mpire™

Identifying and Combating Fraud
to Optimize Ad Network Buys

Marissa Gluck, Radar Research

Catalyst

The promise of aggregated, targeted ad networks and efficient exchanges remains largely unfulfilled, even though up to 80% of online ads are sold and resold through third parties (IAB/Bain, 2008). Advertisers deploying campaigns on many ad networks are vexed by the lack of transparency and the suspicion that there is a significant amount of click and impression fraud taking place, leading to material waste in campaign budgets. Some ad networks, while professing to ensure brand safety and traffic quality, take a laissez-faire approach to policing downstream traffic providers. At the same time, ad agencies and advertisers often turn a blind eye to click fraud, as long as campaign metrics are met. Nonetheless, there is substantial value to be extracted from ad network buys when the proper measures are taken to analyze and optimize the traffic.

Core Finding / Hypothesis

In a difficult economy, the click/impression fraud exigency has created the need for reliable third-party verification, particularly as online advertising competes for budget with other forms of marketing. For marketers to continue to migrate dollars online, the industry needs to offer transparency, brand assurance, measurable user engagement, and verification. Ad networks need to implement solutions that push their accountability towards verifying the legitimacy of not just clicks, but also impressions. Ad networks and third-party technologies that provide such true transparency today are leading the way towards an efficient marketplace where agencies and brands can buy with confidence across broad swaths of remnant inventory.

Click Fraud Continues to Confound Advertisers and Publishers Alike

Ad networks and exchanges initially suffered the ignoble perception as a repository for remnant inventory – the "bottom of the barrel" of advertising inventory. Over the past decade, ad networks have challenged that perception, becoming increasingly legitimate as an alternative media practice to buying directly on premium editorial sites. Yet click fraud remains a problem for both advertisers and networks seeking to shed old perceptions. While the industry acknowledges it exists, click/impression fraud has been difficult to prove or quantify, and as a result has slowed the validation of the ad network and exchange models. Networks need to incorporate or improve on existing processes that identify and excise instances of fraud in order to gain the trust of both advertisers and publishers who contribute their inventory.

In July 2009, campaign verification and optimization solution provider Mpire conducted a test to identify instances of click and impression fraud occurring via some inventory aggregators (ad networks and media exchanges). The results of the test were illuminating, revealing fraud is far more prevalent than previously assumed, and easier to detect with the proper tools and insights.

Key Definitions

- In-view: 50% or more of the online advertising unit area is within a viewable area (i.e., able to be viewed by a user) of the browser, regardless of initially being above the fold or below the fold and subsequently scrolled into view.
- Engagement: Human mouse entry into an online advertising unit

Key Assumptions

Mpire, using its AdXpose technology, executed a test with several key assumptions in mind:

- The baseline for identifying suspected fraud sites were publishers yielding either a click through rate greater than 2% on untargeted, run-of-network (RON) inventory, or a disproportionate amount of clicks to engagements (i.e. if only 10 engagements occur, but 20 clicks occur, fraud is likely).
- The widespread use of I-frames within media servers and on publisher sites on ad networks affects "in-view" data, resulting in higher engagement rates due to "false positives."

Methodology

Mpire conducted 11 RON buys across nine different ad networks (directly and via one exchange) in July. The buys delivered more than 20 million impressions, to ads from 53 different advertisers. These impressions were filled by the initial nine ad networks via downstream daisy-chaining on at least 45 additional ad networks on more than 100,000 sites.

Mpire conducted the test with two distinct types of campaigns at the exchange level. The first was an actively managed campaign where ads were served to reputable site lists, pre-approved by the advertiser, thus reducing concerns around fraud. In the second RON campaign, advertisers executing a RON had no visibility into where the ads will be served and to whom they'll be served. The goal of the test was to verify the following:

1) The URLs where ads were actually served;
2) The location of the ads on each URL;
3) Whether the ads had the opportunity to be viewed (in-view); and
4) User interaction, in order to detect impression fraud and optimize the creative

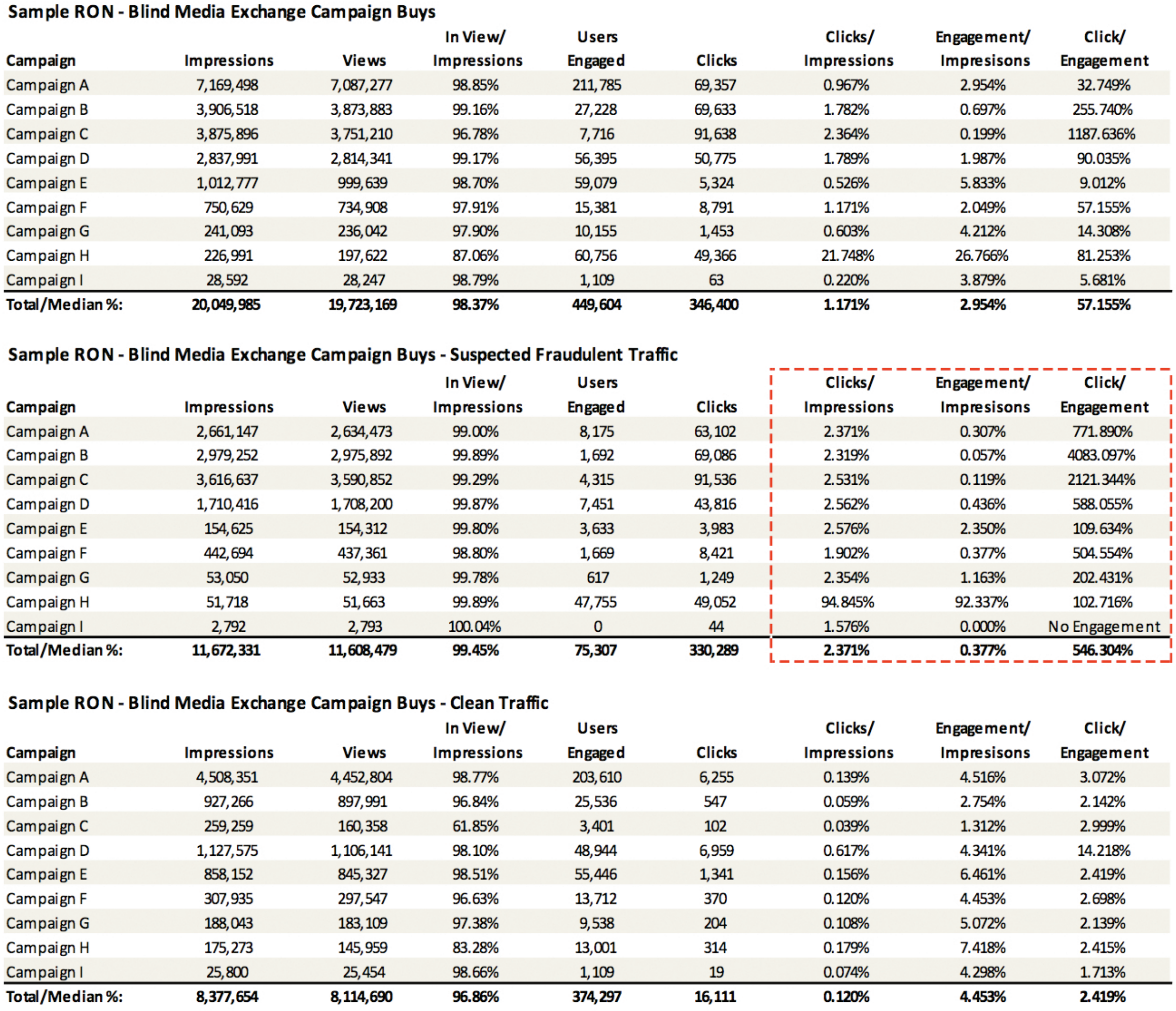
Within the RON campaign, Mpire also tracked the level of user interaction with the creative prior to click events in order to isolate fraud. Finally, Mpire conducted a third test on a major ad network as well, resulting in a "before and after" snapshot of specific campaign performance optimization enabled by Mpire.

Run-of-Network Campaign Revealed Prevalence of Click and Impression Fraud

Mpire ran an additional test on a media exchange, with nine RON buys. The results were startling -- more than half of the impressions delivered and 95% of clicks came from suspected fraudulent sources. Low or non-existent user engagements, combined with high click through rates, are hallmarks of fraudulent traffic. When a click occurs without a mouse entry being registered, the result is highly likely to be a fraudulent click. Fraudulent clicks are, however, simply the symptom of the larger problem of impression fraud; the clicks are generated to lead advertisers to believe the impressions that generated the clicks are valid. This is a much larger problem than simple click fraud, as advertisers are actually paying CPM prices for large tranches of botted impressions.

Further complicating advertiser efforts to track ad network campaigns, many of the sites in these exchanges use multiple layers of i-frames. As a result, nefarious sites are able to hide fraudulent traffic behind numerous layers of nested i-frames, leaving advertisers blind to in-view data.

We believe the default trafficking behavior of many RON buys is to include obviously fraudulent and well-known botted sites. It is not necessarily true that all marketplace/exchange traffic is bad, but rather that exchanges simply include nefarious inventory that could and should be blocked, but for some reason continues to maintain a presence. Most of these campaigns are automatically optimized based on CTR and/or eCPM, so this fraudulent traffic, if left unchecked, will dominate the contracted volume. With simply a little bit of effort to universally exclude these known-fraudulent sites, legitimate marketplace/exchange traffic can begin to be properly trafficked.



Sample RON - Blind Media Exchange Campaign Buys

			In View/	Users		Clicks/	Engagement/	Click/
Campaign	**Impressions**	**Views**	**Impressions**	**Engaged**	**Clicks**	**Impressions**	**Impresisons**	**Engagement**
Campaign A	7,169,498	7,087,277	98.85%	211,785	69,357	0.967%	2.954%	32.749%
Campaign B	3,906,518	3,873,883	99.16%	27,228	69,633	1.782%	0.697%	255.740%
Campaign C	3,875,896	3,751,210	96.78%	7,716	91,638	2.364%	0.199%	1187.636%
Campaign D	2,837,991	2,814,341	99.17%	56,395	50,775	1.789%	1.987%	90.035%
Campaign E	1,012,777	999,639	98.70%	59,079	5,324	0.526%	5.833%	9.012%
Campaign F	750,629	734,908	97.91%	15,381	8,791	1.171%	2.049%	57.155%
Campaign G	241,093	236,042	97.90%	10,155	1,453	0.603%	4.212%	14.308%
Campaign H	226,991	197,622	87.06%	60,756	49,366	21.748%	26.766%	81.253%
Campaign I	28,592	28,247	98.79%	1,109	63	0.220%	3.879%	5.681%
Total/Median %:	**20,049,985**	**19,723,169**	**98.37%**	**449,604**	**346,400**	**1.171%**	**2.954%**	**57.155%**

Sample RON - Blind Media Exchange Campaign Buys - Suspected Fraudulent Traffic

			In View/	Users		Clicks/	Engagement/	Click/
Campaign	**Impressions**	**Views**	**Impressions**	**Engaged**	**Clicks**	**Impressions**	**Impresisons**	**Engagement**
Campaign A	2,661,147	2,634,473	99.00%	8,175	63,102	2.371%	0.307%	771.890%
Campaign B	2,979,252	2,975,892	99.89%	1,692	69,086	2.319%	0.057%	4083.097%
Campaign C	3,616,637	3,590,852	99.29%	4,315	91,536	2.531%	0.119%	2121.344%
Campaign D	1,710,416	1,708,200	99.87%	7,451	43,816	2.562%	0.436%	588.055%
Campaign E	154,625	154,312	99.80%	3,633	3,983	2.576%	2.350%	109.634%
Campaign F	442,694	437,361	98.80%	1,669	8,421	1.902%	0.377%	504.554%
Campaign G	53,050	52,933	99.78%	617	1,249	2.354%	1.163%	202.431%
Campaign H	51,718	51,663	99.89%	47,755	49,052	94.845%	92.337%	102.716%
Campaign I	2,792	2,793	100.04%	0	44	1.576%	0.000%	No Engagement
Total/Median %:	**11,672,331**	**11,608,479**	**99.45%**	**75,307**	**330,289**	**2.371%**	**0.377%**	**546.304%**

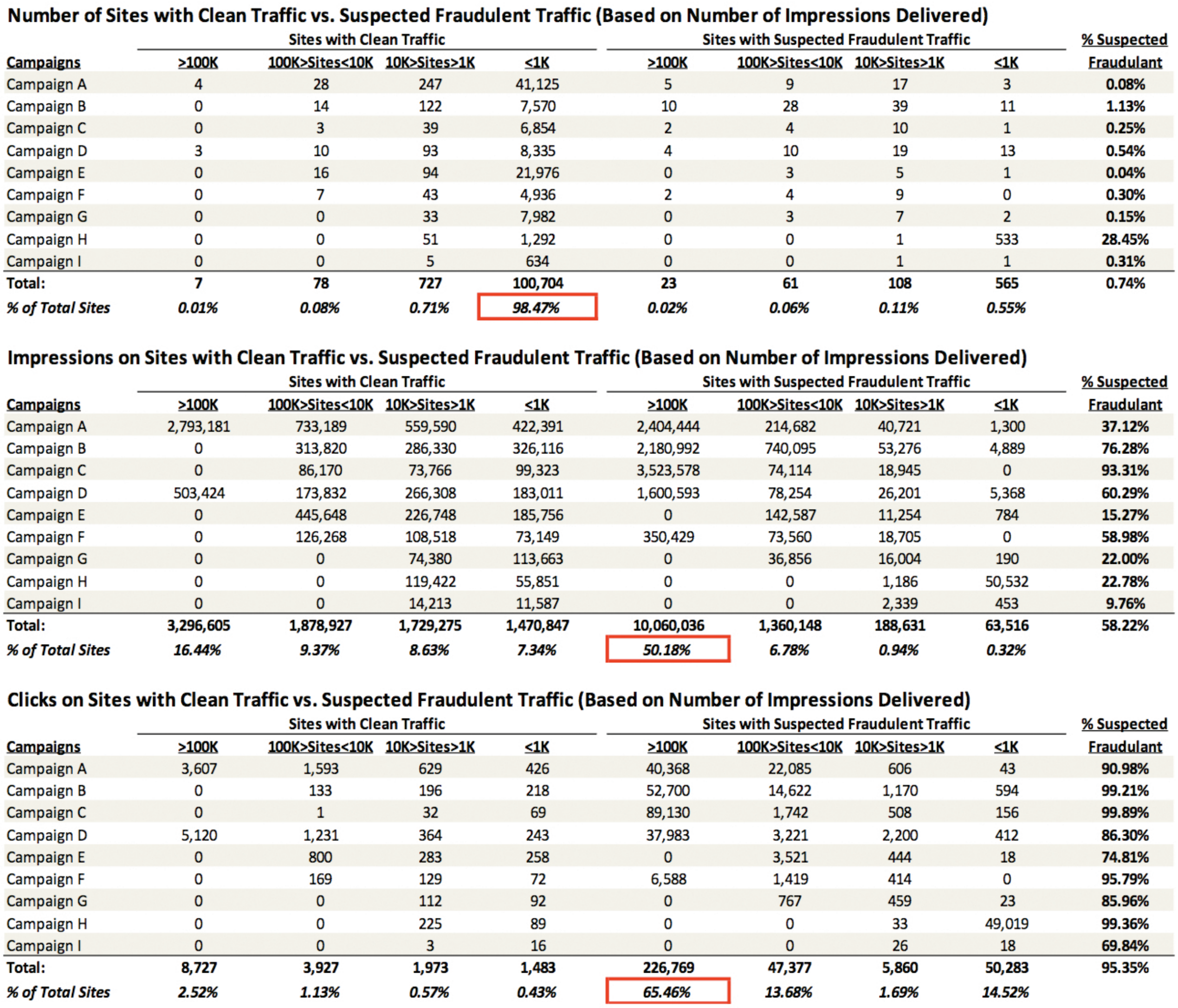
Sample RON - Blind Media Exchange Campaign Buys - Clean Traffic

			In View/	Users		Clicks/	Engagement/	Click/
Campaign	**Impressions**	**Views**	**Impressions**	**Engaged**	**Clicks**	**Impressions**	**Impresisons**	**Engagement**
Campaign A	4,508,351	4,452,804	98.77%	203,610	6,255	0.139%	4.516%	3.072%
Campaign B	927,266	897,991	96.84%	25,536	547	0.059%	2.754%	2.142%
Campaign C	259,259	160,358	61.85%	3,401	102	0.039%	1.312%	2.999%
Campaign D	1,127,575	1,106,141	98.10%	48,944	6,959	0.617%	4.341%	14.218%
Campaign E	858,152	845,327	98.51%	55,446	1,341	0.156%	6.461%	2.419%
Campaign F	307,935	297,547	96.63%	13,712	370	0.120%	4.453%	2.698%
Campaign G	188,043	183,109	97.38%	9,538	204	0.108%	5.072%	2.139%
Campaign H	175,273	145,959	83.28%	13,001	314	0.179%	7.418%	2.415%
Campaign I	25,800	25,454	98.66%	1,109	19	0.074%	4.298%	1.713%
Total/Median %:	**8,377,654**	**8,114,690**	**96.86%**	**374,297**	**16,111**	**0.120%**	**4.453%**	**2.419%**

Figure 1: 95% of clicks on RON exchange buy appear fraudulent

URL Padding is Prevalent

Mpire's test also revealed another common ad network practice – URL padding, or the practice of providing a site list representing the buy but actually delivering the majority of the buy via only a few of the sites. Sites in the media exchange test were distributing impressions across thousands of URLS, yet an 80-20 rule was discovered: a tiny percentage of referring URLs accounted for the majority of the impression volume. The data revealed that 98% percent of the traffic was delivered via just 1.5% of the URLs.



Number of Sites with Clean Traffic vs. Suspected Fraudulent Traffic (Based on Number of Impressions Delivered)

	Sites with Clean Traffic				Sites with Suspected Fraudulent Traffic				% Suspected
Campaigns	**>100K**	**100K>Sites<10K**	**10K>Sites>1K**	**<1K**	**>100K**	**100K>Sites<10K**	**10K>Sites>1K**	**<1K**	**Fraudulant**
Campaign A	4	28	247	41,125	5	9	17	3	**0.08%**
Campaign B	0	14	122	7,570	10	28	39	11	**1.13%**
Campaign C	0	3	39	6,854	2	4	10	1	**0.25%**
Campaign D	3	10	93	8,335	4	10	19	13	**0.54%**
Campaign E	0	16	94	21,976	0	3	5	1	**0.04%**
Campaign F	0	7	43	4,936	2	4	9	0	**0.30%**
Campaign G	0	0	33	7,982	0	3	7	2	**0.15%**
Campaign H	0	0	51	1,292	0	0	1	533	**28.45%**
Campaign I	0	0	5	634	0	0	1	1	**0.31%**
Total:	**7**	**78**	**727**	**100,704**	**23**	**61**	**108**	**565**	**0.74%**
% of Total Sites	***0.01%***	***0.08%***	***0.71%***	***98.47%***	***0.02%***	***0.06%***	***0.11%***	***0.55%***	

Impressions on Sites with Clean Traffic vs. Suspected Fraudulent Traffic (Based on Number of Impressions Delivered)

	Sites with Clean Traffic				Sites with Suspected Fraudulent Traffic				% Suspected
Campaigns	**>100K**	**100K>Sites<10K**	**10K>Sites>1K**	**<1K**	**>100K**	**100K>Sites<10K**	**10K>Sites>1K**	**<1K**	**Fraudulant**
Campaign A	2,793,181	733,189	559,590	422,391	2,404,444	214,682	40,721	1,300	**37.12%**
Campaign B	0	313,820	286,330	326,116	2,180,992	740,095	53,276	4,889	**76.28%**
Campaign C	0	86,170	73,766	99,323	3,523,578	74,114	18,945	0	**93.31%**
Campaign D	503,424	173,832	266,308	183,011	1,600,593	78,254	26,201	5,368	**60.29%**
Campaign E	0	445,648	226,748	185,756	0	142,587	11,254	784	**15.27%**
Campaign F	0	126,268	108,518	73,149	350,429	73,560	18,705	0	**58.98%**
Campaign G	0	0	74,380	113,663	0	36,856	16,004	190	**22.00%**
Campaign H	0	0	119,422	55,851	0	0	1,186	50,532	**22.78%**
Campaign I	0	0	14,213	11,587	0	0	2,339	453	**9.76%**
Total:	**3,296,605**	**1,878,927**	**1,729,275**	**1,470,847**	**10,060,036**	**1,360,148**	**188,631**	**63,516**	**58.22%**
% of Total Sites	***16.44%***	***9.37%***	***8.63%***	***7.34%***	***50.18%***	***6.78%***	***0.94%***	***0.32%***	

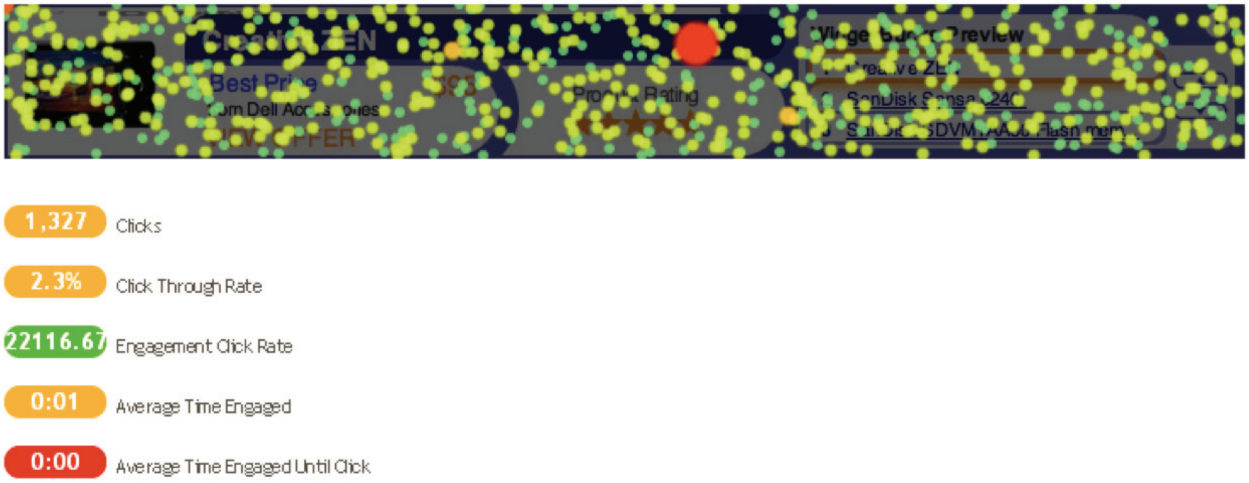
Clicks on Sites with Clean Traffic vs. Suspected Fraudulent Traffic (Based on Number of Impressions Delivered)

	Sites with Clean Traffic				Sites with Suspected Fraudulent Traffic				% Suspected
Campaigns	**>100K**	**100K>Sites<10K**	**10K>Sites>1K**	**<1K**	**>100K**	**100K>Sites<10K**	**10K>Sites>1K**	**<1K**	**Fraudulant**
Campaign A	3,607	1,593	629	426	40,368	22,085	606	43	**90.98%**
Campaign B	0	133	196	218	52,700	14,622	1,170	594	**99.21%**
Campaign C	0	1	32	69	89,130	1,742	508	156	**99.89%**
Campaign D	5,120	1,231	364	243	37,983	3,221	2,200	412	**86.30%**
Campaign E	0	800	283	258	0	3,521	444	18	**74.81%**
Campaign F	0	169	129	72	6,588	1,419	414	0	**95.79%**
Campaign G	0	0	112	92	0	767	459	23	**85.96%**
Campaign H	0	0	225	89	0	0	33	49,019	**99.36%**
Campaign I	0	0	3	16	0	0	26	18	**69.84%**
Total:	**8,727**	**3,927**	**1,973**	**1,483**	**226,769**	**47,377**	**5,860**	**50,283**	**95.35%**
% of Total Sites	***2.52%***	***1.13%***	***0.57%***	***0.43%***	***65.46%***	***13.68%***	***1.69%***	***14.52%***	

Figure 2: 98% of the traffic delivered via 1.5% of the URLs; heavy impression (58%) and click (95%) fraud.

Mpire Maps Click Fraud Patterns

Mpire's technology can represent engagement and click patterns visually, allowing advertisers to easily identify click fraud. In the example below, the creative unit is deployed in Flash, giving users the ability to scroll through offers. Normal viewer behavior would coalesce around specific points in the creative, which correspond with positions for interaction and shopping. Instead, the randomized clicks suggest programmed click fraud. In other types of click fraud, Mpire's mapping technology reveals click fraud via numerous clicks on single pixels within the creative that do not align with offers or action points, or grouped clicks near the borders of the unit.



1,327	Clicks
2.3%	Click Through Rate
22116.67	Engagement Click Rate
0:01	Average Time Engaged
0:00	Average Time Engaged Until Click

Figure 3: A clear example of botted click fraud – Mpire alerts advertisers when the "engagement click rate" exceeds a base threshold

Leveraging Data Yields Positive Results

Mpire also conducted a test directly on a top ten ad network, with the goal of proving the power of referrer and fraud data to unlock the hidden value of horizontal networks. The impressions were bought on a CPM-basis, rather than on a CPC or CPA basis. The results were revealing. While a large percentage (50%) of the impressions were never within view of a user, the click/impression fraud volume, while substantial, was significantly lower than on the exchange based buys. Before optimization based on Mpire's data, there was an in-view to impression ratio of just 53%. After optimization, traffic generated by bots nearly disappeared and user engagement improved dramatically. Overall, the in-view to impressions ratio rose to more than 92%, post-analysis and implementation.

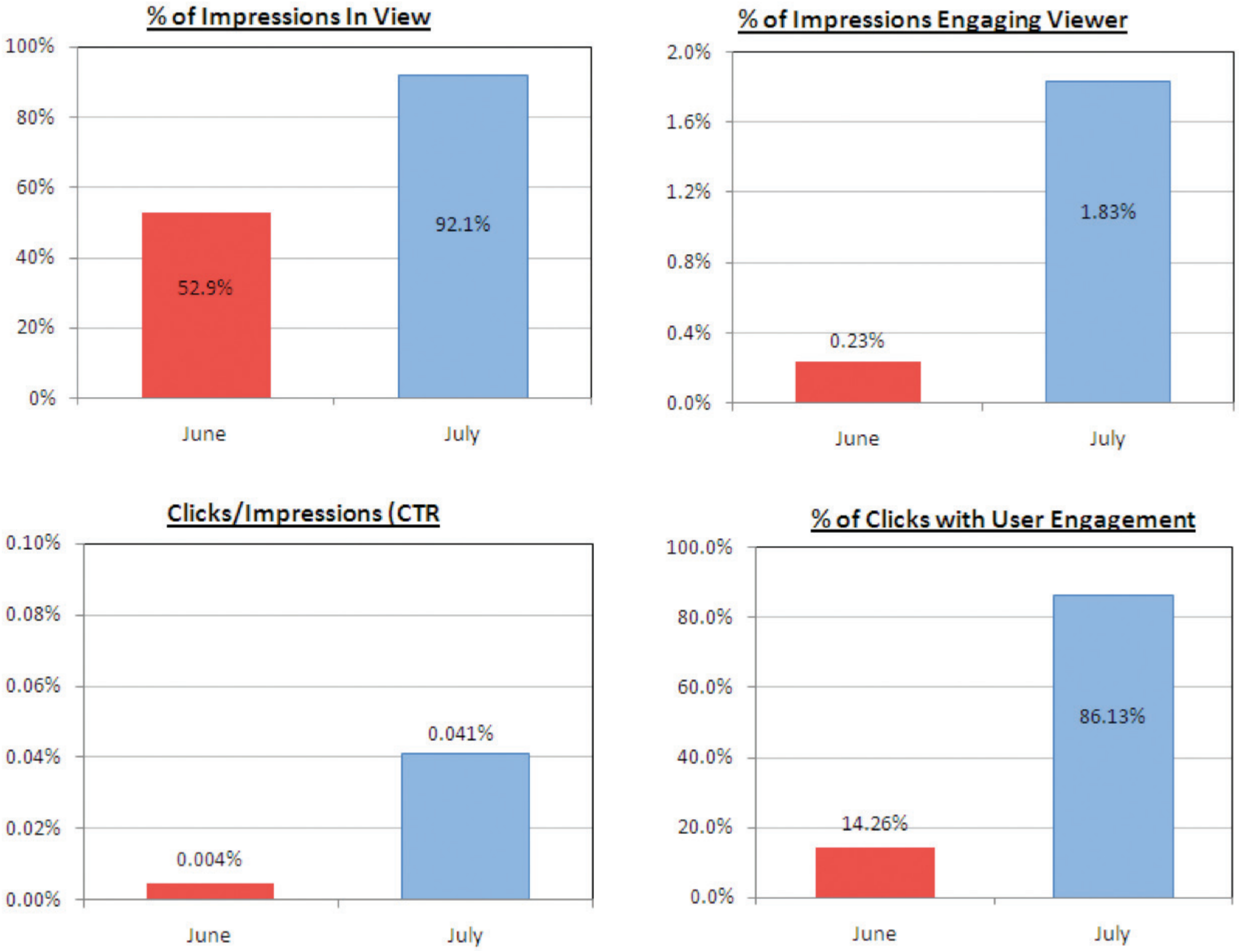
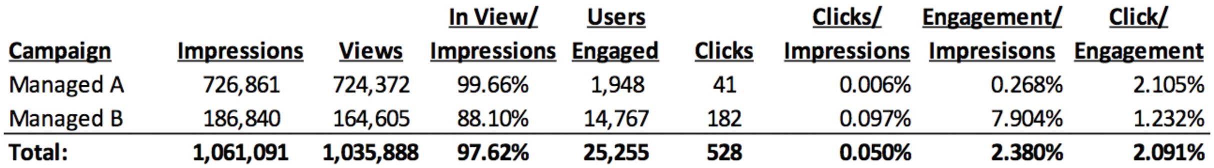


Figure 4: Performance increased substantially after AdXpose learnings were implemented

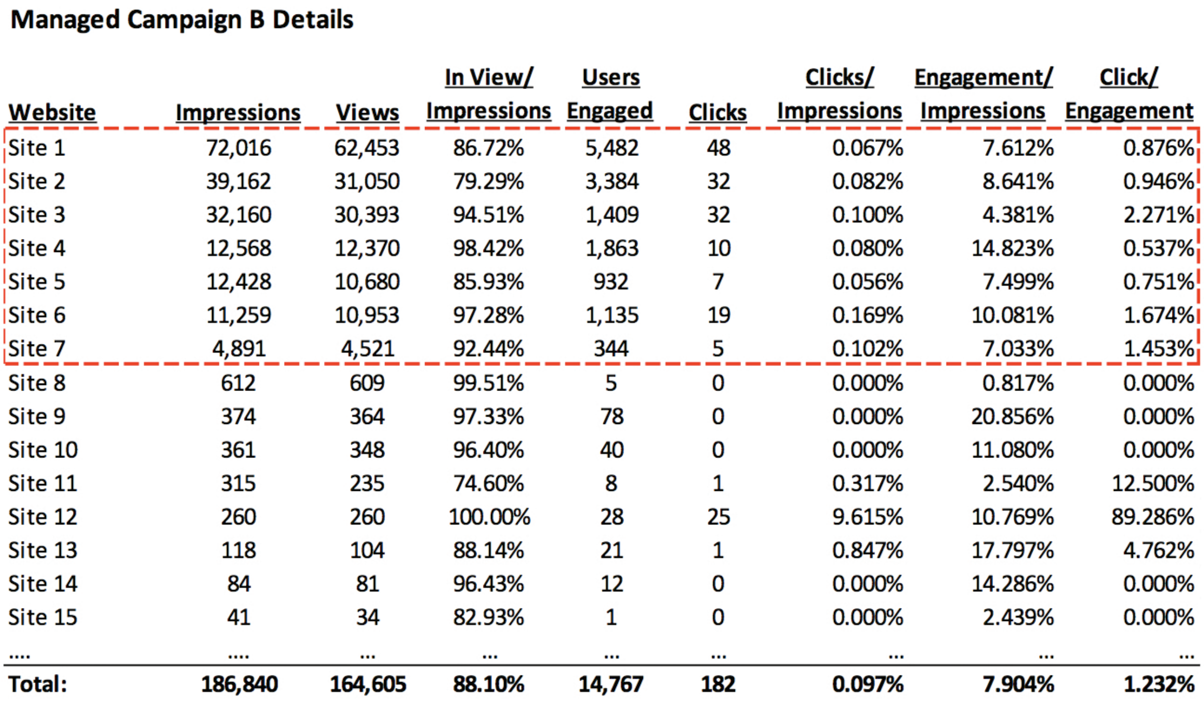
Transparency Drives Dollars to Networks

Within the actively managed campaigns, two test campaigns were conducted. The first, Campaign A, ran on a single URL with above-the-fold placement. Campaign A successfully confirmed that all ads on the single URL being looked into were above the fold. The second, Campaign B, was run on a site list of known and approved URLs with guaranteed above the fold placement. Campaign B detected meaningfully different in-view percentages, click-through rates and engagement statistics across multiple sites on the network. The test confirmed that while in this case, the advertiser was able to justify increased budget allocation to the network based on performance. Ads on many approved network sites were not above the fold and not always viewed by audiences. The overall engagement rate was much higher on the network buy.



Campaign	Impressions	Views	In View/ Impressions	Users Engaged	Clicks	Clicks/ Impressions	Engagement/ Impresisons	Click/ Engagement
Managed A	726,861	724,372	99.66%	1,948	41	0.006%	0.268%	2.105%
Managed B	186,840	164,605	88.10%	14,767	182	0.097%	7.904%	1.232%
Total:	**1,061,091**	**1,035,888**	**97.62%**	**25,255**	**528**	**0.050%**	**2.380%**	**2.091%**

Figure 5a : Site-list based network buy outperforms higer-volume direct buy.



Managed Campaign B Details

Website	Impressions	Views	In View/ Impressions	Users Engaged	Clicks	Clicks/ Impressions	Engagement/ Impressions	Click/ Engagement
Site 1	72,016	62,453	86.72%	5,482	48	0.067%	7.612%	0.876%
Site 2	39,162	31,050	79.29%	3,384	32	0.082%	8.641%	0.946%
Site 3	32,160	30,393	94.51%	1,409	32	0.100%	4.381%	2.271%
Site 4	12,568	12,370	98.42%	1,863	10	0.080%	14.823%	0.537%
Site 5	12,428	10,680	85.93%	932	7	0.056%	7.499%	0.751%
Site 6	11,259	10,953	97.28%	1,135	19	0.169%	10.081%	1.674%
Site 7	4,891	4,521	92.44%	344	5	0.102%	7.033%	1.453%
Site 8	612	609	99.51%	5	0	0.000%	0.817%	0.000%
Site 9	374	364	97.33%	78	0	0.000%	20.856%	0.000%
Site 10	361	348	96.40%	40	0	0.000%	11.080%	0.000%
Site 11	315	235	74.60%	8	1	0.317%	2.540%	12.500%
Site 12	260	260	100.00%	28	25	9.615%	10.769%	89.286%
Site 13	118	104	88.14%	21	1	0.847%	17.797%	4.762%
Site 14	84	81	96.43%	12	0	0.000%	14.286%	0.000%
Site 15	41	34	82.93%	1	0	0.000%	2.439%	0.000%
....	....	...	...	...	...	...	...	...
Total:	**186,840**	**164,605**	**88.10%**	**14,767**	**182**	**0.097%**	**7.904%**	**1.232%**

Figure 5b : Site-level analysis confirms performance.

Why Click Fraud Is Ultimately Damaging to Publishers and Networks

With ad networks reporting 45%-60% operating margins, it is apparent that they have a near-term incentive to maintain the status quo of the online advertising ecosystem. Long term, however, such tacit approval of lazy (at best) and nefarious (at worst) publisher, network and advertisers behavior is ultimately damaging to the industry as a whole.

Clearly, click fraud (as well as its less discussed cousin, impression fraud) is more pervasive than the industry has been willing to admit. The volume of click fraud varies based on the technologies used to detect, measure and analyze the problem. By delving deeper into site-level data, advertisers and agencies can get a better understanding of the impact of click fraud on campaign ROI. Likewise, publishers and networks gain the confidence of advertisers, who have greater reassurance that their spend is not wasted on fraudulent traffic.

Decisive Action Necessary

To truly minimize click and impression fraud, both sellers and buyers need to take vital steps to ensure the validity of campaign data. Advertisers and agencies need to police the activities of networks, exchanges and publishers as part of their overall process of due diligence. Part of that includes requiring campaign verification from a third-party source. Borrowing proven practices of supply chain management, advertisers and agencies can use site-level data and thorough analysis to gain greater leverage over other parties in the ad "supply chain," thereby lowering costs, improving performance and profitability, and illuminating any weaknesses in the chain. At the same time, publishers and networks need to implement policies and technologies that validate the legitimacy of their data. With both parties in the ad equation in agreement, the problem of click fraud can be managed and minimized.

For more information, or to schedule a demo, contact:

Sandy Streim (East Coast)

sandy@mpire.com
(516) 835-6034

Matt Pangrazio (West Coast)

map@mpire.com
(206) 302-2155

Mpire Corporation
1725 Westlake Avenue N.
Suite #203
Seattle, WA 98109