

# On the Bursty Evolution of Online Social Networks

Sabrina Gaito<sup>†</sup>, Matteo Zignani<sup>†</sup>, Gian Paolo Rossi<sup>†</sup>, Alessandra Sala<sup>§</sup>  
Xiaohan Zhao<sup>‡</sup>, Haitao Zheng<sup>‡</sup>, Ben Y. Zhao<sup>‡</sup>

<sup>†</sup>Computer Science, Università degli Studi di Milano

<sup>§</sup>Alcatel-Lucent Bell Labs

<sup>‡</sup>Computer Science Department, U. C. Santa Barbara

{sabrina.gaito,matteo.zignani,gianpaolo.rossi}@unimi.it

alessandra.sala@alcatel-lucent.com

{xiaohanzhao,htzheng,ravenben}@cs.ucsb.edu

## ABSTRACT

The high level of dynamics in today's online social networks (OSNs) creates new challenges for their infrastructures and providers. In particular, dynamics involving edge creation has direct implications on strategies for resource allocation, data partitioning and replication. Understanding network dynamics in the context of physical time is a critical first step towards a predictive approach towards infrastructure management in OSNs. Despite increasing efforts to study social network dynamics, current analyses mainly focus on change over time of static metrics computed on snapshots of social graphs. The limited prior work models network dynamics with respect to a logical clock. In this paper, we present results of analyzing a large timestamped dataset describing the initial growth and evolution of a large social network in China. We analyze and model the burstiness of link creation process, using the second derivative, *i.e.* the acceleration of the degree. This allows us to detect bursts, and to characterize the social activity of a OSN user as one of four phases: acceleration at the beginning of an activity burst, where link creation rate is increasing; deceleration when burst is ending and link creation process is slowing; cruising, when node activity is in a steady state, and complete inactivity.

## Keywords

Online Social Networks, Network Dynamics

## 1. INTRODUCTION

The rapid growth of online social networks (OSNs) has created numerous technical challenges for the providers that

supply the hardware and software infrastructure behind these web services. As one example, the creation of social links between users dramatically change demands on social network infrastructures in terms of access, storage and computation. Depending on the specific configuration of backend servers, for example, changes in the social graph can affect how data is partitioned across clusters, or how much replication is necessary to sustain low query response times.

However, very little is known about how social network dynamics correspond to actual clock time. The large majority of prior work on OSN analysis has focused on analyzing, mining, and modeling static topologies or static snapshots of dynamic processes. Only recently have researchers begun to study dynamic processes in social networks, most often by analyzing how classical graph metrics such as degree, connected components, and shortest paths change over time. This has led to models of underlying processes such as densification and shrinking diameters [15]. These models describe how graphs change and how edges are created with respect to a logical clock, *i.e.* a homogeneous sequence of events.

But how do these events match up to events in real time? Can we better understand how edge creation events relate to each other, and can the occurrence of such events be predicted with respect to a physical clock? This work is an initial effort to answer some of these questions, but analyzing one specific temporal property of *burstiness* in edge creation. Our work is motivated in part by models of human dynamics adopted in a wide range of disciplines, from economics to communications. Recent studies [23, 13] have shown that human dynamics are best described by periods of rapidly occurring events interleaved with long periods of inactivity. Thus we ask the question: *Is link creation in online social networks a bursty process?*

In this paper, we provide an initial answer to this question, by analyzing an anonymized temporal trace of edge creation events over a period of a year in a large Chinese online social network. This online social network has more than 200 million users, and our analysis of its dynamics shows that

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ACM HotSocial'12 August 2012 Beijing, China

Copyright 2012 ACM 978-1-4503-1549-4 ...\$5.00.

edge creation is a highly inhomogeneous and bursty process. We then ask two followup questions: a) *Given a high level bursty structure, does an inner substructure exist, and how can it be characterized;* and b) *How can we detect both the whole burst and its internal phases?*

Understanding the internal structure of edge creation bursts can shed light on the underlying user process, *e.g.* is the user gradually enlarging her circle of friends or has she discovered a new cluster of her offline friends. Known techniques for the analysis and the detection of burst events (gamma-ray, text mining, stock market) focus on locating a burst when it occurs, but they do not consider events inside the detected temporal window. Thus we propose a new methodology able to detect bursts, their internal structure and the transitions between the different phases a node experiences. We perform a second order analysis on the link creation process by computing, for each node, the acceleration of the degree time function to characterize the burst structure.

Finally, we apply our acceleration metric and the detection of bursty phases on our dynamic graph. We find that all nodes exhibit similar patterns over time, characterized by an intense burst of activity following their joining the network. The initial burst is followed by weaker bursts over time, each composed of an acceleration phase, followed by a longer period of slowly vanishing deceleration.

The discovery of highly bursty patterns paves the way for new generative models that not only capture graph dynamics in terms of phases of node activity, but also describes such events with respect to physical time. In addition, burst analysis can reveal further insights into the formation and liveness of individual users, communities, and provide a basic and useful metric to characterize and compare different dynamic behaviors.

## 2. RELATED WORK

**Time evolving OSN Snapshots.** While static features of OSNs are well studied, works on dynamics of online social networks are still ongoing. Among all Leskovec *et al.* in [15] detected two important properties on dynamic OSN data: graph densification, *i.e.* the average degree increases, and shrinking diameter. Several different social graphs has been studied in order to capture the growth of components and communities. Palla *et al.* [19] investigated the time dependence of overlapping communities and Berlingerio *et al.* [4] detected clusters of temporal snapshots of a network, interpreted as eras of evolution. Authors in [16, 25, 1] studied the dynamics of disconnected components. Finally Backstrom *et al.* [2] investigated the structural features which influence people in joining communities and their growth process. Alternatively, the per node dynamics was studied in [14] where the authors captured the evolution of key network parameters, and evaluate the extent to which the edge destination selection process subscribes to preferential attachment. As concerns acceleration, in [9] the authors considered an overall network size growth as a global property

and they modeled this global acceleration for the purpose of predicting the next network stage.

**Interdisciplinary Study of Human Dynamics.** In [3], Barabasi observed that the timing of human activity is inhomogeneous and bursty, disputing the previous hypothesis that human activities are randomly distributed in time. The inhomogeneity idea was extended in [23] and validated on few networks such as a Hungarian news portal, e-mails, library activities in a University and a trade transactions. Similarly, [17] analyzes the activity burstiness of blogs using entropy plots, and shows non-uniformity and self-similarity of the number of posts time sequences. Furthermore, [12] observes that temporal patterns are inhomogeneous or bursty even in mobile phone calls.

More recently, several studies examined user interaction in online social networks, both visible in-network interactions [24, 6], as well as invisible activity such as profile browsing [11, 22, 18]. However, these studies did not consider user interactions in the context of absolute time.

Finally, [13] developed a burst detection algorithm and observed that the appearance of a topic in a stream of documents, such as e-mails or research papers, has a bursty behavior. To the best of our knowledge, burstiness as a metric has never been investigated in the context of online social network dynamics.

## 3. TIMESTAMPED OSN DATASET

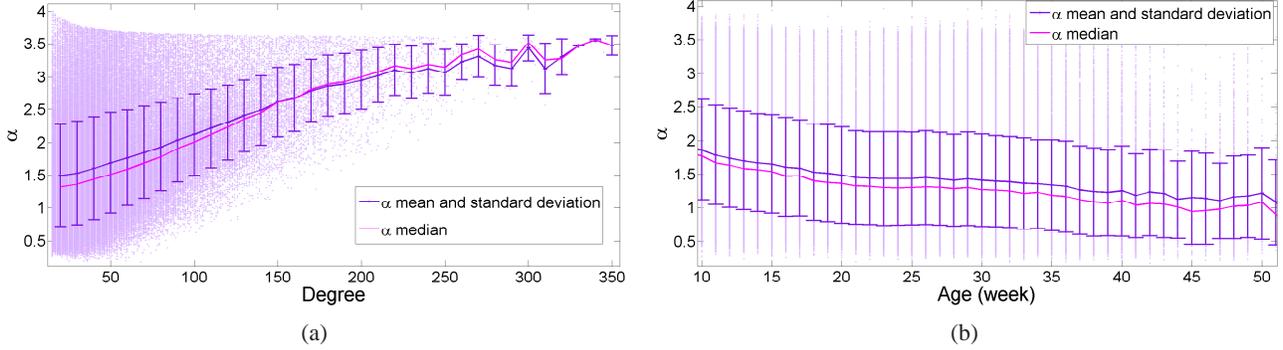
A major obstacle to studying OSN dynamics is the difficulty of obtaining detailed data traces. Our study uses an anonymized dataset that contains the timestamped creation of all users and edges in a large social network in China. This network has functionality very similar to Facebook, and has more than 200 million users.

Our anonymized trace was shared by the social network provider. All information on users and edges have been anonymized to protect user privacy. The trace is a complete dynamic graph that describes the evolution of the social network graph as a sequence of timestamped edge creation events. The first edge created in the network dates back to November 2005. We track the complete evolution over a period of one year from November 2005 to December 2006, including the creation of the first 60000 nodes and 8 million edges created.

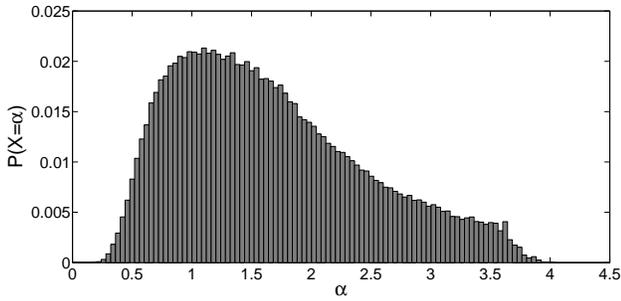
## 4. BURSTY NATURE OF LINK CREATION

Bursty behavior has been observed in various contexts such as WWW traffic patterns [8], emails exchanges, and in general human behavior [23]. But it has not been studied in the context of online social networks. In this section, we study the link creation process as the growth of the neighborhood of each single node, and show that the linking activity of online social networks users is characterized by temporal bursty patterns.

To prove the burstiness of link creation, we consider for each user the event time series where an event represents the



**Figure 2:** a) Mean, median, and standard deviation (error bar) of  $\alpha$  as function of the final degree. Mean and median increase with the final node degree. To compute these values, we group node degrees in bins of 10, and consider their relative  $\alpha$  values. b) Mean, median and standard deviation (error bar) of  $\alpha$  as function of node age measured in weeks. Mean and median decrease very slowly with node age.



**Figure 1:** Distribution of the scale parameter  $\alpha$ , which characterizes the inter-event time distribution between consecutive link creations for a single individual.  $\alpha$  values have been grouped in bins of length 0.05. Values past the peak around 1 decrease much more slowly with respect to the left side.

creation of an edge incident to the considered node. On each time series, we apply the technique proposed in Vasquez *et al.* [23] and extended by McGlohon *et al.* [17], both based on the inter-event time distribution between consecutive events for a single individual. If the edge creation process is a Poisson-like process, *i.e.* homogeneous, then the inter-event time distribution should be an exponential distribution. On the other hand, a bursty arrival process is characterized by a power-law distribution where many short time intervals, each corresponding to intensive activities forming a burst, are separated by relatively fewer but longer periods of low or zero activity.

**Results.** In order to distinguish if the process is homogeneous or bursty, we fit the inter-event time data per node in our dataset using MLE (Maximum Likelihood Estimator), and select the model with the minimum AIC (Akaike Information Criterion). As a representative of the power law distribution family, we choose the Pareto with exponential cutoff  $P(t) = t^{-\alpha} \exp(-t/\lambda)$ , and use the exponential dis-

tribution  $P(t) = \mu \exp(-\mu t)$  to describe the inter-event time Poisson process. Finally, to avoid the impact of outliers, we remove from consideration users who have too few events, *i.e.* nodes with final degree less than 15 (median degree).

Our results show that the Pareto distribution with exponential cutoff exhibits the minimum AIC, meaning that almost all users in our dataset manifest a bursty behavior in link creation. In addition, the Kolmogorov-Smirnov (K-S test) validates the selected hypothesis for almost all users (86% of the population). These measurements offer direct evidence that at the level of a single individual, there is a heavy-tailed activity pattern, also found in other datasets [10].

Having shown that individuals add links in a temporal bursty manner, we analyze the similarity of the bursty process across users, by computing the distribution of the scale parameter  $\alpha$  determined separately for each user. As shown in Figure 1,  $\alpha$  values are scattered around a peak at 1, with an heavy tail in the right side. This partially corroborates the results found in [23], which showed a single group of users with very similar behavior described by the Gaussian distribution of  $\alpha$  centered at 1. However, the heavy tail suggests that users in online social networks cannot be easily grouped in a single category, but have quite different behaviors in adding links.

To understand the reasons behind the observed heavy tail, we take into account two factors: the degree and the age of a node, *i.e.* how long the node has been in the network (in weeks). In Figure 2, we show the relationship between the scaling parameter  $\alpha$  and the two variables, *i.e.* degree and age of a node. Figure 2(a) shows that the mean  $\alpha$  value increases with degree. This fact suggests that nodes with higher degree contribute more to the right tail. This means that, although all the nodes manifest the same bursty behavior, nodes with higher degree have more closely spaced bursts. On the contrary, Figure 2(b) shows that age does not influence the right tail, since the mean value in each age slot is close to the mean of the  $\alpha$  distribution in Figure 1, and remains quite constant for different age values. The small

decrease is due to the fact that older nodes have a greater chance to undergo long periods of inactivity.

In summary, we showed in this section that users follow a bursty process in creating links, where bursts occur more frequently in nodes with high final degree. In the next section, we will describe our new proposal for using acceleration and deceleration as metrics to analyze graph dynamics.

## 5. DEGREE ACCELERATION

Bursty phenomena have been studied in different areas of human activities, such as clicks or queries in search engines [20]. However, these previous investigations focused on bursts resulting from aggregate actions, such as group of users that manifest a common interest at a certain time. These burst detection algorithms are not suitable to investigate per node time series based on their link creation events, or substructures inside bursts.

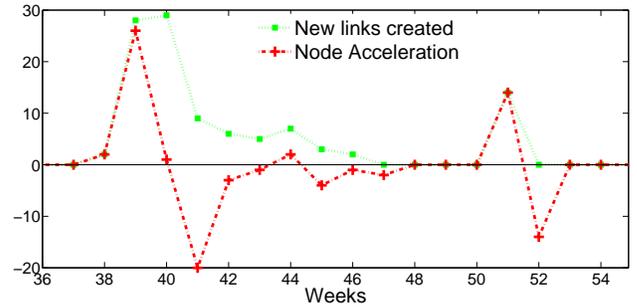
In this section, we propose a new methodology that identifies different phases that make up the bursty nature of the link creation process, and detects when bursts occur. We also identify the role played by each phase during the bursty process. From a high level, we observe that the alternation of activity/inactivity phases determines the burstiness of the event trace. In addition, bursts of activity have a typical internal structure, composed by a rapidly increasing slope and a gradually decreasing phase possibly interleaved by a plateau. An example is shown in Figure 3.

**Degree Acceleration.** Inspired by studies in physics and neuroscience on highly dynamic systems [5], we investigate the phases in bursty processes and detect bursts by measuring significant increments and decrements of new links formed per node. A burst begins when link formation activity rapidly increases, and ends following a decreasing phase. By leveraging the concept of acceleration, it is possible to easily identify and quantify significant changes in link creation activity. Let  $d_i(t)$  be the degree of node  $i$  at time  $t$ , *i.e.* the total number of links incident to node  $i$  at time  $t$ , and let  $\Delta t$  be the time granularity that interleaves each  $d_i(t)$  measures. We can then compute degree acceleration as:

$$a_i^d(t) = \frac{d_i(t) - 2d_i(t - \Delta t) + d_i(t - 2\Delta t)}{(\Delta t)^2} \quad (1)$$

By computing degree acceleration, we can observe the initial start of bursts ( $a_i^d \gg 0$ ) and a burst's decaying phase ( $a_i^d \ll 0$ ). An example is shown in Figure 3. Note that acceleration captures two types of steady state conditions: a period of consistently high activity representing the plateau inside an activity burst (after an initial acceleration phase), and a steady state of low activity outside of activity bursts.

**Defining Phases.** While exploring the burstiness of the link creation process, we found that the growth of each node is characterized by transition phases in which users significantly change their link formation behavior. This led us to identify four different phases that describe the patterns involved in the nodes' growth processes. The differ-



**Figure 3: An example of degree acceleration, computed on a single node from our dataset. The green dotted line represents the number of links created by the node each week. The red dotted line represents the acceleration computed according to Eq. 1. In week 39, the node shows a large acceleration, follows by a plateau. The node decelerates into week 42, when it enters a cruising phase (link creation is stable) for 4 weeks.**

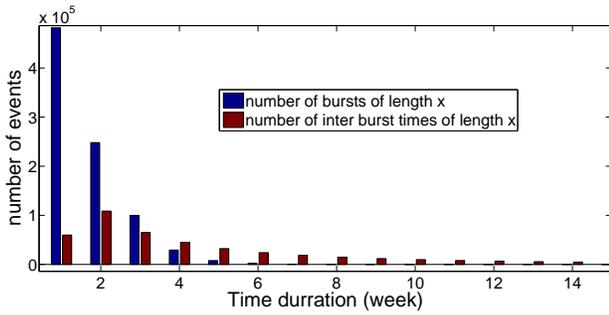
ent phases can be described by defining a time-dependent state variable for each node in the system. More specifically, the acceleration phase is characterized by a large increment in creating new links, *i.e.*  $a_i^d \gg 0$ , and the deceleration phase is described by a strong decay measured by  $a_i^d \ll 0$ . Then we define two intermediate phases: *cruising* and *inactivity*. The first corresponds to a steady state of a node, where the number of links created per week is almost constant. This phase can correspond both to high activity or to small oscillations around inactivity, and is characterized by at least one new edge (captured by the variable  $c_i(t) = 1$ ) and small  $a_i^d$  values. These small  $a_i^d$  values are centered around the value  $a_i^d = 0$ , and are bounded with two thresholds  $\theta_1$  and  $\theta_2$ . The second phase, *i.e.* inactivity, occurs when a node does not create any links for an entire time window. We formalize these four phases by introducing the function  $s_i(t) : \mathbb{R} \rightarrow \{acc, dec, cruise, inact\}$  defined as follows:

$$s_i(t) = \begin{cases} acc & a_i^d(t) \in (\theta_1, +\infty) \\ dec & a_i^d(t) \in (-\infty, \theta_2) \\ cruise & a_i^d(t) \in [\theta_2, \theta_1] \wedge c_i(t) = 1 \\ inact & c_i(t) = 0 \end{cases} \quad (2)$$

where  $c_i(t) = 1$  if and only if node  $i$  creates at least one edge at time  $t$ , otherwise  $c_i(t) = 0$ . Degree acceleration  $a_i^d(t)$  and the related  $s_i(t)$  function represent a general tool to investigate the burstiness structure, and to highlight the detailed properties of each phase.

## 6. EXPERIMENTAL ANALYSIS

In this section, we characterize the link creation process by analyzing our traces using our acceleration methodology. The experimental analysis has been performed with the following settings:  $\Delta t = 1$  week to avoid cyclic fluctuations in acceleration due to increase in user activities over each



**Figure 4: Histogram of the duration (blue) and of the inter-burst times (red)**

weekend, and cruising phase thresholds are  $\theta_1 = 2$  and  $\theta_2 = -2$ . On each user event time series we apply the degree acceleration methodology, thus identifying bursts and their internal phases.

We first calculate the duration of each burst and the inter-burst times, *i.e.* the time between consecutive bursts. By leveraging the  $s_i$  function, we can convert a node time series in a string where a burst corresponds to the string pattern  $acc^+cruise^*dec^*$ . In this way we can extract substring and detect when a burst happens, how long it lasts and how long after the next burst appears. Results on these analyses are presented in Figure 4. As concerns the burst duration we can see that one half of the bursts lasts a week and most of them hold over at most 4 weeks. Therefore, inter burst times exhibit the same behavior of the inter-event time distribution; in fact bursts could also appear with a certain probability after different weeks, as the red bar plot suggests.

## 6.1 The Role of Phases

The role played by each phase along the node lifetime is a key element to understand the network dynamics, and is also crucial when designing generative models based on per-node temporal behavior. To this purpose, we consider two main aspects: (i) the time a node spends in each phase and (ii) the per-node amount of links created in the different phases.

We perform this analysis from two perspectives, by considering the aggregate behavior of all nodes, and on per-node behavior. In order to understand the role of different phases during a node’s lifetime, we define  $\phi^l$  and  $\psi^l(i)$  to compute the percentage of time spent in each phase by all nodes (Equation 3) and by each node (Equation 4).

$$\phi_{phase}^l = \frac{\sum_{i \in N} \sum_{t=1}^T \mathcal{I}_{phase}(s_i(t))}{\sum_{i \in N} life(i)} \quad (3)$$

$$\psi_{phase}^l(i) = \frac{\sum_{t=1}^T \mathcal{I}_{phase}(s_i(t))}{life(i)} \quad (4)$$

where  $life(i)$  represents the lifetime in weeks of a node,  $\mathcal{I}$  is the indicator function and  $phase = \{acc, dec, cruise, inact\}$ .  $N$  indicates the number of nodes at time  $T$ , which represents the last week considered in the dataset.

	acc	dec	cruise	inact
$\phi^l$	0.11	0.14	0.28	0.47
$\phi^e$	0.52	0.17	0.31	0
$\psi_{0.8}^l$	0.25	0.27	0.44	0.66
$\psi_{0.8}^e$	0.74	0.70	0.16	0

**Table 1: In the first two rows we report  $\phi^e$  (definition 5) and  $\phi^l$  (definition 3) values for each phase. In the last two rows the 0.8-quantiles of  $\psi^l$  and  $\psi^e$  distributions.**

The relationship between link creation and phase is quantified by  $\phi^e$ , which corresponds to the percentage of the overall edges created within each phase, and  $\psi^e(i)$ , which is the link generation rate for node  $i$  in each phase:

$$\phi_{phase}^e = \frac{\sum_{i \in N} \sum_{t=1}^T \mathcal{I}_{phase}(s_i(t)) n_i(t)}{2m} \quad (5)$$

$$\psi_{phase}^e(i) = \frac{\sum_{t=1}^T \mathcal{I}_{phase}(s_i(t)) n_i(t)}{d_i(T)} \quad (6)$$

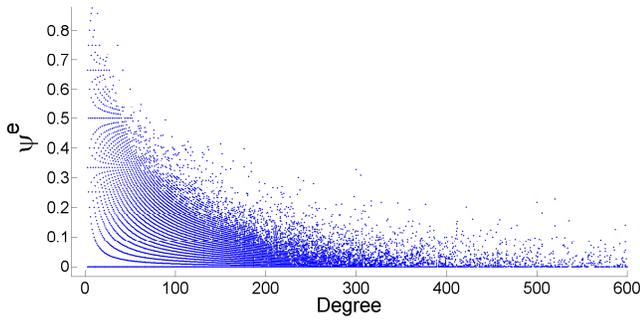
where  $m$  is the number of link at time  $T$  and  $n_i(t)$  is the number of links acquired by the node  $i$  in the time slot  $t$ . The results are reported in Table 1, where  $\psi_{0.8}^l$  and  $\psi_{0.8}^e$  are the 0.8-quantile of the distributions of  $\psi^l$  and  $\psi^e$ , and are discussed below.

**Inactivity phase.** During the inactivity phase, by definition, we do not observe growth since no links are created. However, inactivity acquires importance in the temporal dimension, because it deeply affects the burstiness. The high values of  $\phi^l$  and  $\psi_{0.8}^l$  highlight that node activities are concentrated in few and small periods; thus, for most part of their life, nodes do not influence the network dynamic evolution.

**Acceleration and Deceleration Phases.** Nodes spend only a small amount of their life in these phases, in particular after acceleration events, longer periods of weaker activity follow. However, the amount of links generated in these phases determines the structure of our social network. In fact, a link has very high probability, 69%, to be generated in one of these two phases, in particular 52% in acceleration and 17% in deceleration.

**Cruising Phase.** Cruising periods cover an important portion of nodes’ lifetime. Furthermore,  $\phi^e = 0.31$  and  $\psi^e = 0.16$  would suggest that this phase has a role also in link creation. However, only few cruising periods have relevance in the edge growth. Indeed, it depends on whether the cruising phase is inside a burst or it corresponds to small oscillations around inactivity. A node in a burst cruising phase is creating many links, while in the other case the number of links created is irrelevant. Finally, the cruising phase has a pronounced impact only for nodes with low degree, as shown in Figure 5.

We have shown that acceleration (*acc*) and deceleration (*dec*) phases are those responsible of the growth and dynamics of the network, despite the fact that they represent a very



**Figure 5: Relationship between  $\psi_{cruise}^e$  and the degree.**  $\psi_{cruise}^e$  decreases as the degree raises, so the cruising phase has a pronounced impact only for nodes with low degree.

small part of a node life.

## 6.2 Acceleration and Deceleration Features

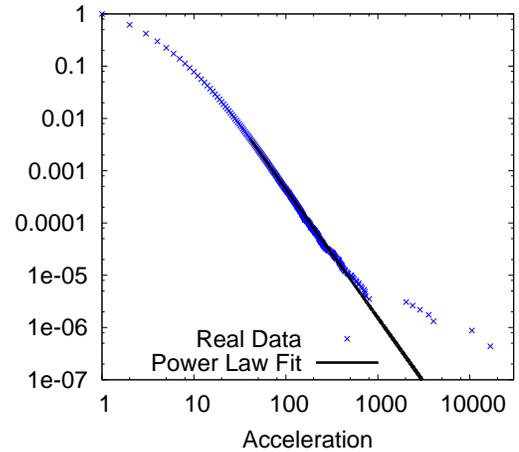
In depth understanding of acceleration/deceleration phases reveals how users operate in the network after they join. This knowledge could be very useful to ensure efficient management of the OSN's resources. This section focuses on acceleration and deceleration by means of illustrating their importance from a network perspective; showing that they follow a power law distribution and finally investigating the impact of node aging on link creation process.

**Network perspective on acceleration/deceleration.** From the network perspective, an estimate of how many and which nodes are changing the graph structure would greatly help in managing the system resources. Figure 7 shows that in each week only a very small number of nodes are in acc/dec phases. Specifically, at the end of the year, they roughly represent the 20% of the network nodes. These nodes can be easily identified as soon as they experience a phase transition from the inactive/cruising to the accelerated phase since their values of acceleration abruptly increase.

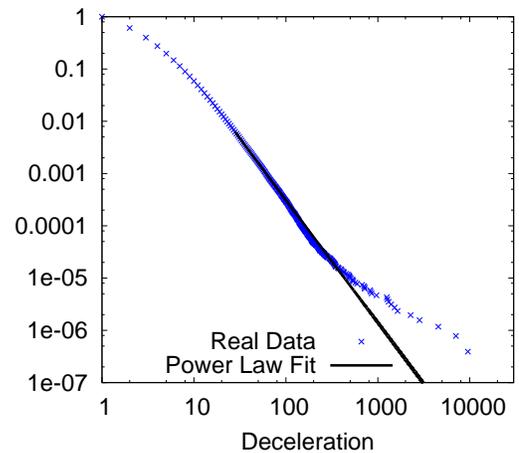
**Acceleration/deceleration probability distributions.** By applying the statistical framework proposed by Clauset [7], we find that acceleration and deceleration distributions are power law, (Figure 6(a) and 6(b)). By considering the overall network, this result implies that half of the acceleration and deceleration events have a small size, but they are very likely to show rapid increase and decrease respectively. The upper tail of the acceleration distribution exhibits such high values of acceleration that can't correspond to normal user. Those events are most likely associated to people with a large amount of followers or accounts for advertisement.

**The impact of aging.** The general behavior of a node is a sequence of acceleration/deceleration phases of constant magnitude, after an initial burst. In general, nodes wait at most for one month before initiating their activity.

We start by defining the  $age(t)$  of a node  $u$  as the time elapsed between the appearance of  $u$  in the network (times-



(a)



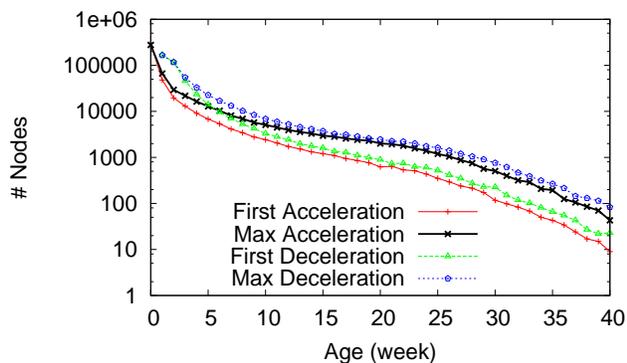
(b)

**Figure 6: a) Acceleration CCDF (complementary cumulative distribution function) and the resulting fitted distribution ( $\alpha = 3.46$ ). b) Deceleration CCDF and the resulting fitted distribution ( $\alpha = 3.34$ ).**

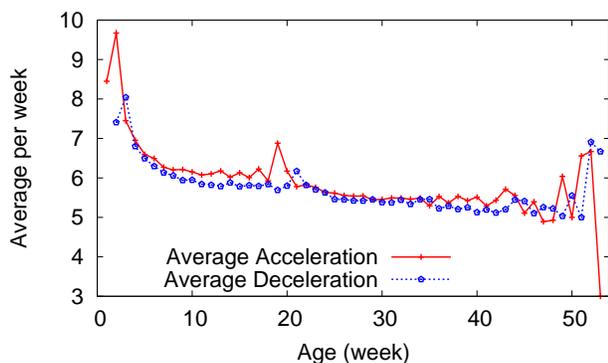
tamp of the first edge incident to  $u$ ), and time  $t$ . The observables whose dependence on age, need to be studied are:  $n_{firstAcc/Dec}(t)$ , the number of nodes showing their first acceleration / deceleration at time  $t$  and  $n_{maxAcc/Dec}(t)$ , the number of nodes manifesting their maximum acceleration / deceleration at time  $t$ . Finally, we calculate the average acceleration/ deceleration  $avg_{Acc/Dec}(age)$ .

Analyzing and comparing  $n_{maxAcc}(t)$  and  $avg_{Acc}(age)$  in Figures 8(a) and 8(b), we observe that most nodes enter the phase of maximum acceleration in the first week. In addition, Figure 8(a) shows that the activity after the first peak does not decrease as fast as its respective acceleration.

Figure 8(b) highlights another interesting behavior of the acc/dec phases. The average acceleration remains constant when age increases. This is consistent with what we found in Figure 3, *i.e.* nodes experience a big burst of acceleration

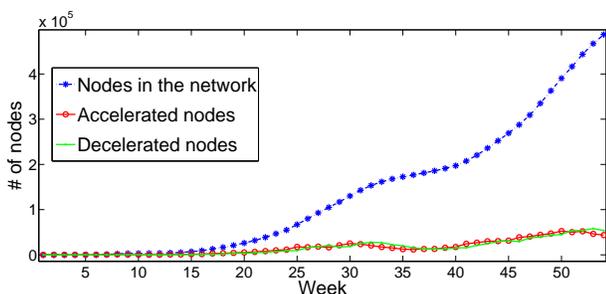


(a)



(b)

**Figure 8:** a) shows the times when nodes experience their first acceleration, maximum acceleration, first deceleration, and maximum deceleration ( $y$ -axis on logscale). b) shows the average acceleration/deceleration with respect to node age.



**Figure 7:** For each week, the number of nodes in the network (network size) and the number of nodes in the acceleration/deceleration phases. In each week only a very small number of nodes are in the acc/dec phases though the network size rapidly grows.

in the first week after joining the network, and subsequent bursts never match the first in intensity.

## 7. CONCLUSION

In this paper we investigated the bursty nature of the link creation process in OSN. We prove not only that it is a highly inhomogeneous process, but also identify patterns of burstiness common to all nodes. In terms of edge creation, users are inactive for most of their lifetimes, and concentrate their link activity in a number of short regular time periods. To characterize node activity, we developed a new methodology based on the acceleration of degree growth, which allows us to highlight the internal structure of link creation bursts.

We believe using acceleration as a general metric to characterize network dynamics prompts future work in studying link generation mechanisms. In particular, defining different phases of edge creation hints at the possibility of characterizing users into distinctive activity levels that correlate with their likelihood of adding social links. Some preliminary results confirm this intuition: when nodes (users) first join the network, they create links based on the preferential attachment mechanism; while in later bursts, nodes seem to explore (acceleration phase) and densify (deceleration) in far regions of the graph that could correlate to spatial areas [26]. These results open the door for new generative models [21] that consider different phases of node activity.

## 8. ACKNOWLEDGMENT

This work was partially funded by the Italian Ministry for Instruction, University and Research under the PRIN PEOPLENET (2009BZM837) Project. The authors would also like to thank DARPA for their support under GRAPHS (BAA-12-01).

## 9. REFERENCES

- [1] AKOGLU, L., AND FALOUTSOS, C. Rtg: a recursive realistic graph generator using random typing. *Machine Learning and Knowledge Discovery in Databases* (2009), 13–28.
- [2] BACKSTROM, L., HUTTENLOCHER, D., KLEINBERG, J., AND LAN, X. Group formation in large social networks: membership, growth, and evolution. In *Proc. of KDD* (2006).
- [3] BARABÁSI, A.-L. The origin of bursts and heavy tails in human dynamics. *Nature* 435 (2005), 207–211.
- [4] BERLINGERIO, M., ET AL. As time goes by: Discovering eras in evolving social networks. In *Proc. of PAKDD* (2010), pp. 81–90.
- [5] CHIAPPALONE, M., ET AL. Burst detection algorithms for the analysis of spatio-temporal patterns in cortical networks of neurons. *Neurocomputing* 65-66 (2005), 653–662.
- [6] CHUN, H., KWAK, H., EOM, Y.-H., AHN, Y.-Y., MOON, S., AND JEONG, H. Comparison of online social relations in terms of volume vs. interaction: A case study of cyworld. In *Proc. of IMC* (2008).
- [7] CLAUSET, A., SHALIZI, C. R., AND NEWMAN, M. E. J. Power-law distributions in empirical data. *SIAM Review* 51 (2009), 661–703.

- [8] CROVELLA, M. E., AND BESTAVROS, A. Self-similarity in world wide web traffic: evidence and possible causes. *IEEE/ACM Transactions on Networking* 5 (1997), 835–846.
- [9] DOROGOVTSSEV, S. N., AND MENDES, J. F. F. Accelerated growth of networks. *CoRR* (2002).
- [10] GAITO, S., PAGANI, E., AND ROSSI, G. P. Strangers help friends to communicate in opportunistic networks. *Computer Networks* 55, 2 (2011).
- [11] JIANG, J., WILSON, C., WANG, X., HUANG, P., SHA, W., DAI, Y., AND ZHAO, B. Y. Understanding latent interactions in online social networks. In *Proc. of IMC* (2010).
- [12] JO, H.-H., PAN, R. K., AND KASKI, K. Emergence of bursts and communities in evolving weighted networks. *PLoS ONE* 6 (08 2011), e22687.
- [13] KLEINBERG, J. Bursty and hierarchical structure in streams. In *Proc. of KDD* (2002).
- [14] LESKOVEC, J., BACKSTROM, L., KUMAR, R., AND TOMKINS, A. Microscopic evolution of social networks. In *Proc. of KDD* (2008).
- [15] LESKOVEC, J., KLEINBERG, J., AND FALOUTSOS, C. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proc. of KDD* (2005).
- [16] MCGLOHON, M., AKOGLU, L., AND FALOUTSOS, C. Weighted graphs and disconnected components: patterns and a generator. In *Proc. of KDD* (2008).
- [17] MCGLOHON, M., ET AL. Finding patterns in blog shapes and blog evolution. In *Proc. of ICWSM* (2007).
- [18] NAZIR, A., RAZA, S., GUPTA, D., CHUAH, C., AND KRISHNAMURTHY, B. Network level footprints of facebook applications. In *Proc. of IMC* (2009).
- [19] PALLA, G., BARABASI, A., AND VICSEK, T. Quantifying social group evolution. *Nature* 446 (2007), 664–667.
- [20] PARIKH, N., AND SUNDARESAN, N. Scalable and near real-time burst detection from ecommerce queries. In *Proc. of KDD* (2008).
- [21] SALA, A., CAO, L., WILSON, C., ZABLIT, R., ZHENG, H., AND ZHAO, B. Y. Measurement-calibrated graph models for social network experiments. In *Proceedings of the 19th international conference on World wide web* (2010), WWW '10.
- [22] SCHNEIDER, F., FELDMANN, A., KRISHNAMURTHY, B., AND WILLINGER, W. Understanding online social network usage from a network perspective. In *Proc. of IMC* (2009).
- [23] VÁZQUEZ, A., OLIVEIRA, J. A. G., DEZSÖ, Z., GOH, K.-I., KONDOR, I., AND BARABÁSI, A.-L. Modeling bursts and heavy tails in human dynamics. *Physical Review E* 73 (2006).
- [24] WILSON, C., BOE, B., SALA, A., PUTTASWAMY, K. P. N., AND ZHAO, B. Y. User interactions in social networks and their implications. In *Proc. of EuroSys* (April 2009).
- [25] ZHELEVA, E., SHARARA, H., AND GETOOR, L. Co-evolution of social and affiliation networks. In *Proc. of KDD* (2009).
- [26] ZIGNANI, M., GAITO, S., AND ROSSI, G. Extracting human mobility and social behavior from location-aware traces. *Wireless Communications and Mobile Computing* (2012).