An Extensive Study on Online and Mobile Ad Fraud

Hala Shaari[1*], Nuredin Ahmed[2]
[1] h.shaari@uot.edu.ly, [2] nuredin@mtit.com.ly
[1] Software Engineering Department, Faculty of Information Technology, University of Tripoli, Libya
[2] Control Department, Technical Computer College, National Board for Technical and Vocational Education, Libya

ABSTRACT

The advertising ecosystem faces major threats from ad fraud caused by artificial display requests or clicks, created by malicious codes, bot-nets, and click-firms. Currently, there is a multibillion-dollar online advertisement market which generates the primary revenue for some of the internet's most successful websites. Unfortunately, the complexities of the advertisement ecosystem attract a considerable amount of cybercrime activity, which profits at the expense of advertisers. Web ad fraud has been extensively studied whereas fraud in mobile ads has received very little attention. Most of these studies have been carried out to identify fraudulent online and mobile ads clicks. However, the identification of individual fraudulent displays in mobile ads has yet to be explored. Additionally, other fraudulent activity aspects such as hacking ad-campaign accounts have rarely been addressed. The purpose of this study is to provide a comprehensive review of state-of-the-art ad fraud in web content as well as mobile apps. In this context, we will introduce a deeper understanding of vulnerabilities of online/mobile advertising ecosystems, the ad fraud's well-known attacks, their effective detection methods and prevention mechanisms.

Keywords:

Online Ad Fraud,
Mobile Ad Fraud,
Advertising Ecosystem,
Detection,
Prevention.

1- Introduction

World wide web today provides consumers with a versatile and easily available platform for listing and viewing advertising compared to more conventional media such as newspapers and printed booklets. Digital advertising has grown into one of the world's largest and most lucrative industries. It is one of the key methods to produce revenues from digital media (e.g., websites and mobile apps) by providing advertisements to customers. Online advertising is unfortunately a fractured industry with a highly inefficient business model which is vulnerable to fraud and abuse. Fraudsters have been able to leverage many vulnerabilities of the online advertisement model and have begun to manipulate the program to make a profit. Such fraudulent activities are popularly known as Advertisement Fraud (Ad Fraud), also known as Invalid Traffic.

Digital advertising is a cornerstone of finance that funds free internet content and services, and free smartphone applications. At a high level of online advertising, the basic concept is to sell space on web pages and apps for advertising. The mechanisms and infrastructure required for online ads are indeed highly diverse and complex. Since web and mobile ads use similar infrastructure, they face the same security issues related to them.

The ad ecosystem can be partitioned into three groups approximately: advertisers, publishers and intermediaries. Advertisers pay publishers to place a specified amount of creative content on websites and applications with embedded links such as text, display or video ads. Intermediaries (e.g. ad servers/networks and ad exchanges) are also used to enable communication between

advertisers and publishers. Intermediaries usually charge advertisers a fee for ad placements and/or ad clicks to publishers. What is instantly evident from this basic explanation is that revenues from publishers and intermediate platforms are directly related to the number of regular visitors to a website or app.

Therefore, publishers and intermediaries are strongly incentives to use whatever means are available to push user traffic to publisher pages. However, there is another form of fraud that does not involve the publishers. Dishonest advertisers aim to simulate demands on their competitor's ads to deplete the advertising budgets of their competitors. These methods for traffic generation have emerged, many of which are deeded as fraudulent by advertisers and intermediaries.

The aim of this research is to comprehensively review recent ad fraud activities in web content and also mobile apps. We intend to provide a better understanding of vulnerabilities of online/mobile advertising ecosystems and the well-known attacks. In Section 2, we first address the online advertising ecosystem model and mobile ecosystem model, and then discuss different revenue models for them. Online/Mobile workflows have also been introduced. The remainder of the paper is organized as follows. A comprehensive description of Ad Fraud types in Online/Mobile advertising provided in Section 3 and 4. Then, Ad Fraud existing prevention mechanisms presented in Section 5. Finally, we conclude the paper in Section 6.

2- Understanding Ad Ecosystem

This section provides an overview of the online digital advertising ecosystem and its main components for better understanding fraud activities, their major characteristics, and corresponding detection mechanisms, which will be described in the following Sections. Then, we describe a number of advertising types pertinent to our discussion. Both a business model and technical framework for delivering advertisements delivering advertisements to publisher websites and apps will introduce.

Main components

According to [1][2]Online ads primarily include four agents:

• Publisher is an organization that publishes content or provides a service through a website or a mobile app.

• Advertiser is an organization that pays the ad networks to show its ads.

• Users are any visitors to the website of the publisher who might be interested and then click on the advertiser.

• Ad networks/servers are companies that manage publishers and advertisers. They are able to buy and sell ad traffic (in the form of ad requests) internally as well as through other ad networks.

In additional to these components; Ad exchange is another prime agent to facilitate the purchase and selling of inventories in real time from various ad networks. An ad exchange serves as a broker to connect buyers and sellers to exchange information for them, enabling buyers and sellers to negotiate rates and deliver ads to end devices in real time.

Types of online advertising

There are many types of online advertising. For example, display-based advertising, search-based advertising, social media marketing, email advertising, chat advertising, classified advertising, affiliate marketing, and content marketing. Related advertising types to the purpose of this paper, have been described as follow:

• Display-based advertising. Display-based advertising visually transmits its promotional messages using text, logos, animations, videos, photographs or other graphics [3]. Display advertisers also

target users with different characteristics to increase the impact of the ads. Online advertisers (typically through their ad networks) Cookies, which are specific identifiers of particular computers, are often used to determine what advertisement to serve a particular user. Cookies will monitor when a user has left a page without purchasing something, so that the advertiser can later retarget ads to the user.

• Search engine marketing (SEM). Search engine marketing, or SEM, is designed to boost the visibility of a website in search engine results. Search engines have sponsored results as well as organic (natural or non-supported) results based on question from a web searcher. Search engines also use visual indicators to separate the sponsored results from organic results. Search engine marketing includes all of an advertiser's actions to make a website's listing more prominent for topical keywords [3].

• Mobile advertising. Advertisements are delivered through mobile devices such as smartphones, tablets or other smart devices (e.g., smart TVs). Mobile advertisements can take the form of static or rich media display ads, SMS (Short Message Service) or MMS (Multimedia Messaging Service) ads, mobile search ads, advertising within mobile websites, or ads within mobile applications or games. Mobile advertising is growing rapidly.

• Social Media advertising. Several social networks show advertisement driven business models. If we take for granted that a social network manager is a huge database, with vast amounts of qualitative data from its users, using those helps brands to micro-segment their promotional activities. Facebook is segmentation king. Its advertising network, Facebook Ads, enables the target audience of each ad to be delimited according to location, age, sex, languages and even interests and behaviors. Facebook is the social network with more data on its users. Facebook advertisement formats are displayed either on the sidebar on the platform's own right or on the user's own timeline (Web and mobile), as well as on the logout page [4]. Facebook advertisement formats are displayed either on the sidebar on the platform's own right or on the user's own timeline (Web and mobile), as well as on the logout page. And they seek to improve interaction, lead users to a website or exclusive deals, get more pages likes, download apps.

• E-mail advertising. E-mail marketing is an online marketing technique which uses email to submit commercial or advertisement information. This is a communication device for attracting new customers or keeping those already loyal to the brand. E-mail is currently the first Internet service to go along with social media. This volume of traffic includes legitimate e-mails and spam. The word "spam" applies to those messages we do not ask for and we do not want them or from an unknown sender, usually sent via mass mail. Although though spam can be used on other channels and apps, such as SMS on cell phones, the most important medium for this activity is e-mail.

Revenue Models

For online advertising or mobile application with advertisements so they can make money through those advertisements where revenue is typically determined by the amounts of impressions and/or clicks. The following revenue models [5] are generally used:

• Cost per mile (CPM) is that advertisers charge publishers with ad networks per thousand impressions. It is often referred to as the cost per thousand (CPT), because it calculates the cost per thousand views of the ads. This measurement is commonly available for advertisements for Android developers.

• Cost per click (CPC) is that the advertisers charge the users per click. It is used for the amount of times a website visitor or user clicks on a banner in an application. This measurement is also common because it can be carried out in a simple way.

• Cost per action (CPA) is that advertisers charge per specific action such as filling a form, signing up for an offer, completing a survey, or downloading software. This model seems beneficial to

advertisers since they only pay for concrete acts directly related to their advertising. Nevertheless, when it comes to complex behavior, implementation is not simple.

Online Ad Workflow

Online advertising is a very complicated process that starts when an internet user visits a web page and sends the web server an HTTP request. It will cause an Ad impression if there is any Ad banner on the requested web page, handled by Ad network/servers. When a user submits an HTTP request to the publisher web server for access to the site, e.g. a Web page containing one or more banners, the web server must contact its Ad network/server to request advertising information to be inserted into the site page. The Ad network/servers must contact server-side platform in most cases (or in real-time bidding scenarios) to prepare for bidding. Server-side platform is essentially an integrated programmatic technology platform that enables ad publishers to control their inventory of advertising spaces, maximize the selling of their online media estate, etc. The main purpose of the server-side platform is to allow publishers to link their inventory to multiple ad exchanges, demand side platforms and networks at once, so that publishers can optimize their benefit and monitor their inventory's selling price for different advertiser groups. This maximization could be overcome by the use of a data management platform that offers server-side platform data support. The server-side platform is now ready to send an auction to selected Ad exchanges after the requisite planning [6].

The details included in the bid request enables Ad Exchange and demand side platform to understand the meaning of the Ad banners to be offered to the audience. IAB OpenRTB specification [7] defines the actual format of the bid request for programmatic ad buying / selling. When an Ad exchange receives an Ad request from publishers, the request will be distributed as an Ad auction to all demand side platforms connected to the exchange. Demand side platform therefore provides broad access to inventory and vertical and lateral targeting, with the ability to serve advertising, bid on advertising in real time, monitor ads and maximize revenue. A demand-side platform will also rely on data management platforms for this purpose to provide efficient data support. Once the Ad exchange receives bid responses from all demand side platform within the time limit (typically less than 100ms from the bid), the one with the highest bid price will be chosen. The advertiser who wins the bid will transfer the details, such as the advertising URL along with the script code, to the server-side platform and then to the Ad network/server publisher. The web content server publisher and the ad network/server then react to the internet user, providing the client devices with the website and advertisements. The above procedure completes a single Ad transaction and this operation happens in real-time with less than 100ms delay in effect. And consumers are not experiencing frustrating latency.

Mobile Ad Workflow

When advertisers choose to promote a product(s), they contact an Ad network/server authority and supply the material to potential clients. The Ad network/server stores these contents and provides custom APIs to the publishers. If a publisher wants to generate revenue using online ads from their website or application, it signs up with an Ad network/server and accesses its APIs to allow mobile advertising. These APIs usually contain a UI component used to view the advertising and the publisher places the UI component in its GUI. When a user initially opens the website or the publisher's application, the Ad API loads the ad UI component with the content of the advertiser and this event is logged as an impression. Now if the user clicks on the Ad component (a click event) another request will be sent to the Ad network/server to redirect the user to suitable URLs [8]. Ad network/servers keep track of each such event. The advertiser must pay the Ad network/server on the basis of their sales model and the publisher earns a certain percentage of the Ad network/server payment. Among the revenue models, Pay-per-click and Pay-per-impression models are most popular.

3- Ad Fraud Types in Online Advertising

By definition, Ad fraud is synonymous with an activity in which views, clicks, acts or data events are misreported to criminally gain revenue, or for other purposes of deceit or malice. Ad-fraud activities aimed at generating revenue are more common, but noise generation and other non-revenue generating activities are also present today in the internet advertisement ecosystem. There are three main types of ad fraud: placement fraud, traffic fraud, and online advertisement action fraud at different rates, showed in figure 1. Reporting in each of these cases validates a visitor to be genuine, but is simply fraudulent. Such fraudulent tourists can be absolutely robotic, human or a mixture of both. This section includes a thorough analysis of the forms of Ad fraud. In addition, for each type of fraud, we will also review detection methods.

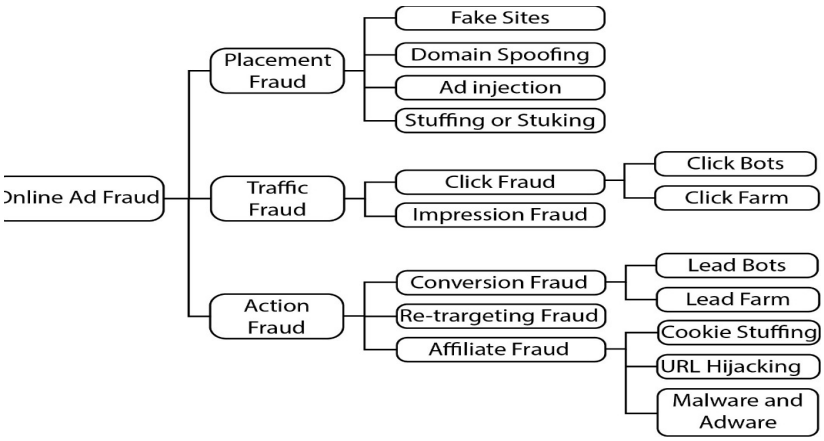

Figure 1: *Online Ad Fraud Types*

Placement Fraud

An Ad placement is always an iframe containing advertisements that contain messages, pictures or videos that are innovative material. Placement fraud is described as fraudulent acts or activities aimed at manipulating or changing the web pages of publishers or changing the web pages displayed on the devices of the user to increase impressions or clicks. Such fraudulent activities involve a variety of acts, ranging from simple keywords stuffing [1], misrepresenting Ad placement position so that a placement is positioned in the invisible frames and is never visible to the public [9] to Malvertising [10] which injects advertisement malware by attracting users to register and then redirecting traffic to malicious sites to generate inflated impressions. According to [6] placement frauds have been categorized into four groups, with each group focusing on one aspect of fraudulent actions, and review solutions to detect each type of placement fraud.

Stuffing or Stacking: Stuffing (whether keyword stuffing or pixel stuffing) is a way to show content that cannot be seen with bare eyes. It is widely used both for keyword stuffing [1] and for placement stuffing. In order to detect stuffing or stacking fraud, Double Verify Inc. [11] suggests many methods of detecting hidden or invisible ads in order to prevent stuffing or stacking fraud. One approach is to compare the ads with graphic images taken from html codes with this web page's snapshot. Any advertisement with the image which is not contained in the snapshot will be labelled invisible using image recognition technology [12]. Another approach is the geometrical analysis. The code snippet contained in the page must measure the position of the advertising, the position of the browser's viewable areas and the size of the open browser window to check whether the advertising is viewable.

Fake Sites: There are two distinct types of fake website fraud. One is to build websites with legal domain names but only contain Ad slots[3]. Then, by entering large Ad networks or Ad exchanges, fraudsters may get significant revenue from fake sites. Another way to deceive tourists is to copy

the content from well-known websites or register a similar domain name. There are two popular ways for detecting fake sites, namely blacklist lookups and recognizing fake sites using machine learning methods. Nearly every browser toolbar contains a blacklist to verify whether the site the user visits is fake or not [13]. Using data mining and machine learning methods to identify fake pages, Abbasi and Chen [14] proposed a classifier system.

Domain Spoofing: Web spoofing [6] is widely known on the Internet, where fraudsters create websites that mimic real websites to conduct fraud activities, for example, stealing identity information or account credentials. Since advertisers are willing to put their advertising on better quality sites and even offer higher rates, fraudsters spoof their domains in order to avoid being placed on the blacklists. Domain spoofing therefore refers to fraudulent activities which attempt to falsify the domain as if the traffic were from publishers in the whitelist. Domain spoofing is normally done in Ad networks via Malware & Tool-bars [15] or Ad Tag Misrepresentation [16].

Ad Injection and Malware: Ad injection and malware are more aggressive fraud activities that explicitly impact client web browsers to either change the ad [17] or view advertising on the current web page that have no Ad placement at all. One form of Ad injection comes from advertisement software, such as malicious adware, a program running on client computers to show unintended ads [18]. Additionally, ad injection can also be done via browser extension software which is often used to increase / enhance web browser functionality [19]. Thomas et al. [17] developed a detection method based on client-side DOM with Google websites to get Ad injection scripts from HTML pages. They search the customer's DOM first to classify suspect elements and fraudulent domains. They then filter scripts that are not affected by ad injection, by excluding usual programs such as browser toolbars and antivirus engines. Finally, they update the scripts manually depending on the quality of the scripts.

Traffic Fraud

The main objective of traffic fraud is to inflate the number of impressions produced by individual sites or placements by manipulating network traffic. However, for CPC-based campaigns, only the users ' click action on the advertising displayed will result in a revenue, so, click fraud [20][21] is often widely seen and is one of the most prevalent fraudulent behaviours.

Impression Fraud: Impression fraud aims to increase the website traffic directly and thereby produce more impressions for auction. This form of fraud has the greatest effect on the CPM-based campaigns, since inflated impressions give advertisers little to no value for their advertising benefits [1]. In addition, it also affects campaigns focused on CPC and CPA, since most impression fraud cannot lead to click or conversion events, and therefore the click-through rate (CTR) will be reduced, because the CTR measurement denominator is the number of page views. In fact, fraud in impression is created through three approaches: hiring human labour to manually view pages, design different types of bots to generate impressions for auctions, and using expired domains to divert users to third-party pages. Since hiring human resources is seen as too costly to produce web traffic, whereas bot also has less intellectual capacity to mimic human trafficking, several hybrid methods seek to increase traffic on websites by incorporating really human behaviour and automated bot functions. For example, publishers gain impressions invisibly in pay-per-view (PPV) networks [22]. If a user views one publisher of this network and clicks anywhere on this page, an invisible frame with other publishers will be activated. The authors [22] implemented three countermeasures, i.e. filtering zero-sized viewports, blocking traffic from PPV networks using referral blacklists, and stopping running advertisement on publishers in blacklist.

Click Fraud: On any advertisement, a click event is a simple indication that a viewer is potentially interested in an advertisement and can thus become a customer. Tap through rate (CTR) is most also used to determine effectiveness at various rates, such as at the placement level, site level, or publisher level etc. Click fraud is probably the most prevalent fraud in the Ad ecosystem, mainly because campaigns focused on CPC dominate Ad networks. Fraudsters use various types of

approaches in a click fraud attack, either manually or with bots, to click on an advertisement. Two parties, publishers and advertisers can root a click fraud, with two motives, respectively, publisher click inflation or advertiser competition [23]. Publisher Click Inflation: Clicking events will carry immediate revenue from the publisher perspective, since publishers are paid on the basis of the percentage of advertising experiences that viewers click. Publishers therefore intuitively accept click fraud attacks, though they do not promote or engage in these activities. To do so, clicks are created either through the use of automated programs or human labours.

Advertiser Competition Clicks: Under the CPC revenue model, a small amount of advertisement budget is consumed per click. So, by making artificial clicks on the advertisements of the competitor, the advertising budget of the competitor may be depleted within a limited period of time. As a result, fraudulent advertiser advertising will have the benefit of targeting legitimate users with a higher chance of user clicking and resulting in a better conversion rate and pleasing the branding company. Some advertisers use a pacing rate control to defend click attacks which specifies daily or hourly advertisement spending cup for smooth budget delivery [24]. It will keep the whole campaign budget from running out in a limited time, whereas clicks of poor quality with no commercial value bring extra pressure to every campaign. Two widely used methods to produce fake clicks are click farms or click bots [25], Where the former is created by human viewers and the latter generated by computer programs.

Action Fraud

Action fraud aims to target relevant business activities of users, such as filling out an online form or survey, placing an online purchase order, or re-targeting important customers through previous acts or behaviours of users [14]. Since advertisers are focused on using cost-per-action (CPA) to determine their promotional costs versus revenue, action fraud affects the ad pricing, campaign preparation and several other major components of the Ad ecosystem directly.

Conversion Fraud: An Ad network conversion refers to one or a series of concrete business actions taken by site visitors that they convert to paid (or future paying) customers. Alternatively, a conversion may also be described as "agreed-upon user action" [26]. For instance, a simple conversion event may be to download a file or fill out a form, or complete an online purchase order. This fraud is also called spam conversion [1]. A conversion usually involves a variety of user activities, and generally a conversion event takes place minutes, hours, or even days after the initial Ad click. It should be noted that a conversion is usually tracked via the placed pixel on the branding sites (or the landing page), while usually a click is tracked on the sites of the publishers. The mapping is also achieved by matching details about the user cookies. A purchase is needed for most clicks via landing pages-based conversions, so users need to provide name, credit card and other important details. As this phase involves sophisticated interactions and financial engagement, there are very few fraudulent activities aimed at this kind of conversion. For conversions based on lead generation landing page users are only required to provide simple details or to take simple actions, such as filling in user name, household address, or downloading a file from the advertisers' site. All these acts can be carried out at minimal or nearly no financial cost. Most conversion fraud thus targets this sort of conversion. Conversion fraud also occurs across two forms of activity [27]: 1) Lead Bots: is a computer agent who fills out lead forms automatically with either randomly generated or partially correct information. 2) Lead Farm: In this situation, the fraudsters are able to employ people to manufacture conversions with lower labour costs from underdeveloped countries.

Re-targeting Fraud: It aims to target valuable customers with precision based on their previous Internet acts, Such as customer purchase history or site surfing history/customer activities [28]. This can be achieved by reviewing past transaction history or monitoring users ' cookies before visiting sites and showing interest in certain products to decide whether or not a user is interested in those products. Commonly, they use techniques like "cookie" or "pixel" as the snippet code. For the re-targeting of fraud, the fraudsters ' main goal is to mimic the particular activities of legitimate

consumers and make them behave like attractive users. This is typically done through the use of computer-generated agents, such as DeceptiBots [29], to mimic the thoughts and actions of a human being and believe that they are interested in a specific product or brand. As a consequence, the bots trick advertisers into believing bots are valuable potential customers, thus, put a higher price on the bots produced auctions/impressions.

Affiliate Fraud: Affiliate marketing is a kind of performance-based approach to marketing, where an associate (i.e. a business entity) collects the benefits from its marketing activities for each visitor or customer. Affiliates use various types of promotional strategies in affiliate marketing, including search engine optimization (SEO), e-mail marketing, or display advertising to attract visitors. Affiliate fraud refers to activities that mislead the framework for reporting commission / revenue not authorized by the affiliate. In most affiliate marketing, commission is only charged if a customer makes a transaction, so affiliate can claim commission only after a conversion has taken place. As a result, affiliate fraud specifically targets consumers who are already on the verge of making purchases [30]. An affiliate fraud [31] is widespread across the following three forms of approaches: 1) Malware and Adware, 2) Cookie Stuffing and 3) URL Hijacking.

4- Ad Fraud Types in Mobile Advertising

Although the literature includes a wide range of ad fraud research in web applications, relatively little attention has been given to such mobile fraud. A recent study [32] proposed a new taxonomy of Mobile fraud as showed in figure 2 that sums up nine different types of ad fraud, which is by far the largest number of types of ad fraud.

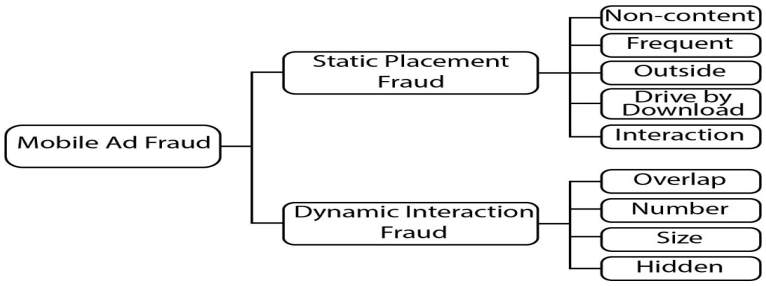

Figure 2: *Mobile Ad Fraud Types*

Static Placement frauds

Many cases of fraud are achieved simply by manipulating the shape and location of the ad view in a UI state. "Static" means that the identification of these frauds can be calculated by static data and takes place in a single UI state. "Placement" means fraudulent conduct exploits aspects of placement, e.g., scale, position, and number of ad views, etc. Four specific types of behaviour related to static placement frauds have been identified:

The Ad Hidden fraud: App developers can hide advertisements (e.g. under buttons) to give users the impression of an "ad-free app" that would hopefully improve user experience. However, these ads are not shown in compliance with the contract with advertisers paying for the advertising role [33][34].

The Ad Size fraud: While ad size advice given by ad networks is not compulsory, and there are no ad size requirements, the size ratio of the ad to the screen must be fair [33], allowing users to display advertisements normally[35]. By extending ad size to the limits, fraudulent activity can be implemented: with incredibly small ad sizes, app developers can have the feeling of an ad-free app,

but it can cheat advertisers; similarly, with an abnormally large ad size, there is a greater risk of attracting or pressuring users to click on the ad in an effort to close it.

The Ad Number fraud: Since advertising must be seen by users as pure extras in addition to the main content of the app, the number of ads must remain fair [34][36]. Unfortunately, developers also provide a large number of advertisements to increase the chance that user interests would be drawn, even if the user experience of the app is deteriorating and even seriously affects regular functionality when ad content exceeds valid product content.

The Ad Overlap fraud: To forces users to trigger unwanted impressions and clicks, App developers can simply show ad views on top of app-relevant views [33][34][36]. By positioning ads in places covering consumer areas of interest in standard experiences with apps, App developers create irritating circumstances in which users will' recognize' the ad.

Dynamic Interaction Frauds

Authors on [32] identified cases of fraud found that go beyond putting ad views on a single UI state, but it requires runtime behaviour, which can occur in an unpredictable scenario of application use. "Dynamic" means that such frauds are identified at runtime. "Interaction" means the fraudulent activity exploits scenarios of user interaction and may include several UI states.

The Interaction Ad fraud: Developers use interstitials in web programming (i.e. web pages shown before or after the intended content of a page) to view advertisements. Translated into mobile programming, when switching between UI states, some ad views are put. Nonetheless, manipulation can be achieved by putting interstitial ads on app load early or by exiting apps that could trick users into unintended clicks, as contact with the app / device is highly likely at this point [33][34][37].

The Drive-by downloads Ad fraud: Ads are intended to provide short advertising content created by marketers to attract the attention of users when they visit an external web page. When app developers are remunerated not by the number of clicks but by the number of users who ultimately become actual customers of the product / service advertised, there is a temptation of fraud. A common example of fraud includes causing unintended downloads (e.g. of advertised APKs) when you click on the ad view [34][37]. This activity also has a significant effect on user experience, and drive-by-downloads are not even readily cancelled in most cases.

The Outside Ad fraud: Advertisements are expected to appear on pages when users use the app. There are, however, illegal activities for ads while applications are running in the background or even outside of the user environment (e.g., Placed ad views on the home screen and covering device icons that users need to hit to launch new apps)[33][34][37][38]. In certain extreme situations, the advertisements appear spuriously and the user must identify them, because these advertisements can only be removed when the user recognizes and launches the application from which they originate.

The Frequent Ad fraud: App developers aim to increase the probability of ad impressions and clicks to gain more revenue. The number of UI states in the device limits this probability. Therefore, developers may enforce deceptive techniques by showing interstitial ads each time the user clicks on the core content of the application (e.g., even when the click is to show a menu in the same page)[33][37].

The Non-content Ad fraud: To boost ad impressions and trick users into accidental clicks, app developers may place advertisements on non-content-based pages such as thank you, error, login, or exit screens. Ads on these types of UI states may confuse a user in the assumption that the ads are actual app content [33].

A lot of research was proposed to classify ad frauds on the web. Such approaches can provide valuable tips for mobile community researchers and practitioners to devise promising methods to

detecting mobile ad fraud. Current mobile ad fraud studies have attempted to classify ad-fraud technologies where the fraudulent activities can be statically identified (the so-called static placement frauds). For examples, Pearce et al. [39] found that 49 percent of Android applications contain at least one advertisement library, and 46 percent of ad-supported applications are affected by these libraries. Shekhar et al. [40] suggested an approach called AdSplit to split applications from its advertisement libraries that could request sensitive privilege permissions. In addition, Liu et al. [41] examined Windows Phone's static placement abuse by examining app layouts. Furthermore, Crussell et al. [42] developed an automated method for the detection of click frauds. Shekhar et al. and Crussell et al. methodology is basically applied in three stages: (1) building HTTP request trees, (2) identifying ad request pages using machine learning, and (3) detecting clicks in HTTP request trees use heuristic rules. Unfortunately, the above-mentioned strategies are unable to classify the new fraudulent behaviors with the evolution of ad fraud, for example, they cannot be used to detect fraud involving dynamic interactions.

5- Ad Fraud Prevention Mechanism

To counter Ad fraud, current approaches typically depend on the following four types of mechanisms [16]:

• **Signature-based Prevention Mechanism**: This type of approach uses predefined features/patterns to identify malicious traffic or impression [43]. For example, research found that when execution of a client-side code is incompatible with established code execution models (such as JavaScript); the traffic is very likely not to be created by real human users, but by a bot [44]. Therefore, checking the execution of code environment, such as support for JavaScript or mouse event test [45], a large portion of fraudulent traffic can be filtered out. These behaviour analysis approaches for Clickbots have been studied extensively [46].

• **Anomaly-based Prevention Mechanism**: This approach uses statistical analysis and historical data to identify suspicious pages, websites or publishers whose traffic is deemed irregular compared to the general traffic of users. For example, as of April 2016, the average probability of click events in advertising display is about 0.17 percent, indicating that there are about 1.7 click events on every 1000 impressions on average [47]. A placement or publisher website that shows substantially higher clicking through rates would be considered anomaly and includes fraudulent activities that deserve further investigation [48].

• **Honeypot-based Prevention Mechanism**: To pinpoint fraudulent activities, Ad servers (such as advertisers) may purposely serve a range of carefully specified bluff ads to publishers where it is understood that bluff / honeypot ads are unidentifiable through individual users, and if bluff advertisements result in events, such as a click event, it will contradict the presumption and therefore involve fraud [49]. Traffic traffickers used such a honeypot strategy to analyse traffic for better Ad service [50].

• **Credential-based Prevention Mechanism**: The credibility of website publishers is strongly associated with possible fraud activities. To determine publisher's credentials, demand side platforms or advertisers may use reverse crawling to find the content of the web pages and test if their content is compatible with the impression-related tags when submitting an auction. However, one can also use the number of impressions produced by a publisher, and compare this value to trustworthy website rankings like Alexa or RageRank. A publisher with a much greater impression than its traffic rating would obviously suggest possible fraudulent activities.

6- Conclusion

Online advertising is now a popular form of business marketing and one of the reasons why free web content or mobile apps are available. The emergence of smart televisions and online content distribution services is expanding rapidly. Consequently, fraudsters are also exploiting the market to drain money from advertisers. This paper provides a detailed description of the current state-

of-the-art online/Mobile advertising fraud. However, more research needs to be addressed to all types of occurring ad fraud. As examples, in the context of online marketing campaigns, hacking ad-campaign accounts need to be studied. Additionally, to explore individual fraudulent displays in mobile advertising as well as dynamic interaction frauds that has been pointed out in this study.

References

[1]. N. Daswani, C. Mysen, V. Rao, S. Weis, K. Gharachorloo, and S. Ghosemajumder, "online advertising fraud crimeware: understanding new attacks and defenses," 2008, vol. 40, no. 20, pp. 1–28.
[2]. R. Oentaryo *et al.*, "Detecting Click Fraud in Online Advertising: A Data Mining Approach," *J. Mach. Learn. Res.*, vol. 15, no. 3, pp. 99–140, 2014, [Online]. Available: http://jmlr.org/papers/v15/oentaryo14a.html.
[3]. D. Chaffey and F. Ellis-Chadwick, *Digital Marketing: Strategy, Implementation and Practice*, 1st ed. Pearson Education, 2012.
[4]. T. Pineiro-Otero and X. Martinez Rolan, *understanding digital marketing basics and actions*. Springer-Verlag GmbH, 2016.
[5]. N. Daswani, C. Mysen, V. Rao, S. Weis, K. Gharachorloo, and S. Ghosemajumder, "online advertising fraud," 2007.
[6]. X. Zhu, H. Tao, Z. Wu, J. Cao, K. Kalish, and J. Kayne, *Fraud Prevention in Online Digital Advertising*, 1st ed. Springer Publishing Company, Incorporated, 2017.
[7]. B. IA, "OpenRTB API Specification Version 2.5." Dec. 2016.
[8]. G. Cho, J. Cho, Y. Song, D. Choi, and H. Kim, "Combating online fraud attacks in mobile-based advertising," *Eurasip J. Inf. Secur.*, vol. 2016, no. 1, pp. 1–9, Dec. 2016, doi: 10.1186/s13635-015-0027-7.
[9]. B. Edelman, "Accountable? The Problems and Solutions of Online Ad Optimization," *IEEE Secur. Priv.*, vol. 12, no. 6, pp. 102–107, Nov. 2014, doi: 10.1109/MSP.2014.107.
[10]. A. K. Sood and R. J. Enbody, "Malvertising exploiting web advertising," *Comput. Fraud &Security*, vol. 2011, no. 4, pp. 11–16, 2011.
[11]. M. McLaughlin, R. K. Rosenfeld, L. M. Abu, and L. Simon, "system and method for identifying hidden content." Google Patents, 2015.
[12]. Y. Zheng, B. Jeon, D. Xu, Q. M. Wu, and H. Zhang, "Image segmentation by generalized hierarchical fuzzy C-means algorithm," *J. Intell. Fuzzy Syst.*, vol. 28, no. 2, pp. 961–973, 2015.
[13]. Y. Zhang, S. Egelman, L. Cranor, and J. Hong, "Phinding phish: Evaluating anti-phishing tools," 2007.
[14]. A. Abbasi and H. Chen, "A comparison of tools for detecting fake websites," *Computer (Long. Beach. Calif).*, vol. 42, no. 10, pp. 78–86, 2009.
[15]. innovation in magazine media 2015-2016, "advertising: digital advertising fraud." .
[16]. B. Stone-Gross, R. Stevens, A. Zarras, R. Kemmerer, C. Kruegel, and G. Vigna, "Understanding fraudulent activities in online ad exchanges," in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, 2011, pp. 279–294.
[17]. K. Thomas *et al.*, "Ad injection at scale: Assessing deceptive advertisement modifications," in *2015 IEEE Symposium on Security and Privacy*, 2015, pp. 151–167.
[18]. S. Emily, "New adware apps bug facebook, google." 2011.
[19]. X. Xing *et al.*, "Understanding malvertising through ad-injecting browser extensions," in *Proceedings of the 24th international conference on world wide web*, 2015, pp. 1286–1295.
[20]. K. C. Wilbur and Y. Zhu, "Click fraud," *Mark. Sci.*, vol. 28, no. 2, pp. 293–308, 2009.
[21]. Q. Zhang, T. Ristenpart, S. Savage, and G. M. Voelker, "Got traffic? An evaluation of click traffic providers," in *Proceedings of the 2011 Joint WICOW/AIRWeb Workshop on Web Quality*, 2011, pp. 19–26.
[22]. K. Springborn and P. Barford, "Impression fraud in on-line advertising via pay-per-view networks," in *Presented as part of the 22nd ${$USENIX$}$ Security Symposium (${$USENIX$}$ Security 13)*, 2013, pp. 211–226.
[23]. N. Vratonjic, M. H. Manshaei, and J.-P. Hubaux, "advertising fraud," 2011.
[24]. K.-C. Lee, A. Jalali, and A. Dasdan, "Real time bid optimization with smooth budget delivery in online advertising," in *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising*, 2013, pp. 1–9.

[25]. P. Pearce *et al.*, "Characterizing large-scale click fraud in zeroaccess," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014, pp. 141–152.
[26]. B. Mungamuru and S. Weis, "Competition and fraud in online advertising markets," in *International Conference on Financial Cryptography and Data Security*, 2008, pp. 187–191.
[27]. Performics, "digital advertising fraud and abuse: strategies and recommendations for mitigation." 2014.
[28]. Z. Li, K. Zhang, Y. Xie, F. Yu, and X. Wang, "Knowing your enemy: understanding and detecting malicious web advertising," in *Proceedings of the 2012 ACM conference on Computer and communications security*, 2012, pp. 674–686.
[29]. P. Nowak, "Deceptibots: when machines go bad," *New Sci.*, vol. 214, no. 2870, pp. 45–47, 2012.
[30]. B. Edelman and W. Brandi, "Risk, information, and incentives in online affiliate marketing," *J. Mark. Res.*, vol. 52, no. 1, pp. 1–12, 2015.
[31]. P. Snyder and C. Kanich, "No Please, After You: Detecting Fraud in Affiliate Marketing Networks.," 2015.
[32]. F. Dong *et al.*, "Frauddroid: Automated ad fraud detection for android apps," in *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 2018, pp. 257–268.
[33]. Google, "google admob & adsense policies." 2017.
[34]. Wandoujia, "wandoujia (ali app) developer policy." 2018.
[35]. Firebase, "banner ads." 2017.
[36]. DoubleClick, "doubleclick ad exchange program policies." 2017.
[37]. C. C. S. Association, "mobile intelligent terminal malicious push information to determine the technical requirements." 2017.
[38]. H. Market, "huawei market app developer policy." 2018.
[39]. P. Pearce, A. P. Felt, G. Nunez, and D. Wagner, "addroid: privilege separation for applications and advertisers in android," in *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*, 2012, pp. 71–72.
[40]. S. Shekhar, M. Dietz, and D. S. Wallach, *Separating Smartphone advertising from applications*. Rice University, 2012.
[41]. B. Liu, S. Nath, R. Govindan, and J. Liu, "${$DECAF$}$: Detecting and Characterizing Ad Fraud in Mobile Apps," in *11th ${$USENIX$}$ Symposium on Networked Systems Design and Implementation (${$NSDI$}$ 14)*, 2014, pp. 57–70.
[42]. J. Crussell, R. Stevens, and H. Chen, "Madfraud: Investigating ad fraud in android applications," in *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*, 2014, pp. 123–134.
[43]. V. Dave, S. Guha, and Y. Zhang, "Measuring and fingerprinting click-spam in ad networks," in *Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication*, 2012, pp. 175–186.
[44]. S. K. Adrian Neal, "quantifying online advertising fraud: ad-click bots vs humans," 2015.
[45]. H. Xu, D. Liu, A. Koehl, H. Wang, and A. Stavrou, "Click fraud detection on the advertiser side," in *European Symposium on Research in Computer Security*, 2014, pp. 419–438.
[46]. B. Miller, P. Pearce, C. Grier, C. Kreibich, and V. Paxson, "What's clicking what? techniques and innovations of today's clickbots," in *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, 2011, pp. 164–183.
[47]. C. Dave, "us, europe and worldwide display ad clickthrough rates statistics summary." Apr. 2016.
[48]. F. Yu, Y. Xie, and Q. Ke, "SBotMiner: Large Scale Search Bot Detection," in *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, 2010, p. 421430, doi: 10.1145/1718487.1718540.
[49]. H. Haddadi, "Fighting Online Click-Fraud Using Bluff Ads," *SIGCOMM Comput. Commun. Rev.*, vol. 40, no. 2, pp. 21–25, Apr. 2010, doi: 10.1145/1764873.1764877.
[50]. blog TT, "using honeypot banners to detect click fraud." 2015.