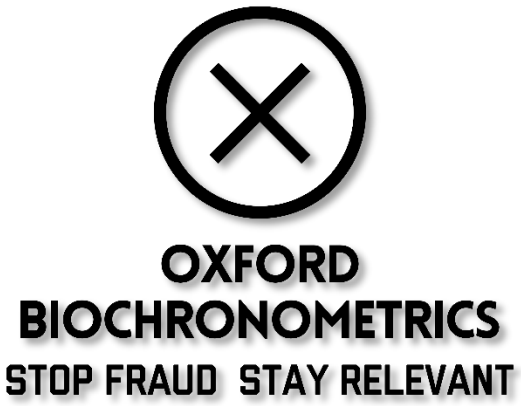

Why Does Ad Fraud Keep Growing?

A Look at the Current and Future Risks to Digital Marketing Budgets

Oxford BioChronometrics Whitepaper Series
September 2021

Executive Summary

In order to determine the exposure to fraud related risks of our clients, Oxford BioChronometrics has analyzed tens of millions of clicks and leads generated by our North American clients in the financial services and insurance industries.

Key findings

- Advertisers carry all the risk in case of fraud
- The total burden of fraud is smeared out over all advertisers in the ecosystem
- Networks use lagging detection technologies that make them and their clients vulnerable to the latest types of fraud
- Networks generally do not refund or give rebates for the fraud that they themselves do not (yet) acknowledge
- Lack of transparency dominates the process.

Recommendations

Marketers should not trust - or be fully dependent upon - the detection methods of large networks. They should measure the quality of traffic with a continuously updated solution that resides on their landing pages that they control. That way, they can set up agreements to get rebates or additional traffic when fraud is detected. In the case of lead generation, they should make agreements to return fraudulent leads in real-time.

When marketers mitigate the risk exposure to fraud, they eliminate direct budget waste, ensure clean campaign data, and increase the overall efficiency of their marketing spend.

How do you run your marketing campaign in the most optimal way? That is one of the most challenging questions any marketer has to face. The digital ecosystem with all its tools and possibilities should have brought tools for easier use, a more accurate audience reach, and full measurability of attribution and viewability. Instead, it brought a technical nightmare!

The promised land of marketing sold by the large publishing networks is that your marketing campaign reaches your desired audience, has accurately measured viewability, is only placed next to brand safe content, and is shown on Alexa top 500 websites. In practice, though, the performance of your marketing campaign is severely affected by all kinds of factors like price volatility on the search keywords and groups, all default configuration settings are in the network's interest, your ads shown on fake news websites pretending to be a premium website, traffic pretending to be in your targeted geo-location, pretending to fit your audience group, pretending to be a human by emulating human behavior, and finally clicking on your advertisements in order to claim the attribution.

Without having the real overview - and without knowing what really happens in the ecosystem - marketers start to optimize their campaigns by changing keywords, adding negative keywords, defining less granular geo-location areas, and subsequently start to A/B test their advertisements, and A/B testing their landing pages hoping to improve their conversion rate. Unfortunately, redefining and/or optimizing a marketing campaign by tweaking the filtering of potential visitors by using inclusion and exclusion of keywords, geo-locations and other demographics does not give you any control over the quality of traffic. Instead, it gives you less traffic for a higher price. But this does not guarantee better quality of the traffic.

The goal of marketing is to attract, engage and convert new customers into leads. This is achieved by running online campaigns, podcasts, newsletters, webinars, etc. targeted to the intended audience. This works well if everybody is honest and plays by the rules. When you think you target a certain audience you expect humans. But, besides humans you also attract invalid traffic generated by botnets and their bots.

Detecting known and unknown fraud is currently achieved by looking at the larger picture. The reason is that a single advertisement or click does not generate enough data-points to determine fraud. This leaves no alternative but sifting through tons of data looking for outliers, non-human characteristics and impossible browsing scenarios for humans without having too many false positives and false negatives. The effect of this methodology is that it is lagging or in other words: just too late. This is what we call 'the detection gap', i.e., the time between the operational start of a botnet and the time of the detection (and takedown) of that same botnet. This gap represents the risk to networks and publishers unknowingly selling invalid traffic as human traffic, resulting in claims, rebates and or refunds. Using current technology, it seems almost impossible to be able to detect an isolated fraudulent browsing session. Even if this browsing session is made using known fraud techniques, let alone unknown fraud. The detection gap unknowingly undermines marketing campaigns due to false data and overcharging for unwanted advertisements.

Large networks have increased the confusion level by playing both sides, over emphasizing the botnets and the bots when the need suits them. Many press releases and subsequent news articles proclaim the findings of new botnets and subsequent rebates to clients. These articles prove the gap in the detection, but these detection gaps are only acknowledged if large clients complain loudly enough and external research cannot be denied and or dismissed. Small and medium sized companies don't have the resources nor the means to determine whether bad performing campaigns were affected by botnets. They are completely dependent on the large networks.

Adtech focuses on botnets and specific types of bots. Crawlers, spiders, scrapers, humanoids, etc. and arguments are made over which are good and which are bad. This overcomplicated classification method confuses marketers and moves the discussion away from what is important. As a marketer you do not need to know which botnet loaded and viewed half of your advertisements, which bots clicked, and which bots filled out your lead generation form. As a marketer, you *only* want to know is this visitor a potential client and likely to purchase my product. If this visitor is a non-human, regardless whether a good bot, bad bot, or click bot, you *don't want to pay for it* and *you don't want the associated data polluting your campaign data and statistics*.

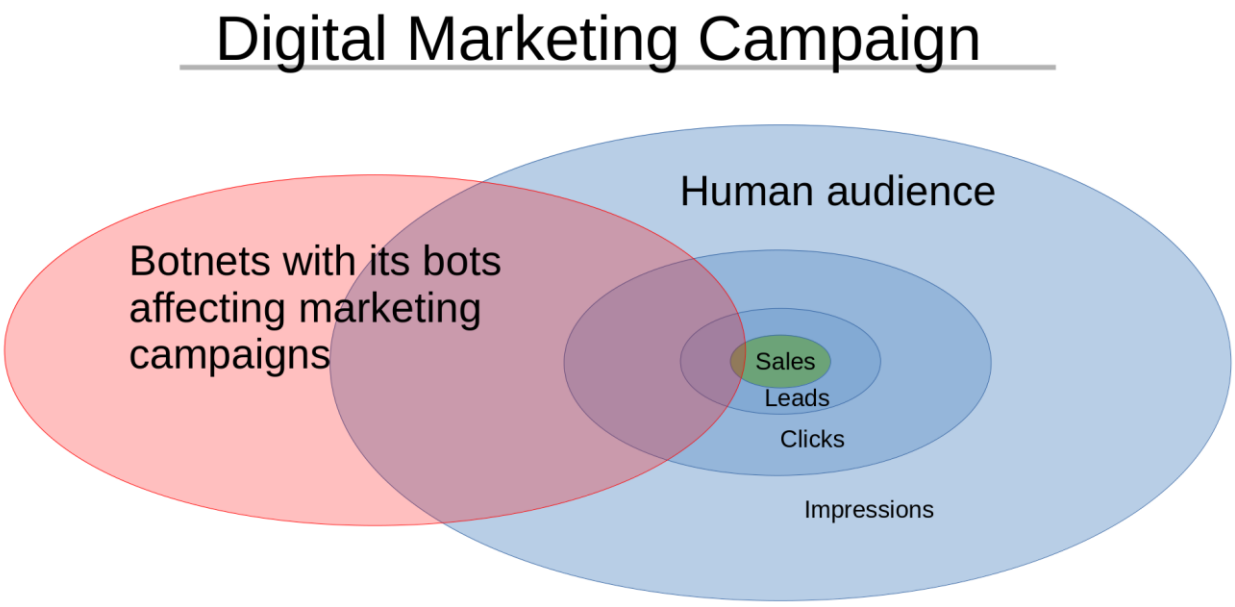

Illustration 1: Botnets will optimize towards those campaigns benefitting their owners. These bots and click-farms are able to view advertisements, generate clicks, fill out lead generation forms and even purchase items.

So, reality has taught us the hard way that, alongside genuinely interested humans, also malicious actors running botnets are attracted to your online marketing campaigns. When such a botnet is discovered only the largest ones get media attention, like methbot[1], 3ve[2] and Xindi[3] and only then when acknowledged you are eligible for a rebate, refund or credit traffic, but how many smaller botnets are silently dissolved, without you and the general public knowing of their existence? The only thing marketing teams know is: Our campaign(s) did not perform as expected, which is often explained as: 'the campaign did not fit the audience', 'after a while ad fatigue occurs', 'the creative was not good enough', 'we need to re-optimize our campaign'. Only a few companies will start to investigate the traffic with 3rd part software such as Oxford BioChronometrics' SecureAd and/or SecureLead. Then and only then, they realize that the performance of their campaigns is severely affected by botnets, its bots and click farms.

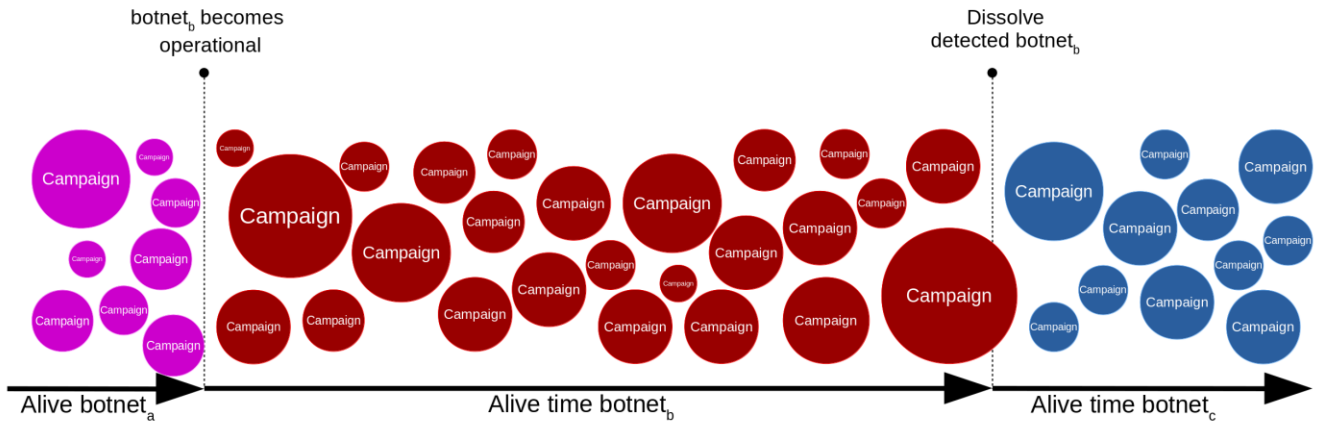

Illustration 2: Without detection a botnet keeps alive and active and multiple campaigns of multiple clients are affected. The duration a botnet is active before being detected is called: 'the detection gap'.

In the past colossal botnets have been created by impersonating established sites, running specific crafted custom browsers and appearing human by generating fake mouse movements, cookie stuffing, and social network logins [1]. When botnets have had enough time to scale in order to scrape the budget from a multitude of marketing campaigns, their owners have earned way more money than the original investment. This means, to the botnet owners, their botnet operation has been a success!

So, how serious of a problem is this detection gap for companies? Since 2021Q1 Oxford BioChronometrics issues a quarterly affiliate benchmark report [4]. This benchmark is based on paid traffic and shows the quality of paid traffic: clicks and the conversion to leads. When looking at the networks generating clicks and removing the top 5 performers and the bottom 5 performers, the 80% middle group remains. The quality of clicks of this middle groups has an invalid traffic rate between 12% and 25%. This means that crooks owning botnets and click-farms are scraping between 12% and 25% of your marketing campaign's budget. Knowing that for an average company the marketing budget is between 6% and 12% of the total company's budget [5], of which online advertising costs are taking a significant part (58% in 2020) [6], fraudsters running botnets are seriously affecting the performance of companies running digital marketing campaigns.

Where does all that fraudulent money go? In May 2020 the ISBA programmatic supply chain transparency study [6] carried out by PwC shows that in the online advertising market publishers

receive half of what advertising spend, with the other half siphoned off by ad-supply intermediaries. In search no intermediaries exist as the complete infrastructure is owned and controlled by a single party. The reasons why fraud is persistent and keeps on evolving is because it is very profitable, and the bad actors are treating it as a regular business case and re-investing the profits in things like fake news sites which show even more advertisements, and buying and or developing the next generation of botnet software with even more advanced bots. Fraudsters even start businesses as part of the ecosystem which do partly legitimate business but also ingest fake traffic into the ecosystem for a premium price. This is possible due to the complexity of the ecosystem and lack of transparency and accountability.

What is the cause of this problem? It lies in the principal-agent problem [7] also known as the agency dilemma. This situation occurs when one company (the "agent") is able to make decisions and/or take actions on behalf of the other company (the "principal"). This dilemma exists in circumstances where agents are motivated to act in their own best interests, which are contrary to those of their principals. For example, all default and recommended configuration settings of provided tools are in the network's (agent's) interest. In case of a dispute, you have to provide full detailed spelled out evidence, where the agent only provides aggregated results. If an agent unknowingly shows advertisements to bots while charging a normal rate, it is impossible for the companies (the principals) to know to whom advertisements have been shown, and who clicked on an advertisement, because the shown results are aggregated.

In search, businesses can show their advertisements when certain keywords or keyword groups are used as search term by the user. If the user clicks on a shown advertisement the business is charged per click. This means that during a detection gap sophisticated fraud is able to search for specific terms and click on this advertisement, while being undetected. It costs businesses real money for each fraudulent click and brings them nothing. That hurts! Even if they show that their campaign's performance goes to zero, claims are dismissed or simply rejected.

The apparent inability of detecting and acknowledging these fraudulent activities, even if they are active for more than a year, keeps on surprising us at Oxford BioChronometrics. How is it possible that large botnets don't show up on any detection radar? How can it be sold that botnets like this can be

active for several years without any team noticing any deviations in their campaign performance data, or anyone else in the ecosystem.

The answer lies in the following points:

- Only a very small subset of the impressions converts. So, simple fraud (view and click traffic) is able to hide in the vast majority which does not convert
- Marketing teams trust that the networks have state of the art fraud detection in place
- The techniques networks are using to detect invalid traffic are based on aggregated statistics and are lagging and or incapable to detect the most sophisticated fraud
- The vast size of the digital advertising ecosystem. It's relatively easy to hide a few hundreds of millions of impressions in an ecosystem with multi-trillion impressions
- The international character of the Internet places ad fraud outside the influence of national courts and slows down prosecution
- Botnets and fraud continuously improve and have to be one step ahead of detection. Once detected their business is gone, the owners will learn from their mistakes, improve and repeat
- The online advertisement ecosystem is unregulated. Proof is difficult to obtain after the fact, you have to continuous monitor in order to catch ad-fraud

This asks for a different approach. An approach enabling the detection of fraudulent traffic in real-time, without vast amounts of data, without models fine-tuned to the most optimal ratio of false negatives and positives. A real-time fraud detection able to detect fraud on the spot, per individual page view, within seconds and if necessary, confirmed by a human supervisor within a few hours. This helps to act immediately and close the gap without losing precious time. Only then this cycle can be broken by detecting new fraud types before they commence. Eventually, it becomes economic unfeasible to deploy a botnet.

Your mobile and or desktop machine needs to update your virus scanner daily in order to be protected using the most recent virus signatures. If your device is subject to a detection gap of several days, let alone several months, your device would be prone to all kinds of attacks, viruses and malware. A similar situation arises in the digital advertising ecospace when detection is too late. This

is why the detection of botnets and its bots should be based on signatures and micro-signatures instead of the current statistical outlier methodology which is lagging.

Oxford BioChronometrics' detection has proven itself to detect both known and unknown fraud on the spot and flag these as fraudulent. Only, if an unknown fraud outlier is found, it needs to be verified by a human supervisor before its micro-signature is added to the regular fraud detection analysis.

To recap, the Oxford BioChronometrics fraud detection enables you to:

- Successfully detect the most sophisticated types of known and unknown fraud
- Detect fraud on the spot, real-time
- No false positives, no false negatives

Oxford BioChronometrics' clients use our real-time detection to validate traffic coming at their landing pages. Fraudulent flagged clicks and or generated leads are excluded from entering the marketing/sales pipeline. Instead, leads are returned, without having to pay for them, fraudulent clicks are disputed and, in most cases, resolved in credit traffic. This keeps their agencies and affiliates sharp and their online sales efficient. This all wouldn't be possible without on the spot real-time detection.

Having implemented a solution with real-time feedback enables you to be in control over the traffic that loads, views and clicks on advertisements and fills out lead generation forms. It will enable you to immediately spot and exclude IVT and SIVT preventing your campaigns being affected with fake traffic, fake clicks and fake leads. It will reduce the detection gap to zero and therefore mitigates any risk, reputation, and financial rebates associated to malicious botnets in the ecospace. To top it off, it prevents any money going to fraudsters which use it for illegitimate purposes.

Third-party References

[1] https://en.wikipedia.org/wiki/Methbot
https://www.whiteops.com/hubfs/Resources/WO_Methbot_Operation_WP.pdf

[2] https://en.wikipedia.org/wiki/3ve

[3] http://www.nbcnews.com/business/business-news/ad-fraud-zombie-army-has-penetrated-fortune-500-companies-report-n465806

[4] https://oxford-biochron.com/affiliate-performance-report-2021/

[5] https://cmosurvey.org/results/

[6] https://www.isba.org.uk/media/2424/executive-summary-programmatic-supply-chain-transparency-study.pdf

[7] https://en.wikipedia.org/wiki/Principal%E2%80%93agent_problem