Detecting Social-Network Bots Based on Multiscale Behavioral Analysis

Francisco Brito, Ivo Petiz
Paulo Salvador, António Nogueira
Instituto de Telecomunicações, Aveiro, Portugal
DETI, University of Aveiro, Portugal
e-mails: franciscobrito@ua.pt, petiz@live.ua.pt,
salvador@ua.pt, nogueira@ua.pt

Eduardo Rocha
Instituto de Telecomunicações, Aveiro, Portugal
Faculty of Computer Science, HTWK Leipzig,
Germany
e-mail: eduardo.rocha@imn.htwk-leipzig.de

***Abstract*—Social network services have become one of the dominant human communication and interaction paradigms. However, the emergence of highly stealth attacks perpetrated by bots in social-networks lead to an increasing need for efficient detection methodologies. The bots objectives can be as varied as those of traditional human criminality by acting as agents of multiple scams. Bots may operate as independent entities that create fake (extremely convincing) profiles or hijack the profile of a real person using his infected computer. Detecting social networks bots may be extremely difficult by using human common sense or automated algorithms that evaluate social relations. However, bots are not able to fake the characteristic human behavior interactions over time. The pseudo-periodicity mixed with random and sometimes chaotic actions characteristic of human behavior is still very difficult to emulate/simulate. Nevertheless, this human uniqueness is very easy to differentiate from other behavioral patterns. As so, novel behavior analysis and identification methodologies are necessary for an accurate detection of social network bots. In this work, we propose a new paradigm that, by jointly analyzing the multiple scales of users' interactions within a social network, can accurately discriminate the characteristic behaviors of humans and bots within a social network. Consequently, different behavior patterns can be built for the different social network bot classes and typical humans interactions, enabling the accurate detection of one of most recent stealth Internet threats.**

Keywords - Facebook user behavior, Human social-networking behavior, social-network bots, bot detection, Facebook interactions model.

I. INTRODUCTION

Together with the growing predominance of social networking services in human communication, a set of scam attacks perpetrated by bots [1], [2] have emerged. We are currently witnessing a major increasing in cybercrime attacks against individuals targeting their personal data and financial assets [3], [4]. Most of the current Internet malware threats disseminate themselves using social engineering and, mainly, using social networks [5], since social networking provides an open field for illicit activities [6]. Social networking sites are always improving their security but this is a constant race behind the leading criminals [7]. Existing botnets [8], [9] can use social networking services to spread themselves, but more importantly, can use social networks to impersonate the owner of the controlled machine in order to obtain valuable personal information or force the person to interact with unwanted individuals or services. One of the better documented examples of social networking services abuse for malicious purposes was Koobface [10], [11]. Being currently the largest social networking service, Facebook is the main vector of attack via social networking services [12]. Current techniques to detect bots within a social network rely on automated algorithms that evaluate social relations. Based on graph-theory techniques, they try to detect unnatural relations in social networks [13], [9]. Another technique used to detect bot activity measures mouse movements and keystrokes produced while interacting in the generation of online contents. In [14] this class of behavioral analysis was applied in blogging activities, but it can also be easily applied to social networks interfaces. The main downside of this approach is that it must rely on software loaded on the client browser, which can be difficult to implement and certainly impossible to generalize to all users due to confidentiality constrains. A viable solution should only rely on ubiquitous statistics that do not compromise the users privacy, namely counting the number of social interactions per time interval (e.g., number of posts, number of likes, number of photo uploads).

It is extremely difficult to program a bot to replicate the characteristic human behavior of social interactions over time. Humans actions have an inherent pseudo-periodicity mixed with random (and sometimes chaotic) actions which are almost impossible to emulate/simulate. Nevertheless, this human uniqueness is very easy to differentiate from other behavioral patterns. Therefore, in this paper, we propose a new methodology that, by jointly analyzing the multiple scales of the users' interactions within a social network, can accurately discriminate the characteristic behaviors of humans and bots within a social network. Consequently, different behavior signatures can be use to accurately detect bots acting with a social network.

The proposed methodology applies the concept of multiscalling analysis based on scalograms to the statistical processes that describe the interaction of a user within the social network. Scalograms reveal much information about the nature of non-stationary processes that was previously hidden, so they are applied to a lot of different scientific areas: diagnosis of special events in structural behavior during earthquake excitation, ground motion analysis, transient building response to wind storms, analysis of bridge response due to vortex shedding, among others [15].

The remaining part of this paper is organized as follows: Section II presents some important background on multiscaling analysis; Section III presents the characteristic behaviors

 ISBN: 978-1-61208-298-1

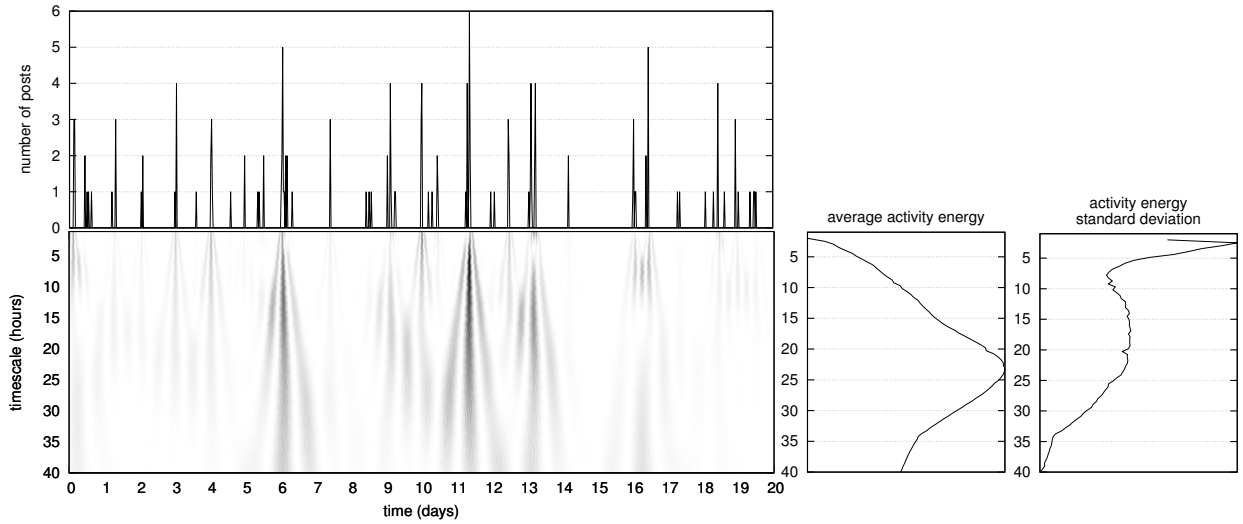

Figure 1. Multiscale analysis example: (top-left) number of Facebook posts in 20-minutes intervals, (bottom-left) scalogram / user activity energy per timescale over time, (bottom-middle) average user activity energy over time and (bottom-right) activity energy standard deviation over time.

of human and bots within a social network and how they differ; Section IV presents the results of a proof-of-concept of the proposed detection methodologies and, finally, Section V presents some brief conclusions about the presented detection methodologies.

II. Multiscaling analysis

The main purposes of a multiscaling analysis is to identify the most important time-scales of (pseudo-periodicity) activity and quantify the constancy of that pseudo-periodicity. In order to achieve that objective, it is necessary to quantify the activity over time for multiple timescales.

Wavelets are mathematical functions that are used to divide a given signal into its different timescales components. Wavelets enable the analysis of each one of the signal components in an appropriate scale. Starting with a mother wavelet $\psi(t)$, a family $\psi_{\tau,s}(t)$ of "wavelet daughters" can be obtained by simply scaling and translating $\psi(t)$:

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{|s|}}\psi(\frac{t-\tau}{s}) \tag{1}$$

where s is a scaling or dilation factor that controls the width of the wavelet (factor $\frac{1}{\sqrt{|s|}}$ is introduced to guarantee the energy preservation, $\|\psi_{\tau,s}\| = |\psi|$) and τ is a translation parameter controlling the time location of the wavelet. Scaling a wavelet simply means stretching it (if $|s| > 1$) or compressing it (if $|s| < 1$), while translating it simply means shifting its position in time.

Given a signal $x(t)$, its Continuous Wavelet Transform (CWT) with respect to the wavelet ψ is a function of time (τ) and scale (s), $W_{x;\psi}(\tau, s)$, obtained by projecting $x(t)$ onto the wavelet family $\{\psi_{\tau,s}\}$:

$$W_{x;\psi}(\tau, s) = \int_{+\infty}^{-\infty} x(t)\frac{1}{\sqrt{|s|}}\psi(\frac{t-\tau}{s})dt \tag{2}$$

By analogy with the terminology used in the Fourier case, the energy components of the signal are given by the square of the CWT components of the signal and the (local) Wavelet Power Spectrum (sometimes called Scalogram or Wavelet Periodogram) is defined as the normalized energy over time and scales:

$$E_x(\tau, s) = 100\frac{|W_{x;\psi}(\tau, s)|^2}{\sum_{\tau'}\sum_{s'}|W_{x;\psi}(\tau', s')|^2} \tag{3}$$

An example of a scalogram can be observed in Figure 1 (bottom-left).

III. Social-Networking Behavior Characterization

A. Single User Behavior Inference

Within the context of this paper, signal $x(t)$ is a counting process that quantifies the number of social network interactions (i.e., number of posts, number of likes and number of posted photos) in a time-interval (e.g., 30 minutes, as used in section IV) over time. An example of a counting process (Facebook posts) can be observed in Figure 1 (top-left). In order to characterize the user multiscale behaviour over time, it is possible to estimate (i) the Average Activity Energy of signal $x(t)$, $\mu E_x(s)$, by averaging the normalized energy of the signal over time for all timescales (see (4)) and (ii) the Activity Energy Standard Deviation of signal $x(t)$, $\sigma E_x(s)$, by calculating the standard deviation of the normalized energy of the signal over time for all timescales (see (5)):

$$\mu E_x(s) = \frac{1}{N}\sum_{i=1}^{N} E_x(\tau_i, s), \forall s \tag{4}$$

$$\sigma E_x(s) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(E_x(\tau_i, s) - \mu E_x(s)\right)^2}, \forall s \tag{5}$$

An example of these metrics can be observed in Figure 1, in the bottom middle and right plots, respectively.

B. User Group Behavior Inference

In order to characterize the behavior of a group of users, it is necessary to define metrics that can quantify their behavior as a group. Assuming a group size of U users and assuming that $x_u(t)$ represents a counting process that describes the activity of user u in the social network, we can quantify the mean and variance values of the (i) group average activity energy (see (6) and (7)) and (ii) group activity energy standard deviation (see (8) and (9)), for all timescales and users within the group:

$$\overline{\mu E(s)} = \frac{1}{U}\sum_{u=1}^{U} \mu E_{x_u}(s), \forall s \quad (6)$$

$$VAR(\mu E(s)) = \frac{1}{U-1}\sum_{u=1}^{U} \left(\mu E_{x_u}(s) - \overline{\mu E(s)}\right)^2, \forall s \quad (7)$$

$$\overline{\sigma E(s)} = \frac{1}{U}\sum_{u=1}^{U} \sigma E_{x_u}(s), \forall s \quad (8)$$

$$VAR(\sigma E(s)) = \frac{1}{U-1}\sum_{u=1}^{U} \left(\sigma E_{x_u}(s) - \overline{\sigma E(s)}\right)^2, \forall s \quad (9)$$

IV. Results

The proposed methodology was applied to two different data-sets. The first data-set comprises Facebook posted between January 1st, 2007 and December 31, 2008 by a group of 160 users, which individually posted more than 300 posts in this two year time frame. The total number of recorded Facebook posts is 72893. This 2007-2008 data-set was extracted from the data presented by Viswanath *et al.* [16] and is freely available online. The second data-set was extracted from a group of Facebook friends of this paper authors using the Facebook Graph API [17]. Facebook Graph API is a low level HTTP-based API used to query data from Facebook's Social Graph. The data queried using Facebook Graph API is returned in JSON format and can be easily post-processed in order to extract its relevant statistics. This data-set contains all social networking activities of 140 users between January 1st, 2011 and December 31, 2012. Only users with more than 300 interactions with the social network in the two years time frame were considered. The total number of Facebook interactions in this data-set are 167171 Facebook posts, 90882 likes and 48755 photo uploads. The second data-set is hereafter referred as 2011-2012 data-set.

Based on the timestamps of each social network interaction, time processes were extracted from the data-sets by counting the number of interactions in 20-minute time intervals. For the 2007-2008 data-set only Facebook posts were considered due the the limited information of the original data. However, for the 2011-2012 data-set we considered separately the Facebook posts, likes and photo uploads.

A. Facebook User Behavior Evolution

We have applied the proposed multiscale user activity characterization to Facebook posts time processes extracted from both data-sets. The obtained results are depicted in Figure 2.

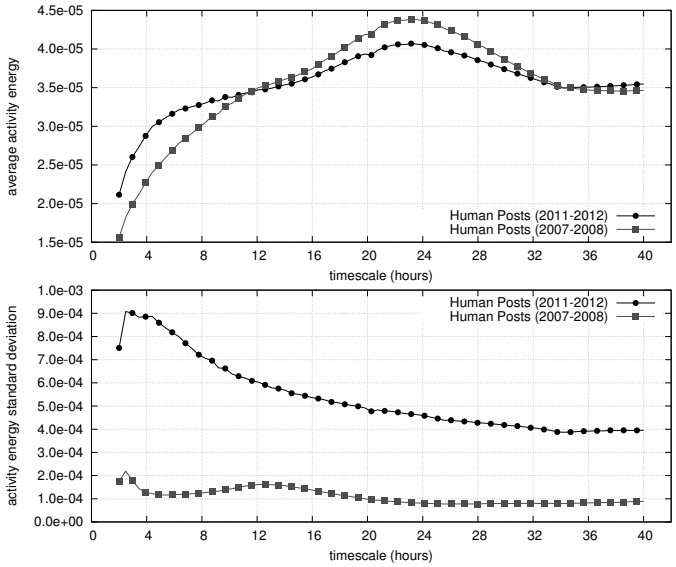

Figure 2. Evolution of multiscale characteristics of human posts on Facebook from 2007-2008 to 2011-2012: (top) average energy activity over time for all scales and (bottom) standard deviation of energy activity over time for all scales.

The average energy activity reveals that pseudo-periodicity with a 24 hours period is predominant for both data-sets. This fact reveals that humans behave and interact with social networks in 24 hour cycles, although humans also spread their pseudo-periodic intervals between activity over a wide range of timescales. Another fact that can be observed is that nowadays users tend to have a more spread usage when compared to the older data-set; this reveals that human users have evolved to a more frequent Facebook interactivity, so their profiles are not so deeply shaped according to the 24-hour cycle. This evolution on the characteristics of the human activities is more visible when analyzing the activity energy standard deviation, which is much higher on the newer data-set. This also show that human users have a very variable behavior in terms of intervals between activities on Facebook.

B. Human and Bot Differentiation

In order to present a proof-of-concept for our methodology as a social network bot detection tool, we created two different bots that emulate Facebook posts, likes and photo uploads according to different behavioral profiles, namely, one periodic bot and one exponential bot. A periodic bot interacts with Facebook in exact intervals (in this paper, we have considered a periodicity of 24 hours) and an exponential bot interacts with Facebook in exponential distributed intervals (in this paper, we considered an average interval of 24 hours).

We have applied the proposed multiscale user activity characterization to human users of 2011-2012 data-set and to the emulated bots. Figures 3, 4 and 5 depict the multiscale characteristics (with 98% confidence intervals) of human and bot Facebook users when making posts, likes and uploading photos, respectively. The results show that the multiscale behaviors are similar when observing the Facebook posts, likes and photo uploads. From the obtained results it is also possible

 ISBN: 978-1-61208-298-1

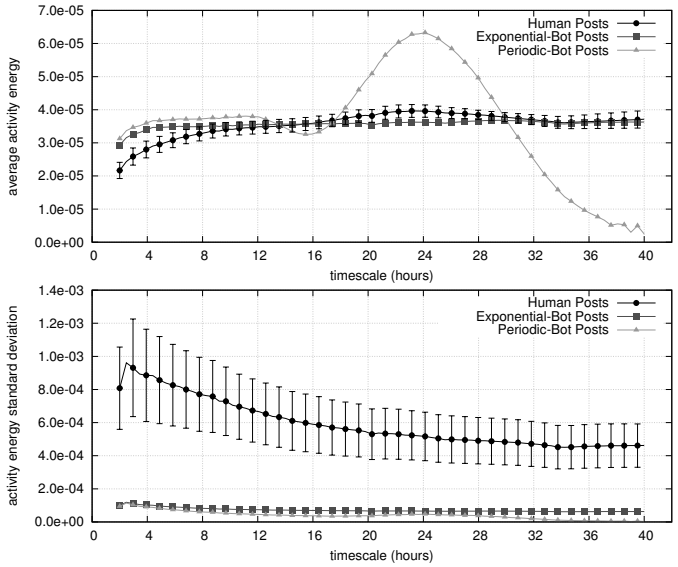

Figure 3. Multiscale characteristics of human and bot posts on Facebook in 2011-2012 dataset: (top) average energy activity over time for all scales and (bottom) standard deviation of energy activity over time for all scales, with 98% confidence intervals.

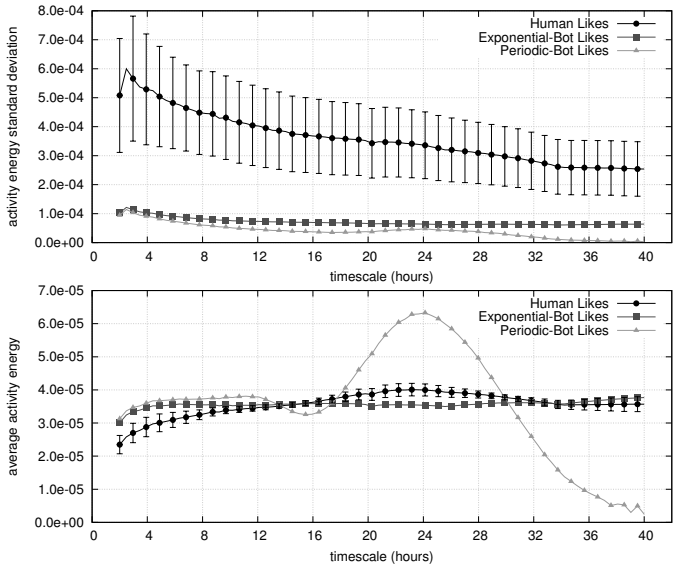

Figure 4. Multiscale characteristics of human and bot likes on Facebook in 2011-2012 dataset: (top) average energy activity over time for all scales and (bottom) standard deviation of energy activity over time for all scales, with 98% confidence intervals.

to observe that bots and humans have distinct multiscale behaviors. Periodic bots have their average activity energy centered on a 24-hour timescale and have low energy variation. Exponential bots and human users have a similar average activity energy distribution over the timescales, however, human users still have slighter higher energy around the 24-hour timescale. However, the variation of activity energy over time is much higher in human users.

The multiscale characteristics of human activities in social networks have a pseudo-periodicity of 24-hour, however, the human analysis reveals an inherent chaotic and unpredictable behavior shown by the much higher variation of activity energy

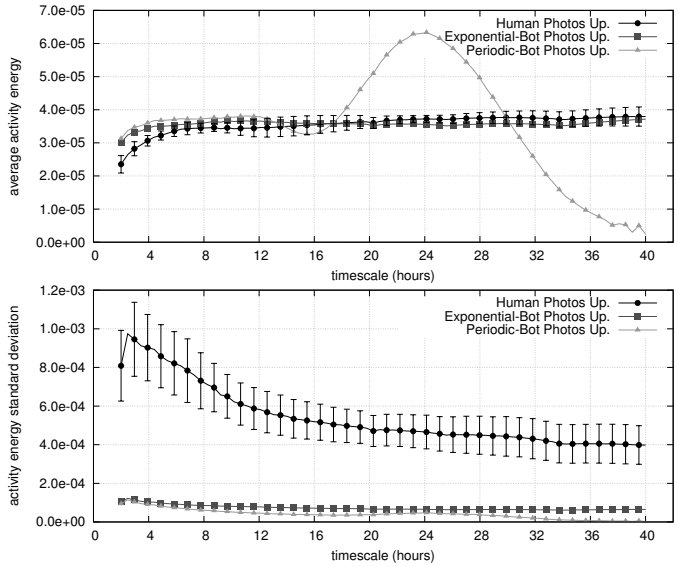

Figure 5. Multiscale characteristics of human and bot photo uploads on Facebook in 2011-2012 dataset: (top) average energy activity over time for all scales and (bottom) standard deviation of energy activity over time for all scales, with 98% confidence intervals.

over time. This inherent chaos of the behavior is extremely difficult to emulate by bots and can be used to differentiate human from bot users within a social network.

V. Conclusion

In this paper, a novel paradigm was proposed to perform the joint analysis of multiple scales of users' interactions within a social network. The presented methodology allows an accurate discrimination between human and bot behaviors within social networks. The results obtained reveal that multiscale behavior signatures can be built for different social network bot classes and typical human interactions, which will enable the development of accurate tools for the detection of social networks bots.

References

[1] S. Sengupta, "Bots Raise Their Heads Again on Facebook," The New York Times - Bits Blog, Jul. 2012.

[2] E. Gamma, "Your Facebook Friends May Be Evil Bots," InfoWorld, Apr. 2013.

[3] E. Kraemer-Mbula, P. Tang, and H. Rush, "The cybercrime ecosystem: Online innovation in the shadows?" Technological Forecasting and Social Change, vol. 80, no. 3, Mar. 2013, pp. 541–555.

[4] W. Kim, O.-R. Jeong, C. Kim, and J. So, "The dark side of the internet: Attacks, costs and responses," Information Systems, Special Issue on WISE 2009 - Web Information Systems Engineering., vol. 36, no. 3, 2011, pp. 675–705.

[5] S. Abraham and I. Chengalur-Smith, "An overview of social engineering malware: Trends, tactics, and implications," Technology in Society, vol. 32, no. 3, Aug. 2010, pp. 183–196.

[6] G. R. Weir, F. Toolan, and D. Smeed, "The threats of social networking: Old wine in new bottles?" Information Security Technical Report, vol. 16, no. 2, May 2011, pp. 38–43.

[7] P. Jagnere, "Vulnerabilities in social networking sites," in 2nd IEEE International Conference on Parallel Distributed and Grid Computing (PDGC 2012), Dec. 2012, pp. 463–468.

[8] S. S. Silva, R. M. Silva, R. C. Pinto, and R. M. Salles, "Botnets: A survey," Computer Networks, Fev. 2013, pp. 378–403 .

 ISBN: 978-1-61208-298-1

[9] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, “Design and analysis of a social botnet,” Computer Networks, Fev. 2012, pp. 556–578.

[10] R. Fergusonnam, “Back to the future,” Network Security, vol. 2010, no. 1, Jan. 2010, pp. 4–7.

[11] K. Thomas and D. Nicol, “The koobface botnet and the rise of social malware,” in 5th International Conference on Malicious and Unwanted Software (MALWARE 2010), Oct. 2010, pp. 63–70.

[12] D. Bradbury, “Spreading fear on facebook,” Network Security, vol. 2012, no. 10, Oct. 2012, pp. 15–17.

[13] G. Yan, “Peri-watchdog: Hunting for hidden botnets in the periphery of online social networks,” Computer Networks, vol. 52, no. 2, Fev. 2012, pp. 540–555.

[14] Z. Chu, S. Gianvecchio, A. Koehl, H. Wang, and S. Jajodia, “Blog or block: Detecting blog bots through behavioral biometrics,” Computer Networks, vol. 57, no. 3, Fev. 2013, pp. 634–646.

[15] K. Gurley and A. Kareem, “Applications of wavelet transforms in earthquake, wind, and ocean engineering,” Engineering Structures, no. 21, 1999, pp. 149–167.

[16] B. Viswanath, A. Mislove, M. Cha, and K. P. Gummadi, “On the evolution of user interaction in facebook,” in Proceedings of the 2nd ACM SIGCOMM Workshop on Social Networks (WOSN’09), Aug. 2009, pp. 37–42.

[17] “Graph API - Facebook Developers,” http://developers.facebook.com/docs/reference/api/, 2013, [Online; accessed April-2013].