

Ziel dieses Whitepapers

Der Markt der Online Werbung sieht drei Herausforderungen, welche die Akzeptanz der digitalen Medien als Werbeträger deutlich beeinträchtigen können: 1. Invalid Traffic (Ad Fraud), 2. Brand Safety und 3. Sichtbarkeit. Die größte dieser Herausforderung ist die Identifikation von Werbekontakten, die nicht den Anforderungen der Buchung entsprechen und welche die Werbekunden daher potentiell nicht bezahlen wollen und sollen.

Diese unerwünschten Werbekontakte lassen sich unterscheiden in solche, die aus einer kriminellen Energie heraus entstehen und daher betrügerisch sind, und solche, die ohne betrügerische Absicht, z.B. wegen fehlerhaften Ad Server oder Targeting-Tool Einstellungen, oder wegen technischer Ursachen nicht den Buchungsanforderungen entsprechen.

In den letzten Jahren wurden von der Online Marketing Industrie daher Verfahren und Systeme entwickelt, diese Werbekontakte zu identifizieren und in Reportings für die Werbekunden abzubilden. Sehr gerne werden diese nicht buchungskonformen Werbekontakte als Ad Fraud bezeichnet.

Das Problem hierbei ist jedoch, dass eine klare Definition des Begriffes "Ad Fraud" nicht existiert. Es gibt auch keine zentralen Vorgaben von Seiten des Internet Advertising Bureau (IAB), welche der ermittelten Metriken zu einer zentralen "Ad Fraud Rate" zusammengefasst werden dürfen. Daher schwanken die Werte, die von den verschiedenen Anbietern sogenannter Ad Verification Systeme am Markt zum Thema Ad Fraud erhoben und veröffentlicht werden, zum Teil deutlich. Das führt im Tagesgeschäft oftmals zu Unstimmigkeiten zwischen den buchenden Agenturen und Kunden und den Vermarktern der Werbeflächen. Zusätzlich führt es zu Verwirrungen und Verunsicherung, wenn die Anbieter mit sehr verschiedenen Zahlen für den Gesamtmarkt an die Öffentlichkeit gehen.

Die Fokusgruppe Digital Marketing Quality (DMQ) im BVDW hat sich daher das Ziel gesetzt, mit diesem White Paper eine Grundlage für ein standardisiertes Reporting und damit für eine Zertifizierung der am Markt angebotenen Systeme für die Erkennung von "Invalid Traffic" zu schaffen.

Im Folgenden werden wir daher sowohl die am Markt gängigen Metriken erläutern, wie auch Definitionen festlegen, die für eine Standardisierung der Reportings genutzt werden können und somit eine Grundlage für eine Zertifizierung der Systeme darstellen.

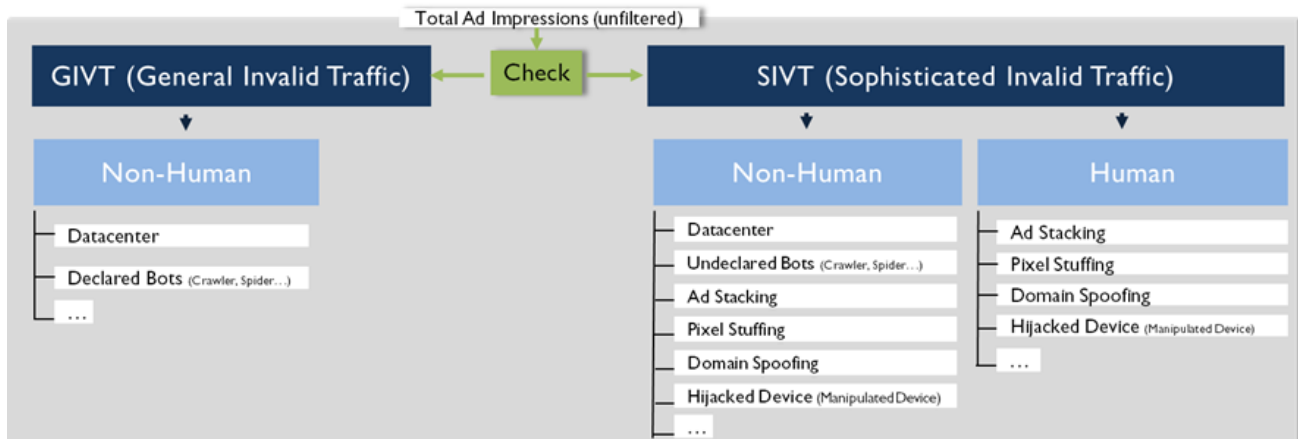
1. Invalid Traffic statt Ad Fraud - denn nicht alles was ungültig ist, muss Betrug sein

Die am Markt verfügbaren technischen Systeme für die Erkennung von "ungültigen" Werbekontakten verwenden in der Regel nicht den Begriff "Fraud", also Betrug. Der Grund hierfür ist einfach, denn Betrug ist laut juristischer Definition (vgl. §263 StGB) eine Täuschung über Tatsachen, die zu einem Irrtum des Opfers führen. Dieser Irrtum muss eine Vermögensverfügung auslösen, die wiederum einen Vermögensschaden entstehen lässt. Ein Betrug wird mit Freiheitsstrafe bis zu fünf Jahren oder mit Geldstrafe bestraft.

Uns stehen in der Online Werbung zum Teil nicht die Mittel für eine korrekte Differenzierung zwischen Betrug und versehentlicher Nicht-Erfüllung der Buchungsbedingungen zur Verfügung. Zudem gibt es zahlreiche Werbekontakte die von sogenannten Bots ohne eine kriminelle Absicht verursacht werden und die ein Werbungtreibender trotzdem nicht bezahlen möchte. Wir sind mit den technischen Systemen in der Lage zu erkennen, dass bestimmte Werbekontakte Indizien aufweisen, welche den Verdacht nahelegen, dass diese

Werbekontakte nicht von Menschen verursacht wurden und / oder nicht in das Muster eines korrekt ausgelieferten Werbekontaktes passen.

Aus diesem Grund soll zukünftig nicht mehr von "Ad Fraud Rate" die Rede sein. Viel mehr werden wir von "Invalid Traffic Rate (IVT Rate)" sprechen, also der Rate der Werbekontakte, die "nicht gültig" sind.



Im Allgemeinen ist es wichtig zwischen zwei Arten von ungültigen Werbekontakten zu unterscheiden. Den sogenannten "Nicht-Menschlichen Kontakten" (Non Human Traffic) und den "Menschlichen Kontakten" (Human Traffic) und dennoch ungültigen, weil manipulierten und somit verdächtigen Werbekontakten, die keine Werbewirkung erzielen können.

1.1 General Invalid Traffic (GIVT):

Alle Werbekontakte, die von "gutartigen" Bots ausgelöst werden, die sich zu erkennen geben und somit den Systemen des Online Marketing (z.B. DSPs, Ad Server, Proxy, usw.) anzeigen, dass sie als nicht abrechnungsrelevant aus den Reportingdaten rausgefiltert werden können. Diese Bots werden in der sogenannten IAB/ABC International Spider and Bots List gesammelt, welche von nahezu allen technischen Systemen für die Bereinigung der gemessenen Zahlen zugrunde gelegt wird. Diese Liste wird laufend in einem festen Turnus aktualisiert.

Werbekontakte, die von Menschen verursacht werden, jedoch z.B. über ein virtuelles, privates Netzwerk (VPN) laufen (z.B. Firmennetzwerke), oder über andere, als sauber einzustufende Proxy-Netzwerke gelenkt werden, gelten als valide Werbekontakte und werden nicht als GIVT eingestuft.

GIVT Rate:

$$\text{GIVT Impression} \times 100 / \text{Total measured Impressions} = \text{GIVT Rate}$$

Neben den gutartigen, nicht-menschlichen Werbekontakten gibt es auch solche, die von sogenannten, bössartigen Bots ausgelöst werden. Diese geben sich nicht zu erkennen, sind nicht Teil der offiziellen IAB Liste und versuchen meistens menschliches Verhalten nachzuahmen. Für die Erkennung solcher nicht-menschlicher Werbekontakte, die potentiell betrügerisch sein können und sich daher geschickt zu tarnen versuchen, hat die Online Marketing Industrie zum Teil sehr komplexe Verfahren entwickelt. Diese geschickt getarnten Impressions reporten die Ad Verification Systeme als Teil des "Sophisticated Invalid Traffic" (SIVT).

1.2 Sophisticated Invalid Traffic (SIVT):

Unter dem SIVT werden alle Werbekontakte erfasst, welche durch zum Teil aufwändige Verfahren als Non-Human, also nicht von Menschen verursacht, oder von Menschen verursacht, aber manipuliert und somit nicht Werbewirkung erzeugend (z.B. weil das Werbemittel versteckt wird) eingestuft werden. Die Verfahren zur Erkennung sind je nach Anbieter verschieden. Daher erkennt der eine oder andere Anbieter für die Erkennung von Invalid Traffic evtl. mehr oder weniger SIVT Impressions.

Die **SIVT Rate** zeigt auf, wie hoch der Anteil der als SIVT eingestuft Impressions in Relation zu den total measured Impressions oder zu den bereits um den GIVT bereinigten Anteil der measured Impressions ist. Das IAB legt hier derzeit keine Regel für die Ausweisung der SIVT Rate fest. Die SIVT Rate kann daher wie folgt im Reporting eines Anbieters für die Erkennung von Invalid Traffic erscheinen. Sofern der Anbieter nicht von sich aus die Formel für die Berechnung der SIVT Rate nennt, sollte diese bei Bedarf erfragt werden.

SIVT Rate (based on total measured impressions):

$$\text{SIVT Impressions} \times 100 / \text{Total measured Impressions} = \text{SIVT Rate}$$

SIVT Rate (based on total measured impressions, abzgl. GIVT impressions):

$$\text{SIVT Impressions} \times 100 / (\text{Total measured Impressions} - \text{GIVT}) = \text{SIVT Rate}$$

1.3 Invalid Traffic (IVT):

Die Summe der Impressions, die als GIVT und SIVT ermittelt wurden, ergibt die Menge der insgesamt als Invalid Traffic einzustufenden Impressions.

IVT Rate:

$$(\text{GIVT Imps.} + \text{SIVT Imps.}) \times 100 / \text{Total measured Impressions} = \text{IVT Rate}$$

2. Formen der ungültigen / verdächtigen Werbekontakte

Im Folgenden geben wir einen Überblick über die verschiedenen Arten der ungültigen oder verdächtigen Werbekontakte und unterscheiden diese wie folgt:

GIVT vs. SIVT

Verdächtig und daher bilateral zu klären

Titel	Kurzbeschreibung	Einordnung
Auto-Refresh	Werbeplätze werden während einer Page Impression oder in einer App (regelmäßig, in verdächtig kurzen Intervallen, ohne Nutzer-Interaktion) neu geladen und erzeugen somit viele Ad Impressions.	verdächtig, bilaterale Klärung
Geo Fraud	Auslieferung ohne Einhaltung des gebuchten Geo Targetings, eine verdächtig hohe Anzahl an Impressions erfüllt nicht die Targeting Kriterien	verdächtig, bilaterale Klärung
Cash for Surfing	Menschliche User, die jedoch für das Werbung ansehen und ggf. klicken bezahlt werden	verdächtig, bilaterale Klärung
Domain oder App Spoofing / Laundering	Vortäuschung von falschen Traffic-Quellen (Domains oder Apps) z.B. beim programmatischen Werbeflächen Einkauf, um bessere Preise zu erzielen oder eine Auslieferung auf Seiten zu erschleichen, die wegen des Targetings oder der Brand Safety Einstellungen von der DSP abgelehnt worden wären	SIVT
Hidden Ads	Werbemittel werden so ausgeliefert, dass sie vom Nutzer nicht wahrnehmbar sind. Beispiele: <i>AdStacking:</i> Das Übereinanderstapeln von Werbeanzeigen in einem iFrame beispielsweise so, dass nur die oberste Anzeige sichtbar ist. <i>Pixel Stuffing:</i> Das Laden von Werbemitteln oder ganzen redaktionellen Seiten in einen 1x1 Pixel großen Bereich.	SIVT
Surfbars	Programme, die beim Surfen auf dem Rechner des Users automatisch weitere Seiten laden, die der User nicht gezielt selber aufruft und daher wahrscheinlich nicht betrachtet	SIVT
Ad Injection	Bots legen Banner, die von betrügerischen Werbenetzwerken stammen, über vorhandene Werbeplätze auf Content-Seiten	SIVT

Click Bots	Automatisiertes anklicken von Werbemitteln durch bössartige Bots.	SIVT
Click Farm	Anklicken von Werbemitteln durch extra dafür bezahlte Personen, ohne Interesse für das beworbene Produkt.	SIVT
Hijacked Device / Adware	Webseiten werden auf dem Rechner eines Nutzers im Hintergrund aufgerufen, ohne dass der jeweilige User es merkt oder das Laden selber initiiert hat.	SIVT
Bot Fraud	nicht menschliche User die sich als menschliche User ausgeben und so Werbekontakte erzeugen	SIVT
Gutartige Bots	nicht menschlicher Traffic, der von bots ausgelöst wird, die zB für Forschung, Preisvergleiche, Suchmaschinen Listings eingesetzt werden, und sich daher den technischen Systemen im online Marketing zu erkennen geben und Teil der IAB Liste sind	GIVT
Ad Clutter / Ad Density	Das Verhältnis zwischen Content und Anzahl der Werbeflächen in einem Standard-/Premiumformat (vgl. werbeformen.de), die zeitgleich im Sichtbereich zu sehen sind, ist so ungünstig, dass sich der Verdacht ergibt, die Website dient alleine dazu Werbeflächen zu laden und so Ad Impressions zu erzeugen. (Richtwert: Mehr als 4 der genannten Formate, die nicht Teil eines zusammengesetzten Werbeformats (z.B. Wallpaper, Fireplace Ad) sind, befinden sich im Sichtbereich)"	verdächtig, bilaterale Klärung
Data Center Traffic	<p>Data Center Traffic wird generell als nicht menschlicher Traffic eingestuft. Er kann in "bekanntem" und "unbekanntem" Data Center Traffic unterschieden werden.</p> <p>Bekannter Data Center Traffic ist jede Form von Traffic (Impressions und Clicks) der durch bekannte IP Adressen von bekannten, offiziell gelisteten Daten Centern (Servern), verursacht wird. Diese Server können z.B. der Ausgangspunkt von bekannten Bots / Crawlern sein, die in der IAB Liste aufgeführt sind.</p> <p>Ausnahmen stellen bekannte Proxy-Server dar, die von Nutzern verwendet werden, um z.B. Internet-Zugang über ein Corporate Gateway zu erhalten (z.B.</p>	<p>Bekannter Data Center Traffic = GIVT</p> <p>Unbekannter Data Center Traffic = SIVT</p>

	<p>VPN Netzwerke). Traffic der über solche Netzwerke erzeugt wird, gilt nicht als Invalid Traffic.</p> <p>Unbekannter Data Center Traffic wird durch Server erzeugt, die nicht bekannt sind und somit nicht von den Adservern als Teil des GIVT gefiltert werden können.</p>	
Ad Tag oder Creative Hijacking	<p>Legitime Ad-Call Codes (Anfragen an den Adserver) bzw. Ad Tags (HTML das die Werbemittel beinhaltet) einer Seite werden kopiert und ohne Zustimmung des Publishers bzw. des Advertisers an anderer Stelle eingebunden.</p>	SIVT

3. Standard Metriken für Ad Verification Systeme:

Der BVDW empfiehlt als Basis für eine Standardisierung des Reportings der Systeme die folgenden Metriken als Mindestanforderung.

- Total measured and unfiltered Impressions
- GIVT Impressions
- GIVT Rate
- Ad Impressions (= Total measured Impressions abzgl. GIVT)
- SIVT Impressions
- SIVT Rate
- IVT Impressions
- IVT Rate
- Total valid Impressions*

*Definition der Total Valid Impressions:

Die Menge der ausgelieferten Impressions, reduziert um die als IVT eingestufted Impressions stellt die Total Valid Impressions dar. Diese können auch ermittelt werden, indem von der Menge der Ad Impressions die Menge der als SIVT eingestufted Impressions abgezogen wird.

Für eine bessere Kontrolle und mehr Transparenz in den gemessenen Werten, wie auch für die gezielte Bekämpfung von vermeintlich betrügerischen Werbekontakten, sollte zudem eine Aufschlüsselung der SIVT Daten erfolgen.

Der BVDW empfiehlt in Anlehnung an das IAB eine Aufsplittung wie folgt:

- Bot / Spider Traffic (masquerading as legitimate users)
- Data Center Traffic (incl. Invalid Proxy Traffic)
- Hidden Ads
- 1x1Pixel Stuffing/stacked Ads/covered or otherwise intentionally obfuscated ad serving
- Adware / Malware
- Incentivized manipulation of measurements
- falsified viewable impression decisions
- Domain Spoofing / Laundering
- Hijacked Devices
- Cookie Stuffing, recycling, harvesting
- Manipulation or falsification of location data

Mitwirkende Personen:

Jonas Kuhnle (Meetrics)
Andreas Hamdorf (pilot)
Tobias Wegmann (Prex)
Claudia Gröner (AGOF)
Diana Czerny (Seven One)

Sascha Kalabuchow (GP One)
Christoph Bauer (ePrivacy)
Stefan Bergmeier (IP Deutschland)
Franziska Gerit John (The ADEX)

Kontakt BVDW: Sabrina Sallach, Referentin Digital Advertising (sallach@bvdw.org)

April 2019

