Bringing Networks Together to Improve Advertising Performance

Luis Miralles Pechuán[1,2] and José M. García Carrasco[1]

[1] Departamento de Ingeniería y Tecnología de Computadores, University of Murcia, Murcia, Spain

[2] Facultad de Ingeniería, Universidad Panamericana, Mexico

lmiralles@up.edu.mx, jmgarcia@ditec.um.es

Abstract. We believe that small networks working together can create a more competitive solution against bigger networks, not only regarding ad performance but also fraud detection. Moreover, we have designed algorithms to uniformly distribute visits over several networks, and we have used the average deviation as a parameter to compare results.

Keywords: Internet marketing, click fraud, ad performance, networks working together, fraud detecting algorithm.

1 Introduction

Internet is one of the most revolutionary inventions in the history of humanity. It evolved from a US Department of Defence project known as Arpanet, which was developed back in 1969. Since then it has allowed us to share photos on Facebook, send email with Gmail, make video calls via Skype, Blog on Wordpress, stream videos on YouTube, sell on eBay or pay online using PayPal. We consider them commonplace in today's world, but barely a few decades ago were they absolutely unthinkable.

It offers endless opportunities to those who use it, such as being able to work from anywhere at any time of day, instantly send information or access resources about anything. Logically, the number of people wishing to enjoy the benefits is constantly growing. There has also been a widespread proliferation of companies offering of huge variety of services and the best way for these companies to blossom is by using online ad campaigns.

The first ad banner ever seen on a webpage was for AT&T in 1994, and from then on its use has continued to grow immensely. In the third quarter of 2013, investments in online advertising reached $10.69 billion dollars [1]. Online advertising offers huge advantages to advertisers as it allows them to modify campaigns at any time. Whilst most channels contract closed packets, online advertising allows campaigns to be cancelled should they not show good results, or areas of high sales can be focused on.

Using the web we can check campaign quality in real time, with parameters like number of products sold or the average time users spend on our page. This kind of publicity allows us to select a segmented public with usable-to-program parameters such as age, gender, geographic zone, likes and much more, thanks to networks holding a huge amount of user data.

The cost of such campaigns can be adapted to any budget, as we can select the number of ads to be shown. Finally, we can establish a bidirectional user channel giving us immediate feedback. Such users act by forwarding messages, bookmarking the page or recommending it to a friend.

2 The Problems with Being Small

The main objectives of any advertising platform are to show users the most relevant ads and reduce the number of faults in fraudulent clicks to zero. The largest ad platforms are at an advantage in respect to the smaller networks given that they have more secure fraud detection systems. This allows them to get more advertisers and publishers, in turn creating higher revenues creating a vicious circle making the small networks even smaller and themselves even bigger.

2.1 Ad Performance

As advertisers make more and more specific campaigns, the number of pages they can be on reduces, but at the same time they are more effective given that advertisers are paying a higher price. This is known as targeting [3]. In order to develop a good targeting campaign, we must filter out a series of parameters such access keywords, age, gender, income level, location and likes from user profiles. Another series of attributes, although not as influential must still be taken into account. They include browser, search engine, operating system or device being used.

When a page is visited by a user fitting the desired characteristics an ad is shown and if it results in a click, the advertiser is charged accordingly. A large platform with a lot of publishers can easily find any page related to and be accepted by an advertisers requirements. On the other hand, if the ad network has very few publishers and works independently, ads will receive much less coverage [1,] or rather they will not be shown as much and will hence be seen by less users. To solve this issue, generic campaigns are created with the disadvantage of lower performance.

2.2 Fraud Detection

First, we have to emphasize that fraud really is a threat. According to experts, 15% of all clicks are fraudulent and out of that 20% go undetected [4]. This means that we

[1] Coverage has a value of 0 & 100 and represents the number of times an ad is shown to a user. Having 50% coverage means half the visits have not created revenue, due to them not meeting any advertisers' requirements.

can discount that 3% (0.15 x 0.2) from what is paid by advertisers. As stated by Tuzhilin, it is statistically possible for some advertisers to be unsatisfied, but if the rest are happy the platform will be successful [5].

The problem arises when the number of fraudulent clicks increases, or the ability to detect competition clicks decreases. Publishers prefer working with the best performing platform for their ad space, and advertisers look to improve campaign results [6].

Large ad networks make millionaire investments and tend to use specialized equipment to continuously improve their fraud detection systems. Their false click detection system is also much more superior to those used by small networks. Google for example knows the CTR[2] of every class of webpage, so should a page have different statistics to the rest, it can be easily detected.

Following Kirchhoff's principal[3], the major platforms should publish the techniques used by scammers and the methods used to detect them, enabling systems to be more secure. Furthermore, there has been research that talks of the convenience of networks working together to improve fraud detection [7]. However, large networks are trusted precisely for their click detecting capability so if all platforms were equally secure, the competitive advantage would disappear.

3 Small Networks Working Together

Some authors affirm that the exchange of ads represent the future of online advertising and the solution for small ad platforms however for such exchanges to be successful firstly the issue of click fraud and the legal questions regarding user privacy need resolving; and an exchange model, that generates benefits for all parties involved, needs developing.

3.1 Working Together to Improve Performance

Advertising exchanges consist of platforms exchanging visits not meeting the requirements of any of their advertisers, or they are simply looking for an advertiser willing to pay more. In this model, advertisers pay for space only if certain requirements are met, and editors leaves a space on their page to be filled by the most profitable ad. Let's imagine there are two small ad platforms, SpainOnline97 and BrazilMarket43. Most of SpainOnline97's advertisers would be Spanish speakers and most of BrazilMarket43's would be Portuguese speakers. If these two networks were

[2] CTR is the number of clicks received by an ad divided by the number of times it has been seen, e.g. if 15 clicks have been received and it has been seen 1000 times, the CTR will be 1.5%.

[3] The success of a cryptographic algorithm should not remain a secret. Any algorithm employed using cryptography is published, and should the system become susceptible to an efficient attack, it automatically improves or stops. This policy has allowed systems to be ever more secure, and is now almost invulnerable.

to exchange ads, a user from Spain visiting a BrazilMarket43 page could be shown ads from Spain and are much more likely to be interested in buying the product and vice versa.

Most platforms follow IAB standards, making exchanges easier. This can seem simple when there are only two networks working together, but when hundreds of networks with thousands of advertisers there are certain factors to taken into account, such as: volume, revenues, fraud committed by advertisers or adequately distributed visits.

In order to make such exchanges possible, we need to develop an algorithm, taking into account the fact the ad is to be shown in split seconds and at the same time be really effective, so it is recommended using parallelization.

3.2 Working Together against Fraud

A cost reducing solution could be for ad platforms to outsource click detection to specialists, although the problem is that these specialists could be tempted to create their own threats to ensure work or ally with cheating publishers to increase revenue.

Advertising platforms face many threats such as Click-bots, illegal traffic or users with bad intentions nonetheless [7], these are no different to each other where a platform is concerned, to when a platform detects a malicious IP and warns the rest, the threat is taken care of [8].

Information sharing is an advantage so that all platforms can offer a better service to advertisers, as well as reducing the number of undetected fraudulent clicks. Such advantages include:

- Awareness of page CTRs from other platforms, so should a user have a page with similar characteristics but a distinct CTR, it will be suspicious.
- Sharing of suspect IPs.
- Updating proxy list [4] to invalidate clicks originating from them.
- Share new click-bot [9] detecting methods.
- Comparing ratios from a specific editor with those from editors of other platforms. This tells us if it differs from the average.
- Calculating percentage of fraudulent clicks in order to apply discounts to advertisers.

3.3 Privacy

Advertising platforms recollect user information when services are used. Both Google Analytics and Webmaster tools allow Google to access many different statistics including how long a user remains on a certain page, number of average pages visited

[4] A proxy is a program or device that connects to the internet from another computer. It is used to maintain anonymity, or better security. In the case of click fraud, it allows clicks to be made without the IP being detected.

and much more. The more user information they have, if used effectively, the more personalised ads shown can be, but privacy must always be strictly respected.

In order to protect themselves from being reported regarding privacy, platforms oblige users to accept these services in their terms and conditions. To guarantee the right to privacy [10] and at the same time segment publicity, platforms create profiles where such information is saved. Theoretically, the profiles are not associated with any particular person, but instead work anonymously. It has been known for these profiles to be tracked by security organisations in the detection of possible terrorist threats and paedophiles.

4 Algorithms to Improve Advertising Management Performance

4.1 Networks Working Together to Increase Ad Coverage

Here we aim to show how to improve coverage; this means the higher the percentage of satisfied visits, the higher the number of collaborating networks. To make this test possible we obtained a total of 104,151 real visits from the site history of buscadoreseninternet.net from 01/06/12 to 01/01/13. Each visit uses a series of fields as seen in Table 1.Visits were exported into an excel table from Google Analytics so they can be seen clearly as shown in Table 2.

To show how to improve we have compared each visit on the table with ad campaign requirements given by X participating networks (where X = 1, 2, 3, 5, 10, 25, 50, 100). Each network has 10 campaigns, giving a total of 1000 campaigns simultaneously, giving X a value of 100.

The values selecting each advertiser for each parameters are randomly established based on probability of occurrence, meaning that should 90% of operating systems in visit history be running Windows, the probability of the advertiser choosing the Windows value as the OS parameter will be 90%. As shown in the table, advertisers select where to show their ads from a series of fields, these could include country, city or page category with the ad platform adding ads to relevant webpages using an algorithm.

For each of the options shown in Table 5 different parameters are configured. The number of parameters to be configured depends on the size of the biggest option, making campaigns more specific as well as more difficult to cover.

In Table 4 the Y axis represents the number of options, and the X axis represents configurable parameters, which were explained earlier in Table 1.

In Table 5 the X axis shows coverage related to the number of networks working together and the Y axis show options to be selected by advertisers upon making campaigns. Parameters to be configured can be seen above in Table 4.

As can be seen in the table, when more networks work together coverage is greatly improved.

4.2 Distributing Visits

Apart from trying to improve coverage by contacting other ad platforms, ads must be distributed as fairly as possible. To do this, three algorithms have been developed and for quality assurance a method of average deviation was developed, the lower the average deviation the better the algorithm to be used will be.

Table 1. User visit parameters.

1	Time	This refers to the time of day of the visit. It ranges from 0 to 24 E.g. 2,4,6 etc
2	Browser	This refers to the user's browser e.g. Internet Explorer, Google Chrome, Mozilla Firefox
3	Browser version	Indicating the browser version being used e.g. in Internet Explorer you can see version 9.0, 8.0 & 7.0 etc.
4	Operating System	This refers to the OS of the computer the webpage is being accessed from. The most common being Windows, but Mac OS X and Linux are also used.
5	OS Version	This refers to OS version e.g. Windows 7, 8 or Mac OS X Lion etc.
6	Flash version	Many browsers have flash preinstalled in order to open certain pages, of which there are many versions e.g. 11.3 r31, la 10.0 r32 or 10.2 r153.
7	Has flash?	This parameter indicates if the browser has flash. The value is YES or NO.
8	Screen bitrate	This refers to number of bits required to show a pixel. The most common is 24-bit or 32-bit.
9	Screen resolution	This is the number of pixels the monitor has, it is usually (shown in width x length) around 1280x1024 or 1024x768.
10	Country	Using the IP we can determine the country of the visit.
11	City	As well as country we can also determine the city the visit is coming from.
12	Language	Language being used on user's system e.g. es-419, es, es-mx.
13	Network address	This refers to the ISP url the user is visiting from e.g. megared.net.mx, cablevision.net.mx, prod-infinitum.com.mx,cableonline.com.mx, maxcom.net.mx.
14	Network name	This refers to the name of the network being used by the user e.g. uninetredirection management, uninet s.a. de c.v…
15	Access page	This is the page where our visit originates from, usually a search engine such asyahoo.com, but it could also be being accessed directly or through a link.
16	Visit type	Visit type could be organic if a search engine is used, or referral is a reference is used.

Table 2. Storing user visits.

	Field 1	**Field 2**	**Field 3**	**Field …**	**Field N-1**	**Field N**
Visit 1	Firefox	16.0	Windows	7	11.4 r402	24-bit
Visit 2	Chrome	22.0.1229.92	Windows	XP	11.4 r31	32-bit
Visit …	I. Explorer	8.0	Windows	7	(not set)	32-bit
Visit N-1	I. Explorer	8.0	Windows	7	11.1 r102	32-bit
Visit N	Chrome	21.0.1180.89	Windows	XP	11.3 r31	32-bit

Table 3. Storing user visits.

	Field 1	**Field 2**	**Field 3**	**Field …**	**Field N-1**	**Field N**
Advertiser 1	Chrome, Firefox	16.0	Windows	XP,7	11.4 r402	24-bit, 32-bit
Advertiser 2	Chrome	22.0.1229.92	Windows	XP	11.4 r31	32-bit
Advertiser N	Internet Explorer	8.0	Windows	XP,7	(not set)	32-bit

Table 4. Parameters selected for each option.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Option 1										X						
Option 2		X								X						
Option 3		X		X						X						
Option 4		X		X						X		X			X	
Option 5		X		X						X	X	X			X	
Option 6		X		X				X		X	X	X			X	
Option 7		X		X				X	X	X	X	X			X	
Option 8	X	X		X				X	X	X	X	X			X	
Option 9	X	X		X	X			X	X	X	X	X			X	
Option 10	X	X	X	X	X			X	X	X	X	X			X	
Option 11	X	X	X	X	X			X	X	X	X	X			X	X
Option 12	X	X	X	X	X			X	X	X	X	X	X	X	X	X
Option 13	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Table 5. Advertising coverage in relation to number of networks and options.

Ad Coverage	1	2	3	5	10	25	50	100
Option 1	0,65901	0,79542	0,85574	0,9092	0,95288	0,98307	0,99184	0,9943
Option 2	0,35315	0,51562	0,63524	0,73938	0,8384	0,92112	0,95214	0,97189
Option 3	0,34136	0,5101	0,60923	0,71234	0,82205	0,8987	0,93521	0,95982
Option 4	0,15334	0,25803	0,33614	0,43064	0,53901	0,65418	0,73267	0,79566
Option 5	0,02476	0,04473	0,06103	0,09263	0,15475	0,25647	0,3468	0,44492
Option 6	0,01351	0,02455	0,03379	0,05396	0,09425	0,17517	0,25696	0,34725
Option 7	0,00169	0,00344	0,00549	0,00882	0,01675	0,03739	0,06599	0,10764
Option 8	9,4E-05	0,00019	0,0003	0,0005	0,001	0,00249	0,00487	0,00955
Option 9	6,9E-05	0,00015	0,00023	0,00036	0,00071	0,00176	0,00343	0,00663
Option 10	1,9E-05	3,5E-05	5,1E-05	8,3E-05	0,00015	0,00036	0,00073	0,00146
Option 11	1,2E-05	2,5E-05	4,3E-05	7,4E-05	0,00014	0,00038	0,00074	0,00147
Option 12	3E-06	7E-06	9E-06	1,4E-05	3,1E-05	7,6E-05	0,00016	0,0003
Option 13	1E-06	2E-06	2E-06	4E-06	0,00001	2,6E-05	5,1E-05	0,0001

Table 6. Average deviation from the simple algorithm.

Simple	2	3	4	5	10	25	50	100
Option 1	25317,5	25074,4	22688,6	20387,7	13193,3	6422,08	3512,67	1862,84
Option 2	8876,84	10684,2	10747,8	10379,5	8419,45	4901,6	2913,74	1634,75
Option 3	10042,8	10521,1	10715,2	10377,8	8224,26	4796,97	2852,07	1603,25
Option 4	4939,88	5217,37	5425,4	5236,4	4307,7	2754,35	1763,37	1064,59
Option 5	620,12	776,33	818,67	819,45	816,09	679,08	525,36	374,93
Option 6	396,52	475,69	499,97	511,62	497,21	428,71	350,36	268,62
Option 7	55,81	67,19	85,37	90,59	92,42	90,14	86,49	74,27
Option 8	5,14	5,72	6,04	5,95	6,63	6,81	6,76	6,92
Option 9	3,44	3,68	4,15	4,02	4,77	5,09	5,18	5,15
Option 10	1,26	1,4	1,41	1,4	1,45	1,52	1,56	1,6
Option 11	1,11	1,34	1,44	1,52	1,54	1,59	1,59	1,55
Option 12	0,24	0,28	0,33	0,34	0,42	0,48	0,47	0,47
Option 13	0,09	0,14	0,17	0,19	0,21	0,2	0,2	0,19
							Total	276,473

The average deviation is the average of the absolute values of the deviations from the mean and is shown as Dm.

$$Dm = \frac{1}{n}\sum_{i=1}^{n}|X_i - X| \quad (1)$$

The Simple algorithm first contacts the number one network and in the case it cannot satisfy the request, number 2 will be contacted and so forth until the last network is reached. Table 6 shows the results obtained.

The Round Robin algorithm first contacts the number 1 network in the first cycle, but the second time it moves on to contacting network number 2. Whenever a visit distributed it starts contacting the following network from the last time it was run. The results are shown in Table 7.

Table 7. Average deviation of Round Robin algorithm.

Round	**2**	**3**	**4**	**5**	**10**	**25**	**50**	**100**
Option 1	4747,32	4666,87	3856,47	3486,34	1925,38	798,02	409,9	207,55
Option 2	2981,15	3936,11	3509,33	3172,09	1916,85	890,94	461,24	233,56
Option 3	3951,06	3976,8	3721,39	3456,14	2009,65	900,78	451,18	230,77
Option 4	2646,23	2643,42	2668,82	2349,33	1599,45	774,59	425,82	224,63
Option 5	634,12	692,18	659,68	729,41	659,47	445,21	277,43	160,23
Option 6	390,75	433,59	453,86	456,33	454,81	316,1	221,94	137,84
Option 7	70,19	78,44	87,07	92,67	95,09	90,39	77,25	61,37
Option 8	4,37	5,32	5,51	5,54	6,44	7,01	7,05	6,82
Option 9	3,19	3,71	4,25	4,8	4,85	5,15	5,21	5,21
Option 10	0,99	1,21	1,28	1,28	1,42	1,55	1,59	1,59
Option 11	1,24	1,23	1,37	1,41	1,46	1,48	1,5	1,49
Option 12	0,31	0,39	0,36	0,37	0,4	0,45	0,44	0,45
Option 13	0,08	0,14	0,16	0,19	0,23	0,2	0,2	0,2
							Total	77,116

The Minimum algorithm always contacts the network with the least satisfied visits. To do this it requires the help of a table showing the number of visits distributed per network. The results are shown in Table 8.

Table 8. Average deviation of Minimum algorithm.

Minimum	**2**	**3**	**4**	**5**	**10**	**25**	**50**	**100**
Option 1	255,76	56,35	22,82	15,74	4,24	0,78	0,43	0,37
Option 2	1290,72	539,18	287,88	76,91	16,09	1,75	0,61	0,41
Option 3	679,6	466,18	178,16	96,93	12,12	2,69	0,88	0,43
Option 4	1127,77	815,04	630,44	387,4	153,21	16,58	2,85	0,95
Option 5	634,4	692,12	647,37	625,64	493,62	208,01	74,87	15,33

Minimum	2	3	4	5	10	25	50	100
Option 6	376,96	442,27	432,25	425,59	344,21	202,55	97,03	32,26
Option 7	62,33	77,08	81,23	86,55	91,51	81,2	66,91	47,92
Option 8	5,13	5,68	5,8	5,96	6,42	6,82	6,75	6,65
Option 9	3,15	4	4,7	4,51	5,1	5,15	5,19	4,93
Option 10	1,09	1,54	1,51	1,51	1,47	1,55	1,52	1,51
Option 11	1,17	1,44	1,56	1,53	1,54	1,58	1,61	1,62
Option 12	0,37	0,4	0,44	0,41	0,43	0,49	0,47	0,47
Option 13	0,1	0,14	0,16	0,17	0,2	0,2	0,2	0,2
							Total	**13,595**

To compare results we have summed up all the tests from each algorithm. Using the lowest sum from each one. The best results were gained by the Minimum (13595.04), followed by Round Robin (77115.58) and finally the Simple (276473.42).

4.3 A Fraud Detecting Algorithm

To test the improvements to fraud detection in a collaborative environment, the captcha technique [11] was used along with irrelevant ads [12]. This helps us detect fraudulent IPs. The captcha technique requires asking users to solve a "captcha", when access to ad content is desired. If captchas were to be put on all ads, users would become frustrated so they are only applied to about 20% of ads. The irrelevant ad technique shows a determined user ads unrelated to their profile, meaning that clicks do not come from user interest but rather by malicious means. The user is not expected to click such ads, so there is a high probability that any clicks being made are from botnets or a group of poorly trained, fraudulent users.

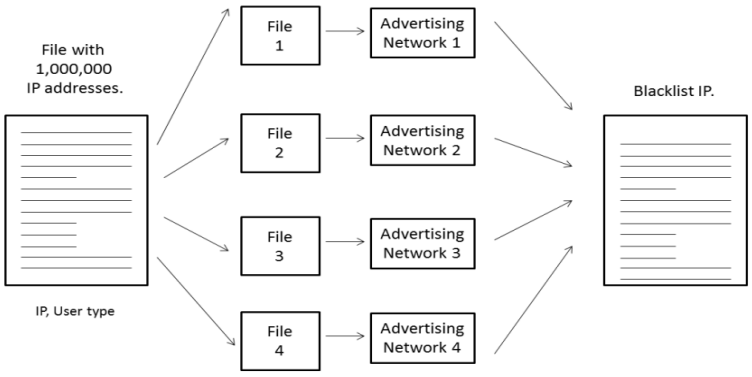

Fig. 1. Network collaboration model for fraud reduction.

If we abuse these, the fraudulent user will feel watched, and will realize that something is not right, provoking them to change techniques. To show the improvement to fraud detection techniques in proportion to collaborating networks, a model where networks exchanged high risk IPs was designed (see Figure 1).

The experiment consisted of creating a catalogue of 100,000 IPs with 10% of them coming from irrelevant ads; of this 10%, 75% come from botnets, 10% fraudulent users and 5% valid users. Of the remaining 90% of visits 80% come from valid users, 15% from botnets and 5% from fraudulent users. 1000 networks took part in the experiment, from which each received 2000 random visits from the original catalogue. To measure fraud detection performance, a check is made using captchas 20% of the time, except in the case of irrelevant ads where checks are always made using captchas. Botnets are unable to able to solve captchas so we will add the IP to a list of suspicious IPs. The detection percentage consists of dividing the number of detected botnets by the numbers of total botnets. Just one network was involved in the first experiment so that there is an empty suspicious IP list. As more networks participated, the number of suspicious IPs on the list increased, so that the 500th network has the fraudulent IPs detected by the previous 499. This explains that as you increase the number of networks their ability to detect fraudulent IPs is higher.

Figure 2 shows the improvements to fraud detection methods against the number of networks working together and number of visits where captchas were applied. The X axis shows the number of networks working together in fraud detection and the Y axis shows the percentage of fraudulent IPs detected.

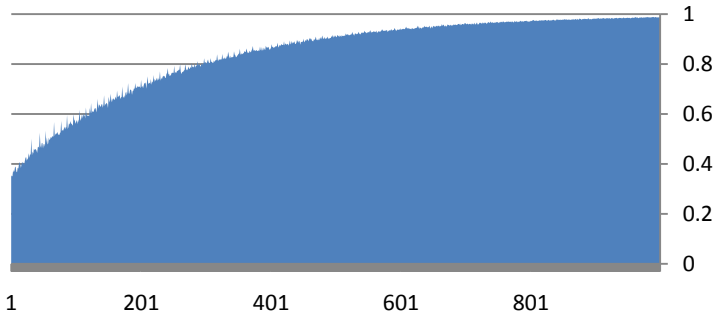

Fig. 2. Percentage of fraudulent IPs detected against number of networks working together in the fight against fraud.

5 Conclusions

The number of internet users has been constantly increasing since the creation of Arpanet back in 1969, the total number of users reached about 2.9 billion in 2014. In developing an online business it is essential to attract visitors, and to do that the most practical way is through online advertising campaigns. The most frequently used payment method is CPC (Cost per Click), where editors pay per every click made by users.

The amount of money circulating online has caught the attention of fraudsters for varying reasons. Such fraud arises from the fact it is relatively easy to commit given that victim and attacker are usually in different countries, and evidence is easy to manipulate.

Some examples of infractions committed in online advertising include click inflation, competition clicks, farmed clicks or the famous click bots. The smaller platforms have two difficult problems to solve, causing them to be less competitive when facing the large platforms, and hence are in danger of disappearing. These problems are fighting fraud and improving advertising performance.

On top of that small platforms lack the financial resources to develop technology to distinguish legitimate clicks from false clicks, by either botnets or humans. On the other hand, advertisers are increasingly focusing on micro-targeting which consists of small groups with similar interests. As publishers have few small networks they are quite often unable to meet the requirements of advertisers using such segmented campaigns.

In this article we have described a collaborative model designed to improve small network performance results as well as increasing their ability to detect fraud is designed. It has been demonstrated that the greater the number of networks cooperating the higher the number of adverts that can be covered.

To ensure adverts are shared fairly amongst networks, so everyone gets an equal gain three algorithms have been used: Simple, Round Robin and Minimum visits. Proving the minimum visits algorithm is the best of the three.

To improve fraud detection we have designed a collaborative environment in which each of the networks informs the rest whenever the IP of a determined click-bot or malicious user is detected, showing that detection is significantly improved when networks work together using captcha and irrelevant advertising techniques.

An interesting line of research and one which could be looked upon further is the optimization of campaign performance. The fact of optimizing campaigns to inform advertisers about parameters allow higher revenues to be gained facilitates advertising campaigns, making it unnecessary to hire an expert to review and analyze results.

Since millions of advertisers can participate in advertising exchanges, it is vital to design an algorithm to find the most relevant advert for every single visitor. Such an algorithm should run in a few tenths of a second, so multiple threads running in parallel will have to be used.

References

1. Goldberg, L., Silber, S.: IAB internet advertising revenue report. Retrieved from http://www.iab.net/about_the_iab/recent_press_releases/press_release_archive/press_release/pr-122313 (2012)
2. Chen, M.: The Effect of Fraud Investigation Cost on Pay-Per-Click Advertising. College of New Jersey, Galloway (2011)
3. Moe, W. W.: Targeting Display Advertising. London, UK: Advanced Database Marketing: Innovative Methodologies & Applications for Managing Customer Relationships (2013)

4. Mills, E.: Click fraud could threaten pay-per-click model. Retrieved from http://www.news.com/Study-Click-fraud-could-threaten-pay-per-clickmodel/2100-1024_3-6090939.html (2006)
5. Tuzhilin, A.: The Lane's Gifts v. Google Report. Mimeo, New York University. Retrieved from http://googleblog.blogspot.com/pdf/Tuzhilin_Report.pdf (2005)
6. Hui, W.: Estimating the Number of Genuine and Fraudulent Clicks in the Pay-Per-Click (PPC) Model. University of Nottingham Ningbo, China (2010)
7. Vidyasagar, N.: India's secret army of online ad clickers. The Times of India. Retrieved from http://timesofindia.indiatimes.com/business/india-business/Indias-secret-army-of-online-ad-clicers/articleshow/654822.cms (2004)
8. Mungamuru, B. W.-M.: Should Ad Networks Bother Fighting Click Fraud? (Yes, They Should). (2008)
9. Goodin, D.: Botnet caught red handed stealing from Google. Retrieved from http://www.theregister.co.uk/2009/10/09/bahama_botnet_steals_from_google (2009)
10. Wikipedia: Right to privacy. Retrieved from http://en.wikipedia.org/wiki/Right_to_privacy (2014)
11. Chow, R. E.: Making CAPTCHAs Clickable (2008)
12. Haddadi, H.: Fighting Online Click-Fraud Using Bluff Ads. ACM SIGCOMM Computer Communication Review, v. 40, no. 2 (2010)