Understanding Fraudulent Activities in Online Ad Exchanges

Andrew Howes and Byungjin Jun
Motivation

● Ad Exchanges
  ○ manage complex relationships between advertisers and publishers

● Security issues in ad exchanges
Related Work

- Various aspects of detecting click-fraud
  - [18, 19, 26]

- How to deal with click-fraud
  - [13, 14, 16]

- Botnet structure
  - [15, 22, 24, 25]
Terminology

- **Ad exchange**
  - marketplace for publishers/advertiser/ad network

- **Ad campaign**
  - how much advertisers pay when their ads are shown
  - CPC, CPA, CPM

- **eCPM**
  - how much 1000 impressions worth to an advertiser
  - \(((\text{Payout per impression}) + (\text{historical CTR}) \times (\text{Payout per click}) + (\text{historical actions to impressions}) \times (\text{Payout per action}))\) \times 1000
Terminology (cont.)

- Auction process
  - how section is populated by a creative process
  - USERs loads a Web page initiating an ad request
  - AD EXCHANGEs (server) receive the request and match to PUBLISHERs base on targeting criteria
  - ADVERTISERs bids in the form of eCPM
  - AD EXCHANGE picks the bid with the highest eCPM
Terminology (cont.)

- Arbitrage
  - done by ad networks
    - after the auction process
    - buy impressions from publishers
    - resell ad slot in a new auction (new section ID)
  - daisy chaining
    - repeating arbitrage across different ad networks
    - takes time to display actual ad to users
Types of Fraud

Click Inflation
- Publishers clicking to increase their profit

Competitor Clicking
- Advertisers clicking their competitors

Conversion Spam
- More specific, requires actions after click

Misrepresentation
- Attempting to get higher value by lying about content
Known Types of Attacks

- **Hired Clickers**
  - Reloading a page to click on ads

- **Keyword Stuffing**
  - Misrepresentation to get high-value keywords

- **Impression Stuffing**
  - Excessive banners on the page
Known Types of Attacks

- **Coercion**
  - Convincing users to click for reasons other than ad content

- **Clickbots**
  - Custom
  - For-Sale
  - Botnet

- **Forced Browser Clicks**
  - Using a script to click the ad
Detection and Prevention

- Signature-based Detection
  - Rule-based decisions to determine invalid clicks.

- Anomaly-based
  - Uses historical data to identify changes in patterns.

- Reverse-Spidering (Auditing)
  - Examining the HTTP Referer to check for misrepresentation.
Detection and Prevention

● **Bluff Ads**
  ○ Uninviting ads used as a control.

● **Popularity and Page Rankings**
  ○ Checking for traffic outstripping ranking.

● **Performance-based Pricing**
  ○ Requires more results than just impressions from the publishers.
Botnet Ad Fraud

Using botnets, a large number of false impressions and clicks can be generated from unsuspecting users.

Two methods: Loading an iframe in the background, or redirecting from popular websites using browser hijacking.

The control servers will send out lists of websites to be used in the Referer section.
Fake Websites

A type of fraud that is based on creating a number of false websites in order to register as publishers.

This allows for distribution of fraudulent traffic across multiple networks, in order to appear less suspicious to a single network.

This also allows for direct loading of ad-containing iframes, in order to reduce traffic needed for loading real pages.
Data Feed

● NetworkX
  ○ ad network of the RightMedia exchange

● Traffic
  ○ local publisher traffic
  ○ arbitrage traffic
  ○ auction traffic

● Custom data feed
  ○ log file
  ○ 30 min / 200MB / 750K impressions
Data Feed (cont.)

- Collected data
  - 514 million impressions / 10 days
  - use local publishers traffic
    - 300 is active out of 1600
    - 1% of 300 - 40% local traffic
    - 10% - 92%

<table>
<thead>
<tr>
<th>Traffic Flow</th>
<th>Traffic (%)</th>
<th>Impressions (per hour)</th>
<th>CTR</th>
<th>Conversion Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auction</td>
<td>37.7%</td>
<td>305,318</td>
<td>0.16%</td>
<td>-</td>
</tr>
<tr>
<td>Publishers</td>
<td>2.0%</td>
<td>15,794</td>
<td>0.56%</td>
<td>0.01%</td>
</tr>
<tr>
<td>Arbitrage</td>
<td>60.3%</td>
<td>489,184</td>
<td>0.12%</td>
<td>0.007%</td>
</tr>
</tbody>
</table>

Table 1: Statistics for each traffic flow.
What's good or What's bad

- **Bad site**
  - does not serve content, or the content that it serves does not render
  - contains entirely/mainly ad content
  - host illegal content, or content against the terms of use of RightMedia
  - stolen content or leftover from an HTML template

- **Good**
  - well-designed, good looking, usable
  - good Alexa ranking
  - contents that supported by users from social network
Identifying Suspicious Traffic

- Reverse spidering
  - crawl the referrer sites of each impression
  - compare section ID of referring page
  - doesn't work
    - 79.2% contains no section IDs
    - instead, use IP-address-based geo-location
    - Ad delivery optimization service not only focusing yahoo! RightMedia
Identifying Suspicious Traffic (cont.)

- Building models for ad requests
  - Impression per Cookie
  - CTR per Cookie
  - Publisher revenue per Cookie
  - Unique IP addresses per Cookie
    - one cookie ID from many IP subnets in short time
  - Impression per IP address
  - CTR per IP address
  - Publisher revenue per IP address
  - Deviation of CTR per IP address
    - consistent CTR regardless change of impressions
Identifying Suspicious Traffic (cont.)

- **Evaluating Models**
  - Impression per Cookie/IP
    - The best indicators
  - CTR per Cookie/IP
    - too explicit
  - Publisher revenue per Cookie/IP
    - hard to generate actual revenue
  - Unique IP addresses per Cookie
    - some sites assigning same cookie ID to users from many different networks
  - Deviation of CTR per IP address
    - works but modest
Identifying Suspicious Traffic (cont.)

● Why cookie is better?
  ○ assumption
    ■ a cookie ID/IP belongs to a user
    ■ if behaviors is inconsistent -> fraud
  ○ RightMedia hide users behind a single IP
    ■ provides the first three octet (e.g. 192.128.168.*)

● NO single model works as a reliable detector
  ○ Significant variety in the ad traffic
    ■ The best detection rate : 60~80%
    ■ incorrectly blame 10% of the legitimate publishers
  ○ Useful to guide a human analyst
Diverse Attack Patterns

Cookie Replay Attacks

- Uses a single cookie ID distributed across several machines to spoof IP addresses.
- Cookie ID is hashed with a server-side key, so keeping it the same is necessary, even though the rest of the values for the cookie are randomized.
- Re-using the cookie can give suspicious patterns, since it's supposed to only represent one browser instance.
Diverse Attack Patterns

Referrer Spoofing
- Refers to using a single section ID and a variety of referer links that may not be part of the fraudster's site.
- Since referers aren't automatically checked, this allows the fraudsters to camouflage their ad requests among a number of sites.
- Searching the referers listed for a single section ID for unfamiliar sites can show publishers that are likely to be bad.
Diverse Attack Patterns

Unrecognized Referrers

May not be related to the section ID that the ad network is assigned.

Missing-in-Action sites

Registering a valid site to get a section ID, but using those section IDs on a different site with the same referer.

Table 4: Top publishers with unknown referrers (April 2011)
Diverse Attack Patterns

Malicious Publishers

Publishers that are actively engaging in fraud, and attempting to increase their own profits.

Often will have a high number of impressions, but a very low eCPM.
Fraud in Ad exchanges

Suppression of Data Fields
Leaving out the last 8 digits of an IP address helps user privacy, but hurts in identifying behavioral patterns.

Consistency vs. Flexibility
With RightMedia, there weren't checks to see if the site registered with the company was the same one that is generating ads.
Fraud in Ad exchanges

Hiding in the exchange

Fraudulent activities can be spread across networks, and are hard to detect if they aren't above a threshold on a single network.

Things that RightMedia does

Provides tools to identify obvious cases of fraud, particularly bot patterns.
Conclusion

We have described how ad exchanges work, and we've examined real-world data for a single ad network.

We've examined how the complexity of the network can mask fraudulent activities, and how flexibility within the network comes at the expense of better security.