Adspect Documentation

Adspect

Jul 09, 2020

Contents:

CHAPTER 1

Overview

Adspect is an easy-to-use cloud-based service for protecting affiliate campaigns (CPA offers, landing pages) from "bad" traffic. By bad traffic we mean:

- click fraud, ubiquitous in display ads and popunder;
- moderators and policy teams of ad networks;
- spy services used by competitors to steal your creatives and landing pages;
- content scrapers;
- credential stuffing bots;
- bots of antivirus companies;
- and other flavors of unwanted or outright hostile visitors.

We work with all traffic sources, both existing and those that will appear in future–our filtering algorithms are perfectly universal and equally efficient across all possible origins of traffic. We support all the largest advertising networks, including:

- Google Ads
- Microsoft Advertising (Bing Ads)
- Facebook
- Instagram
- VK
- Yandex.Direct
- myTarget
- ZeroPark
- ExoClick
- Taboola
- MGID

- PropellerAds
- TrafficStars
- and hundreds of others

We protect your landing pages and offers from various antivirus, security, and ad scoring companies, including:

- GeoEdge
- Adscore
- Google Safe Browsing
- Kaspersky Labs
- Avast
- Forcepoint
- and many others

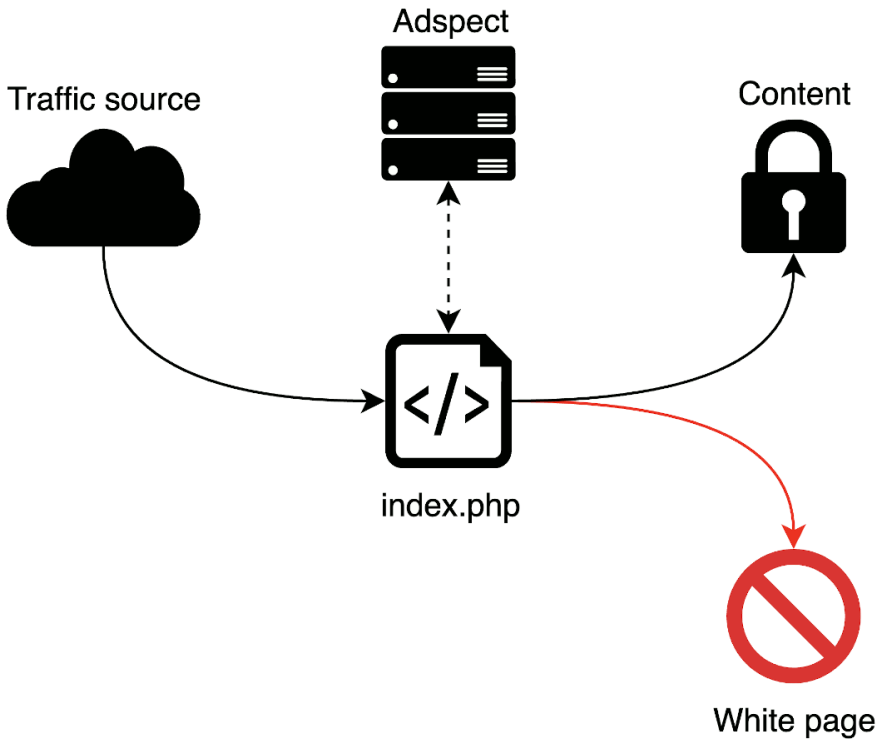
You may find additional information in our FAQ.

We support several types of integration that differ in technical details but all provide equally high levels of protection:

- PHP integration via standalone `index.php` file;
- JavaScript integration via `<script>` HTML tag:
  - Passive mode without cloaking, like Google Analytics–perfect for collecting bot statistics;
  - Cloaking via JavaScript redirect to content page using the `location.replace()` method;
  - Cloaking via iframe overlay without redirecting.

1.1 PHP Integration

In PHP integration filtering is done by a special `index.php` file that you place in your landing page directory or elsewhere accessible via HTTP. This file acts as an entry point for web traffic and is wired to our servers that do the actual filtering. Depending on filtering decision a visitor may be directed to your actual page or to a "white page", that is, a page that contains no sensitive content. In other words, Adspect acts as an intermediary stage in your traffic flow, actively filtering unwanted traffic from legitimate visitors.

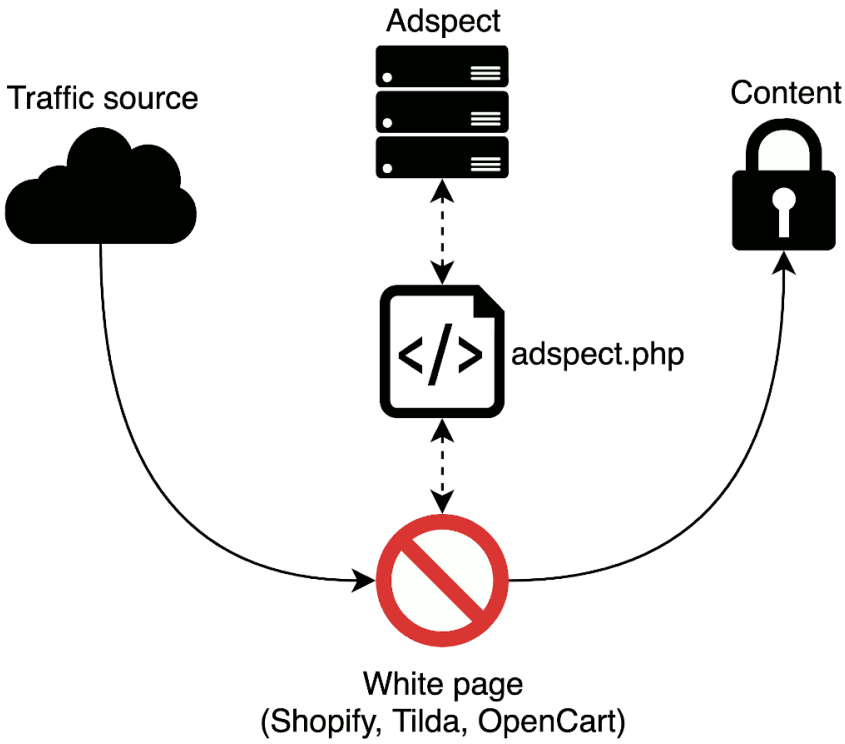


Traffic flow chart

Several copies of the same `index.php` file may be used for protecting several offers or landing pages without interfering with each other.

1.2 JavaScript integration

JavaScript integration is meant to be used with third party services like Shopify, Tilda, or OpenCart, where you cannot upload our `index.php` file for PHP integration. It may also come handy if you want to direct visitors to your white page first and keep them there if they are flagged by Adspect, for extra protection and authenticity, which is especially desirable when working with Facebook and Google Ads in particular.



Traffic flow chart

You will also need to download and host our PHP file `adspect.php` anywhere, but its final location does not matter as it will be linked into the white page via `<script>` HTML tag. When a visitor comes to the white page, the `<script>` tag accesses the remote `adspect.php` file which produces JavaScript code that will do the job. What happens next depends on the mode of operation that you choose during integration:

- In passive mode our statistics will be updated, but no further action will be taken–the visitor will remain on the page. This mode is like Google Analytics, meant for collecting passive insights and blacklists of bot-ridden sources in cases that do not require cloaking.
- In JavaScript redirect mode, legitimate visitors as determined by our filters will be directed to the content page via JavaScript redirect using the `location.replace()` method, i.e. **the URL in the address bar will change**.
- In iframe overlay mode, legitimate visitors will be shown the content page via an iframe overlay without redirecting them anywhere, i.e. the content iframe will be placed over the white page.

1.3 index.php and adspect.php

`index.php` is a PHP script that serves the purpose of a bridge between your premises and our backend servers. The file name `index.php` is just a convention that we use throughout the system, however, you may rename it as you like. The fact that we use a PHP script to filter traffic naturally implies that you need a PHP-enabled web hosting or a tracker with support for landing pages written in PHP. For this reason, Keitaro TDS in particular is not suited for hosting `index.php` as it lacks PHP support for landing pages.

The script is carefully written to be compatible with a wide variety of web hosting environments, ranging from virtual hosting and VPS to dedicated servers and Amazon AWS. Both Windows and Unix-like operating systems are supported, to the extent supported by PHP. PHP 7 is recommended, PHP 5 is also supported.

The only requirement is that PHP has to be built with cURL support. You may check if cURL is supported by examining phpinfo, but cURL is supported by almost every PHP build out there.

The `adspect.php` file is just a different version of the `index.php` file, so everything described above applies.

1.4 Workflow

The common workflow with Adspect for affiliate marketing campaigns consists of the following steps:

1. *Create an Adspect stream* for your campaign and place it in "On Review" mode.
2. Choose an appropriate integration method and follow instructions on the integration page.
3. Create an ad campaign using the link to the `index.php` file if using PHP integration, or to your white page where you put our `<script>` tag for JavaScript integration.
4. Wait for campaign approval and switch the stream into "Filtering" mode.
5. Run traffic and explore statistics in the *Reporting section*.

We will describe caveats of these steps in detail in the next chapters.

CHAPTER 2

Traffic Filtering

There are several approaches to detecting and filtering unwanted visitors in web traffic. In this chapter we will touch upon the basics of all of the major techniques of automatic filtering and explain what makes Adspect innovative and unique on the market.

2.1 Blacklisting

This is the most primitive, naïve, and widespread approach. It normally involves comparing a narrow set of features of a visitor (IP address, HTTP request headers, etc.) against a pre-collected blacklist. A match signals that the visitor should not be allowed further. While popular, this approach suffers from two major flaws:

1. Blacklists are never exhaustive and thus are trivial to circumvent, e.g. by cycling through a very long list of available IP addresses during each campaign review, as often facilitated by specialized proxy services. One cannot blacklist everything, there will always be wide gaps that allow malicious parties to get through. There are entire companies that do their business by maintaining vast pools of clean residential IP addresses ready for use for a fee. Maintaining up to date blacklists of these proxy IP addresses is infeasible.
2. Blacklists may be too broad, yielding false positives. This is especially bad with IPv4 address blacklists. The rather narrow 32-bit IPv4 address space has been exhausted, prompting Internet service providers and carriers to employ NAT (network address translation) to aggregate entire networks of subscribers behind a single shared IP address. This means that blacklisting, say, a single shared residential IP address under suspicion of proxy (yes, there are ways to maintain proxies behind NATs) in a large metropolitan area leads to blacklisting thousands of legitimate potential visitors and high bounce rate.

Blacklisting is the most common–and often the only–approach used by cloaking services in the affiliate marketing domain. While a viable solution in many cases, it is rough and unreliable and cannot be used on its own. Blacklist false negatives are the most common reason of cloaking faults. Adspect maintains massive built-in IP address blacklists of positively bad traffic sources that count up to one billion addresses.

2.2 Fingerprinting

Fingerprinting is, as the name suggests, the process of collecting a fingerprint of a visitor that identifies them. However, unlike human fingerprints that are universally unique, machine fingerprints aren't unique. Depending on implementation, they are composed of varying numbers of features, some of which are very common, like user agent strings of popular browsers. But some of the less common features happen to indicate with high accuracy the exact "bad" traffic that we protect against. And we know which.

Fingerprinting is a much more advanced technique normally used by business-oriented fraud protection companies. You may see their services employed, in particular, by value-added services (VAS) providers, protecting mobile "wap-click" offers from click fraud. Adspect is proud to call itself the pioneer of fingerprint scanning in the adtech industry.

Adspect has great expertise in JavaScript fingerprinting, that is, analyzing fingerprints composed of features of the visitor's JavaScript execution environment. Our fingerprints usually consist of 1600 to 2200 different facts, giving us a very detailed view into every visitor's internal works. We run collected fingerprints against dozens of high-precision tests that allow us to detect malicious visitors with unmatched accuracy. Adspect aims to bring high-end fraud protection into the realm of affiliate marketing.

2.3 Machine Learning

Machine learning (ML) is a broad term colloquially referring to making computers learn and then use what they have learned to do their task. With respect to traffic protection, machine learning is used to analyze the features of each visitor to classify them as either legitimate or malicious. This can be done with great precision, given enough information to teach the learning model.

Machine learning makes a perfect solution for inspecting fingerprints. Adspect is powered by a proprietary machine learning technology called VLA™, constantly trained to detect features of bad traffic well beyond the criteria initially built into it. Please refer to the *VLA chapter* for a detailed discussion.

Machine learning is the rocket science used by very few companies on the market, all of them being the big whales on the anti-fraud market. Adspect is the first vendor to bring machine learning into affiliate marketing.

2.4 Our Approach

Adspect employs all three of these techniques together without relying wholly on any single one. This allows us to make accurate decisions with the lowest rates of false positives and false negatives. We firmly believe that extensive fingerprinting coupled with machine learning appliances will play the leading role in defensive adtech because of the immense potential of both technologies, especially if combined.

CHAPTER 3

VLA™

VLA™ stands for "Virtual Learning Appliance." It is the trademark of our machine learning technology that powers the most advanced filtering capabilities of Adspect. In simple terms, it is a self-adapting mathematical machine that observes incoming traffic and finds suspicious recurring patterns in its fingerprints (thousands of features in every fingerprint) that indicate moderators, fraud, and other malicious activity. VLA constantly teaches itself, evolving and adapting to new types of threats as they emerge. We believe that VLA is our strongest weapon in the race of arms of affiliate marketing as it is able to see well beyond what we initially put into it. What a human analyst may overlook will never escape the mathematically strict scrutiny of a carefully programmed machine.

The concept behind machine learning is best described by analogy. Suppose a policeman at an airport is instructed to detain all passengers with a specific tattoo as they are known to be part of a dangerous gang. The policeman detained ten such persons during the last month, each time noticing that they all were also wearing T-shirts with the same symbol as on their tattoo. Now, the policeman will also stop people wearing those T-shirts under the same suspicion, regardless of whether they have the tattoo.

Whereas fingerprint checks yield a close to 100% confidence in that a given fingerprint belongs to a bot (moderator, spy service, etc.), VLA is inherently probabilistic in nature. The real deal here is that fingerprint checks encompass only those threats that we already know of while VLA detects previously unknown dangers. It takes a fingerprint, inspects every feature encoded in it, and yields a confidence percentage, as if saying, e.g., "I am 97% sure that this fingerprint belongs to someone you better filter out!"

Now, it only remains to determine what confidence is high enough to trigger the filter, and the choice is yours where to draw that line. The VLA section of every stream has a "VLA precision" setting that serves that very purpose: you specify the minimum confidence that you require VLA to have in order to filter out a visitor. For example, if you set VLA precision to 95%, then VLA will filter out all visitors for which it yields certainty of 95% and above, but will let through those that it is less confident about. This single precision parameter lets you fine-tune the system in accordance to your own idea of what is "confident enough". Our tests have shown that 95% is a good value to begin with.

CHAPTER 4

Use Cases

Adspect has a few well-defined use cases that have proven to be consistent and useful. Remember that Adspect has two intertwined but still distinct functions: cloaking and bot filtering. The latter helps a lot in achieving the former. We will describe the benefits of both just below.

4.1 Cloaking

Cloaking is the practice of hiding the real web page, be it a landing page or a CPA offer, from those who should not see it, at the discretion of the one who is in control of that page. We at Adspect firmly believe that if you do not wish to expose your content to a certain party, then you should be able to limit access to it, regardless of your reasons. We give you the power to do so. In particular, this means hiding your landing pages from ad network moderators, spy services, and antivirus robots. Those visitors will never make you any money.

4.2 Detecting Bot Zones

Popular purely web-based ad formats like banners, teasers, native ads, and popunders all suffer from click fraud. Technology they are based on–HTTP, HTML, and JavaScript–allows for relatively easy and cheap automated clicking, just pick any of the programmable headless browsers out there. No wonder these browsers, initially meant for website testing automation, are now widely used to forge clicks in ad networks and make advertisers pay for what will never bring them any revenue.

Adspect can detect all of them with ease. All you have to do is configure a stream to parse the subaccount ID out of the tracking URL as described in the chapter on streams. If you pass a publisher, site, or zone ID (let's call it a *source* from now on) to Adspect via a link parameter, then you will be able to pull per-source reports with exact figures of human ratio in the traffic. The rightmost "Quality" column in reports will let you evaluate and compare different sources, showing the percentage of legitimate traffic among the whole. Just select "Sub ID" in the grouping list to the left of the timezone picker.

Drawing a line at, say, 80% of humans in traffic, you can easily find sources that meet the requirement–just click the "Quality" column header to sort the table by that column. Sources with quality above 80% will give you a whitelist; conversely, sources with quality below 80% will be your blacklist. You will find this simple method invaluable for

determining converting traffic sources in display ads and popunder without spending fortunes on filtering sources by sheer CR (conversion rate.) Filter out bot-ridden sources first, then filter the rest by CR as you would normally do.

4.3 Hiding Traffic Sources

Many affiliate networks have internal media buying teams that would be all too happy to discover your traffic sources and use this information to steal your campaigns. Therefore you would normally want to hide your sources from affiliate networks, or any other parties, for that matter. Adspect does this for you by removing the Referer HTTP request header from all clicks, making sure that observers on the traffic redirect chain will not see the source of your traffic in web server logs.

CHAPTER 5

Configuring Streams

Traffic management in Adspect is organized in terms of streams. A stream is a traffic channel that is managed as a whole, much like a campaign in an ad network or a scheme in TDS.

Streams are managed in the Streams section of the clients area. Use the New Stream button to create new streams. Please be advised that the total number of streams per account is limited to 50. If there's no New Stream button, then you have reached your streams limit. Please contact us if you need more streams, and we will resolve your issue individually.

Each stream has its own `index.php` and `adspect.php` files wired to it that have the stream ID encoded inside. However, you may override that encoded stream ID and send a click to a different stream by putting the destination full stream ID into the `__sid` URL parameter, e.g:

```
https://example.com/index.php?__sid=1ea85c7c-b977-6804-8e69-00162501c2b4
```

Click the short stream ID in the streams list to copy the full stream ID into clipboard.

Below we will visit each stream setting in detail.

5.1 Name

Stream name is just a human-readable identifier that lets you distinguish between different streams. It is a good idea to match stream names with ad campaigns on one-to-one basis to maintain consistency and clarity across your traffic sources and Adspect. We also recommend that you create one stream per GEO (country) to make obtaining per-GEO statistics easier.

5.2 Filter Mode

This is the mode that streams currently operates in. There are four modes:

- **Filtering** – this is primary working mode in which we actively inspect every click coming in the stream and filter legitimate visitors from moderators, click fraud, and other unwanted types of traffic. All filtering technologies of Adspect, including *VLA™*, work only in this mode.
- **On review** – this mode is meant to be used when ad campaign that points to the stream is on review by ad network moderators. Every visitor in this mode will be directed to the white page. There are additional settings that apply in this mode, they will be described below.
- **All money** – auxiliary mode in which all visitors are directed to the money page. Useful for testing accessibility of the money page.
- **All white** – auxiliary mode in which all visitors are directed to the white page. Useful for testing accessibility of the white page. It is also a good idea to put streams into this mode whenever campaigns are paused in ad networks, to prevent unauthorized access to your landing pages or offers during inactivity periods.

On review is the default mode when creating a new stream. You *should* always use it when sending campaigns to moderation. After a campaign is approved, you should change its stream mode to Filtering before actual traffic starts coming.

5.3 Money Page

This is your actual landing page or offer that you are going to advertise. The "money" word is intended to indicate that this is the page that makes you money.

There are two types of values that may be specified: page file name or an URL. Page file name is the advised way of specifying money page–it is the name of an HTML or PHP file of your real landing page that *must* be located in the same directory as where you put `index.php` after stream creation, i.e. in the root directory of your landing page.

The file name should not be easily guessable because it lets determined moderators or competitors figure out the URL of your real landing page. Pick a random long file name.

Do not name your money page `index.html` or `index.htm`! Apart from being easily guessable (i.e. trivially uncloakable), those file names may be in conflict with your existing web server configuration, leading to unforeseen problems.

To put it all together: if you have a landing page directory and the actual landing page file inside named `index.php` (as is most often the case), then you should first rename that `index.php` file to something hard-to-guess like `re3NBX1XtH.php`, then put our special `index.php` file next to `re3NBX1XtH.php` in the same directory after stream creation. Our `index.php` will then display the real file `re3NBX1XtH.php` to approved visitors.

Alternatively, you may use an URL instead, for instance a direct offer URL taken from an affiliate network. This may be optimal for some campaigns, however, external URL implies an additional HTTP redirect, with associated latency and traffic loss considerations, especially on low-quality ad formats like popunder.

You may also use various non-HTTP URLs to achieve specialized tasks on your visitors' devices. Some of the more common examples:

- `mailto:user@example.com` will open up a default e-mail program in compose mode;
- `tel:+08001234567` will dial the number on mobile devices and some desktops with installed telephony software;
- `market://details?id=app` will bring the visitor to a particular app's page in Google Play.

This is particularly useful with the so called "deep links" that link to mobile in-app content.

5.3.1 URL Macros and PHP Variables

Adspect supports several URL macros and their PHP variable counterparts that you can use in redirection URLs and PHP landing pages, respectively:

- `{clickid}` and `$_SERVER["ADSPECT_CLICK_ID"]`: unique click ID–external ID from URL parameter, if specified, or generated by Adspect;
- `{country}` and `$_SERVER["ADSPECT_COUNTRY"]`: ISO 3166-1 alpha-2 country code of the visitor;
- `{os}` and `$_SERVER["ADSPECT_OS"]`: operating system of the visitor with version for Windows and Android (*PHP value may contain special characters and should be encoded for use in URLs with* `rawurlencode()`);
- `{browser}` and `$_SERVER["ADSPECT_BROWSER"]`: name of the browser of the visitor (*PHP value may contain special characters and should be encoded with* `rawurlencode()`);

URL example:

```
https://example.com/?clickid={clickid}&geo={country}&os={os}
```

PHP landing page example:

```
<a href="https://example.com/offer?clickid=<?= $_SERVER["ADSPECT_CLICK_ID"] ?>">
        Link to CPA offer with embedded Adspect click ID
</a>
```

5.4 White Page

This is the safe page to show to moderators, robots, scrapers, etc. It should not contain any sensitive content that may put your affiliate campaign in danger or in violation of any rules. Everything described above for the money page also applies to the white page–you may use an URL or a landing page file. In the latter case, if your money page is also a self-hosted landing page, you will need to effectively merge two landing page directories together to meet the requirement that both page files are located in the root of the common directory.

We strongly advise using a landing page file over external URL for white page. This has to do with suspicion and increased scrutiny from compliance teams of certain ad networks that discourage or outright prohibit the use of redirects in traffic flow.

5.5 Pass URL Parameters to Money/White URL

These settings are applied only when the corresponding money/white page is specified as an URL. When enabled, URL parameters passed to the `index.php` file will be appended to the corresponding money/white URL.

These settings work the same as similar settings available in most affiliate tracking software. For example, consider the stream's money page is configured as follows:

```
https://cpanetwork.test/offer?id=1234
```

A good (non-robot) visitor accesses `index.php` of the stream using the following URL:

```
https://tracker.test/lander/index.php?utm_medium=cpc&utm_content=mycampaign
```

After inspecting the visitor and allowing them through to the money page, they will be redirected to the money URL with URL parameters combined from both of the above:

```
https://cpanetwork.test/offer?id=1234&utm_medium=cpc&utm_content=mycampaign
```

5.6 VLA™

VLA™ stands for Virtual Learning Appliance, the trademark of the machine learning system at the heart of Adspect. It is discussed in detail in the *VLA chapter*. 95% is a good VLA precision value to begin with.

5.7 Sub ID

Sub ID refers to an URL parameter that you want to use for per-subaccount reports, available in the Reporting section by selecting the Sub ID grouping. Please refer to the *Reporting* chapter for details.

The concept is best described by example. Suppose your ad network has a notion of zones for dividing different publishers or ad placements into numbered groups. You would use some form of macro, e.g. `{zoneid}`, to put zone identifiers into your click URL. Your campaign tracking link might look like this:

```
https://tracker.test/lander/index.php?subid={zoneid}
```

For each click, the ad network will replace the `{zoneid}` macro with an actual identifier which can then be taken out of the click link and tracked individually. In this example, `subid` is name of the parameter used to track zone IDs. If you specify `subid` for the Sub ID stream setting, then you will be able to pull per-zone reports in the Reporting section of the clients area as mentioned above. This may come in very handy for building blacklists of bot-ridden publishers, zones, placements, etc.

Sub ID may also be anything else: GEO, hardware platform, OS version, any URL-trackable parameter. You can also combine several parameters into a compound subaccount by using more than one macro in the same parameter:

```
https://tracker.test/lander/index.php?subid={zoneid}-{platform}
```

In the example above each subaccount will be a combination of a zone and a device platform seen within that zone.

5.8 Click ID

Click ID works the same way as Sub ID, but for tracking unique click identifiers often supplied by ad networks or affiliate trackers. If the Click ID setting is specified, click IDs are taken out of that link parameter and recorded in statistics along with other click data. This allows you to find and examine individual clicks in raw CSV reports. One use case would be compiling lists of bot clicks as a proof of click fraud.

If the Click ID setting is omitted, then Adspect will generate its own click IDs for use with its *tracker*. Depending on how your money/white pages are displayed, click IDs may be put into money/white page URLs via the `{clickid}` URL macro or embedded into money/white page files via the `$_SERVER["ADSPECT_CLICK_ID"]` PHP variable.

5.9 Paranoid Mode

The paranoid mode enables additional strict fingerprint checks and vast IP address blacklists (counting up to 2 billion IPv4 addresses) that are considered “paranoid”, that is, with higher false positives chance, but at the same time providing equally higher chance of busting moderators.

We recommend to enable this mode when working with Facebook or Google Ads.

5.10 Allow Traffic From Mobile Apps

This setting tells us to allow traffic that originates from inside mobile applications, e.g. from WebView Android browser. While natural for certain niche ad formats, such traffic is widely seen as click fraud in other formats (automated clicks generated by mobile malware) and should normally be disabled unless your ad format is somehow based on mobile applications.

5.11 Countries, Operating Systems, Browsers, Languages, and Time Zones

These manual targeting options allow you to further restrict your stream to only allow visitors from specified countries, using specified operating systems, browsers, browser language preferences, or time zones. You would normally set them to match your campaign targeting settings. If any of these settings is omitted (the list is empty), then no check will be made for that setting.

Note: time zone settings are restricted to full hour offsets from UTC. If a visitor's time zone is not offset by full hours, then the offset will be rounded.

5.12 Match Browser Time Zone to Location Time Zone

If this setting is enabled, then Adspect will check whether the time zone reported by visitor's browser matches the time zone of the visitor as determined by our geolocation. This check may slightly increase the rate of false positives, but it significantly boosts protection against moderators and bots that use VPN or proxy services. If enabled, the manual time zone list described above is ignored. It is recommended to enable this setting.

5.13 User Agent Filter

This setting allows you to specify a custom Perl-compatible regular expression (PCRE) for filtering visitors by their user agent string. Regular expression matching is case-sensitive. By default, the search is done in any part of the user agent string; you may use anchors to bind matching to the start or the end of the string (see examples below.)

PCRE syntax is very rich and powerful and is well out of scope of this document. Regular expressions can be combined using various syntax constructs to create arbitrarily complex patterns, but please note that the current implementation limits regular expression length by 1023 characters.

Some examples:

```
Firefox|Nexus|Miui
```

This regex will match any user agent that contains words "Firefox", "Nexus", or "Miui", and can be used to filter out visitors that use Mozilla Firefox, Google Nexus, or Xiaomi built-in browser.

```
^Mozilla/4[.]0
```

This regex will match any user agent that begins with "Mozilla/4.0", banning shady visitors that report themselves to be very old browsers yet support contemporary JavaScript features (implied by being able to run our fingerprint collecting code.)

```
^Mozilla/5[.]0$
```

This regex will match user agents that are exactly "Mozilla/5.0", blocking visitors without concrete browser, HTML engine, and platform information, which is very uncommon and suspicious.

All of the expressions above can be combined using logical "or" (i.e. to match the first expression *or* the second *or* the third) this way:

```
Firefox|Nexus|Miui|^Mozilla/4[.]0|^Mozilla/5[.]0$
```

Please be careful! Improperly formed regular expression can lead to erroneous matching and filtering of vast amounts of legitimate traffic. Use this setting only if you know what you are doing.

5.14 Referer Filter

This setting works similarly to the user agent filter described above, but deals with HTTP referer instead. It also takes a Perl-compatible regular expression and filters out all visitors whose referers match it. Regular expression matching is case-sensitive.

One common use case is filtering empty or non-existent referers. This can be achieved with the following regex:

```
^$
```

Please be careful! Improperly formed regular expression can lead to erroneous matching and filtering of vast amounts of legitimate traffic. Use this setting only if you know what you are doing.

5.15 IP/ASN Blacklist

This is the list of IP addresses, IP address ranges, and/or autonomous system numbers (ASN) that should always be shown the white page. Both IPv4 and IPv6 addresses are supported, as well as CIDR and range notations. Examples:

- 192.0.2.1
- 192.0.2.0/24
- 192.0.2.0–192.0.2.255
- 2001:db8::1
- 2001:db8::/112
- 2001:db8::-2001:db8::ffff

Individual entries should be delimited by newlines or whitespaces. Please note that the system will automatically merge adjacent or overlapping ranges in order to optimize storage space and lookup speed.

5.16 Blacklist All IP Addresses In Review Mode

If enabled, this setting instructs Adspect to add IP addresses of all incoming visitors to the IP blacklist if the stream is in Review mode. Since the Review mode is meant to be used only when your ad campaigns are under review by moderators, it is safe to assume that every visitor in this mode is a moderator and should be barred. We recommend you to always enable this setting, but pay attention to the moment your campaign is approved, to switch the stream mode to Filtering in time lest you blacklist IP addresses of legitimate visitors when your campaign goes live.

CHAPTER 6

Tracker

Tracker is an indispensable tool in digital marketing as a whole and affiliate marketing in particular. Its primary function is to register conversion events, that is, orders or sales, and track them back to specific visitors. This in turn allows marketers to gather conversion statistics and analyze it from different angles using various criteria, building what is known as conversion funnels.

Adspect is equipped with a lightweight yet efficient tracker built right into the core of the system. The *Reporting* section of the clients area allows you to explore funnels built with different groupings and filters. Among others, it calculates and displays such important metrics as conversions, cost, revenue, CR (conversion rate), ROI (return of investment), CPA (cost per action), and EPC (earn per click) / EPM (earn per thousand clicks.) These are especially useful combined with subaccounting by source identifiers as described in the previous chapter in the paragraph on Sub ID.

6.1 Postback

In order to use tracking you need to configure postback of conversion events to our postback URL located in your profile. We accept postback via any HTTP method: `GET`, `POST`, `PUT`, etc. The postback URL takes three parameters:

1. `aid` – Adspect account ID, which is pre-filled and normally will not change;
2. `cid` – unique click ID that identifies particular click that made a conversion;
3. `sum` (optional) – payout sum of the conversion in case of CPA or revenue share tracking.

Most affiliate programs and networks support postback and provide various macros that can be used to fill variable portions of the URL (`cid` and `sum` parameters.) If you need to fire postback manually, then you could place a conversion pixel somewhere, e.g. on a "Thank you for your order" page. For example, assuming that click ID is contained in the `clickid` link parameter:

```
<script>
(function () {
  const cid = new URLSearchParams(location.search).get("clickid");
  const url = "https://rpc.adspect.net/v1/postback?aid=1ea704aa-d0d3-6262-bf65-
↪ac1f6b95a853&cid=" + cid;
```

(continues on next page)

(continued from previous page)

```
  fetch(url, {mode: "no-cors"});
})();
</script>
```

This code will send a postback request immediately when the page is loaded. For a successfully registered conversion, the postback URL will return the HTTP status code 200 and `OK` in plain text.

6.2 Click IDs

For a conversion to be registered and processed it is required that Adspect has a record of the corresponding click in its statistics database as determined by its click ID. You may either use externally generated click IDs passed to Adspect from outside in a link parameter, in which case you should put the name of that parameter into the Click ID stream setting, or omit the Click ID stream setting and let Adspect generate click IDs automatically. The postback URL will accept previously unregistered click IDs, but they will be discarded at later stages of conversion processing.

CHAPTER 7

Reporting

Our reports are a comprehensive and valuable source of analytical information on your affiliate campaigns performance and traffic quality. You can use reports to evaluate traffic quality of different sources, publishers, ad spots, etc. Reports come in two flavors: raw and aggregate.

Please note that statistical data is not real-time and is updated once every minute.

7.1 Raw Reports

Raw reports are per-click, that is, they contain information on every click that was processed by Adspect. They are available for download in the CSV format via the "Get CSV" button menu. There you have two options: get full report or only for those visitors that were filtered by Adspect because of triggering one of more checks. Report will be limited by the selected date range. Downloaded CSV files may then be imported into Microsoft Excel or similar spreadsheet software.

Please do not select too broad date ranges as it will lead to the formation of huge CSV files and additional strain on our servers. We limit the total number of rows that will be included in report, and this limit is subject to change at the sole discretion of our systems administrators.

7.2 Raw Report Columns

Raw reports can have either one or two rows per each click. The first row corresponds to serving of our fingerprint collector script for the client browser to execute. The second row, if present, corresponds to fingerprint scanning and making the decision: allow or block. The second row may be missing if the visitor failed to produce or submit a fingerprint.

Raw reports consist of the following columns:

- timestamp – date and time of the event;
- ip_address – IP address of the visitor in IPv6 format (IPv4 addresses are represented via standard IPv4-to-IPv6 mapping);

- stream_id – ID of the stream that the event happened in;
- country_code – ISO 3166-1 alpha-2 country code of the visitor;
- os – name and release of the visitor's operating system;
- browser – name of the visitor's browser;
- cost – cost of the click, if passed via URL parameter;
- sub_id – sub ID of the click, if passed via URL parameter;
- click_id – unique ID of the click, if passed via URL parameter;
- mode – stream mode at the moment of the event;
- sequence – click processing stage: 0 for fingerprint collecting, 1 for fingerprint scan;
- valid – 1 if the click has been allowed through, 0 otherwise (meaningful only if sequence = 1);
- tags – list of mnemonic tags, mostly for internal use, that represent particular filtering reasons.

The exact nature of click tags is a trade secret–we do not disclose our filtering techniques. However, we do give out information about some of them that can be used as proofs of bot traffic (e.g. for demanding refunds from ad networks) or for debug purposes:

- `REVIEW`, `MONEY`, `WHITE` – decision made by customer via stream mode;
- `GEO`, `OS`, `BROWSER` – decision made by customer via stream targeting;
- `IP`, `IP*` – IP address blacklisted by us: proxies, VPN and hosting providers, antivirus companies, ad scoring companies, security companies, known moderator origins, etc;
- `IPSLB`, `IPSB` – IP address blacklisted by the stream blacklist;
- `GOOGLE` – visitors with user agents that belong to Google or their affiliates;
- `BOT`, `GENBOT` – visitors with user agents that identify them as bots;
- `EMU` – clicks from known device emulators and virtualized environments;
- `GEO`, `OS`, `BROWSER`, `LANG`, `TZ`, `IPTZ` – visitors blocked by manual stream filters;
- `UARE` – visitors whose user agent matched customer-supplied regular expression;
- `REF` – visitors whose referer matched customer-supplied regular expression.

7.3 Aggregate Reports

Aggregate reports are produced by splitting raw reports into groups and summing metrics on per-group basis. Grouping may be set set in the field to the left of the timezone selector and defaults to "Stream", meaning that values will be computed on per-stream basis. Other grouping options include "Date" for per-date grouping and "Sub ID" for per-subaccount grouping (refer to this paragraph for an idea of what may constitute a subaccount.)

Grouping may be nested, e.g. "Date" followed by "Stream"–this will result in first splitting reports by date, and then by stream on each date. You may combine grouping criteria in any way you need to inspect different funnels.

Press the "Get Report" button to produce an aggregate report in the form of a table that will be displayed to you right below the report settings panel.

In addition to browsing reports in our clients area you can also download them in the CSV format using the gray "CSV" button in the bottom left corner of the report window.

7.4 Aggregate Report Columns

Each aggregate report consists of the grouping columns on the left site followed by a number of statistical columns. Some of the statistical columns have a gray percentage value following a slash–that is percentage of the total clicks, displayed for more comprehensivess.

The list of statistical columns, explained:

- Clicks – the total number of clicks that accessed the `index.php` file.
- Uniques – approximate number of unique visitors are per uniqueness of their IP addresses.
- FP – the number of visitors that successfully executed our JavaScript fingerprint collector code and submitted their fingerprints for analysis. This figure may be lower than the total number of clicks for various reasons, most often it being the inability of dumb click bots to run JavaScript.
- Money hits – how many visitors have been shown the money page. This is the best metric for accounting legitimate traffic. Please note that this also includes all visitors when the stream works in the "All money" mode.
- White hits – how many visitors have or would have been shown the white page. This metric is calculated as clicks minus money hits and includes those dumb bots that would have been show the white page if they were able to execute JavaScript (there's a fallback "meta refresh" mechanism to deal with them.)
- GIVT – general invalid traffic, which is computed as the number of visitors that failed to produce a fingerprint. As mentioned above, these are often dumb bots with limited JavaScript support. Another common reason is network latency, especially evident in traffic with slow connection rates when visitors manage to close the tab or window before their fingerprint is submitted. Currently, this column also accounts all clicks received when a stream was in All money, All white, or On review mode with disabled fingerprint collection because fingerprint scanning is not performed in these modes. We plan to change this logic in future to make this column reflect real GIVT more precisely and transparently.
- SIVT – sophisticated invalid traffic, that is, the number of fingerprints that Adspect consciously filtered out as bad traffic. This may serve as a rough traffic quality metric with respect to the more advanced types of click fraud that get more spread today. This metric also includes visitors blocked by manual stream filters (country, OS, browser, regular expressions, IP address blacklist.)
- Cost – total traffic cost computed as a sum of costs of each click, if passed via URL parameter.
- Bots cost – cost of the traffic that was directed to the white page, which is the precise metric of your budget loss.
- Quality – percentage of money hits in the whole click volume. This is the the best metric for evaluating traffic quality as a whole and may be used to compare different traffic sources, publishers, ad spots, etc. Especially useful with grouping by sub ID for compiling blacklists of bot-ridden zones, as described in a dedicated paragraph.

There's also the "CSV" button in the bottom left corner of the report box. You can use it to download CSV versions of aggregate reports for further analysis, printing, or sharing.

CHAPTER 8

Recommendations

Below you may find a compiled list of general recommendations that we give to all of our customers. We highly encourage you to follow them in order to achieve the best results with Adspect.

1. Domains and hosting
    1. **Do not use** domain names in cheap zones like `.site`, `.club`, `.world`, etc. as they attract more thorough and frequent scrutiny from antivirus and ad scoring companies, effectively being tainted from the beginning. **Only use** domains in `.org`, `.net`, and `.com` zones, in the order of priority.
    2. **Do not use** domain names that contain questionable or stop words: “sex”, “xxx”, “win”, “diet”, “health”, “meet”, “date”, and so forth. Be creative and choose domain names that look and sound like brands.
    3. **Always use** Cloudflare to hide IP addresses of your servers. Many ad networks flag IP addresses behind domains in banned accounts, however, they will never flag addresses that belong to a huge CDN company serving 10% of the World Wide Web. Adspect fully supports Cloudflare in proxy mode.
    4. **Do not use** technical domains of hosting companies. They both raise suspicion and uniquely identify particular servers of the hoster.
    5. **Do not use** Namecheap virtual hosting because they employ WAF (web application firewall) that by default blocks POST requests that we rely upon, causing 403 Forbidden errors.
    6. **Do not name** any of your money or white page files `index.html` because it is likely to take priority over our `index.php` if the file name is omitted in the URL after `/`, thus exposing your real pages. Always use distinct and hard to guess names for money and white pages.
    7. **Make sure** that your white pages do not have broken links, non-loading images, or scripts with errors. Check the Console and Network tabs of developer tools in your browser, it will highlight potential problems red.
    8. **Do not use** web resources (images, styles, scripts) from third party domains unless they are well-known CDNs (content delivery networks) like those of jQuery, Bootstrap, Font Awesome, Google Fonts, etc. Better download them and host locally.
    9. **Always set up** your pages for HTTPS. Cloudflare provides free SSL/TLS certificates for domains in proxy mode. Let’s Encrypt also provides an infrastructure for obtaining free SSL/TLS certificates.

10. **It is recommended** to use hosting providers in proximity to your target audience, ideally in the same country. This is especially important when working with popunder ad format.

2. Cloaking tips

    1. **Never reuse** domain names in the same network. If an account is banned, then write down all the domains used by its campaigns and don't use them again. Register new domain names for new accounts.
    2. **Always use** the most strict manual filters by country, OS, browser, and languages. Match them to targeting settings of your campaigns.
    3. **Always use** zero redirect (file-based) display mechanism for white pages, if possible. This rule is **mandatory in Facebook and Google Ads** with PHP integration!
    4. **Always make sure** that your white pages are convincing and relevant to your ad campaigns (creatives, languages, targeting options.) **Never use** obviously bogus white pages like redirections to Google.
    5. **Use site constructors** like Shopify, Wix, Tilda, etc. for white pages when working with Facebook or Google Ads as they come with highly trusted domains. We support JavaScript integration for them. **One exception is Tilda**: Google seems to preemptively flag sites in the `tilda.ws` zone for cloaking, which has been reported several times recently.
    6. When using site constructors, **check availability** of your site periodically–it is not uncommon for such services to ban sites for suspected use in various cloaking schemes.
    7. **Put links** to legitimate-looking terms of use, privacy policy, and cookies policy on your white pages when working with the more strict ad networks like Facebook, Google Ads, Microsoft Advertising, etc. **The EU Cookie Law also requires** websites to obtain informed consent from visitors if using cookies.
    8. **Put** `robots.txt` and `sitemap.xml` files into the root of your domain, especially when working with search engine-based ad networks (Google Ads, Microsoft Advertising, Verizon Media Native, etc.)
    9. **Do not use** copied landing pages as is, always alter them in one way or another to make them overall unique. This practice helps with disarming signature-based flagging.
    10. **Do not use** copyrighted materials on your white pages, they may be detected and rejected.
    11. **Do not use** overly simple, single page, skewed/broken, non-mobile-optimized, or low quality white pages. Always remember that a white page must look like a legitimate, useful website with authentic content.
    12. **Do not use** landing pages of supposedly "white" offers as white pages–advertising networks tend to have a wildly different idea of "white". **Never use** direct affiliate links as white pages.

CHAPTER 9

Tips and Tricks

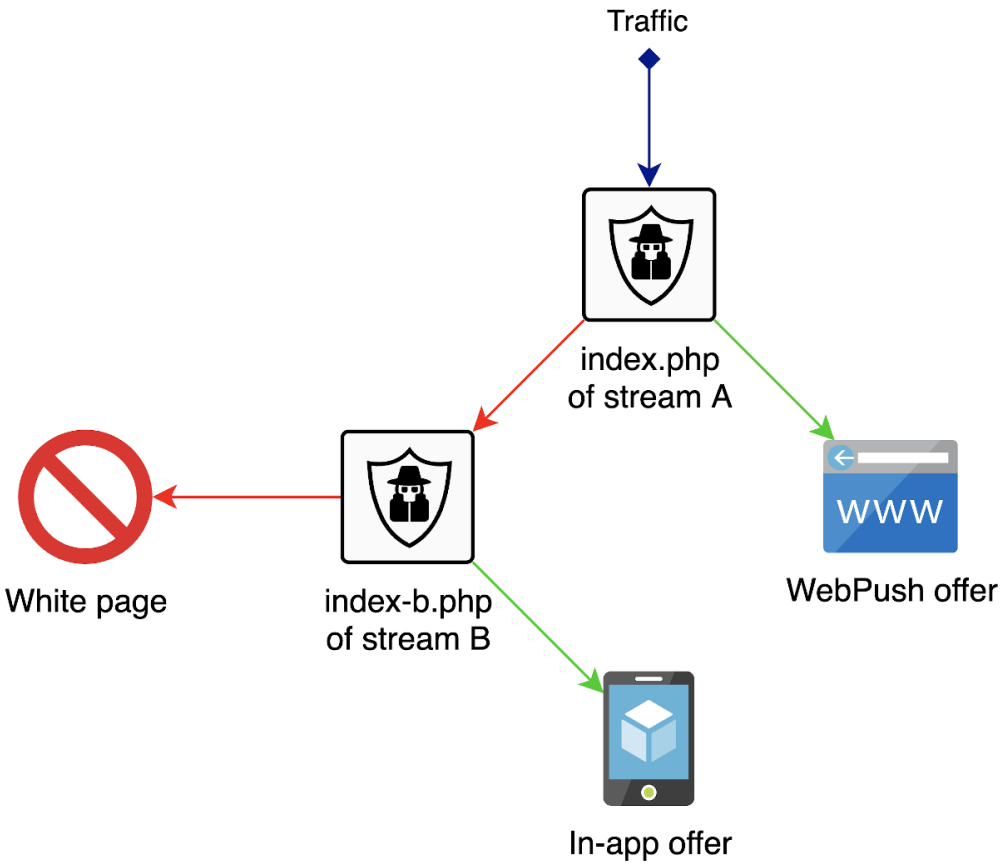
The flexible nature of `index.php` files employed by Adspect combined with file-based page display mechanism (which uses the `require()` language construct of PHP) allows for more sophisticated setups of streams. In this chapter we will describe several advanced scenarios that may be very useful to you.

9.1 Stream Chaining

Since `index.php` is a regular PHP script, it may be used as a money page or a white page of a stream, that is, one stream may redirect visitors to another stream, letting you chain them in different ways. Typical setup of stream chaining is best described by example of a real-world application.

Suppose you have an ad campaign in a source that supplies both browser-based and mobile app-based visitors in a mix without providing any option to split them. Such networks exist–push notifications networks that have both WebPush subscribers and subscribers for in-app or PWA notifications. You would like to split these traffic types and send them to different affiliate offers: `https://example.com/webpush-offer` for WebPush-based visitors and `https://example.com/inapp-offer` for mobile app-based visitors.

You can accomplish this by chaining two streams that have different settings for mobile app-based visitors. The first stream A accepts incoming clicks and filters out app-based ones to its white page. The second stream B is attached to the first one as its white page and filters out bots and moderators, letting legitimate app-based visitors through to the separate offer.



Traffic flow chart

Here are the relevant settings of the entryway stream A:

- Money page: `https://example.com/webpush-offer`
- White page: `index-b.php`
- Allow traffic from mobile apps: *disabled*

The relevant settings of the rear stream B:

- Money page: `https://example.com/inapp-offer`
- White page: `https://google.com/`
- Allow traffic from mobile apps: *enabled*

The mobile apps setting of stream A is what makes it consider app-based visitors as malicious and send them all to stream B along with actual bots and moderators. Stream B will in turn re-analyze the traffic and send good app-based clicks to the in-app offer while still filtering out unwanted visitors to Google.

Finally, you would place `index.php` files of both streams into the same directory: leave the `index.php` file of stream A named as is and rename the file of stream B to `index-b.php`, which will act as a surrogate white page of stream A.

9.2 Dedicated IP Blacklist Stream

Streams have a useful "Blacklist all IP addresses in "Review" mode" setting for collecting IP blacklists of moderators during the "Review" phase of stream lifecycle. This setting may be used to create a dedicated stream for the purpose of collecting a single blacklist of all moderators, bots, etc.

The process is as follows:

1. Create a separate stream that will be used for accumulating IP addresses in its blacklist. Set and leave it in the "Review" mode and enable the "Blacklist all IP addresses in "Review" mode" setting. Effectively, this means that the stream will blacklist IP addresses of every visitor.
2. Use the stream's `index.php` file as a white page for other streams as described in Stream Chaining above. This will direct all bad visitors to the blacklisting stream, making it collect their IP addresses. Alternatively, you may use the stream as a white page only during the "Review" phase of other streams to exclude IP addresses of regular bots from the blacklist as it may possibly lead to false positives (please read this chapter for an explanation of why this may happen.)
3. Watch the IP blacklist of your dedicated stream being collected and copy-paste it into other streams every once in a while (yes, this isn't very convenient, we are working on a solution for blacklist sharing.)

9.3 Combining Cloakers

In case you have access to other cloaking and traffic protection solutions, you may use them together with Adspect to attain potentially higher levels of protection at the expense of additional processing latency. Since most of our competing solutions have a notion of money and white (safe) pages, you should always put Adspect at the rear side of the cloaker chain and create a special setup with two separate streams:

- One stream will be used as a money page of your front-side cloaker, taking the baton from it and inspecting supposedly good visitors that the cloaker has allowed through. The money page of this stream will be your final traffic destination whereas the white page will be the real white page that your intend to cloak with. This stream should be set to the "Filtering" mode.
- Another stream will be used as a white (safe) page of your front-side cloaker, accepting traffic from it in order to collect an IP address blacklist of visitors that the front-side cloaker deems dangerous and additionally train our *VLA* machine learning system on their results. This data lets us absorb their filtering techniques and make Adspect more comprehensive and precise. You should always enable the "Blacklist all IP addresses in "Review" mode" setting as described in the section above in order to populate the IP address blacklist automatically. Set both money and white pages of the stream to the real white page that you intend to cloak with. Leave the stream in "Review" mode.

CHAPTER 10

Drawbacks and Pitfalls

As with every sophisticated system, Adspect has its drawbacks and pitfalls. You should be aware of them. This chapter will educate you how to avoid common mistakes lest you put your affiliate success at stake.

10.1 Do Not Stand Out!

Most ad networks have a routine practice of reviewing all advertising campaigns every once in a while. Apart from helping you pass the initial review after launching a new campaign, the first and foremost task of any cloaking service is to protect your running campaigns from these recurring reviews. Adspect does this very efficiently, proven by many successful campaigns in different ad networks.

But there's a catch: if your advertising activity stands out compared to an average advertiser as perceived by a particular network, it will eventually attract attention of their policy-enforcing team, invoke manual and detailed scrutiny, and will inevitably lead to "piercing the cloak" and suspension. We can guarantee solid protection from routine reviews, but no service can protect you from determined investigators driven by suspicion.

Remember: *do not stand out!* If you slip somehow and put network staff on alert, then you're done. There's no coming back from that one. Here are some common preventive guidelines:

- Do not run ridiculous amounts of traffic from a single account. High volume indicates consistent profitability and thereby raises interest in the nature of your campaigns.
- Keep the number of active campaigns per account low. More campaigns = more things to inspect and uncover.
- Always use a tracker. Continuous affiliate campaigns run without any tracking prompts an observer to wonder how it is possible to sustain them on profitable level.
- Always use tracking parameters with macros supported by the network. This tip is related to the one above and is especially crucial for campaigns with wide targeting settings.
- Use postback feature, if supported by the network.

These considerations outline the principle of division of responsibility: Adspect is responsible for protecting you from regular campaign reviews, and you are responsible for not attracting attention.

10.2 Long Redirect Chains

One notable drawback of cloud-based services like Adspect is that they contribute to overall latency of click processing because of round-trip network lags between your tracker and the service's backend servers. If you observe high technical loss in the "Reporting" section, then it may indicate a network latency problem.

We highly recommend to keep your redirect chains as short as possible. Host your own landing pages with file-based display mechanism instead of redirects to external URLs (detailed here.) Put tracker either before or after our `index.php` file in the traffic flow, but not at both sides. Place your tracker geographically close to your target country or region, if possible.

On a side note, we are developing a self-hosted filtering solution that will allow our clients to cut network latency by performing real-time filtering right at their tracker side, i.e. without having to contact our backend servers on every incoming click. Additional information will be released later.

10.3 False Positives

False positives occur when a filter bars a legitimate visitor, falsely classifying them as malicious. No solution can guarantee 100% precise results, but we are confident enough to state that our rate of false positives is the lowest on the market, sometimes dramatically lower compared to certain competitors. Some solutions on the market yield lower amounts of filtered traffic than Adspect, but that is in fact their fault and not an achievement–they let malicious visitors slip through. Such misclassifications are called false negatives and can be easily proved on click-by-click analysis of decisions. Adspect provides decision logs in the form of raw CSV reports.

10.4 False Negatives

False negatives occur when a filter fails to detect a malicious visitor, letting them through. This is the very reason why affiliate campaign cloaking fails every once in a while. False positives stem from the practical impossibility to detect and filter out every malicious visitor. Any and all measures taken can be circumvented, given enough determination and proper technical expertise. What we described in the Do Not Stand Out! paragraph above is how such determination occurs in the heads of those who you are protecting from; and be sure that they do have all the technical skills needed.

Remember: *false negatives and their consequences result from drawing attention!*

CHAPTER 11

Referral Program

Adspect has a referral program that allows our customers to earn money for bringing new clients to us. The referral program works on a revenue share basis, that is, the referral fee is credited to your balance each time a customer that you brought purchases a subscription.

The referral fee is 5%. However, it may be increased individually if you manage to bring new clients consistently. Please contact us if you have considerable presence on affiliate marketing or SMM forums, blogs, or in Telegram groups, and we will find a mutually profitable agreement.

The referral fee is credited to your account balance. Each time an invoice is created, funds on your balance are used first and deducted from the total amount that remains to be paid. This means that if you have enough funds on your balance, then you may use them to purchase a subscription wholly. For each credited fee, you will receive a referral fee receipt in the Invoices section of the members area.

You may find your referral link at the bottom of your profile, along with a list of your referral clients. Use that link to promote Adspect, and each new member registered by it will be accounted as your referral. On the technical side, we use both a cookie with one month expiration time and a parameter in the link that we e-mail to new customers in order to complete registration, which means that even if they do so on a different device, they will still be accounted to you.