This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2011 proceedings.

Examining Social Dynamics for Countering Botnet Attacks

Ziming Zhao, Gail-Joon Ahn, and Hongxin Hu
Laboratory of Security Engineering for Future Computing (SEFCOM)
Arizona State University, Tempe, AZ 85281, USA
{zmzhao, gahn, hxhu}@asu.edu

***Abstract*—Even though promising results have been obtained from existing research on bots and associated command and control channels, there is little research in exploring the ways on how bots are created and distributed by adversaries. Consequently, innovative methods that help determine the linkage between the rogue programs and adversaries are imperative for mitigating and combating botnet attacks. Recent study discovers that rogue programs are sold in black markets in online social networks and adversaries use online social networks to coordinate attacks. Correlation of botnet attacks and activities in online underground social networks is crucial to tactically cope with net-centric threats. In this paper, we take the first step toward adversarial behavior identification by modeling social dynamics of underground adversarial communities and tracing the origin of certain malwares and attack events in underground communities. We also describe our evaluation to demonstrate the effectiveness of our approach.**

I. Introduction

The risk of malware infection becomes greater than ever and it was estimated that a quarter of the networked computers were compromised by one or more malicious programs [12]. With sophisticated social engineering and signature-evading technologies, adversaries are capable of circumventing anti-malware systems, and can then eventually contaminate production computers. Malware-infected computers are deliberately facilitated as large scale destructive botnets to steal information and disrupt, deny access to, degrade or destroy critical net-centric information systems [11], [13].

Given the significance of this problem, huge research efforts have been invested in capturing, understanding and analyzing the malwares and their command and control (C&C) communications in wild [8], [9], [11]. Promising results have been obtained from collection and analysis of malwares and their communications. Also, preventive solutions against thousands of known malwares have been deployed on networked systems. However, the majority of adversaries who engineered malicious tools and coordinated attacks are still at large. As adversaries keep threatening the Internet by developing more sophisticated penetration tools and launching more net-centric attacks, it is critical to identify the linkage between the rogue programs and adversaries for mitigating and combating botnet attacks.

Recent research efforts [3] indicate that adversaries use online social networks (OSNs) to share news, release malwares or penetration tools, and coordinate attacks. However, the organizational structures and relationships in the online underground social networks (OUSNs) are not yet well-studied and understood. Therefore, investigations of the relationships between OUSNs and botnet attacks are imperative to tactically cope with net-centric threats.

In this paper, we discuss why bots and C&C analysis are not sufficient enough for countering botnet attacks and propose a systematic analysis approach to identify adversarial behaviors by examining social dynamics of underground adversarial communities, tracing the origin of certain malwares and attack events. To the best of our knowledge, this is the first attempt to study online underground social networks for analyzing and understanding botnet attacks.

The rest of this paper is organized as follows. Section II overviews the background and motivation of our work. Section III describes our high-level view on social dynamics for identifying adversarial behaviors. We formulate characteristics of online underground social networks and discuss our systematic ranking analysis approach followed by the evaluation in Section IV. Section V concludes the paper.

II. Background and Motivation

In this section, we analyze the workflow of cybercrime and the destructive nature of botnets. We summarize the research efforts in combatting botnet attacks and address the reasons why bots and associated C&C analysis are not enough for countering this emerging threat.

A. Cybercrime Workflow and Destructive Nature of Botnets

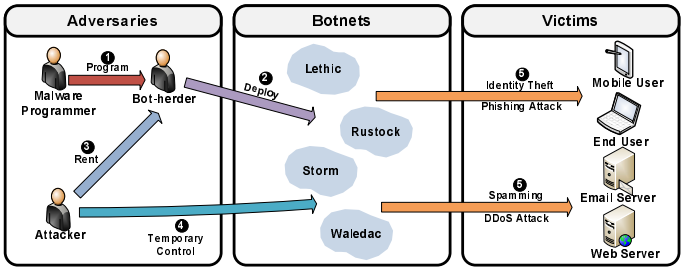


Fig. 1: Cybercrime Workflow

Figure 1 shows a typical workflow of cybercrime. In *Step 1*, malware programmers develop crafted attack tools. The most prevalent and destructive tool developed to carry out various attacks is a set of bots. Malware programmers turn bots to bot-herders through online black markets or offline channels. In

978-1-4244-9268-8/11/$26.00 ©2011 IEEE

Step 2, bot-herders deploy a botnet through social engineering, drive-by-download or other possible vectors. In *Step 3*, bot-herders rent a botnet out to other adversaries, from which bot-herders and malware programmers profit, that have targets in mind but do not have the technological expertise to design or administer the botnet. In *Step 4*, attackers, such as spammers, take control of the botnet. A rented botnet may result in a variety of attacks launched by multiple adversaries who might have different intents. In *Step 5*, attackers coordinate bot nodes to perform multiple attacks such as spamming, identity theft, DDoS, phishing attacks, etc.

The power of botnets relies on their coordination and the volume of the responses from the bot nodes. In a typical botnet, hundreds to thousands of bot nodes respond to botmaster's commands. When these nodes are instructed to connect to one webpage at the same time, the aggregated volume of the network traffic would be tremendous for most companies to handle, causing denial-of-service to the targeted servers. When these nodes are instructed to download banking credentials, the botmaster receives credentials from each bot which can be thousands or even millions in some botnets. Another critical problem caused by botnets is e-mail spamming. Nowadays a spam causes not only a network-clogging problem, but also a means for adversaries to distribute additional malwares.

B. Analyzing Bots and C&C Channels

Most research in the area of botnets focuses on finding bots and associated C&C channels, and shutting the botnets down in a timely manner [2]. In other words, existing approaches concentrate on *Step 2*, *Step 4* and *Step 5* in Figure 1. Although promising results from collection and analysis of bots and their communications have been utilized to build preventive solutions on networked systems against thousands of known malwares, these analysis processes still have limitations and fail to identify adversaries and their behaviors. We summarize several characteristics behind this phenomenon:

- **Complexity in binary analysis**: Evidence acquisition from malicious files is difficult. Static analysis methods are not suitable for malwares that are devised by deliberate obfuscation techniques and dynamic analysis may not cover all the control paths. Moreover, there exist dynamic defense techniques to detect whether the running environment of a malware is suspicious [4]. Therefore, even though there exist concise evidences in malicious files, it is hard for analysts to extract them successfully.
- **Lack of clear evidence**: There exist evidences that do not explicitly disclose the identity of adversaries in malicious files and behaviors. Even if analysts can extract every bit of information from malicious files and behaviors, the information that would point out to adversaries behind the scene may be limited and obscure.
- **Obsolete evidence**: Adversaries design sophisticated algorithms and approaches to change their identities and credentials that have to be stored in malicious files. Even if investigators successfully extract some information from malicious files, the information may be obsolete.
- **Evolution in botnets**: Most significant botnets today constantly change and evolve. They are evolved by adding bots, deleting bots, changing to new channels, being upgraded, etc. Attempting to discover their C&C servers may bring immediate benefits but stimulate the evolution of botnets. Park and Reeves [7] claimed that it is also important to monitor botnets for an extended time to learn the purpose of the botnets and to develop more effective countermeasures.

III. Identifying Adversarial Behaviors through Social Dynamics

While there exist some knowledge on how bots operate, there is little research in exploring the ways on how bots are created and distributed by adversaries. A recent study reveals that these programs are sold in online black markets on social networks [3]. Individuals who control existing botnets also sell access to their infected machines for a variety of attacks including spamming and denial of service. As a consequence, these markets enable a great deal of unskilled adversaries and innocent computer users to engage in cybercrime.

We propose a solution to systematically examine the creation, distribution and trend of bots that are being circulated online. In other words, we attempt to fully grasp bot-herders' communities and social activities. This vital information will be used to determine specific social communities related to adversarial threats. In order to address this issue we focus on discovering characteristics of the botmasters over time which allows us to discover not only the means to shut the botnets down, but also information to identify the attackers to prevent the creation of potential botnets. Our approach on online underground societies complements existing research efforts in terms of *Step 1* and *Step 3* in Figure 1.

Our approach consists of two stages. First, we model an online underground social network considering its social dynamics and user-generated contents. Second, we develop a systematic ranking analysis mechanism by introducing several indices indicating the influence of adversaries, relevance of adversaries to certain events, and the ongoing trend of underground society.

A. Modeling Online Underground Social Networks

Adversaries choose online social networks which meet their special requirements to form online underground social communities. OUSNs are used to share technical articles and trade malicious tools, rather than photo-sharing or video-sharing, making them different from normal OSNs in the following aspects:

- OUSNs provide a blog-like article-sharing mechanism, which has less constraints on the length of articles a user can post. Length limitation of posts adopted in traditional OSNs, such as Twitter and Facebook, is unlikely suitable for well-explained technical articles in OUSNs.
- OUSNs have less access and write constraints on posted articles. For instance, Facebook adopts strict policies to protect its users' privacy, in which one user has to be in the others' trust circles to access and comment on their posts. However, in

OUSNs, a user does not need to be a friend of the article author to read the article or give comments on it. This characteristic allows OUSNs to disseminate more knowledge and technical discussions than OSNs.

• OUSNs do not require users to provide their real world identities. Adversaries prefer not to associate their real world identities with their online profiles, therefore OUSNs do not claim themselves as *real* social networks. However, OSNs such as Facebook requires users to provide their real names, education backgrounds, and relationship statuses.

Based on the above-identified characteristics of OUSNs, we now formally model and define online underground social networks. Different from previous work on modeling OSNs [5] which mainly focus on users, groups and relations, our model considers user-generated contents shared in OUSNs as well:

Definition 1: (**Online Underground Social Network**). An online underground social network is modeled as an 11-tuple $OUSN =< U, G, A, C, P, S, \Re_{UU}, \Re_{UG}, \Re_{UP}, \Re_{AC}, \Re_{PS} >$, where

- U is a set of registered users who have the rights to post articles and join groups in an OUSN. Users are identified by system-generated unique identifiers and user-chosen nicknames. Users can make their profiles (birthday, residence, interests) open to the public but have no obligation to verify their authenticity;
- G is a set of groups to which users can belong. A group could be formed based on common interests or any other similarity among its users. Groups are identified by system-generated unique identifiers and creator-chosen nicknames;
- A is a set of articles which are posted by users who want to share them with the society. In OUSNs, articles might introduce latest technologies, analyze recent vulnerabilities, call for participation of network attacks and trade newly developed and deployed botnets;
- C is a set of comments which are the subsequent posts to articles. Comments represent the reactions from the society to posted articles;
- P is a set of posts. Posts are the union of articles and comments where $P = A \cup C$;
- S is a set of strings which are the elementary components of articles and comments. Strings are not necessarily meaningful English words. They could be names, URLs and underground language such as, *c4n* as *can*, and *sUm1* as *someone*;
- $\Re_{UU} : U \times U \rightarrow RT_{UU}$ is a function to assign relationships among users in the OUSN, where RT_{UU} is a set of user-user relationship types supported by the OUSN. The most common relationship among users is $followerOf$;
- $\Re_{UG} : U \times G \rightarrow RT_{UG}$ is a function to assign relationships between users and groups in the OUSN, where RT_{UG} is a set of user-group relationship types supported by the OUSN. The most common relationships between users and groups include $memberOf, subscriberOf$;
- $\Re_{UP} : U \times P \rightarrow RT_{UP}$ is a function to assign relationships between users and posts in the OUSN, where RT_{UP} is a set of user-post relationship types supported by the OUSN. The most common relationship between users and posts is $authorOf$;
- $\Re_{AC} : A \times C \rightarrow RT_{AC}$ is a function to assign relationships between articles and comments in the OUSN. RT_{AC} is a set of article-comment relationship types supported by the OUSN. The most common relationship between articles and comments is $hostOf$; and
- $\Re_{PS} : P \times S \rightarrow RT_{PS}$ is a function to assign relationships between posts and strings in the OUSN. RT_{PS} is a set of post-string relationship types supported by the OUSN. The most common relationship between posts and strings is $containerOf$;

B. Systematic Ranking Analysis

In this section, we present a systematic ranking analysis approach to facilitate the understanding of OUSNs. In order to achieve this, we first provide several functions based on our OUSN model. We introduce N as the set of natural numbers to support counting and T as the set of time to support temporal pattern analysis. We include several core functions in Table I.

TABLE I: Core Functions

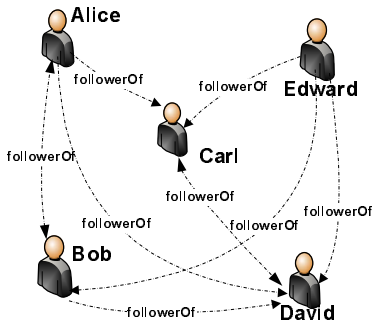
Function	Category
$atlpost : U \times T \times T \rightarrow N$	$atlpost(u, t_1, t_2) = n$, if u posted n articles from time t_1 to t_2
$cmtpost : U \times T \times T \rightarrow N$	$cmtpost(u, t_1, t_2) = n$, if u posted n comments from time t_1 to t_2
$following : U \times T \times T \rightarrow N$	$following(u, t_1, t_2) = n$, if u started to follow n users from time t_1 to t_2
$follower : U \times T \times T \rightarrow N$	$follower(u, t_1, t_2) = n$, if u was followed by n users from time t_1 to t_2
$affiliated : U \rightarrow N$	$affiliated(u) = n$, if u is the member of n groups
$subscribing : U \rightarrow N$	$subscribing(u) = n$, if u subscribes n groups' news

To understand OUSNs, we also introduce several indices to model the activeness and influence of users and groups, prevalence of topics and ongoing trends. Due to the page limit, we only discuss four major indices here.

Index 1: User Influence Index (UII) represents the influence of a user in a given time period.

$$UII(u, t_1, t_2) = w_a \sum_{i=1}^{n} AII(a_i, t_1, t_2) + w_f follower(u, t_1, t_2) \quad (1)$$

where u is the user identifier, t_1 and t_2 denote the start and end of given time period, respectively, and w_a and w_f are the weights that can be used to adjust the contribution of posts and social relations. One user's influence can be split into two parts. One is the impact of the user's opinions or remarks, which is modeled by article influence index (AII). Another related part is the user's social relationships. For a more influential user, more people will follow her/him. We consider an additive model to combine these two parts together. In order to reduce the bias caused by different amount of posted



Alice					
Stage→	1	2	3	4	5
Articles	4	3	0	1	2
Comments Received	19	13	0	4	9
Comments Posted	5	5	5	10	3
Follows	3				
Followers	1				

Bob					
Stage→	1	2	3	4	5
Articles	2	4	4	2	1
Comments Received	6	13	23	8	2
Comments Posted	13	6	12	20	9
Follows	2				
Followers	2				

Carl					
Stage→	1	2	3	4	5
Articles	2	5	3	5	3
Comments Received	5	15	9	19	9
Comments Posted	7	8	13	18	7
Follows	1				
Followers	3				

David					
Stage→	1	2	3	4	5
Articles	2	1	3	4	2
Comments Received	4	6	18	12	11
Comments Posted	11	10	15	2	5
Follows	1				
Followers	4				

Edward					
Stage→	1	2	3	4	5
Articles	6	1	4	5	4
Comments Received	2	1	19	13	5
Comments Posted	20	19	5	6	12
Follows	3				
Followers	0				

Fig. 2: An OUSN Scenario

Rank	Name	Index
1	Alice	69.326
2	Bob	23.988
3	Carl	21.543
4	David	13.678
5	Edward	5.105

(a) UII according to our solution

Rank	Name	Index
1	Edward	29
2	Bob	17
3	David	14
4	Alice	12
5	Carl	10

(b) UAI according to our solution

Rank	Name	Index
1	Carl	1.111
2	David	0.371
3	Alice	0.075
4	Bob	0.024
5	Edward	0.000

(c) PageRank Index

Fig. 3: Indices according to our solution and PageRank

articles, we only consider the top n articles from a user in terms of AII.

Index 2: User Relevance Index (URI) represents the relevance of a user to some evidences in a given time period. In reality, these evidences could be in the format of text, picture, video, audio or any other forms. Yet representing multimedia content like picture and video in a machine-understandable way is still difficult. Our model, acting like a modern web search engine, takes keyword-based queries.

$$\begin{aligned} &URI(u, s_1, ..., s_n, t_1, t_2) \\ &= w_r \textstyle\sum_{i=1}^{n} ARI(a_i, s_1, ..., s_n, t_1, t_2) \\ &+ w_c \textstyle\sum_{i=1}^{n} CRI(c_i, s_1, ..., s_n, t_1, t_2) \end{aligned} \tag{2}$$

where $s_1, ..., s_n$ are n given keywords, and w_r and w_c are the weights for article and comment respectively. The relevance of a user to evidences is measured by combining her/his article relevance index (ARI) and comment relevance index (CRI). We only consider the top n articles and comments to reduce the bias caused by different amount of posts.

Index 3: User Activeness Index (UAI) represents the activeness of a user in a given time period. We call those users who energetically participate in OUSN *active* users. These activities include posting articles and comments, following other users or groups, and joining groups.

$$\begin{aligned} &UAI(u, t_1, t_2) = w_p(altpost(u, t_1, t_2) + cmtpost(u, t_1, t_2)) \\ &+ w_w(following(u, t_1, t_2) + affiliated(u, t_1, t_2) \\ &+ subscribing(u, t_1, t_2)) \end{aligned} \tag{3}$$

where w_p and w_w are the weights for post activity and social activity respectively.

Index 4: Prevalence Index (PI) represents the popularity of a word in a given period of time.

$$PI(s_i, t_1, t_2) = \sum_{p_j \in P} tfidf_{s_i, p_j} \tag{4}$$

We use *term frequency-inverse document frequency* (TF-IDF) as the basis to estimate the prevalence of words. TF-IDF evaluates how important a word is to a document in a collection, which is widely used in text mining and information retrieval [10]. TF-IDF model defines 1) term frequency $tf_{w,d}$, which is the number of occurrence of a word w in a document d; 2) document frequency df_w, which is the number of documents in the collection that contains the word w; and 3) inverse document frequency $idf_w = \log \frac{N}{df_w}$, where N is the total number of documents. TF-IDF of a word w_i to a document d_j is denoted as $tfidf_{w_i,d_j} = tf_{w_i,d_j} \times idf_{w_i}$. For each string s_i, we calculate its $Prevalence\ Index$. Then we sort $PI(s_i, t_1, t_2)$ in descending order. By calculating most prevalent words in different time, we can derive a topic trend of online underground social communities.

IV. Evaluation

We compared our approach with a PageRank-based solution [6] to evaluate the effectiveness of our mechanism. PageRank uses numerical weights of elements that are linked together to measure their relative importance. Although PageRank is successfully deployed in commercial search engines, as discussed in [1], it is not very suitable to rank sparsely linked elements.

For simplicity, we only included a case study considering user influence and activeness, as shown in Figure 2. In

this scenario with five users, the left-hand figure shows the relationships between users and the right-hand tables show information from user-generated contents in five different stages. *David* and *Carl* are two of most popular users with four and three followers, respectively. While *Edward* seems the most eloquent user in the stage one with six articles and twenty posted comments. Figure 3(a) shows the user influence index sorted in descending order in the stage one. *Alice* is ranked as the most influential user, mainly because her four articles received nineteen comments from the society. Note that, although *David* has four followers in the society, his influence is limited due to the fact that he initiated few attractive conversations. Although, *Edward* has no followers, our model considers he still has some influence in the community because his contributions have attracted other's attention. Figure 3(b) shows the user activeness index in the stage one. The least influential user *Edward* is ranked as the most active user not only because he is eloquent, but also because he follows everyone in the community. *Bob* is ranked as the second most active member since he contributed to many conversations. We notice that the most influential user *Alice* is ranked as the second least active user which verifies that a user does not need to speak much to make a difference. Figure 3(c) shows the PageRank-based ranking analysis results. Note that PageRank ignores user-generated contents, and only considers user relationships for ranking analysis. One reason for *Carl* being more influential than *David*, even if in the case that *David* has more followers than *Carl*, is because in PageRank analysis the value of link-votes is divided among all outbounds. The fact that *David* only follows *Carl* indicates *Carl*'s influence in this model. PageRank fails to identify *Edward*'s influence as well merely because he has no follower. Another drawback of PageRank analysis is that it cannot generate temporal patterns of the influential users since relationships do not change.

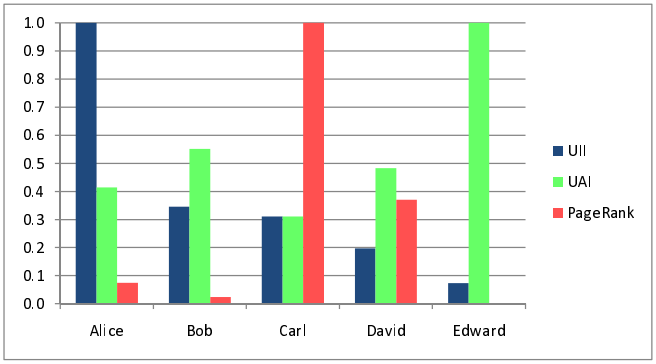


Fig. 4: Comparison of Normalized Indices for Each User

We normalize three different indices calculated for each adversary and show their comparison in Figure 4. The significant difference between our solution and PageRank-based solution shows the importance of considering user-generated contents as well as social relationships.

V. Conclusion and Future Work

In this paper, we have presented a novel approach to identify adversarial behaviors through social dynamics, complementing existing bots and C&C analysis for understanding botnet attacks. We evaluated the effectiveness of our approach by showing the influential and active users in underground society and comparing with a PageRank-based solution, which could not take user-generated contents into account. For analyzing and predicting adversarial behaviors in the real world, we are currently developing a tool to experiment our solution on an online underground community dataset from Livejournal.com.

Acknowledgments

This work was partially supported by the grants from National Science Foundation (NSF-IIS-0900970 and NSF-CNS-0831360) and Department of Energy (DE-SC0004308).

References

[1] N. Agarwal, H. Liu, L. Tang, and P. Yu. Identifying the influential bloggers in a community. In *Proceedings of the 1st International Conference on Web Search and Web Data Mining (WSDM)*, pages 207–218. ACM, 2008.

[2] G. Gu, V. Yegneswaran, P. Porras, J. Stoll, and W. Lee. Active Botnet Probing to Identify Obscure Command and Control Channels. In *Proceedings of the 26th Annual Computer Security Applications Conference (ACSAC)*, pages 241–253. IEEE, 2010.

[3] G. W. B. Holt, Thomas J. and A. M. Bossler. Social Learning and Cyber Deviance: Examining the Importance of a Full Social Learning Model in the Virtual World. *Journal of Crime and Justice*, page 33, 2010.

[4] T. Holz and F. Raynal. Detecting honeypots and other suspicious environments. In *Proceedings of the 6th Annual IEEE SMC, Information Assurance Workshop (IAW)*, pages 29–36. IEEE, 2005.

[5] A. Mislove, M. Marcon, K. Gummadi, P. Druschel, and B. Bhattacharjee. Measurement and analysis of online social networks. In *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement (IMC)*, pages 29–42. ACM, 2007.

[6] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999. Previous number = SIDL-WP-1999-0120.

[7] Y. Park and D. Reeves. Identification of Bot Commands by Run-Time Execution Monitoring. In *Proceedings of the 25th Annual Computer Security Applications Conference (ACSAC)*, pages 321–330. IEEE, 2009.

[8] P. Porras, H. Saidi, and V. Yegneswaran. A foray into Conficker logic and rendezvous points. In *Proceedings of the 2nd USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)*, 2009.

[9] M. Rajab, J. Zarfoss, F. Monrose, and A. Terzis. My botnet is bigger than yours (maybe, better than yours): why size estimates remain challenging. In *Proceedings of the 1st Usenix Workshop on Hot Topics in Understanding Botnets (HotBots)*, pages 5–5. USENIX Association, 2007.

[10] G. Salton and C. Buckley. Term-weighting approaches in automatic text retrieval. *Information processing & management*, 24(5):513–523, 1988.

[11] B. Stone-Gross, M. Cova, L. Cavallaro, B. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna. Your botnet is my botnet: Analysis of a botnet takeover. In *Proceedings of the 16th ACM conference on Computer and Communications Security (CCS)*, pages 635–647. ACM, 2009.

[12] W. Sturgeon. Internet guru warns of botnet pandemic, http://www.zdnet.co.uk/news/networking/2007/01/29/internet-guru-warns-of-botnet-pandemic-39285665/.

[13] K. Thomas. The Koobface botnet and the rise of social malware. In *Proceedings of the 5th IEEE International Conference on Malicious and Unwanted Software (MALWARE)*, pages 1–8, 2010.