

ISPs and Ad Networks Against Botnet Ad Fraud

Nevena Vratonjic, Mohammad Hossein Manshaei,
Maxim Raya, and Jean-Pierre Hubaux

Laboratory for computer Communications and Applications (LCA), EPFL, Switzerland
{nevena.vratonjic, hossein.manshaei, jean-pierre.hubaux}@epfl.ch
maxim.raya@gmail.com

Abstract. Botnets are a serious threat on the Internet and require huge resources to be thwarted. ISPs are in the best position to fight botnets and there are a number of recently proposed initiatives that focus on how ISPs should detect and remediate bots. However, it is very expensive for ISPs to do it alone and they would probably welcome some external funding. Among others, botnets severely affect ad networks (ANs), as botnets are increasingly used for ad fraud. Thus, ANs have an economic incentive, but they are not in the best position to fight botnet ad fraud. Consequently, ANs might be willing to subsidize the ISPs to do so. We provide a game-theoretic model to study the strategic behavior of ISPs and ANs and we identify the conditions under which ANs are likely to solve the problem of botnet ad fraud by themselves and those under which the AN will subsidize the ISP to achieve this goal. Our analytical and numerical results show that the optimal strategy depends on the ad revenue loss of the ANs due to ad fraud and the number of bots participating in ad fraud.

Keywords: Ad Fraud, Botnets, ISP, Ad Network, Security, Game Theory.

1 Introduction

Today, botnets are a very popular tool for perpetrating distributed attacks on the Internet. Botnets are a serious threat for a number of entities: end users, enterprises with online businesses, websites, Internet Service Providers (ISPs), advertisers and ad networks (ANs). Botnets usually consist of compromised end users' PCs. Thus, depending on the malware, the consequences for end users can be severe (e.g., stolen credentials). Very often botnets are used for sending spam, which creates problems for ISPs, enterprises and end users. Botnet operators (aka bot masters) also use botnets to extort money from websites' owners under the threat of Distributed Denial of Service Attacks (DDoS). Lately, it is becoming more and more popular to use botnets for ad fraud [4], which creates a loss of ad revenue for advertisers, associated websites and ad networks and security threats for end users (e.g., fraudulent ads that lead to phishing attacks).

Consequently, thwarting botnets would benefit everyone and would reduce the level of online crime on the Internet. However, the problem of botnets in general cannot be solved exclusively by users (lack of know-how), ISPs (too expensive to fight botnets alone), ad networks, advertisers, websites and enterprises (lack of tools and resources).

Recent initiatives propose that ISPs should perform the detection of botnets and remediation of the infected devices [20] [24]. Indeed, it is the ISPs that are in the best

position to detect the presence of a botnet and to take measures against it. Yet, the revenue of ISPs are not (directly) affected by the botnets and ISPs would probably welcome some external funding in the efforts to fight botnets. One possible approach is a government-sponsored program, as in Australia [7] and Germany [10]. In the case governments are unwilling to fund these initiatives, ISPs need to find a way to make them, at the very least, cost neutral if not cost positive.

Over the last decade, online advertising has become a major component of the Web, leading to annual revenues expressed in tens of billions of US Dollars (e.g., \$22.4 billion in the US in 2009 [5]). The business model of a fast growing number of online services is based on online advertising and much of the Internet activity depends on that source of revenue. Unsurprisingly, the ad revenue has caught the eye of many ill-intentioned people who have started abusing the advertising system in various ways. In particular, click fraud has become a phenomenon of alarming proportions [4]. Recently, a new type of ad fraud attack has appeared, consisting in the on-the-fly modification of the ads themselves. A prominent example is the *Bahama botnet*, in which malware causes infected systems to display altered ads, as well as altered results for Google or Yahoo searches to the end users [17]. Another example of such a botnet is Gumblar [16]. If the modification of ads is successful, users see ads that are different from what they would otherwise be. Consequently, users' clicks on the altered ads generate a revenue for the bot master instead of the AN. Thus, the modification of the ads negatively affects the revenues of the "legitimate" advertisers and undermines the business model of the ANs.

Considering the increasing trend of botnet ad-fraud attacks and the consequently increasing loss of ad revenue for ad networks, ANs have economic incentives to fight botnets. However, ANs are not in the best position to thwart botnets themselves and thus ANs might be willing to subsidize the ISPs to achieve that goal. In this paper, we investigate whether ad fraud botnets alone are cause enough for ISPs and ANs to cooperate. Such cooperation would help ISPs deploy detection and remediation mechanisms and would be a first step towards fighting all botnets.

The contributions of our paper are threefold. First, we identify two potential countermeasures that ANs could use to address the problem of botnet ad fraud and we propose a cooperation scheme in which ISPs and ANs jointly fight botnets. Second, we provide a game-theoretic model to study the interactions between ISPs and ANs, as well as to identify an optimal countermeasure strategy of ANs and ISPs under different conditions. Finally, we apply the results to a real data set to study the practical impact. To the best of our knowledge, this paper is the first to model the behavior of ISPs and ad networks facing botnet ad fraud.

The rest of the paper is organized as follows. After a brief presentation of the state of the art in Section 2, we describe the impact of botnets on the online advertising business in Section 3. We then address the various threats and countermeasures in Section 4 and provide a case study of a botnet ad fraud in Section 5. In Section 6, we present a game-theoretic model with two players, the ISP and the AN, and identify equilibrium outcomes of that game. We provide a numerical example to study the practical impact of the obtained results in Section 7 and conclude the paper in Section 8.

2 Related Work

There are two main categories of literature that are relevant to our work: research on fraud in online advertising and analyses of security investments on the Internet.

Research on online advertising fraud is mostly focused on click fraud [12] [15] [21]. Many problems that stem from online advertising and security gaps, especially the consequences for the end users, are addressed in [13]. The economics of click fraud are briefly addressed in [21]. In [8], the economic analysis based on a game-theoretic model of the online advertising market, shows that ad networks that deploy effective algorithms for click fraud detection gain a significant competitive advantage. If it is the case that some ad networks do not fight click fraud, mechanisms are proposed in [14] to protect online advertisers from paying for fraudulent clicks. In comparison, our model does not address click fraud detection mechanisms but introduces a new strategic player - the ISP - in addition to the traditional players in online advertising (i.e., ad networks, advertisers and publishers). Our results show that this player can yield significant implications for the security of the Internet.

In [30], the authors investigate novel man-in-the-middle attacks on online advertising systems, which can be perpetrated by access networks (e.g., an ISP) to exploit online advertising systems. The authors propose a collaborative secure scheme that relies on web servers and ad networks to fix the identified vulnerabilities of online advertising systems. This solution also relies on the fact that most of online advertising networks own digital authentication certificates and can become a source of trust. The authors explain why the deployment of this solution would benefit the Web browsing security in general. In this paper we propose new approaches to thwart distributed ad fraud attacks, where we address the possibility of collaboration between ISPs and ad networks.

Related to the second issue - finding the right incentives to increase the security on the Internet - there are several contributions in the literature. The game-theoretic approach of [18] models how users choose between investments in security (e.g., firewalls) or insurance (e.g., backup) mechanisms. The positive effect of cyberinsurance on the investment of agents in self-protection is analyzed using a game-theoretic model in [23]. The main conclusion of this work is that cyberinsurance without regulation does not provide sufficient incentives for self-protection. Another line of work proposes a centralized certification mechanism to encourage ISPs to secure their traffic and analyzes the resulting scheme using game theory [33]. In contrast to these works, our analysis shows that Internet security can be increased, under given conditions, without any central oversight and thanks to self-interested decisions by only a few key players (namely, the ad networks and the ISPs).

In [31], the authors investigate the recent problem of ISPs becoming strategic participants in the online advertising business. They propose a game-theoretic model of this problem to study the behavior and interactions of the ISPs and ANs. Their results show that if the users private information can improve ad targeting significantly and if ad networks do not have to pay a high share of revenue to the ISPs, ad networks and ISPs will cooperate to jointly provide targeted online ads. Otherwise, ISPs will divert part of the online ad revenue for themselves. In that case, if the diverted revenue is small, ad networks will not react. However, if their revenue loss is significant, the ad networks will invest into improving the security of the Web and protecting their ad revenue. Our work

also concludes that, when facing ad fraud, ANs are willing to collaborate with ISPs in order to protect their ad revenue. However, since we consider a distributed threat (in contrast to the centralized model in [31]) we propose new collaborative approach that takes into consideration the economic incentives of the ANs and ISPs.

3 System Model

We consider a system consisting of an *online advertising system*, a number of *bots* that attempt to exploit the online advertising system and an *ISP*, as depicted in Figure 1.

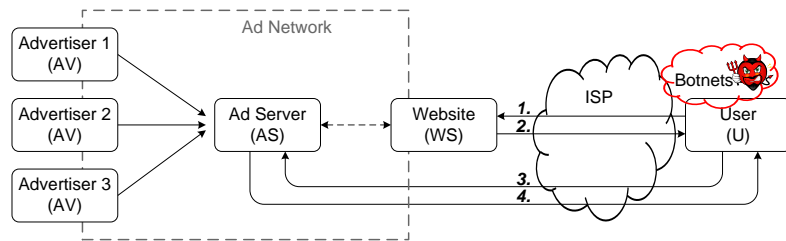


Fig. 1. System model: Online advertising system, ISP and bots exploiting the advertising system.

3.1 Online Advertising Systems

The most prevalent model of serving online ads to end users is depicted in Figure 1. To have their ads appear with the appropriate web content, Advertisers (AV) subscribe with an ad network (AN) whose role is to automatically embed ads into web pages. Ad networks have contracts with Websites (WS) that want to host advertisements. When a User (U) visits a website (Figure 1, step 1) that hosts ads, while downloading the content of the web page (step 2), the user's browser will be directed to communicate with one of the Ad Servers (AS) belonging to the ad network (step 3). The AS chooses and serves (step 4) the most appropriate ads to the user, such that users' interests are matched and the potential revenue is maximized. Throughout the rest of the paper, we use the terms "user" and "user's browser" interchangeably.

In the most popular ad revenue model [5], advertisers pay a *cost-per-click* (CPC) to the ad network for each user-generated click that directs the user's browser to the advertised website. The AN gives a fraction of the ad generated revenue to the WS that hosted the ad. Popular websites that attract more visitors create more traffic towards advertised websites, thus generating more revenue for themselves and for the associated ad networks. Since we consider a single AN in our system model, we assume that all the websites that host online ads are associated to that AN.

3.2 Botnets

A botnet is a collection of software robots, or *bots*, that run autonomously and automatically. Bots are typically compromised computers running software, usually installed via drive-by downloads exploiting Web browser vulnerabilities, worms, Trojan horses or backdoors, under a common command-and-control infrastructure. Recently, a botnet of compromised wireless routers has been detected [25]. Such a botnet has the advantage of having the bots almost always connected to the Internet (compared to the typical end-user machine that is connected to the Internet only from time to time). In addition, it is more difficult to detect that a device has been compromised, due to the lack of security software for such devices (e.g., no anti-virus software) or by a user.

A bot master controls the botnet remotely, usually through a covert channel (e.g., Internet Relay Chat) and usually for nefarious purposes. According to Click Forensics, a company that produces tools to detect and filter fraudulent clicks, for the third quarter of 2009, 42.6% of fraudulent clicks came from bots (Figure 2)¹ [4]. The number is the highest in four years (since Click Forensics has been producing the reports). For the same period in 2008, botnets accounted for 27.5% of fraudulent clicks. The data show that using botnets for ad fraud is becoming more and more popular. This creates a problem for advertisers, ad networks and websites as they lose a part of the ad revenue. In the system model, we consider a number of compromised devices that run a malware that causes infected machines to participate in an advertising fraud.

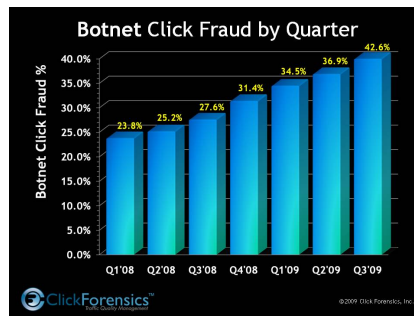


Fig. 2. Significance of botnet ad fraud: Botnet Click Fraud by Quarter.

3.3 Internet Service Providers

The traditional role of an ISP is to provide Internet access to end users and to forward users' traffic in compliance with the Network Neutrality policy [11]. However, recently, ISPs have begun taking on additional roles. In the EU, ISPs have to obtain and keep the records about their users' online activities and provide them upon request to law enforcement agencies [1].

¹ Permission for inclusion obtained from ClickForensics Inc.

A new IETF initiative focuses on how ISPs can manage the effects of devices used by their subscribers, detect those that have been infected with malicious bots, notify the subscribers and remediate the infection via various techniques [20]. The Internet Industry Association (IIA) has also drafted a new code of conduct that suggests ISPs should detect malware-infected machines of their subscribers and actually take the action to address the problem [24]. Complying with these initiatives, ISPs would make it more difficult for botnets to operate, thus helping to reduce the level of online crime on the Web. However, the problem is that ISPs have to find funding for those initiatives.

One possible approach is a government-sponsored program, such as the Australian Internet Security Initiative, in which a third-party helps identify malware-infected devices, notifies the appropriate ISP which then notifies and helps the subscriber to remedy the problem. About 90% of Australian ISP subscribers are covered by this initiative. A similar program is ready to be launched in 2010 in Germany, where ISPs are cooperating with the German Federal Office for Information Security [10]. In the case governments are unwilling to fund the initiative, ISPs need to find a way to make it, at the very least, cost neutral if not cost positive. In our model, we consider an ISP that is willing to comply to the initiative, if doing so is at least cost neutral.

4 Ad Fraud: Threats and Countermeasures

Due to the immense revenues generated by online advertising, the temptation to exploit the online advertising system is high. The loss of revenue for ad networks due to ad fraud is substantial. Based on the report from Click Forensics, the overall click-fraud rate was around 14.1% in the third quarter of 2009 [4], which means that 14.1% of the clicks on the ads were bogus. Thus, click fraud alone creates a significant loss of revenue for ad networks, advertisers and publishers. In addition, ANs lose ad revenue due to new types of ad fraud, such as injecting ads into the content of webpages on-the-fly between a web server and a user [19, 26, 29].

One possible approach for ad networks to protect their revenue is to improve the security of the online advertising systems, thus making it more difficult for an adversary to successfully exploit those systems. In [31], the authors use game theory to model AN's economic incentives and show that when facing ad fraud attacks securing ad systems may maximize the revenue of a rational AN. For example, ad fraud can be reduced if webpages and ads are served over HTTPS instead of HTTP. The cost of implementing HTTPS at a web server includes the cost of obtaining a valid X.509 authentication certificate. Usually, website owners are not willing to bear this cost. Thus, if an ad network wants the secure protocol to be deployed, it should cover the costs itself. As explained previously, websites are not of the same value to the ad network, because of the different ad revenue they generate, but the cost of securing the ad revenue from a website is the same for all websites. Therefore, the ad network may decide to selectively secure only the websites that generate sufficient ad revenue that would compensate the costs.

Another possible approach for ad networks to protect their revenue is to cooperate with ISPs and eliminate the major cause of the revenue loss, botnets. They can do so by funding the existing initiatives for ISPs to detect and remove botnets, since ISPs are in

a privileged position to fight botnets. As removing botnets would benefit ad networks, they have economic incentives to subsidize ISPs to fight botnets.

Thus, we can envision the following two scenarios of ad networks fighting ad fraud: (i) improving the security of the online advertising systems or (ii) funding ISPs to fight botnets involved in ad frauds.

5 Botnet Ad Fraud: A Case Study

Consider the system as described in Section 3, in which N_B devices (e.g., end-users' computers or routers) have been infected by a malware and participate in ad fraud. We consider exclusively the types of ad fraud: (i) that has been the most prominent lately [16, 17], in which malware causes infected devices to return altered Search Engine Result Pages (SERPs) or altered content of the ads in web pages, due to DNS poisoning and (ii) in which subverted users' routers modify ad traffic on-the-fly between a web server and a user [29]. In the example of Bahama botnet, malware uses DNS poisoning by modifying HOSTS files on infected machines to redirect traffic to rogue Google servers which return altered results [9]. Thus, affected users see ads and links that are different from what they would otherwise be. When users click on the altered ads, the clicks generate revenue for the bot master instead of the ad network. Thus, the bots divert a part of the ad revenue from the ad network. For simplicity of treatment, we assume that each bot diverts an equal part of the revenue and in aggregate, all the bots together divert $\lambda \in (0, 1]$ fraction of the total ad network's revenue P . Thus, the revenue of the AN in the case of ad fraud is $P(1 - \lambda)$.

The popularity of websites, and consequently the number of user-generated clicks on ads, follow a heavy-tail distribution [6]. We infer the generated volume of clicks on ads on the 1000 most popular websites, based on the data of page views on each website in 2009, obtained from *Compete.com*. The exposure of users to online ads has been evaluated extensively in [22], showing that 58% of the top 1000 websites host advertisements and there are 8 ads per web page on average. The probability that a click occurs on an advertisement is 0.1% [2]. Consequently, to convert the number of page views into the number of clicks on ads on each website, we use the following formula: $Q(n) = (\text{Page views on the website } n) \times 0.58 \times 8 \times 0.001$. Figure 3 shows the annual number of clicks $Q(n)$ on ads, where $n \in \{1, 2, \dots, 1000\}$ is the popularity rank of a website.

Applying curve fitting to the data set, we obtain that the distribution of clicks on ads across websites corresponds to the power law $Q(n) = \alpha \cdot n^{-\beta}$, where $Q(n)$ is the annual number of clicks on ads that occurred at the website with the n -th rank. The obtained parameters of the power law are $\alpha = 3.18 \cdot 10^9$ and $\beta = 1.044$ as shown in Figure 3. In general, we assume that the number of clicks on ads follows the power law distribution $Q(n) = \alpha \cdot n^{-\beta}$, where $Q(n)$ is the annual number of clicks on ads that occurred at the website with the n -th rank and $\beta > 1$ [6]. Note that the value of parameters α and β are a characteristic of a given AN and depend on the number and the type of associated websites. In order to extend our analysis and investigate what would be the effect on the entire Web (i.e., for all websites), we extrapolate the data set we have obtained from *Compete.com* with the obtained power law.

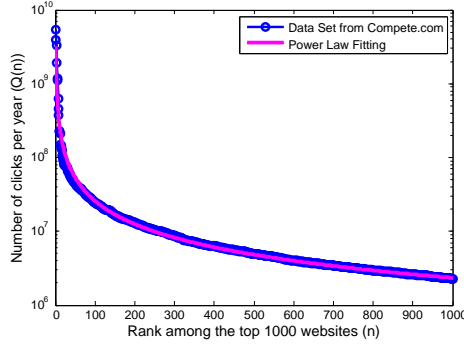


Fig. 3. Annual number of clicks for the 1000 most popular websites and the power law fitting curve, $Q(n) = \alpha n^{-\beta} = 3.18 \cdot 10^9 n^{-1.044}$.

Given the power law distribution of the clicks, the ad revenue generated by the top x websites can be estimated by²

$$k \int_1^x \alpha n^{-\beta} dn = k \frac{\alpha}{(\beta - 1)} (1 - x^{1-\beta}),$$

where k is the amount of revenue that each click on ads generates for the ad network³. If P is the total revenue of the ad network, generated by all the websites (i.e., when $x \rightarrow \infty$), then per-click revenue can be calculated by $k = \frac{P(\beta-1)}{\alpha}$. According to the reports [5], the total ad revenue P in 2009 in the US is 22.4 billion dollars.

In the following two subsections, we analyze the two proposed strategies (i.e., improving the security of the online advertising system and cooperation between the AN and the ISP) to fight botnet ad fraud. Table 1 shows the used notation.

5.1 Securing Websites

As a countermeasure to the considered type of ad fraud (i.e., rogue servers delivering altered ads due to DNS poisoning attack on users machines or on-the-fly traffic modifications by compromised users' routers), the AN can secure the communication between users and web servers as well as between users and ad servers. For example, secure communication can be provided by the HTTPS protocol. Deploying HTTPS requires web servers to obtain an authentication certificate from a trusted third party. In the case when websites and ad servers deploy HTTPS with valid authentication certificates, even if an adversary successfully mounts a DNS poisoning attack and redirects users' communication to rogue servers, the rogues servers cannot serve valid authentication certificates that correspond to the domain names users originally wanted to visit,

² Due to the impossibility of obtaining closed-form expressions in the discrete domain, we perform computations in the continuous domain. The upper bound of the error is 8% [32].

³ Modeling auctions and different per click revenue for ad networks is out of the scope of this paper, thus we assume that all the clicks are of the same quality.

Table 1. Table of symbols.

| Symbol | Definition |
|----------------------|---|
| N_B | Number of bots |
| λ | Fraction of diverted ad revenue by the botnet |
| P | Total online advertising revenue of the AN |
| k | Amount of generated revenue for each click |
| $Q(n)$ | Number of clicks per year for the top 1000 websites |
| n | Popularity rank of the websites |
| α and β | Estimated parameters of power law distribution for $Q(n)$ |
| c_S | Cost of securing a website |
| N_S | Optimal number of secured websites with S strategy |
| N_{SC} | Optimal number of secured websites with $S + C$ strategy |
| P_D | Fraction of bots detected by the ISP |
| c_D | Cost of the botnet detection system |
| c_R | Cost for the ISP per remediated infected device |
| R | Cost for the AN per remediated infected device |
| N_R | Optimal number of remediated infected devices |
| C | Cooperation strategy (employed by the ISP or the AN) |
| S | Secure websites strategy by the AN |
| $S + C$ | Simultaneous Secure and Cooperation strategy by the AN |
| A | Abstain Strategy (employed by the ISP or the AN) |

thus browsers will detect security issues. HTTPS also prevents on-the-fly modifications of the content. Consequently, users would receive unaltered links and ads and the clicks on unaltered ads would generate revenue for the intended AN, not the adversary.

As discussed in Section 4, website owners usually lack incentive to bear the cost of obtaining a valid certificate. Thus, to secure the communication, and consequently the ad revenue, the AN would have to pay a cost of securing the website. The cost of deploying HTTPS at ad servers can be considered negligible, given that the AN already has a valid certificate and that there are typically only a few ad servers (compared to the number of web servers).

Let c_S be the cost of securing a website, i.e., the cost of obtaining a certificate and deploying HTTPS at a web server. Then the AN should pay $N_S \cdot c_S$ to secure N_S websites. N_S is the optimal number of websites that AN secures to maximize its payoff in the presence of N_B bots diverting fraction λ of the revenue. It can be calculated by the following lemma.

Lemma 1. *If the ad network fights botnet ad fraud by securing the websites, the optimal number of those secured websites is equal to $N_S = \left(\frac{P}{c_S} \lambda (\beta - 1)\right)^{\frac{1}{\beta}}$.*

Proof. The total amount of revenue for the ad network (u_{AN}) when it secures x websites, due to the attack of N_B bots diverting fraction λ of the revenue, can be estimated by

$$u_{AN} = k \int_1^x \alpha n^{-\beta} dn + (1 - \lambda)k \int_x^\infty \alpha n^{-\beta} dn - c_S x.$$

Recall that k is the revenue generated per each click and can be calculated as $\frac{P(\beta-1)}{\alpha}$. The first term in the revenue equation represents the revenue that the AN obtains from clicks generated on secured websites. The second term shows that the AN obtains only the remaining fraction $(1 - \lambda)$ of the revenue from clicks generated on unsecured websites, as the bots divert the fraction λ of the revenue.

After simplifications we obtain: $u_{AN} = P(1 - \lambda x^{1-\beta}) - c_S x$, which is a concave function of x . We can obtain the optimal N_S by finding the root of the first derivation of u_{AN} with respect to x , that is $\left(\frac{P}{c_S} \lambda (\beta - 1)\right)^{\frac{1}{\beta}}$. \square

5.2 ISP and Ad Network Cooperation

In addition to the just described countermeasure of securing websites, the AN can offer the ISP to cooperate in the fight against botnets. The AN has an economic incentive to fund the ISP to perform detection of the botnets and remediation of the infected devices, as discussed in Section 3. To detect bots in the network, the ISP must deploy a detection system [20, 24]. We note the deployment cost of the detection system as c_D and we assume that such a system can successfully detect a fraction P_D of the bots in the network. The proposed initiatives [20, 24] envision an online help desk where all the subscribers whose devices have been detected as bots can obtain instructions on how to remediate the problem and restore the functionality of their devices. Thus, the ISP has a cost per each remediated infected device, which we note as c_R .

For the ISP to cooperate with the AN, the AN has to provide a sufficient reward such that the detection and remediation is at least cost neutral for the ISP. Let R represent the reward the AN should pay to the ISP for the remediation of each infected device.⁴

If the AN and the ISP agree to cooperate, the outcome is that the ISP remediates N_R infected devices and the AN pays $N_R \cdot R$ to the ISP. The optimal N_R that maximizes both revenues, of the ISP and the AN, can be calculated by the following lemma.

Lemma 2. *The cooperative ISP and the cooperative AN can maximize their revenues by remediation of $N_R = P_D N_B$ infected devices.*

Proof. The total amount of revenue that the ISP can obtain by cooperation and remediation of x infected devices is $x(R - c_R) - c_D$ which is a linear function of x . Therefore, the ISP can maximize its revenue by remediating all of detected bots $P_D N_B$. Remediation of x infected devices reduces the aggregate power of the bots in the network, and together they can divert only a fraction $\lambda(1 - \frac{x}{N_B})$ of the revenue. The total amount of revenue that the AN can obtain by cooperation is then $P(1 - \lambda(1 - \frac{x}{N_B})) - xR = \left(\frac{P\lambda}{N_B} - R\right)x + P(1 - \lambda)$, which is a linear function with respect to x and will be maximized at $x = N_R = P_D N_B$, i.e., for all of the detected bots. \square

In summary, the ad network can use one of the above two actions to fight botnet ad fraud in the Internet. Each strategy has different benefits and costs for the ISP and the

⁴ Our model also applies to the case when ISPs and ANs jointly bear the costs (i.e., when it is cost negative for ISPs to thwart the botnets) by adapting the values R or c_R .

Table 2. Static game: *ISP* chooses an action from $\{A, C\}$; *AN* from $\{A, C, S+C, S\}$. Strategy profiles (C, A) and $(S+C, A)$ are not applicable unless when *ISP* plays C .

| | | ISP | |
|----|-----|---|--|
| | | A | C |
| AN | A | $(0, P(1 - \lambda))$ | $(-c_D, P(1 - \lambda))$ |
| | C | N/A | $(N_R(R - c_R) - c_D, P(1 - \lambda(1 - \frac{N_R}{N_B}))) - N_R R$ |
| | S+C | N/A | $(N_R(R - c_R) - c_D, P(1 - \lambda(1 - \frac{N_R}{N_B})N_S^{1-\beta}) - N_S c_S - N_R R)$ |
| | S | $(0, P(1 - \lambda N_S^{1-\beta}) - N_S c_S)$ | $(-c_D, P(1 - \lambda N_S^{1-\beta}) - N_S c_S)$ |

AN. In the next section, we use game theory to model this situation and consequently predict the behavior of the AN and the ISP in different situations.

6 Game-Theoretic Model

In this section, we introduce a static game \mathbf{G} to analyze the interaction between the ISP and the AN. Our model considers potential strategies of the ISP and the AN to protect the systems against the above defined threats. Considering the benefits and the costs of different strategies we also present the equilibria for the defined game. The key points of our game-theoretic analysis is that by using the computed equilibria it is possible to choose the optimal countermeasure protocol for different situations. Note that our game is a perfect and complete information game. We assume that the players have common knowledge about their strategies and payoffs and can observe the actions of each other.

6.1 Game Model: Strategies and Payoffs

Table 2 shows the normal form of the proposed static game \mathbf{G} . In this game, the players play simultaneously. The ISP can choose between the following two actions: *Abstain* (A) and *Cooperate* (C). The *Abstain* action models the behavior of the ISP that is not willing to participate in the detection and remediation of the bots. Hence the payoff of the ISP is 0, when it plays A . The cooperative ISP (that plays C) first detects the bots and then remediates the infected devices. In return, the ISP receives a reward $N_R R$ from the AN. Recall that the cost for the ISP to remediate all detected devices is $c_R N_R$. Consequently, when the ISP and the AN cooperate, the payoff of the ISP is $N_R(R - c_R) - c_D$.

In our model, the AN can choose one of the following four possible actions: *Abstain* (A), *Cooperate* (C), *Secure and Cooperate* ($S+C$), and *Secure* (S). With the *Abstain* action we model the behavior of the AN that is not willing to perform any countermeasures. In this case, the payoff of the AN will decrease to $P(1 - \lambda)$. Recall that $\lambda \in [0, 1]$ is the fraction of diverted ad revenue by the bots.

If the AN cooperates with the ISP, its utility will increase to $P(1 - \lambda(1 - \frac{N_R}{N_B}))$, where N_R is the optimal number of infected devices remediated by the ISP, which can

be calculated by Lemma 2. However, the AN should pay $N_R R$ to the ISP for N_R remediated devices. As a result, the total payoff of the AN when both players are cooperative is $P \left(1 - \lambda \left(1 - \frac{N_R}{N_B} \right) \right) - N_R R$.

The AN can also secure the websites by choosing the action S , as discussed in Section 5.1. The AN should pay $N_S c_S$ to secure N_S websites. The benefit of the AN will then increase to $P(1 - \lambda N_S^{1-\beta})$. Consequently, the total payoff of the AN when it plays S is $P(1 - \lambda N_S^{1-\beta}) - N_S c_S$, independently of whether the ISP plays C or A .

Finally, the AN can choose to simultaneously secure some of the websites and cooperate with the ISP to remediate some of the infected devices. This action is represented by $S + C$ and the total payoff of the AN in this case is $P(1 - \lambda N_{SC}^{1-\beta} (1 - \frac{N_R}{N_B})) - N_{SC} c_S - N_R R$, where N_{SC} is the optimal number of secured websites when the AN plays $S + C$ and can be obtained by the following lemma.

Lemma 3. *If the AN fights botnet ad fraud with both countermeasures (action $S + C$), the optimal number of secured websites is equal to $N_{SC} = \left(\frac{P}{c_S} \lambda (\beta - 1) \left(1 - \frac{N_R}{N_B} \right) \right)^{\frac{1}{\beta}}$.*

Proof. The proof is similar to Lemma 1. We can obtain the optimal N_{SC} , by maximizing the total payoff of the AN when it plays $S + C$. \square

Lemma 3 shows that when the AN plays $S + C$ a smaller number (N_{SC}) of websites is secured, compared to the number (N_S) of secured websites when the AN plays S (i.e., $N_{SC} = N_S \left(1 - \frac{N_R}{N_B} \right)^{\frac{1}{\beta}} < N_S$).

6.2 Game Results

In order to predict and choose the optimal action for the ISP and the AN, we investigate all Nash equilibrium strategy profiles of the defined game. In other words, we are interested in finding the strategy profiles, where neither the ISP nor the AN can increase their payoffs by unilaterally changing their strategies. We will check the existence of Nash equilibria by comparing the payoffs obtained in the game \mathbf{G} .

The following theorem states conditions when the AN does not provide sufficient incentive to the ISP, such that the ISP will abstain at the Nash equilibrium.

Theorem 1. *In \mathbf{G} , if $R < \frac{c_D}{N_R} + c_R$, the best response of the ISP is to play action A .*

Proof. By comparing the ISP's payoff when it plays C (i.e., whether $-c_D$ or $N_R(R - c_R) - c_D$) with that of A (i.e., 0) we obtain that the best response of the ISP is A if $N_R(R - c_R) - c_D < 0$ or $R < \frac{c_D}{N_R} + c_R$. \square

This means that if the reward for remediation of the infected devices is small, the ISP will not be willing to cooperate with the AN to fight the bots.

The following theorem states when the revenue loss due to ad fraud is not significant enough to cause the AN and the ISP to perform any countermeasure against the bots.

Theorem 2. *In \mathbf{G} , if $R < \frac{c_D}{N_R} + c_R$ and $\lambda \leq \frac{N_S c_S}{P(1 - N_S^{1-\beta})}$, the action A by the ISP and the AN result in a Nash equilibrium.*

Proof. Considering Theorem 1, the ISP chooses A as its best response. The AN also plays A if its payoff when playing A (i.e., $P(1 - \lambda)$) is bigger than its payoff when playing S (i.e., $P(1 - \lambda N_S^{1-\beta}) - N_S c_S$). Comparing these two payoffs results in the second condition of this theorem, i.e., $\lambda \leq \frac{N_S c_S}{P(1 - N_S^{1-\beta})}$. \square

In other words, if the reward provided by the AN does not generate sufficient incentives for the ISP to cooperate, and the amount of revenue diverted by the bots is smaller than a given threshold, both the ISP and the AN choose A to be at Nash equilibrium.

Theorem 3 shows when the AN fights the bots alone by securing some of the websites.

Theorem 3. In \mathbf{G} , if $R < \frac{c_D}{N_R} + c_R$ and $\lambda > \frac{N_S c_S}{P(1 - N_S^{1-\beta})}$, action A by the ISP and action S by the AN result in a Nash equilibrium.

Proof. The proof is similar to Theorem 2. \square

This result shows that the amount of diverted ad revenue is significant such that a countermeasure should be deployed, but the ISP does not have enough incentive to cooperate and fight bots at this equilibrium. Consequently, the AN secures some of the websites.

Let us assume that λ is very small. Considering all the possible actions and the corresponding payoffs for the AN, the *Abstain* results in maximum payoff for the AN. In fact, action A avoids unnecessary costs for the AN, such as $N_S c_S$ or $N_R R = P_D N_B R$. These results are also in line with Theorem 2 meaning that playing A by both players results in a Nash equilibrium when λ is very small.

When λ increases (i.e., more ad revenue is diverted by the bots) the AN should deviate from A and select one of the three remaining actions as its best response. The following lemma states when the AN should begin securing N_S websites.

Conjecture 1. In \mathbf{G} , the AN should start securing the websites (Play S) when $\lambda > \frac{N_S c_S}{P(1 - N_S^{1-\beta})}$, which corresponds to the equilibrium presented by Theorem 3.

Proof. We should compare the payoffs of the AN when it plays S or C , with the one obtained by playing the action A . The AN should then play C if $\lambda > \frac{R N_B}{P} = \lambda_1$ and should play S if $\lambda > \frac{N_S c_S}{P(1 - N_S^{1-\beta})} = \lambda_2$. One can show that $\lambda_1 > \lambda_2$ when λ , and consequently N_S , is small enough. This means that the AN switches from A to S at equilibrium, when λ increases. \square

Note that the AN does not switch from the A to the $S + C$, when λ increases, because the AN can protect the revenue first by playing S . In other words, the AN does not need to pay $N_R R$ to the ISP, since the cost would exceed the revenue loss. Consequently, the equilibrium of \mathbf{G} corresponds to the one presented by Theorem 3. Finally, the following conjecture shows when the AN plays $S + C$ in the response to the cooperative ISP.

Conjecture 2. In \mathbf{G} , if the ISP is cooperative, the best response of the AN is action $S + C$ if $\lambda > \frac{N_R R - N_S c_S G}{P N_S^{1-\beta} G}$, where $G = 1 - (1 - \frac{N_R}{N_B})^{\frac{1}{\beta}}$.

Proof. The above threshold can be obtained by comparing the payoffs of the AN when it plays $S + C$ and S . \square

Conjecture 2 shows that if the bots divert even more revenue from the ad network, the AN will cooperate with the ISP and pay $N_R R$ to the ISP to remediate N_R bots. It will then secure a smaller number of websites compared to the case when it plays S .

7 Numerical Analysis

In order to understand implications of the analytical results (presented in Section 6) in reality, we simulate the game using the real data. We compute numerically the payoffs of the static game (Table 2), identify the resulting equilibria and present conclusions. To investigate the effect on the entire Web (i.e., for all websites), we extrapolate the data set we have obtained from *Compete.com* with the obtained power law, as explained in Section 5.

We use the following estimated costs in our evaluations: (i) the cost of deploying HTTPS at a web server is $c_S = \$400$ [28]; (ii) the cost of remediating an infected device $c_R = \$100$ (given that it is done via online support [20], it is the estimated cost of human labor for remediating one device per hour); (iii) the cost of the intrusion detection system $c_D = \$100k$ [27].

We take into account different values of the fraction $\lambda \in (0, 1]$ of the ad revenue that the AN loses due to botnet ad fraud and the number of bots N_B . Given that the largest botnets detected so far [3] had several million bots each, we consider the total number of infected devices that participate in the ad fraud considered in our case study to be up to 100 million (regardless of whether they form a single or multiple botnets).

We represent the outcomes of the game for $N_B = 10^4$ in Figure 4. Figure 4(a) shows the number of secured websites depending on the level of threat λ . When the AN cooperates with the ISP, the fraction of remediated devices depending on the level of threat λ is shown in Figure 4(b). We consider three scenarios (the three curves in Figure 4), for three different efficiencies of the detection system employed by the ISP (i.e., when the fraction of detected bots is $P_D = 0.1$, $P_D = 0.5$ and $P_D = 0.9$).

When the threat of the botnet ad fraud is very small, $\lambda < 2 \cdot 10^{-6}$, the AN does not perceive the need to perform any countermeasure against bots. Thus, there are no websites that are secured ($N_S = 0$ in Figure 4(a))⁵ and no devices are remediated ($N_R = 0$ in Figure 4(b)). This result corresponds to Theorem 2.

⁵ Absence of curves in Figure 4(a) signifies $\log(0)$, i.e., that zero websites are secured.

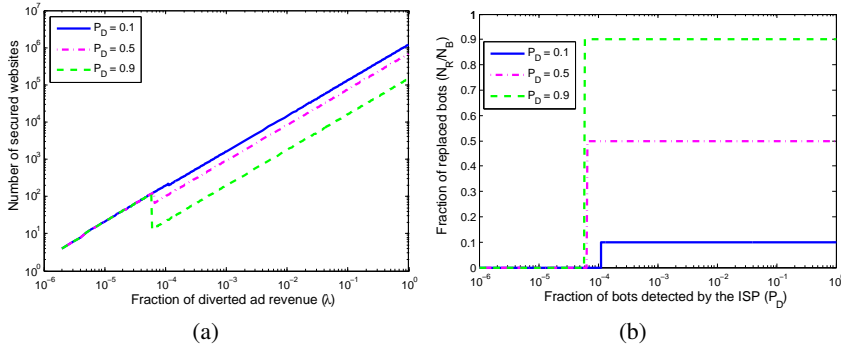


Fig. 4. Outcomes of the game applied to real data when $N_B = 10^4$: (a) Number of the most popular websites that should be secured; (b) Fraction of infected devices remediated by the ISP.

When the bots divert a higher fraction of ad revenue, $\lambda > 2 \cdot 10^{-6}$, the AN first secures a number of websites (Figure 4(a)). As there is no cooperation with the ISP ($N_R = 0$ in Figure 4(b)) the number of secured websites does not depend on P_D , thus it is the same in all three scenarios. The result corresponds to the finding of Theorem 3, i.e., the best choice for the AN is to play *Secure* and for the ISP to *Abstain*. The intuition behind this result is that the relatively small threat λ is distributed over N_B infected devices, thus each bot diverts a small amount of ad revenue. The cost of remediating the infected device would be higher than the loss of ad revenue the bot causes, thus it does not pay off for the AN to cooperate with the ISP. However, the loss is significant enough that the AN has to deploy a countermeasure, hence it secures some of the websites. The number of secured websites corresponds to the Lemma 1.

We observe that the higher λ is, the higher is the number of websites to be secured (Figure 4(a)), until λ reaches a threshold value ($\lambda_1 = 1.12 \cdot 10^{-4}$, $\lambda_2 = 6.6 \cdot 10^{-5}$ and $\lambda_3 = 6 \cdot 10^{-5}$ for $P_D = 0.1$, $P_D = 0.5$ and $P_D = 0.9$, respectively). At the threshold values the AN starts cooperating with the ISP (N_R becomes greater than zero, Figure 4(b)). Thus, the threshold value of λ represents the level of threat after which it is not enough to only secure the websites, but the AN will also cooperate with the ISP to fight bots (i.e., plays $S + C$). This result corresponds to Lemma 2.

When the AN plays $S + C$, each countermeasure protects a given part of the revenue that is otherwise diverted by the bots. The total loss of revenue for the AN due to ad fraud committed by N_B bots is $P\lambda$. The remediation of N_R infected devices reduces the loss of revenue to $P\lambda(1 - \frac{N_R}{N_B})$. As the part of the revenue loss is now eliminated by the ISP, the remaining part is smaller and consequently the AN secures a smaller number of websites. This explains the drop in the number of secured websites (Figure 4(a)), which happens at the threshold value of λ when the AN starts cooperating with the ISP. When λ increases (for values of λ greater than the thresholds), since N_R is constant for a given P_D (Figure 4(b)), in order to eliminate the increasing loss, the AN secures an increasing number of websites for the increasing λ (Figure 4(a)).

In Figure 4(b), we observe that the number of remediated devices is equal to $P_D N_B$, which confirms analytical results stated by Lemma 2. The higher the P_D is, the bigger the benefit of cooperation is, because a larger number of devices is remediated. Consequently, the AN secures a smaller number of websites for a higher P_D (Figure 4(a)).

In summary, the obtained results illustrate that: (i) For a very low level of threat λ , no countermeasures will be taken against bots; (ii) When the fraction λ of the diverted revenue increases, the AN secures a number of websites; (iii) Securing websites is not sufficient for an even higher level of threats, thus the AN will cooperate with the ISP to remediate infected devices.

Next, we analyze the effect of the number of bots N_B in the system on the equilibrium outcomes of the game.

Figure 5 represents the outcomes of the game, in the case of $N_B = 10^7$. Figure 5(a) shows the number of secured websites depending on the level of threat λ . The fraction of remediated devices depending on the level of threat λ is shown in Figure 5(b). As before, the three curves in Figures 5(a) and 5(b), correspond to the three scenarios ($P_D = 0.1$, $P_D = 0.5$ and $P_D = 0.9$).

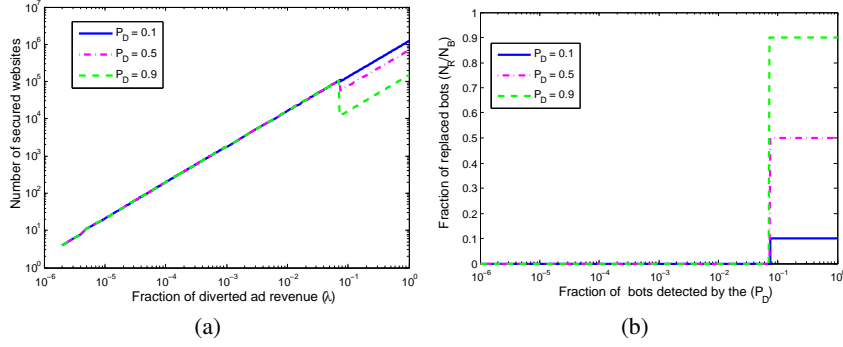


Fig. 5. Outcomes of the game applied to real data when $N_B = 10^7$: (a) Number of the most popular websites that should be secured; (b) Fraction of infected devices remediated by the ISP.

We observe the same behavior as in the case of $N_B = 10^4$ bots in the system. The difference in the results for the case of $N_B = 10^7$ (Figure 5) compared to results for the case of $N_B = 10^4$ (Figure 4 in Section 7) is that the threshold values of λ , for which the AN begins to cooperate with the ISP, are higher. The explanation for this results is the following.

When cooperating, the ISP remediates $P_D N_B$ devices, and the AN pays $P_D N_B \cdot R$ to the ISP. Therefore, the cost of cooperation for the AN is higher when N_B is higher. Whereas, the benefit for the AN, due to remediation of $N_R = P_D N_B$ devices is $P \lambda \frac{N_R}{N_B} = P \lambda P_D$, which does not depend on N_B . For a given P_D , the cooperation benefit for the AN is higher only for the higher threat λ . Hence, when the number of bots N_B is high, the AN agrees to cooperate and pay the high cost $P_D N_B R$, only when the fraction λ of the revenue bots divert is high. Because only for the high λ the cooperation benefit $P \lambda P_D$ is high enough to justify the costs of cooperation.

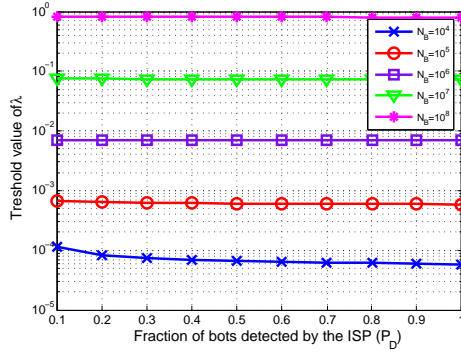


Fig. 6. Threshold values of λ for which the AN begins cooperating with the ISP, in addition to securing the websites.

Figure 6 illustrates the threshold values of λ for different numbers of bots N_B in the system and for different efficiencies P_D of the detection system. For example, in the system with $P_D = 0.5$ and $N_B = 10^4$ the AN is cooperative when $\lambda > 6.6 \cdot 10^{-5}$. Whereas, if N_B is much higher, $N_B = 10^8$, the AN is cooperative only if the fraction of diverted revenue is much higher, $\lambda > 0.8$. The results confirm that for a system with a given P_D , when the number of bots is high, the AN is cooperative only when the revenue loss is very high. Based on the results in Figure 6, we also observe that the threshold value of λ does not vary much for different values of P_D . Hence, we can conclude that the value of N_B is the dominant factor in the decision of the AN whether to cooperate with the ISP or not. These results are also confirmed by Lemma 2.

8 Conclusion

In this paper, we have investigated the novel situation of ISPs and ad networks behaving as strategic participants in the efforts to fight botnets. Due to the revenue loss caused by botnet ad fraud, ad networks have economic incentives to protect their revenue by either: (i) improving the security of the online advertising systems or (ii) fighting the major cause of the revenue loss, botnets. To fight botnets, ad networks might need help from ISPs, who are in a better position to deploy detection and remediation mechanisms. We have proposed a game-theoretic model to study the behavior and interactions of the ISPs and ad networks. We have applied our model to the real data to understand the meaning of the results in practice. Our analysis shows that cooperation between the AN and the ISP could emerge under certain conditions that mostly depend on: (i) the number of infected devices (ii) the aggregate power with which bots divert revenue from the ad network and (iii) the efficiency of the botnet detection system. The cooperation is a win-win situation where: (i) users benefit from the ISP's help in maintaining the security of users' devices; (ii) the AN protects its ad revenue as the botnet ad fraud is reduced; (iii) it is at least cost neutral, if not cost positive for the ISP to fight botnets. Cooperation between the AN and the ISP would help to reduce the level of online crime and improve the Web security in general.

Acknowledgements

We would like to thank Wojciech Galuba, Julien Freudiger, and Mathias Humbert for their insights and suggestions on earlier versions of this work, and the anonymous reviewers for their helpful feedback.

References

1. Directive 2006/24/EC of the European parliament and of the council. Official Journal of the European Union (2006)
2. 2008 Year-in-Review Benchmarks. DoubleClick Research Report (2009)
3. Biggest, Baddest Botnets: Wanted Dead or Alive. PC World (2009), http://www.pcworld.com/article/169033/biggest_baddest_botnets_wanted_dead_or_alive.html
4. Click Fraud Index. ClickForensics Inc. (2009)
5. Internet Advertising Revenue Report. Interactive Advertising Bureau (2009)
6. Adamic, L.A., Huberman, B.A.: The Web's hidden order. Communication ACM (2001)
7. Australian Internet Security Initiative (AISI), A.C., Media Authority: (2010), http://www.acma.gov.au/WEB/STANDARD/1001/pc=PC_310317
8. B. Mungamuru, S.W., Garcia-Molina, H.: Should Ad Networks Bother Fighting Click Fraud? (Yes, They Should.). Technical report, Stanford InfoLab (2008)
9. Click Forensics Discovers Click Fraud Surge from New Sophisticated Bahama Botnet: (2009), <http://www.clickforensics.com/newsroom/press-releases/144-bahama-botnet.html>
10. Constantin, L.: German Government to Help Rid Computers of Malware (2009), <http://news.softpedia.com>
11. Crowcroft, J.: Net Neutrality: The Technical Side of the Debate: A White Paper. SIGCOMM Computer Communication Review (2007)
12. Daswani, N., Stoppelman, M.: The Anatomy of Clickbot.A. In: Hot Topics in Understanding Botnets (HotBots) (2007)
13. Edelman, B.G.: Securing Online Advertising: Rustlers and Sheriffs in the New Wild West. SSRN eLibrary (2008)
14. Edelman, B.G.: Deterring Online Advertising Fraud Through Optimal Payment in Arrears. SSRN eLibrary (2009)
15. Gandhi, M., Jakobsson, M., Ratkiewicz, J.: Badvertisements: Stealthy Click-Fraud with Unwitting Accessories. Digital Forensic Practice 1(2) (2006)
16. Viral Web infection siphons ad dollars from Google: http://www.theregister.co.uk/2009/05/14/viral_web_infection/
17. Botnet caught red handed stealing from Google: (2009), http://www.theregister.co.uk/2009/10/09/bahama_botnet_steals_from_google
18. Grossklags, J., Christin, N., Chuang, J.: Secure or insure?: a game-theoretic analysis of information security games. In: International conference on World Wide Web (WWW) (2008)
19. Growing number Of ISPs Injecting Own Content Into Websites: (2008), <http://www.techdirt.com/articles/20080417/041032874.shtml>
20. J. Livingood, N. Mody, M.O., Communications, C.: Recommendations for the Remediation of Bots in ISP Networks. Internet-Draft Version 3, IETF (2009)
21. Jakobsson, M., Ramzan, Z.: Crimeware. Addison-Wesley, Reading, MA (2008)
22. Krishnamurthy, B., Wills, C.E.: Cat and Mouse: Content Delivery Tradeoffs in Web Access. In: International conference on World Wide Web (WWW) (2006)
23. Lelarge, M., Bolot, J.: Economic Incentives to Increase Security in the Internet: The Case for Insurance. In: INFOCOM (2009)
24. Livingood, J., Mody, N., O'Reirdan, M., Comcast Communications: ISP Voluntary Code of Practice for Industry Self-regulation in the Area of e-security. Internet industry code of practice, Internet Industry Association (2009)
25. Network Blueprint: Stealth Router-based Botnet: (2009), <http://dronebl.org/blog>

26. Reis, C., Gribble, S.D., Kohno, T., Weaver, N.C.: Detecting In-Flight Page Changes with Web Tripwires. In: USENIX Symposium on Networked Systems Design & Implementation (NSDI) (2008)
27. Cisco Intrusion Detection Systems: <http://www.google.com/products?q=cisco+intrusion+detection+system&aq=3&oq=cisco+in>
28. VeriSign Inc: <http://www.verisign.com/ssl/buy-ssl-certificates/secure-site-services/index.html>
29. Vratonjic, N., Freudiger, J., Felegyhazi, M., Hubaux, J.P.: Securing Online Advertising. Technical report 2008-017, EPFL (2008)
30. Vratonjic, N., Freudiger, J., Hubaux, J.P.: Integrity of the Web Content: The Case of Online Advertising. In: Usenix CollSec (2010)
31. Vratonjic, N., Raya, M., Hubaux, J.P., Parkes, D.C.: Security Games in Online Advertising: Can Ads Help Secure the Web? In: Workshop on the Economics of Information Security (WEIS) (2010)
32. Weisstein, E.: Euler-maclaurin integration formulas. MathWorld (2010), <http://mathworld.wolfram.com/Euler-MaclaurinIntegrationFormulas.html>
33. Zhao, X., Fang, F., Whinston, A.B.: An economic mechanism for better Internet security. Decision Support Systems 45(4) (2008)