クロスデバイストラッキング技術に関する調査：2018年版

齋藤 孝道[1]　細谷 竜平[2]　森 達哉[3]

概要：ある調査 [17] によると，90%以上の米国の家庭では 3 台以上のデバイスを所持しているとのことである．さらに，別の調査 [14] によると，ある Web サイトに初めて訪問してから購入に至る利用者の 36%は，複数デバイスを使うとのことである．そのような背景の中，2015 年 11 月に米国連邦取引委員会（FTC）により開催されたワークショップでの議論，および，同年同月に開催された国際会議 ICDM にて併設開催された競技コンテストもあり，クロスデバイストラッキング技術が広く知られるようになった．複数のデバイスを所持しそれを使い分ける現在において，インターネット空間における端末識別技術として，クロスデバイストラッキング技術が学術的にも多く発表され，商業的にも利用されるようになってきたと言える．クッキーを用いないトラッキング手法として，デバイス間だけでなく，ブラウザ間，アプリケーション間，それらのハイブリッドしたものなど様々ある．本論文では，クロスデバイストラッキングの実践についての類型をまとめ，代表的な実現手法の分析，および，それらに関連する議論について整理しまとめる．

Survey of Cross-Device Tracking: 2018 edition

Takamichi Saito[1]　Ryohei Hosoya[2]　Tatsuya Mori[3]

Abstract:
According to some survey[17], roughly 90 percent of American households have three or more devices connected to the internet. 36% consumers selectively use multiple devices to buy items in the internet shopping [14]. On November in 2015, the U.S. Federal Trade Commission (FTC) hosted a cross-device workshop to discuss the privacy issues of cross-device trackings. Moreover, Drawbridge Inc. hosted the Drawbridge Cross-Device Connections competition in ICDM 2015. Because of the events, people know that the cross-device trackings are getting popular on the internet. Currently, there are some types of cross-device trackings, developed by academic or commercial ones. In this study, we explore cross-device trackings to make a taxonomy of them.

1. はじめに

フランスの Web 広告配信企業 Criteo による 2016 年下 4 半期クロスデバイス・コマースレポート [14] によると，ある Web サイトに初めて訪問してから購入に至る利用者の中で，1 つのデバイスのみの利用者は 64%である一方で，複数デバイスの利用者は 36%であり，複数のデバイスを利用する人が少なからずいることが報告された．クロスデバイスでのトランザクションの 30 %はスマートフォンで開始されデスクトップでコンバージョンし，24 %はデスクトップで開始されスマートフォンでコンバージョンしているとのことである．日中，多くの利用者はデスクトップ PC を用いる傾向があり，夜や週末はモバイルを使う傾向にある．すなわち，一般消費者は，複数のデバイスを使い分けながら，インターネットショッピングを行なっていると言える．

これらの事実は，（HTTP）クッキーを使った従来型のトラッキングでは，サードパーティーを含む事業者[*1]は，利用者[*2]を，正確に識別できないことを如実に物語っている．例えば，サードパーティクッキーを用いて広告効果の

[1] 明治大学
Meiji University
[2] 明治大学大学院
Graduate School of Meiji University
[3] 早稲田大学
Waseda University

[*1] 本論文では，トラッキングにより利用者を識別するサードパーティーの主体を，事業者と呼ぶ．
[*2] 本論文では，トラッキングされる対象を，利用者と呼ぶ．

計測を試みる際，適切に計測できない．そこで，事業者らは，メールアドレス，広告 ID など用いて，PC やモバイルといった端末を跨いだトラッキング手法，いわゆる，クロスデバイストラッキング技術を開発し，それを用いるようになってきた．

2015 年以降，一般利用者が所有するさまざまなデバイス間の動作を紐付けするクロスデバイストラッキングと呼ばれる技術が開発され普及してきた．代表的な利用例としては，以下がある：

- ソーシャルメディアのアカウントへのログインにより，複数のデバイスにおいてシームレスなサービスの提供
- 複数のデバイスを利用する者の識別による行動追跡
- デバイス間のトラッキングによって，不正なアクセスの防止

クロスデバイストラッキング技術を使用することで，事業者は．同一人物により使用される複数のデバイスを，当該同一人物に関連付けることができる．

本論文では，デバイス，ブラウザ，および，アプリケーションを跨いで利用者を紐づける行為を，広義のクロスデバイストラッキングと捉え，インターネット空間における利用者の識別という観点から，サードパーティーで利用されるクロスデバイストラッキング技術を網羅的に調査・整理した結果を報告する．

2. 背景知識

2.1 フィンガープリンティング

フィンガープリンティングはデバイスで採取可能な情報を用いてアプリケーション，ブラウザやデバイスなどを，サーバ側で識別する技術である．ここで，デバイスから採取可能な個々の情報の種類を特徴点，採取された情報（値）の組み合わせをフィンガープリントと定義する．特に，ブラウザを識別する技術を，ブラウザフィンガープリンティングという．

フィンガープリンティングにおける識別は，限られた少数種の情報（フィンガープリント）により，確率的に主体を識別でき得ることが特徴である．たとえば，4 箇所の情報と所在した日時の情報があれば，アメリカ人の 95 %の氏名を特定できる [36] といった事実が知られているが，少ない情報の組み合わせを用いて，数億のうちの 1 つを特定できる．フィンガープリンティングにおける識別もこのような性質に依拠している．

Englehardt ら [26] は，2016 年 1 月時点での Alexa トップ 100 万サイトのブラウザフィンガープリンティング使用状況を調査した．その結果，ブラウザフィンガープリンティングを実施しているサイトは，100 万サイトの内 14,371（1.6%）あることのことだった．

フィンガープリンティングは，その手法に基づき，2 つに分類できる：

(1) パッシブ・フィンガープリンティング
(2) アクティブ・フィンガープリンティング

パッシブ・フィンガープリンティングとは，ブラウザなどの通常のアクセス通信に伴ってサーバ側で受動的に採取可能な情報，たとえば，HTTP ヘッダ，UserAgent や IP アドレスなどの情報のみで行うフィンガープリンティングのことを言う．高橋ら [38] によれば，パッシブ・フィンガープリンティングでも，比較的高精度で識別できるとのことである．

アクティブ・フィンガープリンティングとは，JavaScript などのスクリプトをブラウザなどで実行して情報を得るフィンガープリンティングのことを言う．JavaScript を用いたフィンガープリンティングが広く知られているが，CSS のみで行うものもある [34]．かつては，Flash を利用するものもあった．ブラウザにて採取したフィンガープリントは，フィンガープリンティングを行うサーバにて集約し，独自のアルゴリズムにて識別する．ただし，ブラウザにおいてアクティブフィンガープリンティングの際，JavaScript 等の実行は，サンドボックスにより，ネイティブアプリのように，デバイス識別情報等を取得できない制約がある．モバイルアプリにおいては，事業者よりアプリ開発者向けキットという形態で，採取のためのライブラリが提供される事例がある．そのライブラリをアプリに組み込み，当該ライブラリを通して，事業者が用意するサーバにフィンガープリントを集約し，デバイスを識別する．

2.2 クロスデバイストラッキングの手法

クロスデバイストラッキングの主なアプローチには，以下の 2 つがある [9], [29], [32]：

(1) 決定論的（クロスデバイス）トラッキング
(2) 確率論的（クロスデバイス）トラッキング

決定論的トラッキングとは，クレデンシャル情報やデバイス識別情報によって，複数のデバイスを紐付ける手法である．クッキーを用いた（クロスデバイス）トラッキングもこれに分類する．利用する情報として，利用者がアプリに提供する電子メールアドレス，電話番号や，認証 API（Facebook，OpenID，Google など）のアプリにサインアップする際に提示する ID などを用いる．そのほかに，Apple 固有のデバイス ID，Google の Android ID，MAC アドレス[*3]などのデバイス固有の識別子を用いることもある．精度はほぼ 100%と高いが，その仕組み上，一つのサービス内でのトラッキングや，メールアドレスを用いたトラッキングとなるため，スケールアップが難しい．

確率論的トラッキングとは，IP アドレスや UserAgent などのフィンガープリントを用いて，確率モデルを用いて，複数のデバイスを紐付ける手法である．ここで適用する確

[*3] 現状，Android，iOS において，アプリおよびブラウザ共に，MAC アドレスの取得・利用は実質不可能である．

率モデルには様々あり，以下に示すような経験則を組み合わせて行うものや，教師あり機械学習を用いるものなど多種存在する．たとえば，同じ IP アドレスを使う違うデバイスは，ある家庭内の 1 つの Wi-Fi ルータを経由している可能性があるという経験則などにより，デバイス間を紐付けてグラフを作成することで，デバイス間の紐付けを実現する．この時のグラフをデバイスグラフ [2] と呼ぶ．

通常，PC もしくはスマートフォンに限らず，デバイス内であってもブラウザ間でのデータのやり取りは実質不可能であるので，決定論的トラッキングを大規模に実施することは難しく，確率論的手法によりスケールアップを狙う．

2.3 クロスデバイストラッキング

狭義のクロスデバイストラッキングとは，同一の利用者が使用している複数のデバイスを紐付けることである（図 1 参照）．たとえば，モバイルデバイスからのアクセスと，デスクトップデバイスからのアクセスは，同一のサーバへのアクセスとは限らず，別のサーバへのアクセスをそれぞれ別個に採取し，それらをフィンガープリントを用いて紐付けることもある．

利用するフィンガープリントの例としては，一般に，以下がある：

- IP アドレス（IP と表記）
- タイムスタンプ（TS と表記）
- UserAgent（UA と表記）
- 広告 ID（モバイルアプリの場合）
- （サイトへの）アクセス履歴
- クッキー

ここで，広告 ID は，モバイルデバイスだけで取得される．広告 ID やクッキーは，ブラウザ間で共有できないだけでなく，利用者に更新や削除されることを想定する．このケースでは，広告 ID を除き，いずれもパッシブフィンガープリントである．

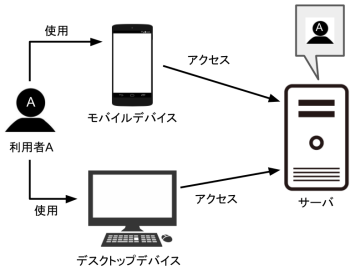

図 1 クロスデバイストラッキングの概念図

2.4 クロスアプリケーショントラッキング

クロスアプリケーショントラッキングとは，利用者が使用しているデバイスの中で，複数のアプリケーションを跨いでそれらを紐付けることである（図 2 参照）．

クロスデバイスのケースと同様に，2 つのモバイルアプリからのアクセスは，同一のサーバへのアクセスとは限らず，別のサーバへのアクセスをそれぞれ別個に採取し，それらをフィンガープリントを用いて紐付けることもある．

利用するフィンガープリントの例としては，クロスデバイストラッキングのそれらと同じである．特に，モバイルアプリなどの場合，広告 ID が利用できれば，トラッキングは容易であるが，利用者に拒否されることや更新されることを想定して実施する必要がある．

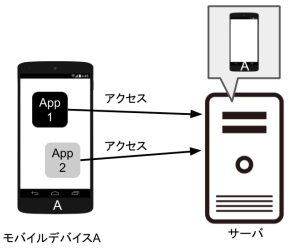

図 2 クロスアプリケーショントラッキングの概念図

2.5 クロスブラウザトラッキング

クロスブラウザトラッキングとは，クロスアプリケーショントラッキングの一種だが，利用者が使用している端末の中でブラウザを跨いでそれらを紐付けることである（図 3 参照）．

クロスデバイスのケースと同様に，2 つのブラウザのアクセスは，同一のサーバへのアクセスとは限らず，別のサーバへのアクセスをそれぞれ別個に採取し，それらをフィンガープリントを用いて紐付けることもある．

利用するフィンガープリントは，主に，以下を用いる：

- IP アドレス（IP と表記）
- タイムスタンプ（TS と表記）
- UserAgent（UA と表記）
- （サイトへの）アクセス履歴
- クッキー

ここで，クッキーは，ブラウザ間で共有できないだけでなく，利用者に削除されることを想定する．このケースでは，いずれもパッシブフィンガープリントである．

3. クロスデバイストラッキングの関連動向

3.1 業界の動向

2012 年 3 月に，FTC は，消費者プライバシーの向上のため，事業者が保有しているデータについての透明性などを求め，業界全体へ注意喚起した [15]．しかし，業界の反

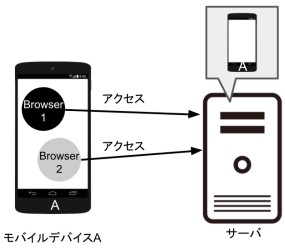

図 3 クロスブラウザトラッキングの概念図

応は鈍かったとのことである．

さらに，クロスデバイスでのトラッキング技術の導入の実情を知るために，FTC は 2015 年 11 月に，Cross-Device Tracking Workshop を開催した [32]．

同年同月に開催されたデータマイニング分野の国際学会である ICDM2015 にて，DrawBridge 社が主催するクロスデバイストラッキングの競技コンテスト [3] が開催された．このコンテストでは，Kaggle によって与えられたデータ [6] を用いてクロスデバイスの紐付け精度を競った．このコンテストに併せて，クロスデバイストラッキングのワークショップが開催され，クロスデバイストラッキングに関する論文が多数発表された．

2017 年には，クロスデバイストラッキング技術に関する透明性を確保するベストプラクティスとして，FTC のレポート [32] が公開された．

3.2 ガイドライン

FTC のレポート（2017 年）[32] において，クロスデバイストラッキングを実践および利用する企業へ，消費者の保護の観点から守るべき指針が示されている．

FTC のレポートに加え，業界の自己規制の取り組みとして DAA によるガイドライン（2015 年）や NAI のガイドライン（2017 年）[5] によれば，クロスデバイストラッキングの実践に対して，以下のような指針が示されている．

- 透明性：クロスデバイストラッキングを利用する企業は，トラッキング活動を開示する必要がある
- 選択肢：ユーザに対して，オプトアウトの仕組みや Web，アプリ，サービスの連携をさせるさせないの選択を与える必要がある．シングルオプトアウトの提供が望ましい

その他の指針として，FTC のレポートには，クロスデバイストラッキングにおける，機微情報の取り扱いや，セキュリティの確保についても示されている．

サードパーティの広告プラットフォームが確率論的トラッキングをする場合，特に，一般消費者（利用者）からはその行為を特定することは不可能である．よって，トラッキングを実施する事業者は，その実践について自己開示が求められる．

しかしながら，Brookman ら [23] によると，サードパーティのクロスデバイストラッキングを実施していることを明示するサイトは，人気のある 100 のうち 3 つのサイトしかなかったとのことであった．

一般消費者（利用者）への選択肢の提示についても，オプトアウトの仕組みを用意しつつも，意図的に誤解をさせている，もしくは，オプトアウトが一般消費者（利用者）にとって容易ではないケース，いわゆる，ダークパターンの指摘 [19] もあり，昨今では，事業者側の対応によっては非難を受けることがある．

3.3 クロスデバイストラッキング実施状況の調査研究

2009 年，Kane ら [27] は，Web 利用者による複数の端末を用いた場合の Web ブラウジングの状況を調査した．その結果，大半の利用者が複数のデバイスにおいて Web の利用情報を共有していることが判明した．また，追実験として PC とモバイルデバイスを併用する Web 利用者を実際にトラッキングし，経験サンプリング法を用いてデバイスの使い分け方を調査した．その結果，最初の実験と同様に，大半の利用者は PC とモバイルデバイスで同じような Web サイトを閲覧していることが判明した．その上，Web 利用者は PC で Web サイトを閲覧したのち，その Web サイトの付加的な部分をモバイルデバイスで閲覧していることが判明した．

Su ら [33] は，2017 年，Web サイトの閲覧履歴と Twitter や Facebook などの SNS で公開されているプロフィールを結び付けることによって，個人の特定が十分可能であることを示した．また，これらの情報を利用することで，サードパーティによるクロスデバイストラッキングが可能であると述べられている．

Arp ら [21] は，2017 年，モバイル端末における超音波ビーコンを使った用いたクロスデバイストラッキングの利用状況を調査した．その結果，超音波ビーコンによるクロスデバイストラッキングを可能にするためにマイクへのアクセス承認を求めている Android アプリが 234 件あることが判明した．また，調査チームが訪れたドイツ国内の 35 店舗のうち 4 店舗で超音波ビーコンを発する装置が店の入り口に設置されていたことが判明した．そして，超音波ビーコンを実現するために SilverPush [12]，Lisnr [8] および Shopkick [11] と呼ばれる 3 種類の SDK が用いられており，SilverPush はクロスデバイストラッキングを，Lisnr と Shopkick は位置情報のトラッキングを行っていることが判明した．その中でも，SilverPush が最も多く使われていることが判明した．

Mavroudis ら [30] は，2017 年，超音波ビーコンを用いたクロスデバイストラッキングのメカニズムおよび，それにまつわるプライバシーの懸念について調査した．

Brookman ら [23] は，2017 年，100 の Web サイトを調査し，以下の結果を得た．

(1) クロスデバイストラッキング事業者が，100 のうちの 87 存在した
(2) クロスデバイストラッキングに関与するサードパーティが 861 あった
(3) 100 ののうち 96 が利用者にユーザ名または電子メールアドレスをサイトに提供させていた．そのうちの一部は，サードパーティとの共有のためと推定される

さらに，100 のうち 16 のサイトで，サードパーティとのユーザ名または電子メール（raw またはハッシュ）を共有していた．また，彼らは，クロスデバイストラッキング事業者が，利用者にトラッキング実施状況をどのように告知しているかを調査した．

4. 決定論的手法によるトラッキングの類型

ここでは，決定論的な手法によるトラッキングの類型についてまとめる．

4.1 クッキーを用いたトラッキング

事業者が，対象とするパブリッシャーのサイトに，タグを埋め込めさせ，事業者サイトからのクッキーを発行し，それによるトラッキングを行う手法が広く普及した．2011 年の Purra らの論文 [31] によると，Alexa トップ 500 の Web サイトの内 79%のサイトにおいてサードパーティートラッキングが確認されたとのことである．

その他のクッキーを用いたトラッキングには，次のようなものがある．モバイル端末 α 内のブラウザ X において，利用者が広告クリック時に取得した際にクッキー A を事業者サーバ側で発行する．その後，利用者が，モバイル端末 α 内のアプリケーション Y を起動時に，何らかの方法で，クッキー A をブラウザ X からアプリケーション Y に渡し，事業者サーバにおいて，クッキー A を照会することで，ブラウザ X とアプリケーション Y との紐付けを行う．

2017 年実施の調査 [18] によると，クッキーの 64%がブロックまたは削除されていたことがわかった．特に，モバイルでの拒否率は 75%である．Apple Safari においては，ITP[*4]により，サードパーティークッキーの利用はできない．これらのことより，クッキーを用いたトラッキングは，今後，減少することが予想される．

4.2 クレデンシャル情報によるトラッキング

決定論的トラッキング手法として，メールアドレスなどクレデンシャル情報を用いたトラッキングがある．この手法では，会員制のファーストパーティの Web サイトに登録されたメールアドレスやそのハッシュ値などのクレデンシャル ID を，サードパーティの事業者サーバに登録することにより，事業者サーバが一元的に会員（利用者）をトラッキングすることが可能になる．また，メールアドレスは利用者が複数の端末で共通に利用するものであることから，事業者サーバによる利用者のクロスデバイストラッキングが実現できる．この手法はサービスとして普及しており，その一例として米 TARGETO 社の Email Retargeting [4] が挙げられる．

2018 年 7 月に，Google Analytics の新機能として，自動クロスデバイストラッキング機能が公開された [16]．Google アカウントから収集された情報を用いて，デバイス間を紐付けるとのことである．

4.3 デバイス内情報を用いたトラッキング

2018 年現在，デバイス固有情報はトラッキングに実質利用ができないので，広告 ID を利用することがある．デバイス固有情報との違いは，広告 ID の場合，利用者が，事業者に対して，利用の可否を選択できることにある．また，利用者は任意のタイミングで，広告 ID をリセットおよび拒否できる．

広告 ID を用いたトラッキングでは，モバイルのアプリケーションにおいて，利用者が広告クリック時に取得した際に，広告 ID を事業者サーバに送る．その後，利用者が，別のアプリケーションの起動時に，取得した広告 ID を保存されている広告 ID と照会することで，紐付けを行う．

Google のみが提供する機能として，Google Play ストアへ遷移した際のリファラ情報および付加情報を使うことで，トラッキングする手法もある．たとえば，アプリに広告を埋め込み，その広告経由でアプリをダウンロードする際に，アプリとブラウザを紐付ける．

4.4 超音波ビーコンを利用したトラッキング

2014 年，SilverPush 社は，「オーディオビーコン」と呼ばれる超音波のビーコンを使用したクロスデバイストラッキング手法を発表した [1], [21]．

超音波ビーコンとは，人間の耳には聞こえない高い周波数を持つビーコンのことであり，大半のモバイルデバイスはマイクにより超音波ビーコンを認識することが可能である．この技術では，テレビの広告の音声や，特定の Web サイトに超音波ビーコンを再生する仕組みを埋め込んでおき，利用者がテレビ CM を閲覧したり Web サイトを訪れた際に超音波を発信する．そして，モバイルデバイスはその超音波ビーコンを認識し，利用者と複数のデバイスを紐付ける．どんな CM を見たか，どの程度の時間にわたって広告を見たか，そして広告を見たことで利用者がどのような行動を取り，何を買うに至ったのか，という一連の行動が全てトラッキングされるということにつながる．なお，BlackHat にて対策手法が公開された [13]．

[*4] ITP2.0 [7] においては，User-Agent 固定化により，フィンガープリンティング自体も困難となる．

5. 確率論的手法によるトラッキングの類型

ここでは，確率論的手法によるトラッキングの類型についてまとめる．これらについて，表 1 に比較をまとめた．

5.1 クロスブラウザトラッキング

確率論的手法ではあるが，実験規模も小さく，あくまでコンセプトの提案という位置付けとして，次の研究がある．

Boda ら [22] は，異なるブラウザ間でのトラッキングに言及した最も初期の研究を 2012 年に発表した．ブラウザから取得可能な IP アドレス，フォントセット，タイムゾーン，画面解像度から，利用者を識別する手法を議論した．利用者を一意に識別でき得ることに言及した．

5.2 BLE を用いたクロスアプリトラッキング

これも確率論的手法ではあるが，実験規模も小さく，あくまで実証実験という位置付けとして，次の研究がある．Korolova ら [29] は，モバイルデバイスにて，Bluetooth Low Energy (BLE)[*5] により，近接エリア探索を行い，ペリフェラルなどの種類などをフィンガープリントとして採取し，モバイルアプリ間での結果を比較して，アプリ同士の紐付けができることを示した（図 4 参照）．

URI Advertising Mode と Configuration Mode で送信される BLE パケットに含まれる，Access Address や Peripheral UUID の情報などを用いる．Access Address は，Android の場合は，`8E 89 BE D6` と固定されている．Apple iOS では，Peripheral UUID は，ランダム値になっているので，デバイスごとに値が変わる．BLE アドバタイズスキャンによる取得した値の組をフィンガープリントとして識別する．特に，Android 6.0(API level 23) と iOS のいずれのデベロッパー API でも，BLE アドバタイズスキャンには，利用者からの許可を必要としない．

100 人の参加による実験により，Cosine 類似度を計算して，紐付けを行なった．Android および iOS のいずれの場合でも，10 分に 1 度の頻度であればほぼ 100 %で紐付けができた．アプリを 1 日に 1 回しか使用しない場合でも，利用者の半分以上が正しく一致できた．

5.3 ICDM 発表のクロスデバイストラッキング研究

ここでは，確率論的手法の例として，2015 年の ICDM にて併催された競技コンテストでの発表研究をまとめる．

Anand ら [20] は，クロスデバイストラッキングの手法として，Feature Engineering と機械学習のアルゴリズムを用いた手法を提案した．機械学習のアルゴリズムには Extreme Gradient Boosting (xgboost) や Follow the Reguralized Leader Proximal (FTRL-Proximal) のような手法が利用された．

[*5] Bluetooth Version 4.0 で新規に追加された低消費電力の通信モード．BLE の最大通信距離は 100m 長である．

図 4 BLE を用いたクロスアプリケーショントラッキング

Kejera ら [28] は，クロスデバイストラッキングの手法として，Gradient Boosting Decision Trees (GBDT) とランダムフォレストのアンサンブル学習を用いた手法を提案した．

Díaz-Moralesl [25] は，クロスデバイストラッキングの手法として，半教師あり機械学習の技術を用いた手法を提案した．

Walthers [35] は，クロスデバイストラッキングの手法として，Learning to Rank を用いた手法を提案した．結果として，この手法が最も高い精度を誇ることになった．

5.4 デバイスグラフによるクロスデバイストラッキング

Drawbridge 社 [2] は，相互接続されたデバイスグラフと呼ばれるもので構成される Connected Consumer Graph の相関関係を利用する技術を利用している．特許（US9514248B1）も取得している．いくつかの確率モデルを組み合わせて，デバイスについての予測を行い，precision は 97.3%と示されている．

当該技術は，以下に示すいくつかの経験則および仮説を複合的に用いて，デバイスグラフを作成する．同じパブリック IP アドレスを使用する PC，スマートフォン，および，タブレットがあった場合，同じ家庭に属していると推測する．ただし，複数のデバイスが同じ IP アドレスを使う場合であっても，家庭で利用するデバイスの数より大きければ，そこはネットカフェであるとするような推測をする．また，対象者のスマートフォン A が営業時間中に仕事用 PC B と同じパブリック IP アドレス C を使用し，営業時間外に家庭用コンピュータと同じパブリック IP アドレス X を使用する場合は，スマートフォン A，家庭用 PC D は同じ利用者が利用していると推測する．

5.5 クロスデバイストラッキング

次も確率論的手法ではあるが，実験環境下でのサンプル採取であり，規模も小さく，あくまで実証実験という位置付けとして，次の研究がある．

Zimmeck ら [37] は，被験者として利用者 126 名から収集した 27,000 のサンプルを用いて，クロスデバイストラッ

表 1　確率論的手法を用いたクロスデバイストラッキングの比較

	利用する FP	サンプル数	手法	精度（metrics）	備考
Anand ら [20]	IP, クッキー	244M	xgboost, FTRL-Proximal	0.809 ($F_{0.5}$)	ICDM2015
Kejera ら [28]	IP, クッキー	244M	ensemble 学習 (GBDT と Random forest)	0.855 ($F_{0.5}$)	ICDM2015
Walthers[35]	IP, クッキー	244M	binary classification	0.89($F_{0.5}$)	ICDM2015
Cao ら [24]	IP, クッキー	244M	ensemble learning	0.86 ($F_{0.5}$)	ICDM2015
Tapad 社 [10]	N/A	N/A	Device Graph	0.91 (Prec)	パンフレット
Drawbridge 社 [2]	N/A	N/A	Connected Consumer Graph	0.97 (Prec)	パンフレット
Zimmeck ら [37]	IP, Dom(web, app)	32K(IP 4K, Dom 28K)	Bhattacharyya 距離を多層利用	0.88(Prec) 0.91(F_1)	USENIX2017
齋藤ら [39]	IP，UA，TS, クッキー	376M(App 22M, Dom 354M)	深層学習・ランダムフォレスト	0.98(Prec) 0.98(F_1) 0.99($F_{0.5}$)	CSS2018

キングを実施した．この実験により，モバイル端末からデスクトップ端末へのクロスデバイストラッキングの精度として，F_1 値が 0.91 となったことを示した．これは，IP アドレス，Web サイトの閲覧履歴の類似性（Bhattacharyya 距離を多層的に利用）により達成した．また，彼らは広告配信を観察することによって，クロスデバイストラッキングの有無を検知することが出来ることを示した．

しかしながら，サンプル数が少なく，一般性のある結論かどうかについては断定できない．

5.6　深層学習を用いたトラッキング

次も確率論的手法ではあるが，実環境下でのサンプル採取であり，規模も比較的大きく，実践的な試みとして，次の研究がある．

齋藤ら [39] は，深層学習とランダムフォレストをそれぞれ用いて，モバイルアプリの通信とブラウザの通信との紐付け，いわゆる，クロス App-Web トラッキングを行なった．

これは，同一デバイスからのアクセス情報のうち，モバイルアプリから採取できるフィンガープリント（以降，アプリ由来データと呼ぶ）と，ブラウザから Web サーバへのアクセスの際に採取できるフィンガープリント（以降，Web 由来データと呼ぶ）の紐付けのために，教師データを用いた深層学習により判別器を作成し，その判別器により，紐付けを行う．

利用したフィンガープリントは，IP アドレス，UserAgent 情報，タイムスタンプおよび，クッキーである．また，サンプル総数は，約 376M 個であり，Web 由来データのサンプル数は 22M 個，また，アプリ由来データのサンプル数は 354M 個である．提案手法の精度として，precision，recall，accurarcy，F_1，F_2，および，$F_{0.5}$ の値は，それぞれ，0.980，0.999，0.989，0.989，0.984，および，0.995 であった．大規模なサンプルにおいて，高精度な結果を得たと言える．

Drawbridge 社 [2] のクロスデバイストラッキングの精度として，precision が 97.3%と示されているので，同社の手法を用いたクロスデバイストラッキングをそのまま適用し同じ値を取ると仮定をしても，この結果は同程度の精度があるといえそうである．

6.　まとめ

クロスデバイストラッキングの手法が様々提案され導入されていることがわかった．その精度は手法によらず高く，precision および recall などいずれも，概ね 90%以上と高精度であることがわかった．

クロスデバイストラッキングにおいては，パッシブ・フィンガープリントを用いた確率的トラッキングを行うことも多い．その場合，利用者には，事業者によるクロスデバイストラッキングの実施の有無を特定することは不可能である．クロスデバイストラッキングの精度はとても高いので，今後，事業者によるクロスデバイストラッキングの実施状況の自己開示といった透明性の確保が期待される．

参考文献

[1] Beware of ads that use inaudible sound to link your phone, TV, tablet, and PC, https://arstechnica.com/tech-policy/2015/11/beware-of-ads-that-use-inaudible-sound-to-link-your-phone-tv-tablet-and-pc/.

[2] Cross-Device Consumer Graph, https://go.drawbridge.com/rs/454-ORY-155/images/Drawbridge-Cross-Device-Consumer-Graph.pdf.

[3] Drawbridge Cross-Device Connection Challenge, https://icdm2015.stonybrook.edu/content/call-contest/.

[4] Email Retargeting - TARGETO, https://targeto.io/email-retargeting.

[5] Guidance for NAI Members: Cross-Device Linking, https://www.networkadvertising.org/pdfs/NAI_Cross_Device_Guidance.pdf.

[6] ICDM 2015: Drawbridge Cross-Device Connections - Kaggle, https://www.kaggle.com/c/icdm-2015-drawbridge-cross-device-connections.

[7] Intelligent Tracking Prevention 2.0 — WebKit, https://webkit.org/blog/8311/intelligent-tracking-prevention-2-0/.

[8] Lisnr, http://lisnr.com/.

[9] Mobile Technology Tracking Methods other than cookies, http://www.allaboutcookies.org/mobile/mobile-tracking.html.

[10] Nielsen Study Finds Tapad's Device Connections 91.2 Percent Accurate, https://www.tapad.com/press-release/nielsen-study-finds-tapads-cross-

device-connections-91-2-percent-accurate.
[11] Shopkick, https://www.shopkick.com/.
[12] SilverPush, https://www.silverpush.co/.
[13] TALKING BEHIND YOUR BACK: ATTACKS AND COUNTERMEASURES OF ULTRASONIC CROSS-DEVICE TRACKING , https://www.blackhat.com/eu-16/briefings/schedule/#talking-behind-your-back-attacks-and-countermeasures-of-ultrasonic-cross-device-tracking-4864.
[14] The State of Cross-Device Commerce, https://www.criteo.com/wp-content/uploads/2017/07/Report-criteo-state-of-cross-device-commerce-2016-h2-UK.pdf.
[15] Protecting Consumer Privacy in an Era of Rapid Change: Recommendations For Businesses and Policymakers, https://www.ftc.gov/reports/protecting-consumer-privacy-era-rapid-change-recommendations-businesses-policymakers (Mar 2012).
[16] Better understand and reach your customers with new Cross Device capabilities in Google Analytics, https://www.principle-c.com/column/ga/ga-cross-device-tracking/#__1156673465.1533265079 (Nov 2015).
[17] A third of Americans live in a household with three or more smartphones, http://www.pewresearch.org/fact-tank/2017/05/25/a-third-of-americans-live-in-a-household-with-three-or-more-smartphones/ (2017).
[18] 64% Of Tracking Cookies Are Blocked, Deleted By Web Browsers, https://www.mediapost.com/publications/article/316757/64-of-tracking-cookies-are-blocked-deleted-by-we.html?edition=108287 (Mar 2018).
[19] DECEIVED BY DESIGN, https://fil.forbrukerradet.no/wp-content/uploads/2018/06/2018-06-27-deceived-by-design-final.pdf (2018).
[20] Anand, T. R. and Renov, O. Machine Learning Approach to Identify Users Across Their Digital Devices, 2015 IEEE International Conference on Data Mining Workshop (ICDMW) (Nov 2015).
[21] Arp, D., Quiring, E., Wressnegger, C. and Rieck, K. Privacy Threats through Ultrasonic Side Channels on Mobile Devices, 2017 IEEE European Symposium on Security and Privacy, EuroS&P 2017, Paris, France, April 26-28, 2017 (2017).
[22] Boda, K., Földes, A. M., Gulyás, G. G. and Imre, S. User Tracking on the Web via Cross-browser Fingerprinting, Proceedings of the 16th Nordic Conference on Information Security Technology for Applications, NordSec'11, Berlin, Heidelberg (2012), Springer-Verlag.
[23] Brookman, J., Rouge, P., Alva, A. and Yeung, C. Cross-device tracking: Measurement and disclosures, *Proceedings on Privacy Enhancing Technologies*, **2017**, 2 (2017), 133–148.
[24] Cao, X., Huang, W. and Yu, Y. Recovering Cross-Device Connections via Mining IP Footprints with Ensemble Learning., ICDM Workshops, IEEE Computer Society (2015).
[25] Díaz-Morales, R. Cross-Device Tracking: Matching Devices and Cookies, 2015 IEEE International Conference on Data Mining Workshop (ICDMW) (Nov 2015).
[26] Englehardt, S. and Narayanan, A. Online Tracking: A 1-million-site Measurement and Analysis, Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, CCS '16, New York, NY, USA (2016), ACM.
[27] Kane, S. K., Karlson, A. K., Meyers, B. R., Johns, P., Jacobs, A. and Smith, G. Exploring Cross-Device Web Use on PCs and Mobile Devices, Human-Computer Interaction – INTERACT 2009 (eds.Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R. O. and Winckler, M.), Berlin, Heidelberg (2009), Springer Berlin Heidelberg.
[28] Kejela, G. and Rong, C. Cross-Device Consumer Identification, 2015 IEEE International Conference on Data Mining Workshop (ICDMW) (Nov 2015).
[29] Korolova, A. and Sharma, V. Cross-App Tracking via Nearby Bluetooth Low Energy Devices, Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy, CODASPY '18, New York, NY, USA (2018), ACM.
[30] Mavroudis, V., Hao, S., Fratantonio, Y., Maggi, F., Kruegel, C. and Vigna, G. On the Privacy and Security of the Ultrasound Ecosystem, *Proceedings on Privacy Enhancing Technologies*, **2017**, 2 (2017), 95–112.
[31] Purra, J. and Carlsson, N. Third-Party Tracking on the Web: A Swedish Perspective, 2016 IEEE 41st Conference on Local Computer Networks (LCN) (Nov 2016).
[32] Ramirez, E., Ohlhausen, M. K. and McSweeny, T. Cross-Device Tracking, An FTC Staff Report, FEDERAL TRADE COMMISSION (2017).
[33] Su, J., Shukla, A., Goel, S. and Narayanan, A. De-anonymizing Web Browsing Data with Social Networks, Proceedings of the 26th International Conference on World Wide Web, WWW '17, Republic and Canton of Geneva, Switzerland (2017), International World Wide Web Conferences Steering Committee.
[34] Takei, N., Saito, T., Takasu, K. and Yamada, T. Web Browser Fingerprinting Using Only Cascading Style Sheets, 2015 10th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA) (Nov 2015).
[35] Walthers, J. Learning to Rank for Cross-Device Identification, 2015 IEEE International Conference on Data Mining Workshop (ICDMW) (Nov 2015).
[36] Montjoye, Yves-Alexandre de M. V. . V. D. B., Csar A. Hidalgo Unique in the crowd: The privacy bounds of human mobility, http://www.nature.com/srep/2013/130325/srep01376/full/srep01376.html (2013).
[37] Zimmeck, S., Li, J. S., Kim, H., Bellovin, S. M. and Jebara, T. A privacy analysis of cross-device tracking, 26th USENIX Security Symposium (USENIX Security 2017) (2017).
[38] 高橋和司, 安田昂樹, 種岡優幸, 田邉一寿, 細谷竜平, 野田隆文, 齋藤祐太, 小芝力太, 齋藤孝道, HTTP ヘッダのみを用いた Browser Fingerprinting の考察, 暗号と情報セキュリティシンポジウム (SCIS) 2018 (2018).
[39] 齋藤祐太, 細谷竜平, 齋藤孝道, 森達哉, クロスアプリケーションフィンガープリンティング—同一端末上のアプリケーション間の紐付け技術—, 2018 Computer Security Symposium (2018).