配信システムにおいて、ユーザーが自分の好みを発信できる仕組みが求められる。

3.4.3 ブラックリスト属性の発信

アドネットワークが抱える、属性の推測に誤りが生じてしまう点、ブラックリスト属性を判別出来ない点、を解消するため、ブラックリスト属性の発信が必要である。また、3.2 の調査の結果を見ても、被験者の%がブラックリスト属性に追加したいトピックを抱いており、ユーザーエクスペリエンスの観点からもブラックリスト属性の配信が求められる。

3.4.4 属性の切り替え

現在の Web 広告システムにおいて、全く達成されていない項目が、ユーザーの実世界での状況に合わせた、属性の切り替えである。Web 広告は、ユーザーが置かれている状況次第で、配信される広告が適切かどうかが変わることは、3.2.4 から明らかである。そこで、Web 広告におけるユーザーエクスペリエンスを改善するために、ユーザーの状況に合わせて、属性を切り替えることが求められる。

3.5 提案手法

提案手法の概要について述べる。3.4 で述べた要件を満たすために、本研究では、ユーザーが自ら決めた属性を元に Web 広告を配信し、ユーザーの実世界での状況によって属性を使い分けることができる、新しい Web 広告配信手法を提案する。これを実現するために、本研究では、従来はアドネットワーク側で管理・保持されていた属性情報を、ユーザーのブラウザ内で管理し、HTTP の通信におけるリクエストヘッダに埋め込み、受け渡された属性情報を用いて、Web 広告を配信する。

以下が提案手法の概略図である。

■ 提案手法

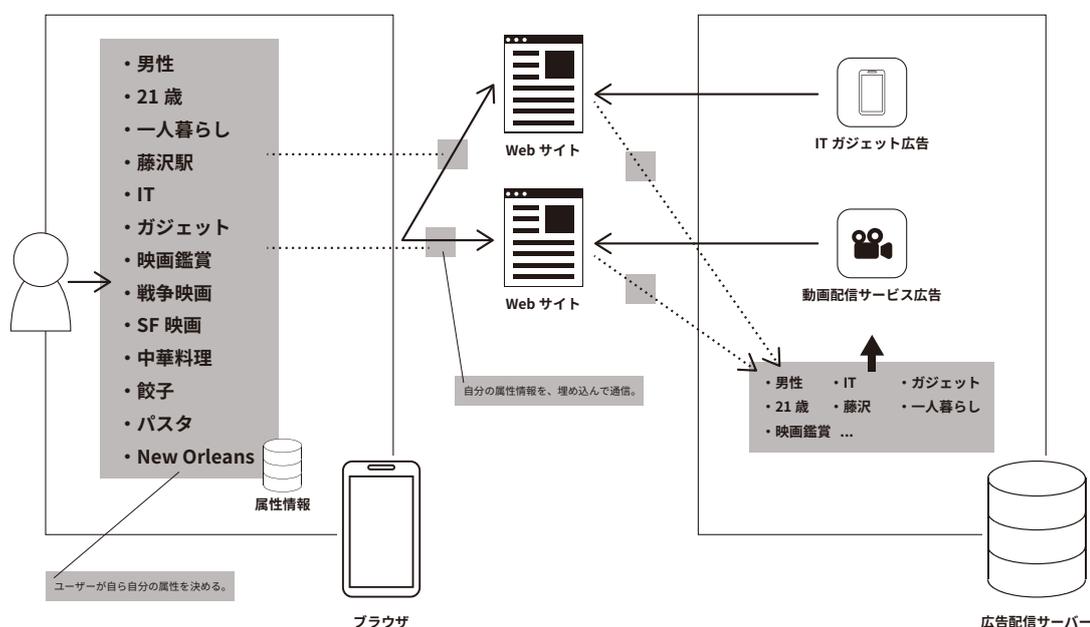


図 3.9: 提案手法概略図

第4章 システムの設計と実装

本章では、3.5節で述べた提案システムの実験として行った実装に関して説明する。本システムは、ブラウザアプリケーションと、サーバーサイドアプリケーションから構成される。さらに、ユーザーが設定する属性値の規格についても説明する。

4.1 システム概要

実装システムの概要について説明する。今回の実装においては、ブラウザ側でユーザーによる属性情報の管理、属性情報のリクエストヘッダへの埋め込みを行い、サーバー側で属性情報の抽出とHTMLデータのレスポンスを行うよう設計した。また、3.9で示しているように、提案手法では属性情報の認識と配信する広告の決定はアドネットワーク側で行う。そのため、抽出した属性情報を必ずHTML文書に受け渡し、Webサイト側で扱えるようにすることを心がけた。

ブラウザ側はModHeader[30]というChromeエクステンションアプリケーションを用いた。サーバー側はnode.jsで実装した。

■ システム構成図

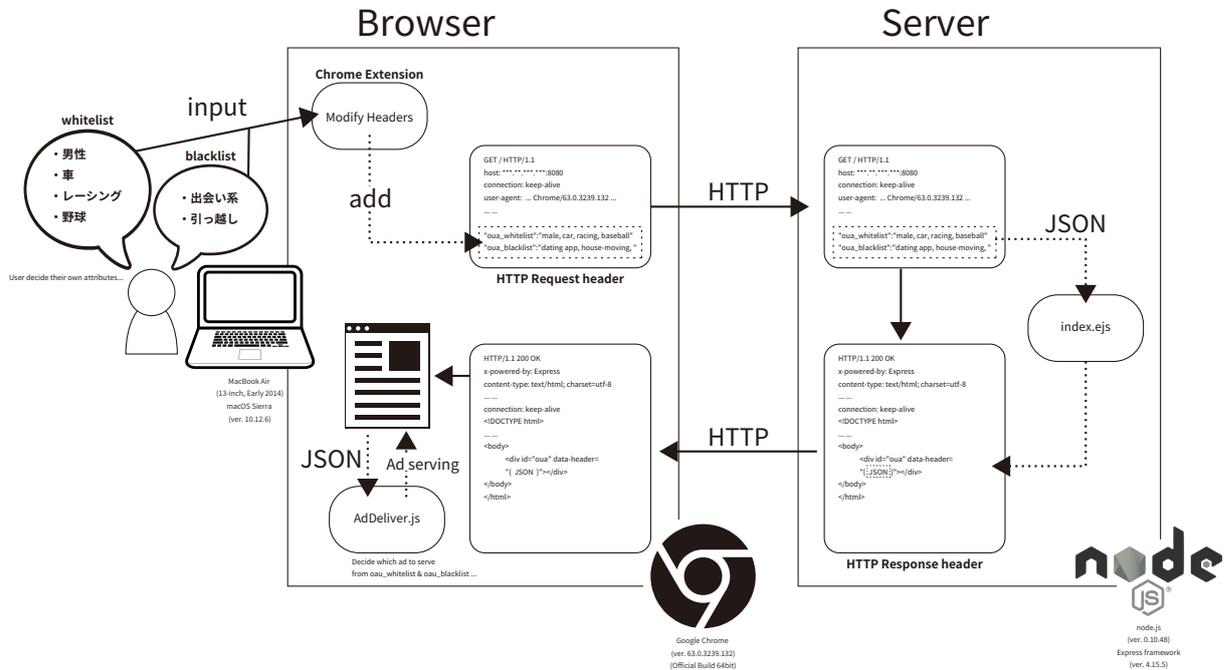


図 4.1: システム構成図

4.2 本手法における属性値の規格

本手法において、HTTP リクエストヘッダに埋め込む属性を、Open User Attribute (以後、OUA) と命名した。この OUA は以下の規格で設計される。

表 4.1: Open User Attribute の規格

項目	ヘッダ名	パラメータの記法
ホワイトリスト	oua.whitelist	属性, 属性,, 例) male,design,shoes
ブラックリスト	oua.blacklist	属性, 属性,, 例) porno,job hunting

これを HTTP リクエストヘッダに挿入すると以下ようになる。(***は伏せ字)

1 GET / HTTP/1.1

```

2 host:localhost:3300
3 connection:keep-alive
4 upgrade-insecure-requests:1
5 user-agent:Mozilla/5.0
6 accept:text/html,application/*
7 accept-encoding:gzip, deflate, br
8 accept-language:ja,en-US;q=0.9,en;q=0.8
9 cookie:gsScrollPos-1506=0;***
10 oua_whitelist:male,design,shoes
11 oua_blacklist:porno,job hunting

```

4.3 実装環境

本研究で実装するシステムを構成するためのハードウェアおよびソフトウェアについて説明する。以下、表に詳細なバージョンを示す。

表 4.2: 使用ソフトウェアおよびハードウェアのバージョン

ハード・ソフトウェア	実装環境	バージョン
サーバーサイドプログラム	node.js	ver. 0.10.48
フレームワーク	express	ver. 4.15.5
テンプレートエンジン	EJS	
ブラウザ	Google Chrome	ver. 63.0.3239.132 Official Build 64bit
属性情報の管理と リクエストヘッダへの追加	ModHeader	ver. 2.1.2
広告配信プログラム	javascript	ver. 1.9

4.3.1 サーバーサイドの実装

HTTP 通信の制御、及びリクエストヘッダ内の whitelist と blacklist を受信し、そのデータを HTML 文書に受け渡すための、サーバーサイドプログラムとして、node.js を用いる。また今回の実装では、Node.js 上で動作する Web アプリのフレームワークである Express[31] と、Javascript のテンプレートエンジンである EJS[32] を用いることで、oua ヘッダの中身を HTML に受け渡すことを実現した。

Express 側で HTTP リクエストヘッダの内容を JavaScript オブジェクトとして取得し、oua_whitelist と oua_blacklist だけを抜き出す。また、EJS の記法である、`!%= %!` を用いることで、中にある JS 変数の値をエスケープ込みで展開することができる。そこで本実装においては、抜き出した oua ヘッダの情報を JSON に変換し、EJS ファイルに受け渡した。受け渡した値をサイド JavaScript オブジェクト化して広告配信プログラムに展開することで、oua の属性情報を元に、配信する広告を決定可能にした。

4.3.2 クライアントサイドの実装

本実装において、ブラウザは Google Chrome を使用し、属性情報を管理・編集し、HTTP リクエストヘッダに埋め込むためのアプリケーションとして、ModHeader[30] という Chrome エクステンションを用いた。

ModHeader の中で、oua_whitelist と oua_blacklist を定義し、各々属性値を設定した。また、3.2 の結果を元に、4 つのシチュエーションを想定し、各々属性値を設定、これを切り替え可能な状態で準備した。

表 4.3: 本実装において想定したユーザーのシチュエーションと属性情報

シチュエーション	ユーザーの状況	oua_whitelist	oua_blacklist
alone, Wi-Fi	一人, 在宅, Wi-Fi 環境	game,movie,drama, anime,porno	job
share	画面共有中	high brand	porno,housing, geographic, ecommerce
moving	移動中	news,buzz,	porno,game,apps
searching	探し物中	scarf,female,lady, black,mother	male,girl,white,

今回設定した 4 つのシチュエーションは、3.2 の調査で、調査 5 : 過去に Web 広告が表示された際、その広告の内容で困ったエピソード、で何うことができた内容を元に、各シチュエーションごとに典型的な属性情報を設定した。

シチュエーション 1 : 一人, 在宅, Wi-Fi 環境

このシチュエーションで想定されているのは、一人暮らしの男性が在宅中にインターネットを利用している、という状況である。仕事終わりや休日に、映画やアニメなどを見てリラックスし、Wi-Fi 環境にいる間にアプリのアップデートや容量の大きいデータのやり取りをしている。このような状況を踏まえ、シチュエーション 1 では、動画コンテンツに関するワードをホワイトリストに設定する一方、休日気分を壊してしまう仕事関連のワードをブラックリストに設定した。

シチュエーション 2 : 画面共有中

このシチュエーションで想定されているのは、他人と一緒に同じ画面を閲覧している、という状況である。PC の画面を一緒に見る、ディスプレイに移したものを一緒に見る、プレゼン中で自分の PC の画面を皆が見ているなど、様々な状況が想定される。これらの状況を踏まえ、他人に不快感を与える可能性がある広告や、ユーザーのプライバシー情報が含まれる広告をブラックリストに設定した。男性向けの性的な広告や、マッチングサービスの広告、ユーザーの住所や移住を検討している地域が表示されてしまう物件の広告、

EC サイトの商品広告などが代表例である。また、他人から好意的に見られる心理が働くと想定し、高級ブランドの広告をホワイトリストに設定した。

シチュエーション 3 : 移動中

このシチュエーションで想定されているのは、ユーザーが徒歩や電車で移動をしている、という状況である。このシチュエーションでは、周りの乗客に画面を覗かれる可能性があるため、シチュエーション 2 における、他人に不快感を与える可能性がある広告は、同じく適さない。また、ユーザーは Wi-Fi 環境下にはないことが予想されるため、スマートフォンにおけるゲームやアプリのダウンロードを促す広告も適さない。

シチュエーション 4 : 探し物中

このシチュエーションで想定されているのは、ユーザーが特定の商品の購入を検討している、という状況である。このシチュエーションでは、ユーザーが自分の探している商品について、名前、値段、大きさ、などを設定する。贈り物の際は、贈る相手の性別や年齢を設定する。このシチュエーションを利用することで、ユーザーが自分の意中の商品を見つけられていない際に、日常的に Web サイトを利用しながら、商品と出会う機会を得ることができる。

Name	Value
Request Headers	
<input checked="" type="checkbox"/> oua_whitelist	game,movie,drama,anime,porno
<input checked="" type="checkbox"/> oua_blacklist	job

図 4.2: シチュエーション 1

Name	Value
Request Headers	
<input checked="" type="checkbox"/> oua_whitelist	high brand
<input checked="" type="checkbox"/> oua_blacklist	porno,housing,geographic,ecommerce

図 4.3: シチュエーション 2

Name	Value
Request Headers	
<input checked="" type="checkbox"/> oua_whitelist	news,buzz,
<input checked="" type="checkbox"/> oua_blacklist	porno,game,apps

図 4.4: シチュエーション 3

Name	Value
Request Headers	
<input checked="" type="checkbox"/> oua_whitelist	scarf,female,lady,black,mother
<input checked="" type="checkbox"/> oua_blacklist	male,girl,white.

図 4.5: シチュエーション 4

4.3.3 広告配信プログラムの実装

ユーザーの属性情報からどのように広告を配信するかについては、本研究の対象外であるため、属性情報自体の違いで配信広告が変わるよう、簡素なプログラムを実装した。

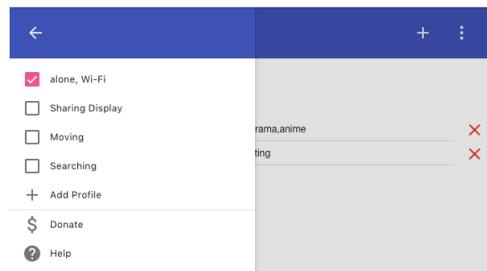


図 4.6: シチュエーションの切り替え

1つの広告データに対し、広告の内容に関連する3つのワードを設定。まず、`oua_whitelist`の属性情報のワードと照らし合わせ、これと1つでも合致した広告データを選び出す。そして、`oua_blacklist`の属性情報のワードと照らし合わせ、こちらと1つでも合致した広告データを除外する。このプロセスを通して残った広告データを配信するプログラムである。

第5章 評価および考察

本章では，4章で実装した提案システムの動作検証の結果と，評価についてまとめる．

5.1 動作検証

実装した提案手法が正しく動作しているか検証するために，4.3.2で述べたシチュエーションに応じて適切な広告が配信されるか，シチュエーションを切り替えた際に配信広告が切り替わるかどうか検証し，問題なく動作することを確認した．

5.2 定性評価

本研究で実装したシステムの評価は定性評価でおこなう．表5.1に3.4で設定した要件の達成の可否と，その評価をまとめる．

表 5.1: 本研究の課題解決のために満たすべき要件

	要件	評価
要件 1	透明性	○
要件 2	ホワイトリスト属性の発信	○
要件 3	ブラックリスト属性の発信	○
要件 4	属性の切り替え	○

5.2.1 要件 1：透明性

既存手法では，アドネットワークがユーザーに対してトラッキングされた内容は開示されず，そこから推測した属性情報が開示されている例は Google など一部に留まる．また，開示されているアドネットワークにおいても，広告の配信過程が明らかでないため，全ての属性情報が開示されている保証は無い．本実装では，ユーザーが属性情報を決定し，ローカル側で保持されているその属性情報を使って広告配信を可能にした．これにより，既存手法に比べ，明瞭な配信過程で広告を配信することを達成したため，システムのユーザーに対する透明性は向上したと言える．

5.2.2 要件 2 : ホワイトリスト属性の発信

これまで Web 広告配信のシステムにおいて、アドネットワークがユーザーの好みを調査する、一方向的な仕組みが中心であった。しかし本実装において、OUA 属性を用いることで、ユーザーが興味のあるものをホワイトリスト属性として発信することが可能になり、本要件は達成された。

5.2.3 要件 3 : ブラックリスト属性の発信

アドネットワークが抱える問題の一つが、無関心のトピックと好みではないトピックを区別出来ないことだった。またそれにより、ユーザーの好みではない広告を配信してしまうことは、広告効果を下げってしまう要因である。本実装においては、OUA 属性を用いることで、ユーザーの好みではないものをブラックリスト属性として発信することが可能になり、本要件は達成された。また、4.3.3 で示した広告配信プログラムのように、ブラックリスト属性に合致した広告を配信から外すことで、配信されることを防ぐことができた。

5.2.4 要件 4 : 属性の切り替え

現在の Web 広告手法において、ユーザーの実世界での状況に合わせて、配信される広告を変えることはできなかった。これはアドネットワーク側が時間ごとのユーザーのシチュエーションの変化を認識することができなかったからである。本実装では、ユーザーのシチュエーションごとに OU A 属性を設定し、シチュエーションを切り替えることで、それを元に配信される広告の内容を、シチュエーションごとに切り替えることができた。よって、本要件は達成できた。

第6章 結論

本章では、本研究のまとめと、今後の課題、展望について述べる。

6.1 本研究のまとめ

本研究では、Web サイトに配信される広告の配信過程において、ユーザーに不信感を与えず、刻々と変わるユーザー実世界での状況に対して、その状況に最適な広告を配信するために、ユーザー属性埋込み型 HTTP 通信を利用した Web 広告配信手法を提案した。Web 広告においてユーザーに不信感を与えず、シチュエーションごとに最適な広告を配信するためには、透明性、ホワイトリスト属性の発信、ブラックリスト属性の発信、属性の切り替えの4点が求められた。そこで提案手法として、クライアント側でユーザーが自らの属性情報を決定し、それを HTTP リクエストヘッダの中に埋め込んで通信、またクライアント側でユーザーの状況に対して属性情報を切り替えられるシステムを実装した。4点の要件に対して、実験システムがそれぞれを満たしているかを検証した。広告配信に利用する属性情報をユーザーが自ら決定し、ローカル環境で保持していることから、広告配信の透明性が担保されていることを確認した。属性情報の規格をユーザーの好み属性のリストであるホワイトリストと、好みでない属性のリストであるブラックリストの2つに分けて定義し、HTTP リクエストヘッダに埋め込んで通信、Web サイトに受け渡し広告配信にもちいたことで、配信される広告に反映されていたことから、ホワイトリスト属性の発信・ブラックリスト属性の発信が達成されたことを確認した。ユーザーの実世界での状況を、4つのシチュエーションに定義し、シチュエーションごとに最適な属性情報を設定した。このリストを切り替えることで、配信される広告が設定した属性情報に合わせて切り替わったことから、属性の切り替えが達成されたことを確認した。

本システムを用いることで、アドネットワークはユーザーに敬遠されがちな Web トラッキングを用いずとも、ユーザーに最適な広告を配信することができる。また、これまで実現できなかった、ユーザーの実世界での状況に合わせた広告を配信が可能になるので、Web 広告に対するユーザーエクスペリエンスが向上される。

6.2 本研究の課題と展望

本節では、本研究で提案したシステムの課題と展望について述べる。

6.2.1 ユーザーの興味における潜在性

本研究において提案したシステムにおいて、最も懸念される問題は、ユーザーの興味における潜在性である。ユーザーは、自身の興味のあるものを全て把握しているわけではない。ユーザーが、自分の好きなもの、興味のあるもの、と認識しているもの以外にも、他人から言われたり、実際に目にすることで初めて気づく自分の興味が存在する。この潜在性の発掘は、Web トラッキングから得た行動情報の分析を用いて、広告配信を行う利点の1つである。トラッキングして得られたユーザーの行動履歴と近い履歴を持つ別のユーザーの属性を参照することで、潜在的な興味を発掘することができる。

この潜在性の発掘を提案手法において実現するためには、既存手法においてアドネットワークが行っているユーザーの行動分析を、ブラウザ側で行う必要がある。行った行動分析を元に、OUA 属性を変更する、もしくは変更をユーザーに提案するなどして、潜在的な興味を反映されていく仕組みである。これはユーザーがブラウザを閲覧する機器内で行動情報を取得するので、Web トラッキングを用いた場合より、多くの情報を容易に入手できる。しかし、ブラウザ側に負荷をかける処理であるため、ブラウザのアップデートに依存する問題でもある。また、ユーザーの機器内で行動情報の取得から分析までが完結するので、広告配信過程の透明性が保たれる一方、既存手法で行われている他のユーザーの行動情報と合わせた、ビッグデータ分析は行うことができない。これにより、分析の精度についても懸念が残る問題である。

6.2.2 属性の切り替えの高度化

本研究において、ユーザーの状況に応じた属性の切り替えの機能は、特徴的なシチュエーションを定義し、各々シチュエーションに適した属性を設定した。このシチュエーションを切り替える形で実装を行い、属性の切り替えを実現した。この機能においては、シチュエーション切り替えの自動化の点で今後の発展が見込まれる。今回の実装において、シチュエーションの切り替えはユーザーが手動で行っていた。これに接続する Wi-Fi のネットワークにより滞り場所の判定をすることで、シチュエーションを切り替えることが可能だ。ここに GPS による位置情報を用いた移動中の判定や、インカメラの映像を用いた画面を閲覧中の人数の判定など、切り替えの精度の向上が見込まれる。この手法の素晴らしい点は、GPS 情報の利用やインカメラの映像を利用しても、全てユーザーの機器内で情報が利用されるので、透明性が担保される点である。

6.2.3 レコメンドシステムへの転用

本研究で述べてきた Web 広告配信の手法は、広告以外の分野でも応用されている。特にデモグラフィックターゲティングの手法は、EC サイトやニュースサイトなどで利用されている。EC サイトにおいては、過去の購買履歴から、ユーザーの購買意欲をそそる商品がトップページに表示されている。ニュースサイトにおいては、ユーザーが過去に読んだ記事の履歴から、興味のある記事がトップページに表示されている。このコンテンツをそのユーザーにとって最適化する行為はサジェストと呼ばれるものである。またこのような、ブラウザや Web サイトに賢さを持たせる取り組みは Web Intelligence と呼ばれ、研究が進められている。[33][34]

本研究で提案した OUA 属性は、このレコメンドシステムに転用可能だと考える。本研究での Web 広告の配信手法のように、発信された OUA 属性を元にコンテンツをサジェストする仕組みは、容易に想像可能である。そこで、Web におけるサジェストに OUA 属性が活用されることで、解決される課題と、展望についても述べる。

初めて訪れたサイトにおけるサジェスト

既存のサジェストシステムの問題点の 1 つとして、ユーザーが初めて訪れたサイトやサービスでは、最適なサジェストが行うことができないというものがある。これは、初めて訪れたサイトでは、ユーザーのアカウントとそれに紐づく情報がないため、サジェストするための材料が足りないためである。そこで、OUA 属性を活用することで、どの Web サイトにおいても、サイト側はアクセスしてきたユーザーの属性情報を得ることができる。これにより、初めて訪れたサイトでも、OUA 属性を元にサジェストが可能になるのである。

コンテンツ以外のサジェスト

現在、Web サイトにおけるサジェストシステムで最適化されているのは、商品ページやニュースの記事など、サイトのコンテンツである。しかし Web サイトにおいて、ユーザーエクスペリエンスに関わる要素は、コンテンツだけではない。サイトのデザインやレイアウトはその最たる例である。既存の Web サイトにおいて、一部のコンセプトチュアラルなサイトを除き、そのサイトにアクセスした全てのユーザーに対して、画面サイズに合わせて同じデザインの Web サイトが表示される。しかしデザインやレイアウトにおいても、当然ユーザーごとの好みがある。代表的な例では、メニューバーの位置、メニューの構成、カラム数、文字の縦書き横書き [35] などが想定される。OUA 属性を活用すれば、あらゆるサイトでアクセスしているユーザーの属性情報を得ることができる。OUA 属性のサジェストへの利用が広く認知され、ユーザーが Web のデザイン・レイアウトに対する好みを設定しておけば、Web サイトのデザインもユーザーごとに最適化して、表示することが可能になる。

このように、OUA 属性によるユーザーの好みの発信は、Web サイトにおいて様々な活用が見込まれる。

謝辞

この論文を執筆するにあたりご指導頂きました，政策・メディア研究科委員長 村井純博士，環境情報学部客員教授 徳田英幸博士，環境情報学部教授 中村修博士，環境情報学部教授 楠本博之博士，環境情報学部教授 三次仁博士，環境情報学部教授 武田圭史博士，環境情報学部准教授 高汐一紀博士，環境情報学部准教授 植原啓介博士，環境情報学部准教授 中澤仁博士，環境情報学部准教授 Rodney D. Van Meter 博士，政策・メディア研究科特任准教授 鈴木茂哉博士，政策・メディア研究科特任准教授 佐藤雅明博士，環境情報学部講師 齊藤賢爾博士，政策・メディア研究科特任講師 Achmad Husni Thamrin 博士，政策・メディア研究科特任助教 工藤紀篤博士，政策・メディア研究科特任助教 佐藤貴彦博士，政策・メディア研究科特任講師 空閑洋平博士，政策・メディア研究科特任講師 松谷健史博士，政策・メディア研究科特任講師 大越匡博士，政策・メディア研究科研究員 伊藤友隆博士，政策・メディア研究科 中島博敬氏に厚く感謝いたします。

特に貴重な研究の機会を与えてくださった村井純博士，中村修博士に重ねてお礼申し上げます。私が2年間取り組んで参りました縦書き Web に出会ったきっかけは，中村先生の紹介でありましたし，TPAC2015 への参加や出版の未来の授業運営をさせて頂く機会を下されたのが村井先生でした。

植原啓介博士，中澤仁博士に重ねてお礼申し上げます。お二人の下で延べ3回参加させて頂いた，ORF NOC チームとしての活動は，一学生では味わえない達成感を与えられる，私にとってかけがえのないものでありました。

鈴木茂哉博士，Achmad Husni Thamrin 博士，中島博敬氏に重ねてお礼申し上げます。見込みが甘くいつもギリギリまで動き出せないわたしに対し，時に厳しく，時に厳しく優しく，指導して頂きました。皆様のご指導がなければ私は本日までこの研究室での活動を続けてこれなかったと思います。中島氏にはプライベートでも非常にお世話になりました。本当にありがとうございました。

慶應義塾インフォメーションテクノロジーセンター本部 助教 近藤賢郎氏に感謝致します。近藤氏には ORF NOC チームの活動に多大なご協力を頂きました。また私に対して常に励ましの言葉をかけて頂きました。本当にありがとうございました。

村井研究室 Kumo 研究グループの皆様には感謝致します。Kumo の先輩の皆様には，日頃至らぬ私に対し常に叱咤激励を頂きました。また，本論文の執筆に当たっても多くのご協力を頂きました。研究の傍ら，常に食への探求を続けた Kumo での3年間は私の村井研生活の全てです。本当にありがとうございました。

徳田・村井・楠本・中村・高汐・バンミーター・植原・三次・中澤・武田 合同研究プロジェクトの皆様には感謝致します。研究室の先輩であります，永山翔太博士，五十嵐祐貴氏，水谷伊織氏，高橋俊成氏，塚越広彬氏，池田貴匡氏，江頭和輝氏，重田桂子氏，阿部涼介氏，

大戸浩司氏、森重浩直氏、靱山奈々子女史、黒米祐馬氏、木下舜氏、鈴木恒平氏、沖幸太郎氏に重ねて御礼申し上げます。研究室に所属を始めた当初から、多くの素晴らしい先輩方に囲まれたお陰で、研究室での3年半を乗り切ることができました。本論文の執筆にも多くの方に相談に乗って頂き、ご協力して頂きました。本当にありがとうございました。

辛く長い残留生活を共に過ごした、河口綾摩氏、尾周也氏、東海林晃氏、豊田安信氏、城一統氏、菅藤佑太氏、安井瑛男氏、栗原祐二氏、深川祐太氏、森島隆成氏、熊谷啓孝氏、鈴木雄祐氏に重ねて御礼申し上げます。皆様と過ごした日々は驚きと楽しさと狂気に満ち溢れたかけがえのないものです。本当にありがとうございました。また尾崎氏とは研究室・Kumo・CNS コンサルタントと非常に多くの時間を共有する中で、その勤勉さと優しさにとってもお世話になりました。改めて感謝致します。

石川達敬氏、山田真也氏に感謝致します。二人とともに過ごした研究室合宿はとても楽しい思い出です。

本研究の実験にご協力頂いた皆様に感謝致します。締め切りまで時間の無い中で急遽ご連絡したにもかかわらず、2日間で50名程の皆様が参加してくださいました。皆様のおかげで本研究における課題を明確にすることが出来ました。本当にありがとうございました。

慶應義塾大学環境情報学部非常勤講師 古屋 友章氏に感謝致します。私が常に秘めていたデザインやコミュニケーションへの興味に対し、先生の「情報デザイン基礎」授業の中で勉強と実践の場を頂きました。また2度のSAを含め3年に渡り多くの助言を頂きました。先生の下での経験が本研究の礎となっています。本当にありがとうございました。また、情報デザイン基礎で出会った皆様に感謝致します。この授業を履修する皆様はいずれも個性豊かで、自分なりの表現を突き詰める方々でした。皆様と一緒に授業を受ける日々はとても刺激的で私を成長させてくれるものでした。

志伯健太郎氏に感謝致します。氏の下で働かせて頂いた2年間で多くのことを学ばせて頂きました。また、私の個人的な相談にも親身に乗って頂き、私の人生における格言となる言葉を頂きました。頂いた言葉を胸に、悔いのない人生を歩んでいきたいと思えます。本当にありがとうございました。

籠球倶楽部 Kagers の皆様に感謝致します。あまり練習に行けなかった私に対し、常に好意的に接し輪に加えて頂きました。Kagers での思い出は、間違いなく私の人生における1ページであります。Kagers という帰る場所があったからこそ、他の団体や学外での活動など、自分の興味の赴くままに4年間を走りきることができました。Kagers の今後の発展と後輩達の幸せな学生生活を願っております。

青木佑介氏に感謝致します。浪人時代に氏の紹介でORFを訪れなければ、私がこのキャンパスに来ることはありませんでした。共にSFCを志し入学後もお互いの信念の下に突き進む、切磋琢磨し合う関係でした。また一昨年から昨年にかけて氏が敢行した世界一周の旅は、私を勇気付け奮い立たせるものでした。青木氏の今後の活躍を願っております。

山本海斗氏に感謝致します。2015年に会って以降、公私に渡って非常にお世話になりました。ふざけながらも周囲に気を使い、同期のように接してくださる山本氏と過した時間は本当に楽しい時間でした。無事に卒業した暁には恩返しに伺います。今後ともどうぞよろしくお願い致します。

SFC の友人達に感謝致します。こんな私でありますが幸いなことに多くの友人に恵まれました。常に誰かと出会い様々なことが起きるこのキャンパスと皆様が大好きです。本当にありがとうございました。今後ともどうぞよろしくお願い致します。

最後にここまで育ててくれた両親に感謝致します。自分の興味に突き進む私を常に応援してくれました。本当にありがとうございます。

付 録 A 付 録

3.2 の調査における，ユーザーの回答データを添付する．

	gender	設定されている トピックの総数	興味のない トピックの数	興味のないトピックの内容
user1	male	なし		なし
user2	female	65	6	スポーツ,ローン,格闘技,紳士服, 自動車販売,バレンタインデー
user3	female	84	10	アウトドア,オートバイ,キャリア設計,クラシック音 楽,スキンケア・ネイルケア,ヨガ・ピラティス,交際, 着メロ,フットウエア,スパ・美容サービス
user4	male	62	7	クラシック音楽,テレビドラマ,フォーク音楽,ブル ース,紳士服,舞台芸術,東アジアの音楽
user5	male	20	2	フィットネス,マナー
user6	male	65	9	オフィス・ビジネスソフトウェア,キャリア設計,チェ ス,ヘアケア,下着,政治,統計,芸能ニュース,女性向け
user7	male	なし		なし
user8	male	26	3	ランニング・ウォーキング,センター試験,小中学校
user9	female	49	13	アセットファイナンス,コンピュータ・電化製品,セン ターテスト・入学試験,コンピュータモニタ・ディス プレイ,ブルース,モバイル・ワイヤレス,銀行,住居・ 庭,政治,鉄道,舞台芸術,郵便・宅配便,路線図・時刻 表・運行情報
user10	male	20	5	電子音楽,ブルース,ロック,金融,野球
user11	male	なし		なし
user12	male	81	7	クラシック音楽,フォーク・伝統音楽,礼服,舞台芸術, 東アジアの音楽,着メロ,モバイル・ワイヤレス
user13	male	66	10	女性向け,服飾雑貨,抵当,辞書・百科事典,引越し・移 転
user14	female	34	7	テクノロジー関連のニュース,知的所有権,チェス・ア ブストラクトゲーム,舞台芸術,留学,ビジネスニュー ス,ワールドミュージック
user15	male	69	4	SIP,栃木,東アジアの音楽,政治
user16	male	ブロック		ブロック
user17	female	59	15	アメフト,ウェブデザイン,カントリーミュージック,コ ミックアニメ,センター試験,ブルース,リフォーム, ロックミュージック,会計,財務プランニング,自動車, 子育て,住居,着メロ,金融
user18	male	64	5	センター試験・入学試験,コピー・印刷サービス,アウ ター,クラシック音楽,ジャズ
user19	female	ブロック		アウトドア,子育て,事務用品,アクションゲーム,サッ カー,ロールプレイングゲーム

user20	female	65	9	オートバイ,オフィス・ビジネスソフトウェア,ゲーム,モータースポーツ,科学,数学,東アジアの音楽,舞台芸能,自動車
user21	female	68	10	ウィンタースポーツ,オートバイ,スポーツ,ビジネス・産業,モータースポーツ,金融,自動車,搾乳用品,離乳食用品,宝石・アクセサリ
user22	male	20	5	フード・ドリンク,ブルース,ポップミュージック,金融,芸能ニュース
user23	female	43	3	個人ローン,小中学校,コンピューター・電化製品
user24	female	15	6	コンピューター・電化製品,ワールドミュージック,ホッケー,フォーク・伝統音楽,自動車,モバイル・ワイヤレス
user25	male	46	8	オフィス・ビジネスソフトウェア,タイル並べ,ビジネス・産業 プリンタ・コピー機・femaleAX機,行政機関の求人情報,作曲・音楽理論,政治,企業向けテクノロジー
user26	male	76	13	ゴルフ,ゴルフグッズ,アーバン・ヒップホップ,ウォータースポーツ,クラシック音楽,ブルース,ポップミュージック,企業向けテクノロジー,金融,商業用不動産,住居・庭,東アジアの音楽,小中学校
user27	female	89	15	ブルース,礼服,政治,辞書・百科事典,子供服,ジャズ,事務用品,企業向けテクノロジー,ロックミュージック,モバイル・ワイヤレス,プレゼンテーション・ソフトウェア,ビジネス・産業,オフィス・ビジネスソフトウェア,¥¥ クラシック音楽,キャリア設計・プランニング
user28	female	なし		なし
user29	female	57	8	ISP,オートバイ,クリップアート,コンピューター・電化製品,センターテスト,ロールプレイングゲーム,家電,婦人服
user30	female	28	2	ホッケー,ゴルフ
user31	female	32	5	アメリカンフットボール,フォーク・伝統音楽,料理・レシピ,ブルース,スポーツ
user32	male	56	17	キャリア設計・プランニング,金融,芸能ニュース,ロールプレイングゲーム,オンラインコミュニティ,格闘技,教育,野球,礼服,腕時計,舞台芸術,シューティングゲーム,ワールドミュージック,ロックミュージック,ポップミュージック,フォーク・伝統音楽,ビジネスニュース
user33	male	58	4	プレゼンテーション・ソフトウェア,femaleeed Aggregation & Social Bookmarking,芸能ニュース,ポップミュージック

user34	female	ブロック		ブロック
user35	male		72	4 ISP,ポート,小中学校,美容・フィットネス
user36	male		63	10 保険,舞台芸術,小中学校,女性向け,金融,ラグビー,ブルース,チェス・アブストラクトゲーム,タイル並べ,オートバイ
user37	male	ブロック		ブロック
user38	male		60	6 フォーク・伝統音楽,センターテスト・入学試験,ブルース,引越し・移転,健康保険,金融
user39	male		49	5 クラシック音楽,ブルース,モータースポーツ,舞台芸術,ダンス・電子音楽
user40	female		36	4 フォーク・伝統音楽,ビジネス・産業,オートバイ,Resumalees & Portfemaleolios
user41	male		46	6 タイル並べ,伝統音楽,舞台芸術,ポート,モータースポーツ,電話回線サービス
user42	male		28	8 オートバイ,サッカー,サイクリング,確定申告・税務,会計・財務ソフトウェア,自動車販売,婦人服,宝石・アクセサリー
user43	female		50	2 チェス,クラシック音楽
user44	female		48	3 ブルース,ホッケー,金融
user45	female		69	12 BMW,フォーク・伝統音楽,ブルース,リンカーン,会計・財務ソフトウェア,企業向けテクノロジー,自動車販売,ラジオ,住居・庭,紳士服,着メロ,モバイル端末関連
user46	female		33	5 金融,政治,ビジネスニュース,ビジネス・産業,アニメ・漫画
user47	female		61	6 オーストラリアンフットボール,ファーストフード,コンピュータ・電化製品,オートバイ,オンラインコミュニティ,ビジネス・産業
user48	male		44	1 ウォータースポーツ
user49	female	なし		なし
user50	female		38	3 ホッケー,ブルース,インターネットクライアント
user51	male	ブロック		ブロック
user52	male		74	11 キャリア設計,オートバイ,ビジネス産業,ブルース,ヨガ,金融,教育,化学,結婚,子育て,婦人服
user53	female		43	8 IT・技術系の求人情報,キャリア設計・プランニング,チェス・アブストラクトゲーム,テレビドラマ,辞書・百科事典,表計算ソフトウェア,オフィス・ビジネスソフトウェア,デスクトップパソコン

	ブラックリストに自分が追加したいワード	タマイズ ページの存在を知って	広告の内容で困ったエピソード	属性を自分で管理する必要があるか
user1	仮想通貨,英会話,アダルト系	いいえ	他人という時にエロサイトの広告が出てきて恥ずかしい。	いいえ
user2	動画配信サービス,漫画	いいえ	ゲームのサイト・漫画のサイトで好みでない・不適切な格好をしたキャラクターが出てくると見ていて不愉快。漫画の広告で不気味な絵のものが特に不愉快な思いをする。	はい
user3	美容系,ゲーム	いいえ	押したいものが押せない、プレゼントを探していたら送る本人の前でその商品の広告が出てしまった	はい
user4	スノボ	いいえ	スムーズにエロサイトが見れない	はい
user5	消費者金融,フィットネス,仮想通貨,アプリ	いいえ	友達に見せたときにエロサイトに誤って移動してしまう。	はい
user6	出会い系,仮想通貨,自己啓発本	いいえ	誤って押して困る	はい
user7	保険	いいえ	動画のスタートが遅れた、違うページに飛んでしまう、	はい
user8	消費者金融,物件,仮想通貨	いいえ	エロい系全般/音声付きの広告で静かな場所で音声が出た	はい
user9	モバイル・ワイヤレス,銀行	いいえ		はい
user10	グロい系	いいえ	グロい広告は嫌な気持ちになる。	はい
user11	情報商材,アダルト系,出会い系,通販	いいえ	異性という時に、エロ広告が出てしまう。彼女にサプライズでプレゼントを買おうと探し	はい
user12	賃貸物件,仮想通貨,資格の広告,出会い系	いいえ	見たいサイトに飛ばうとすると別のサイトに飛んでしまう。	はい
user13	出会い系	いいえ	出会い系の広告が音量がでる状態で再生されてしまった	はい
user14	アダルト系,旅行,物件	いいえ	間違っって広告をクリックしちゃった、	はい
user15	家具,VR,アダルト系	いいえ	邪魔 広告を間違っってクリックしてしまったことで行きたいページにいけなかった	はい
user16	住宅,漫画,転職,アプリ,ゲーム	いいえ	エロ	はい
user17	アニメ,漫画,ゲーム	いいえ	うざい、誤タップする、わざわざ広告を消しています。	はい
user18	漫画,アニメ,アクセサリ	いいえ	・うっとおしい ・女の子とyoutube見てて、エロい広告が出てきた時 ・広告の時間が単純に長い	はい

user19	お笑い番組,フィットネス,芸能ニュース	はい	間違えてクリックして、前のページに戻れない	はい
user20	オートバイ,オフィス・ビジネスソフトウェア,モータースポーツ,科学,数学,東アジアの音楽,舞台芸能,自動車,トラック・SUV	いいえ		はい
user21	就活	いいえ	出所の分からない、信憑性の薄い情報を当たり前のようポンポン出してくるのに苛立つ	はい
user22	ポップミュージック,アダルト系,消費者金融,ゲーム,投資ビジネス	いいえ	研究会で画面をプロジェクターで映しているときにすごい真面目なページ開いても広告のところに最近ヤフオクでチェックしたものができて恥ずかしかった。	はい
user23	ゲーム,広告CM	いいえ	ゾンビが出てくる広告は見て嫌になる	いいえ
user24	ホラー,アプリ	いいえ	間違えて押してしまうとそのサイトに飛ぶだけではなくapp storeまで開いてしまうやつ本当にイライラして困ります。フリーフォントのサイトでフォントデータをダウンロードしたいのに異なるもののダウンロード広告を表示してリンクを踏ませるやつ。	はい
user25		いいえ	Adblock使用	はい
user26	出会い系,ゴルフ,情報商材,消費者金融,パチンコ,ゲーム	いいえ	誰かと見てる時にエロ広告、ミーティングで画面共有中に出会い系、Wi-Fiじゃないのにアプリダウンロード	はい
user27	インターン,就活,アダルト系,アプリ	いいえ	なし	はい
user28	仮想通貨,政治家,アダルト系,就活,物件	いいえ	なし	はい
user29	出会い系,コラム	いいえ	邪魔、	いいえ
user30	アダルト系,物件,出会い系,保険	いいえ	18禁のページを開いているのを友人に覗かれた。誤ってクリックした。鬱陶しい。	はい
user31	漫画,消費者金融,アダルト系,ゲーム	いいえ	見てもないのに、ちょっと怪しいサイトに行くと18禁ばい内容が出てくる。	いいえ
user32		いいえ	アダルト広告で、漫画のアダルト広告がドンピシャで購入してしまったが、冷静になったときに無駄遣いしてしまったと感じた。	はい
user33	出会い系,受験,資格系,レイパン,不動産,ゲー	いいえ	電車でエロいのが出る、割と困る。あと、ソシャゲの広告がうざい	はい
user34	服,出会い系,美容	いいえ	リンクを踏んじゃってページを閉じるまでにタイムロス	はい

user35	アダルト系,電化製品	いいえ	すでに買ったものが複数回表示される、サプライズで購入した商品がばれる	はい
user36	ゲーム	いいえ		いいえ
user37	ゲーム,バスケットボール,アニメ	いいえ	音量が大きい広告、データ読み込みが重い広告	はい
user38		いいえ	特になし。	はい
user39		いいえ		はい
user40		いいえ	電車とかでなんかエロいやつ出てきた時。	はい
user41	出会い系,仮想通貨	いいえ	特になし	はい
user42	バレーボール,化粧品	いいえ		はい
user43	かき氷	いいえ		はい
user44	アダルト系,出会い系	いいえ	電車とかで自分が見たいサイトのわきとかにエロ漫画の広告があると、普段読んでいるのかなと思われそうだから嫌。出会い系も同様	はい
user45	アダルト系,出会い系	いいえ		いいえ
user46	農業,モバイル,学習塾,TOEIC	いいえ	特にSNSにおいてどエロい広告表示が多く、目のやり場に困る。そういうものに関心があると認識されているようで、屈辱を感じる。	はい
user47	健康サプリ	いいえ	需要がなくなった広告が出続ける	はい
user48	アダルト系	いいえ	出会い系など頻繁に出てきて、友達にタイムラインが見せられない。	はい
user49	アダルト系,ゲーム,仮想通貨,アダルト系	いいえ	アダルト系のサイトをクリックしてしまった	はい
user50	仮想通貨,ゲーム	いいえ	間違っってクリックしてしまい、いやらしいサイトに飛んでしまった	いいえ
user51	ゲーム,スポーツ	はい	下着をネットショッピングしていると広告に自分の下着と似たものが出てきてしまうので、知人に画面を見せる際などは不適切なときがあ	はい
user52		いいえ		いいえ
user53	転職・求人,プログラミング学習	いいえ	特になし	いいえ

参考文献

- [1] 株式会社電通. 2016年日本の広告費. <http://www.dentsu.co.jp/news/release/pdf-cms/2017027-0223.pdf>, 2016.
- [2] グレン・グリーンウォルド／著, 田口俊樹／訳, 濱野大道／訳, and 武藤陽生／訳. 暴露:スノーデンが私に託したファイル. 新潮社, 2014.
- [3] Thaddeus Swanke, Michael S Montalto, and John E Morse. Web tracking apparatus, August 17 1976. US Patent 3,974,952.
- [4] Stephen C. Rumsey Joseph Brian J. Web tracking apparatus. <https://docs.google.com/viewer?url=patentimages.storage.googleapis.com/pdfs/US4572417.pdf>, 1986.
- [5] Douglass L Blanding and Terry N Morganti. Web tracking apparatus, January 3 1989. US Patent 4,795,070.
- [6] Jamzadeh Fereidoon S. Method and apparatus for web tracking with predictive control. <https://docs.google.com/viewer?url=patentimages.storage.googleapis.com/pdfs/US4961089.pdf>, 1990.
- [7] Robert H Ducoffe. Advertising value and advertising on the web. *Journal of advertising research*, 36(5):21–21, 1996.
- [8] Rex Briggs and Nigel Hollis. Advertising on the web: Is there response before click-through? *Journal of Advertising research*, 37(2):33–46, 1997.
- [9] Robbin Lee Zeff and Bradley Aronson. *Advertising on the Internet*. John Wiley & Sons, Inc., 1999.
- [10] Pierre Berthon, Leyland F Pitt, and Richard T Watson. The world wide web as an advertising medium. *Journal of advertising research*, 36(1):43–54, 1996.
- [11] Gary B Robinson. Automated collaborative filtering in world wide web advertising, June 29 1999. US Patent 5,918,014.
- [12] Richard Shambare, Robert Rugimbana, and Nkosinathi Sithole. Social networking habits among students. *African Journal of Business Management*, 6(2):578, 2012.

- [13] Hanna Krasnova, Natasha F Veltri, and Oliver Günther. Self-disclosure and privacy calculus on social networking sites: The role of culture. *Business & Information Systems Engineering*, 4(3):127–135, 2012.
- [14] 株式会社 メディックス. ターゲティング広告の全手法と最も効果的な活用方法. https://webbu.jp/targeting_advertising-238.
- [15] Dwight Allen Merriman and Kevin Joseph O’connor. Method of delivery, targeting, and measuring advertising over networks, September 7 1999. US Patent 5,948,061.
- [16] Myung Jong Lee. Method and system for providing targeting advertisement service in social network, February 20 2009. US Patent App. 12/389,504.
- [17] Jonathan R Mayer and John C Mitchell. Third-party web tracking: Policy and technology. In *Security and Privacy (SP), 2012 IEEE Symposium on*, pages 413–427. IEEE, 2012.
- [18] Attila Narin, Keith A Kegley, and David A Sobeski. Tracking usage behavior in computer systems, May 2 2006. US Patent 7,039,699.
- [19] panopticlick. <https://panopticlick.eff.org/>.
- [20] Károly Boda, Ádám Máté Földes, Gábor György Gulyás, and Sándor Imre. User tracking on the web via cross-browser fingerprinting. In *Nordic Conference on Secure IT Systems*, pages 31–46. Springer, 2011.
- [21] Martin Mulazzani, Philipp Reschl, Markus Huber, Manuel Leithner, Sebastian Schrittwieser, Edgar Weippl, and FC Wien. Fast and reliable browser identification with javascript engine fingerprinting. In *Web 2.0 Workshop on Security and Privacy (W2SP)*, volume 5, 2013.
- [22] Lance Cottrell. Cookieなしでユーザー特定が可能、ブラウザー・フィンガープリントとは. <http://itpro.nikkeibp.co.jp/atcl/idg/15/030200019/030200001/?ST=cm-network&P=1>.
- [23] Canvas api. <https://developer.mozilla.org/ja/docs/Web/HTML/Canvas>.
- [24] Keaton Mowery and Hovav Shacham. Pixel perfect: Fingerprinting canvas in html5. *Proceedings of W2SP*, pages 1–12, 2012.
- [25] Meet the online tracking device that is virtually impossible to block. <https://www.propublica.org/article/meet-the-online-tracking-device-that-is-virtually-impossible-to-block>.
- [26] Google 広告設定. <https://adssettings.google.com/authenticated>.

- [27] Roy Fielding, Jim Gettys, Jeffrey Mogul, Henrik Frystyk, Larry Masinter, Paul Leach, and Tim Berners-Lee. Hypertext transfer protocol–http/1.1. Technical report, 1999.
- [28] 竹下隆史, 村山公保, 荒井透, and 荻田幸雄. マスタリング TCP/IP 入門編 第 5 版. オーム社, 2012.
- [29] ネットワークエンジニアとして. <http://www.infraexpert.com/study/tcpip16.html>.
- [30] Modheader. <https://chrome.google.com/webstore/detail/modheader/idgpnmonknjnojddfkpgkljpfnnfcklj>.
- [31] Express. <http://expressjs.com/>.
- [32] Effective javascript templating. <http://ejs.co/>.
- [33] Jiming Liu, Ning Zhong, Yiyu Yao, and Zbigniew W Ras. The wisdom web: new challenges for web intelligence (wi). *Journal of Intelligent Information Systems*, 20(1):5–9, 2003.
- [34] Ning Zhong, Jiming Liu, and Yiyu Yao. Envisioning intelligent information technologies through the prism of web intelligence. *Communications of the ACM*, 50(3):89–94, 2007.
- [35] 縦書き web 普及委員会. <https://tategaki.github.io/>.

卒業論文

題名：ユーザー属性埋込み型 HTTP 通信を利用した Web 広告手法の提案

氏名：桑原誠尚

親氏名：鈴木茂哉

桑原誠尚君の卒論本綴じ提出を認めます。

親：



コメント：

少し記述に不自然なところが多い（口語表現が多い）ところとか、細かい体裁の点でなっていない点が多々ある。こういう所は課題だと思う。一方、全体的には良く頑張って面白い形で書けたと思う。