botnet uses tens of hundreds of IP addresses. Detectives detects coalition hit inflation attacks by their similarity seeker algorithm; it discovers coalitions made by pairs of fraudsters, which is then enhanced in [31] by finding groups of fraudsters. All these approaches apply only to botnet and malware driven click-spam, which is dwarfed by other sources of click-spam in our data.

Premium Clicks [27], access control gadgets (ACG) [37] and CDN fraud prevention [30] focus on mitigation strategies that go beyond botnets. Premium clicks employs economic disincentives that devalue clicks from non-gold-standard users. ACGs ensure authentic UI interactions by users clicking a link. CDN fraud prevention proposes a heavy-weight challenge-response protocol for publisher-payee CDN models. While the first assumes an alternate ad economy, the second and third (applied to ad networks) require re-architecting the browser, or the ad network infrastructure. None of these approaches apply to click-spam in existing ad networks.

Focusing squarely on existing ad networks, Camelot [28] is Google's click-fraud penetration system. It can test the susceptibility of the network to known click-spam signatures, but does not itself detect new signatures. [39] describes the invalid click detection system inside Google, without identifying the specific heuristics that are used to identify invalid clicks. No heuristic is perfect. Our data shows click-spam is still an open problem despite these deployed systems.

7. CONCLUSION

In this paper, we take a systematic look at click-spam. We propose the first methodology for advertisers to independently measure click-spam rates on their ads. We also develop an automated methodology for ad networks to proactively fingerprint different simultaneous click-spam attacks. We validate both methodologies using data from major ad networks. We then conduct a large-scale measurement study of click-spam across ten major ad networks and four types of ads. In the process, we identify and perform in-depth analysis on seven ongoing click-spam attacks not currently caught by major ad networks. We conclude that even for the *largest* ad networks, click-spam is a serious problem, and is especially rampant in the mobile advertising context. Given the evolving nature of click-spam, we believe that click-spam is an open problem that requires a concerted effort from the research community to tackle. To this end we have publicly released the data gathered for this paper to aid other researchers in the design of novel click-spam defense techniques.

Acknowledgments

We'd like to thank Jigar Mody, Matt Graham, our shepherd Kirill Levchenko, and our anonymous reviewers. This paper is much improved thanks to their valuable feedback and suggestions.

8. REFERENCES

[1] AdSense for domains program policies. http://support.google.com/adsense/bin/answer.py?answer=96332.
[2] The adsense revenue share. http://adsense.blogspot.com/2010/05/adsense-revenue-share.html.
[3] Click Fraud Falls in Q4 2010. http://searchenginewatch.com/article/2050117/Click-Fraud-Falls-in-Q4-2010.
[4] Click fraud rampant in online ads, says bing. http://www.theaustralian.com.au/media/click-fraud-rampant-in-online-ads-says-bing/story-e6frg996-1226056349034.
[5] Cloaking and Faking the Referrer. http://kbeezie.com/view/cloaking-and-faking-referrer/.
[6] For Impatient Web Users, an Eye Blink Is Just Too Long to Wait. http://www.nytimes.com/2012/03/01/technology/impatient-web-users-flee-slow-loading-sites.html.
[7] Google AdSense for Domains. http://www.google.com/domainpark/index.html.
[8] Google Click Fraud Inflates Conversion Rates and Tricks Advertisers into Overpaying. http://www.benedelman.org/news/011210-1.html.
[9] Google Redirect Virus: How to Remove. http://www.pcmag.com/article2/0,2817,2370676,00.asp.
[10] International Cyber Ring That Infected Millions of Computers Dismantled. http://www.fbi.gov/news/stories/2011/november/malware_110911.
[11] Malware connection report. http://www.malware-control.com/statics-pages/03aaf7c8e47ef32e8de23dfe9215d4a5.php.
[12] Stealing Clicks. http://www.forbes.com/2007/09/21/google-click-forensics-tech-secure-cx_ag_0924fraud.html.
[13] Uncovering an advertising fraud scheme. Or "the Internet is for porn". http://www.behind-the-enemy-lines.com/2011/03/uncovering-advertising-fraud-scheme.html.
[14] Upcoming changes in Google's HTTP Referrer. http://googlewebmastercentral.blogspot.com/2012/03/upcoming-changes-in-googles-http.html.
[15] L. V. Ahn, M. Blum, N. J. Hopper, and J. Langford. Captcha: using hard ai problems for security. EUROCRYPT'03, 2003.
[16] Click Quality Team, Google Inc. How Fictitious Clicks Occur in Third-Party Click Fraud Audit Reports. http://www.google.com/adwords/ReportonThird-PartyClickFraudAuditing.pdf.
[17] G. Cormode and S. Muthukrishnan. What's hot and what's not: Tracking most frequent items dynamically. In *Proceedings of ACM PODC*, July 2003.
[18] G. Cormode and S. Muthukrishnan. Improved data stream summaries: The count-min sketch and its applications. *Journal of Algorithms*, 2004.
[19] N. Daswani, C. Mysen, V. Rao, S. Weis, K. Gharachorloo, and S. Ghosemajumder. *Crimeware:Understanding New Attacks and Defenses*, chapter Online Advertising Fraud. 2008.
[20] N. Daswani and M. Stoppelman. The anatomy of clickbot.A. In *Proceedings of HotBots*, 2007.
[21] J. R. Douceur. The Sybil Attack. In *Proceedings of IPTPS '02*.
[22] C. Estan and G. Varghese. New Directions in Traffic Measurement and Accounting. In *Proceedings of ACM SIGCOMM*, Aug. 2002.
[23] A. M. Eugene Rodionov. The evolution of tdl: Conquering x64. http://go.eset.com/us/resources/white-papers/The_Evolution_of_TDL.pdf.
[24] B. Geddes. Search arbitrage: Web blight or brilliant marketing strategy? http://searchengineland.com/search-arbitrage-web-blight-or-brilliant-marketing-strategy-10768.
[25] H. Haddadi. Fighting online click-fraud using bluff ads. *SIGCOMM Comput. Commun. Rev.*, 2010.
[26] N. Immorlica, K. Jain, M. Mahdian, and K. Talwar. Click Fraud Resistant Methods for Learning Click-Through Rates. In *Proceedings of the Workshop on Internet and Network Economics (WINE '05)*.
[27] A. Juels, S. Stamm, and M. Jakobsson. Combating click fraud via premium clicks. In *Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium*, 2007.
[28] C. Kintana, D. Turner, J.-Y. Pan, A. Metwally, N. Daswani, E. Chin, and A. Bortz. The goals and challenges of click fraud penetration testing systems. International Symposium on Software Reliability Engineering, 2009.
[29] H. Lieberman. Letizia: An agent that assists web browsing. In *International Joint Conference on Artificial Intelligence*, 1995.
[30] S. Majumdar, D. Kulkarni, and C. V. Ravishankar. Addressing click fraud in content delivery systems. In *In Proceedings of the 26th IEEE INFOCOM International Conference on Computer Communications*, 2007.
[31] A. Metwally, D. Agrawal, A. El Abbad, and Q. Zheng. On hit inflation techniques and detection in streams of web advertising networks. ICDCS '07, 2007.
[32] A. Metwally, D. Agrawal, and A. El Abbadi. Detectives: detecting coalition hit inflation attacks in advertising networks streams. WWW '07, 2007.
[33] A. Metwally, F. Emekçi, D. Agrawal, and A. El Abbadi. Sleuth: Single-publisher attack detection using correlation hunting. *VLDB*, 2008.
[34] B. Miller, P. Pearce, C. Grier, C. Kreibich, and V. Paxson. What's clicking what? techniques and innovations of today's clickbots. In *Proceedings of the 8th international conference on Detection of intrusions and malware, and vulnerability assessment*, DIMVA'11, 2011.
[35] B. Mordkovich and E. Mordkovich. Click fraud and how to counteract it in ad campaigns. In *Pay-Per-Click Search Engine Marketing Handbook*, 2005.
[36] E. Rodionov and A. Matrosov. The evolution of TDL: Conquering x64. Technical report, 2011.
[37] F. Roesner, T. Kohno, A. Moshchuk, B. Parno, H. Wang, and C. Cowan. User-driven access control Rethinking permission granting in modern operating systems. IEEE Symposium on Security and Privacy, 2012.
[38] B. Stone-Gross, R. Stevens, A. Zarras, R. Kemmerer, C. Kruegel, and G. Vigna. Understanding fraudulent activities in online ad exchanges. Internet Measurement Conference, 2011.
[39] A. Tuzhilin. The lane's gift v. google report. http://googleblog.blogspot.in/pdf/Tuzhilin_Report.pdf.
[40] F. Yu, Y. Xie, and Q. Ke. Sbotminer: large scale search bot detection. WSDM '10, 2010.
[41] Q. Zhang, T. Ristenpart, S. Savage, and G. M. Voelker. Got traffic?: an evaluation of click traffic providers. In *Proceedings of the 2011 Joint WICOW/AIRWeb Workshop on Web Quality*, WebQuality '11, 2011.