WE IDENTIFY AND BLOCK FRAUD AND MARKETING ABUSES OF THE TYPE

INVALID CLICKS AND LEADS

Technical Specification of the Solution

Description of Methodology

www.trafficwatchdog.pl

Spark DigitUP Ltd.

Plac Wolnica 13 lok. 10

31-060 Krakow

NIP 6762496391

1. TrafficWatchDog (TwD) system – description of functionalities implemented

TrafficWatchDog (TwD) - a system operating in the field of fraud detection and online/mobile marketing abuse for advertising forms such as KLIK and LEAD. TwD collects and analyzes the parameters of clicks and/or leads (impersonal data) from individual paid sources, and based on the collected parameters, makes an assessment of the clicks/leads:

1) CLICK Scanner - Identifying 'invalid clicks' according to IAB - Interactive Advertising Bureau Click Measurement Guidelines standards: "Invalid Clicks arising from suspected "click fraud" are a sub-component of Invalid Clicks and originate from a user, program or automated agent (e.g., Internet robot or spider) that accesses a URL for the purpose of manipulating click measurement activity or click-based advertising payments, having no intention of legitimately browsing site content, making a purchase or performing any other type of legitimate conversion action. Suspected click fraud can arise from both human-initiated and application-initiated automated activity; also, suspected click fraud can arise from invalid Ad Impression activity. Click Fraud also includes situations where a user is unwillingly, or tricked into, accessing information (for example, user "virus" infected activity, or auto-clicking functions)."

2) GOOGLE ADS Scanner - identifying 'invalid clicks' in a Google Ads campaign, according to IAB - Interactive Advertising Bureau Click Measurement Guidelines standards: "Invalid Clicks arising from suspected 'click fraud' are a sub-component of Invalid Clicks and originate from a user, program or automated agent (e.g., Internet robot or spider) that accesses a URL for the purpose of manipulating click measurement activity or click-based advertising payments, having no intention of legitimately browsing site content, making a purchase or performing any other type of legitimate conversion action. Suspected click fraud can arise from both human-initiated and application-initiated automated activity; also, suspected click fraud can arise from invalid Ad Impression activity. Click Fraud also includes situations where a user is unwillingly, or tricked into, accessing information (for example, user "virus" infected activity, or auto-clicking functions)."

   Google Ads Scanner system - in addition, it allows you to automatically block clicks of Client's Google Ads by the same users and block invalid clicks in Client's Google Ads campaigns.

3) LEAD Scanner - identifying 'invalid leads' according to IAB - Interactive Advertising Bureau Online Lead Generation standards: "Lead fraud occurs when leads are submitted with malicious intent or simply for financial gain. These leads should be deemed invalid, and advertisers and agencies should not pay for these leads. Although uncommon, there have been cases when offers are filled out by an artificial, automated system to generate a large quantity of leads. Consumers or companies may also fraudulently fill out offers."

Main functionalities implemented:

1. Full 24/7 monitoring of all sources providing paid clicks/leads for the client.

2. Automatic analysis of ads and clicks in Google Ads.
3. Detection and analysis of the user's device - based on the so-called virtual imprint of the device (DEVICE FINGERPRINT).
4. Identify country of origin of click/lead and IP providers.
5. Automatically block clickable Google Ads by suspicious/fruad IP addresses.
6. Automatic blocking of clickable Google Ads by suspicious/fraudulent devices (DEVICE FINGERPRINT identification).
7. Automatically block clickable Google Ads from suspicious/fraudulent 'cookies'.
8. Ability to set and customize individual rules for automatically blocking clicks from Google Ads campaigns.
9. Claim reports of invalid clicks and leads for sources providing paid clicks/leads.
10. Report claiming invalid clicks for Google Ads.
11. Cyclical reports sent to a designated e-mail address.
12. Detailed information for each verified click/lead - including analysis of potential user/bot behavior on the monitored site.
13. Online access to the Customer Panel.

2. Technical architecture of the solution TwD

The system consists of the following components:

1) TwD tracking script - an application responsible for collecting parameters for evaluation in terms of clicks/leads.
2) Webservice - an application installed on the TwD server, to which the data collected by the script will be sent
3) Database
4) Server
5) AI - neural network that performs data analysis
6) Customer Panel - the front of the customer in the presentation of the results of the TwD system.

3. How TwD codes work

The operation of dedicated tracking codes is matched to the target structure of the monitored Client's website - on which the appropriate fragments of HTML codes containing Javascript scripts and pixels are implemented, which optimally if placed in the indicated target areas of the monitored site - directly in the code of the monitored site (so-called 'body' of the page). The scripts are executed only on the user's side in his browser.

Implementation of the prepared Javascript/pixel tracking code on the target monitored page of the client, can also be done through the GTM (Google Tag Manager) system. The scripts are executed only on the user's side in his browser. If TwD codes are implemented via GTM - loading TwD codes will depend on the user's browser's support of GTM (some browser versions - block GTM). This may mean 'losses' in terms of the number of monitored clicks/leads.

TwD scripts and code elements run as the web page loads. They also run in the background while the user is working on the web page using event handling, generated by web page elements. Data sent

when a completed form (lead) is sent via the POST method, or via the GET method when an image is posted on the page.

The data is sent to the target server, where it is analyzed. The results of the analysis are available in the TrafficWatchDog service's dedicated client panel.

4. Technical implementation of TwD codes

The implementation process and system architecture consists of the following elements:

On the TrafficWatchDog side:

1. Development of target TwD code (Javascript/pixel file) - responsible for data collection, tailored to the specifications of the monitored website
2. Webservice - running an application installed on the TrafficWatchDog server, to which the information collected by the TwD code will be sent via HTTPS protocol
3. Server - record of collected evaluation parameters
4. Neural network - calibration of model and algorithms for data analysis and record evaluation
5. Issue a dedicated client panel - where you can analyze online the results of analysis/create claim reports for e.g. Adwords, etc.

On the part of the client - the owner of the monitored site:

1. Pasting dedicated TwD code (Javascript/pixels) into the target website
2. It is up to the Client to decide on which sites TwD code will be implemented - the code may be placed only on selected sites/so-called micro-sites to be analyzed (e.g. selected campaign Landing Pages - so-called marketing micro-sites). Pages treated specifically for security reasons, such as the Client's administration panel, shopping or transactional pages may be excluded from analysis (no TwD code inserted).
3. Depending on the client's decision - using the JavaScript script located on the TrafficWatchDog server or placing the Javascript file on the monitored site's own server.

5. Impact of TwD codes on the architecture of the monitored landing page

The implemented TwD codes do not affect the architecture of the target monitored website. JavaScipt/pixel scripts are executed only on the user's side in his browser. No impact on user experiance in terms of page loading speed/itd. (tests with the tool: https://developers.google.com/web/tools/lighthouse)

The posted scripts in no way affect the login path to the Client/itd transaction systems

6. IT Security

Posted TwD codes on the client's target website collect parameters that are generated when the website is used. It should be noted that the data are collected anonymously, it is not possible to associate them with a specific user on the part of the TwD application provider.

All data is transmitted and stored on internal TwD servers - multi-level security is used. Between the script running in the user's browser and the server collecting the data, communication takes place over an encrypted HTTPS channel (SSL TLS 1.2 with 2048 bit key). The data is then processed on the application server, which is the only one with a direct connection to the public Internet. It is protected by a Stateful Packet Inspection (SPI) firewall with Login/Intrusion Detection and protection against BruteForce attacks. After processing, the data is stored in an internal database that can only be accessed from the internal network. Logging to all servers is exclusively with individual 4096 bit RSA keys.

Applied implementation standards in the security techniques standard PN-ISO/IEC 27002:2014-12 for information security organization, access control, cryptography and communication security.

7. Parameter categories analyzed

Categories of parameters analyzed:

I. URL details
II. Visit details
III. IP parameters
IV. Browser parameters + Browser Fingerprint
V. System parameters
VI. Device Parameters + Device Fingerprint
VII. Parameters user bahavioral
VIII. Script handling
IX. Google reCaptcha v.3 parameters (optional)
X. Session data

8. Analyzed parameters - benchmark

Examples of parameters taken by TwD monitoring codes. In order to protect the security of TwD monitoring processes from breaches or attempts at circumvention - the full list of monitored parameters in each area under investigation is not presented.

url/page and visit data

- website url
- elements on the page
- loading page data

- cookie handling
- visit badges
- number of visitors

- data transfer time
- page refresh

IP analysis

- identification of ISP
- geo-location
- connection identification - proxy, VPN, TOR network, data center infrastructure
- GSM network
- WebRTC with local IP numbers

Browser parameters

- browser name
- browser version
- user agent
- browser engine

- error handling
- function handling

- window size
- installed plug-ins

System parameters

- system name
- system version

- resolution
- image settings
- system settings

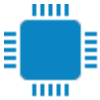
- font set

Device parameters

- microphone
- speakers
- sound card

- camera
- WiFi
- bluetooth

- location of the device



- processor
- platform

- kind of screen

Fingerprint of the device

- Unique identifier to determine the uniqueness of the device

Browser Fingerprint

- Unique identifier to determine the uniqueness of the viewer
- Audio context fingerprint
- Canvas Fingerprint

User behavioral analysis

- cursor coordinates

- scroll

- type and number of keys

- time of events on the site
- interaction with elements

- time on form fields
- number of characters in the fields
- field handling

Script support - site rendering

- how to download and run the script
- script errors

- upload time
- methods of sending data

- script checkpoints

Google Score

- Google reCaptcha v3 evaluation

Session data

- Identification of visits based on cookie ID or fingerprint
- Number of visits
- Number of pages
- Frequency and repetition of visits

9. Click audit methodology

Methodology for auditing clicks through the TwD system - in accordance with the guidelines of the IAB organization (document "Click Measurement Guidelines", Version 1.0, Final Release May 12, 2009 [1]).

Description of the click audit process

For the audit of clicks and leads - the TwD system performs CLICK SCANNER, GOGLE ADS SCANNER and LEAD SCANNER services, verifying and evaluating each click and lead recorded by the TwD system for potential irregularities.

The TwD system works on the basis of technologies:

1. System/Browser Parameter Analysis - Analysis of bots/programs/scrapers impersonating various types of systems and browsers that, when checked, do not match the attributes of a particular type of system or browser. Verification of bot-lists.
2. identify advanced browser automation tools such as Selenium and PhantomJS.
3. Device Fingerprinting - unique device identifiers
4. canvas Fingerprinting - unique identifiers of browsers
5 Machine learning - classifiers and artificial intelligence algorithms.
6. Bot 'honeypots' - 'traps' that identify bots operating on the site.
7. user behavioral analysis (UBA) - behavioral analysis to identify unnatural user behavior patterns on the monitored site.
8 IP analysis - analysis of geo-locations, addresses and IP providers. Identification of IP addresses generated by data centers, TOR services, VPN, proxy servers. Identification of IP addresses on IP black lists.
9. Google reCAPTCHA v3 API - identifying bot vs human behavior.
10. visit data analysis - analysis of URL parameters, cookies.
11. tamper proofing - 'proof of work' algorithms that perform tasks, possible during the natural use of the browser.
12. site rendering monitoring - analysis of JavaScript support on the monitored site in terms of rendering the target page.

The TwD system measures clicks known as Resolved Click (see Section 2.4 [1]) occurring when a click received launches the Advertiser's target website or other page designed to provide the Advertiser with a commercial interaction with a potential user.

No personal information is collected in the click audit - no information is collected about the characters typed in the forms posted on the website, etc. The content of the page viewed by the user is not analyzed. The collected data is used only for the purpose of click classification and is not in any way linked to a specific user (e.g., to create a user profile for marketing or other purposes).

Methodology for measuring clicks

Clicks audited by the TwD system include only inputs from defined sources. According to the guidelines in [1], a method described as One-Click-Per-Impression Method is adopted as the basic counting methodology. Recording single clicks from sources gives flexibility at the time of settlement between Advertiser and Publisher. Reports can be adjusted for the appropriate time intervals of counted clicks to avoid, for example, counting repetitions in certain time intervals.
A diagram of the data flow and sequences sequentially executed from redirection after a click is presented in the diagram below:

1 Clicks are measured from the moment the test pixel is downloaded and the JavaScript posted on the monitored page is downloaded.
(2) In the next steps, the script executes in the user's browser, establishes a connection with the TrafficWatchDog application server, and starts downloading more data.
(3) Each of the listed activities, along with the received data, is saved in the database.

The process of data acquisition and processing consists of the following:

1. A Javascript file whose task is to collect data. The script is placed on the monitored site of the client. Beforehand, it is properly adjusted to the specifications of the monitored website.
2. webservice, a service that retrieves encrypted data collected by Javascript code placed on the Customer's site.
3. database, the purpose of which is to store data collected on the Customer's site,
4. click classification algorithms and AI, act as ongoing processes to analyze and evaluate incoming data.
5. The Customer Panel allows online analysis of classification results and has the ability to generate claim reports.

Data from the monitored customer site is collected on an ongoing basis. The code is placed only on the pages selected by the clients to be ultimately monitored.
Measuring a user's subsequent visits is done through two, independent mechanisms: the information contained in a cookie and/or the unique device identifier Fingerprint [2]. The first solution, as long as it is not deactivated, provides opportunities for direct tracking of user presence. The second solution is used when cookies are disabled or removed from the browser for some reason (for example, to remove presence traces).

Click classification methodology

The TrafficWatchDog system uses click classification techniques based on identifiers, a survey of user activity and developed patterns of data taken during user activity on the Client's website. The

classification used is intended to identify an invalid click (as defined in the document used [1]) involving accessing a URL for the purpose of manipulating click measurement activities or click-based advertising payments, without a legitimate intent to view site content, make a purchase or perform another type of legitimate conversion. The classification also includes click fraud in situations where a user unknowingly or fraudulently accesses information (for example, activity caused by virus infection or auto-click features).

The parameters monitored and taken during the click audit are divided into 10 main areas subject to individual evaluations - and making up the final click rating.
List of audited areas:

1. The way the page is rendered - this area includes the analysis of the correctness of JavaScript and test pixels implemented on the monitored page. Also set in the script are so-called honeypots or 'traps' identifying bots operating on the page. If irregularities are detected in the loading of the client's target page - the area will be classified as incorrect.
2. System parameters - a set of data identifying the operating system retrieved during JavaScript execution. Data extracted from HTTP headers, in particular from the User-agent, are also collected. If non-standard system parameters are detected - the area will be classified as invalid.
3. Browser parameters - data retrieved by the JavaScript API specific to a particular visit to a website. The parameters are part of the generated browser Fingerprint. They serve to verify the correctness of the User-agent entry and the possibility of its intentional substitution. Custom parameter values or bot tags - allow to identify bots, including bots using the so-called webdriver interface, i.e. the browser control mechanism implemented in Selenium or PhantomJS libraries, among others. In addition, a proof-of-work task is implemented, which performs tasks to identify the real browser.
4. User behavioral - a set of data collected over a finite period of time to determine human behavior on a website. Data provided from peripheral devices are analyzed: mouse and keyboard in the case of desktop computers or touch in the case of mobile device touch screens. The time intervals between performed actions and the locations of activity on the page are also measured. The collected parameters make it possible to identify unnatural behavior and exclude a human as a browser user - or confirm the lack of activity of a potential user on the monitored client's site. If unnatural behavioral patterns of a potential user are detected, the lack of user activity on the monitored Customer site - the area will be classified as abnormal.
5. Device parameters - within the collected parameter values, among other things, tests are made related to the canvas and WebGL interfaces provided in HTML 5, as well as the so-called audio context and WebRTC. Some of these parameters then affect the generated Fingerprint of the device and are used during IP address analysis.
6. IP address analysis. Using external databases of IP numbers, a range of data related to a browser user's IP number is identified, such as ISP name and geo-location. Information related to a potential user's use of a TOR network, VPN connection or proxy server is also checked. This allows the detection of potential abuse and changes in IP numbers to hide traces of abuse.
7. Device Fingerprint is a distinctive device identifier generated from selected device parameters. With its unique value, it allows to replace the cookie to identify the user. It consists of Fingerprint generated from canvas, webgl and audio context, among others. If an abnormality is detected in the generation of the device's Fingerprint - the area will be classified as abnormal.
8. Browser Fingerprint is a unique browser fingerprint generated from parameters provided by the browser API. Along with the Fingerprint of the device allows you to replace the cookie and identify repeated clicks. If an irregularity is detected in the generation of the browser fingerprint - the area will be classified as invalid.
9. Google Score. Google's external tool reCAPTCHA v3 API is used, which allow to identify bot vs human behavior (rating scale is adopted: null - 100). Google scores indicating invalid traffic - classify a given area of the clicks score as invalid

10. Session data. Verification of individual visits, their number, pages visited and frequency of visits. If anomalies in session data are detected - the area will be classified as invalid.

If irregularities are detected in a minimum of one of the analyzed 10 areas of the evaluation of a click - a click can be classified as an invalid click with an indication of the areas in which the invalid parameters were tracked.

Click assessment reporting

The TrafficWatchDog system provides the Client's administration panel, where the Client can track real-time monitoring and ratings of clicks/leads conducted on his monitored site. Parameters available for viewing and possible inclusion in the generated report may include the following information:
1. information about the classification of the click and the values of the individual components that make up the final score.
2. time to enter the client's site.
3. IP number, user locations, ISP information
4. unique device and canvas Fingerprint
5. user behavioral visualizations of the potential user's behavior on the site.
6. Clicks/leads reports (current day, date range, monthly, etc.) by source,
7. information about the browser, operating system, device
8 Google Score.

For reporting and complaint purposes, the data is tailored and provided to Clients depending on their needs (structure and scope of TwD reports). Clients' pages may additionally be monitored for correctness of clicks after clicking on the target link to the Client's page - located on the publishers' page. Test clicks will be marked as invalid clicks in reports to the Client. Data collected during click monitoring is stored for a minimum of 3 months including the month under study.

Additional materials

Interactive Advertising Bureau, "Click Measurement Guidelines," Version 1.0 - Final Release May 12, 2009.

IAB Online Lead Generation: Lead Quality Accountability Best Practices for Advertisers and Publishers Released December 2008