

Microscopic Modelling Circadian and Bursty Pattern of Human Activities

Jinhong Kim¹, Deokjae Lee², Byungnam Kahng^{2*}

1 Complex System Analysis Team, NHN Corp., Seongnam City, Gyeonggi-do, Korea, **2** Department of Physics and Astronomy, Seoul National University, Seoul, Korea

Abstract

Recent studies for a wide range of human activities such as email communication, Web browsing, and library visiting, have revealed the bursty nature of human activities. The distribution of inter-event times (IETs) between two consecutive human activities exhibits a heavy-tailed decay behavior and the oscillating pattern with a one-day period, reflective of the circadian pattern of human life. Even though a priority-based queueing model was successful as a basic model for understanding the heavy-tailed behavior, it ignored important ingredients, such as the diversity of individual activities and the circadian pattern of human life. Here, we collect a large scale of dataset which contains individuals' time stamps when articles are posted on blog posts, and based on which we construct a theoretical model which can take into account of both ignored ingredients. Once we identify active and inactive time intervals of individuals and remove the inactive time interval, thereby constructing an *ad hoc* continuous time domain. Therein, the priority-based queueing model is applied by adjusting the arrival and the execution rates of tasks by comparing them with the activity data of individuals. Then, the obtained results are transferred back to the real-time domain, which produces the oscillating and heavy-tailed IET distribution. This microscopic model enables us to develop theoretical understanding towards more empirical results.

Citation: Kim J, Lee D, Kahng B (2013) Microscopic Modelling Circadian and Bursty Pattern of Human Activities. PLoS ONE 8(3): e58292. doi:10.1371/journal.pone.0058292

Editor: Petter Holme, Umeå University, Sweden

Received: December 7, 2012; **Accepted:** February 1, 2013; **Published:** March 11, 2013

Copyright: © 2013 Kim et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: National Research Foundation grant awarded through the Acceleration Research Program (Grant No. 2010-0015066). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: JK is an employee of NHN Corp. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: bkahng@snu.ac.kr

Introduction

In the information age, a large scale of databases containing information on human activities on the Web are easily accessible. Understanding the emerging patterns from those datasets is a new interdisciplinary research subject [1,2]. Since individuals behave through complex and sometimes random decision-making processes, one may wonder whether it is indeed possible to predict human behaviors quantitatively. However, it was recently revealed that digital records left at the media behind one's activities make it possible to predict human activities up to 93% [3]. Accordingly, it has become an attractive subject to investigate emerging patterns from such large-scale data bases. Power-law or heavy-tailed behavior in the distribution of inter-event times (IET) between two consecutive human activities is one example of such emerging patterns. This example can be seen in various systems such as email [4–9] or surface mail communications [10], Web browsing [7,11], library loans [7], financial trades [7,12], on-line movie watching [13], file downloads [14–16], printing requests [17], and various actions on the Web [18]. This power-law behavior indicates that human activities proceed in a bursty manner during a short time interval, which is separated from other such intervals by long intermittent periods [19,20].

Several theoretical models have been proposed to explain such heavy-tailed behaviors in the IET distribution. One interesting model is the priority-based queueing model [4,7,21,22], in which the human activity of uploading articles is regarded as task executions in a queue, where tasks are performed based on the order of priorities assigned to each task. The use of this priority-

based model leads to a power-law or heavy-tailed behavior in the waiting time distribution of tasks in the queue [9,13,18,23–25]. The waiting time distribution was interpreted as the IET of human activities. However, the priority-based queueing model ignores important ingredients, such as the circadian pattern of human life [26] and the diversity of individual activities. Indeed, the empirical data recently collected exhibit an oscillating pattern with a one-day period in the IET distribution [6,13,18,27], which cannot be produced in the queueing model. Moreover, the decay behavior of the IET distribution in the long-time regime depends on the activities of individuals. Here, the activity of an individual is defined as the average number of posted articles in unit time. In this paper, we obtained a large-scale dataset containing high-resolution data, and found a new pattern in the IET distribution that exhibits a power-law behavior when the IET is smaller than one day, where the exponent is insensitive to the activities of individuals. However, when the IET is longer than one day, the IET distribution exhibits a heavy-tailed behavior, in which the tail part depends on the activities of individuals. These empirical results are reproduced by developing a theoretical model below.

Methods

We analyze a large scale of dataset from the largest portal site in Korea, NAVER (<http://naver.com>) during more than five years. The dataset consists of individuals' time stamps when articles were posted on blog posts, which were recorded in the unit of seconds. There are 520,771,167 postings contributed by 9,878,904 distinct bloggers. Among them, we only select the data that were written

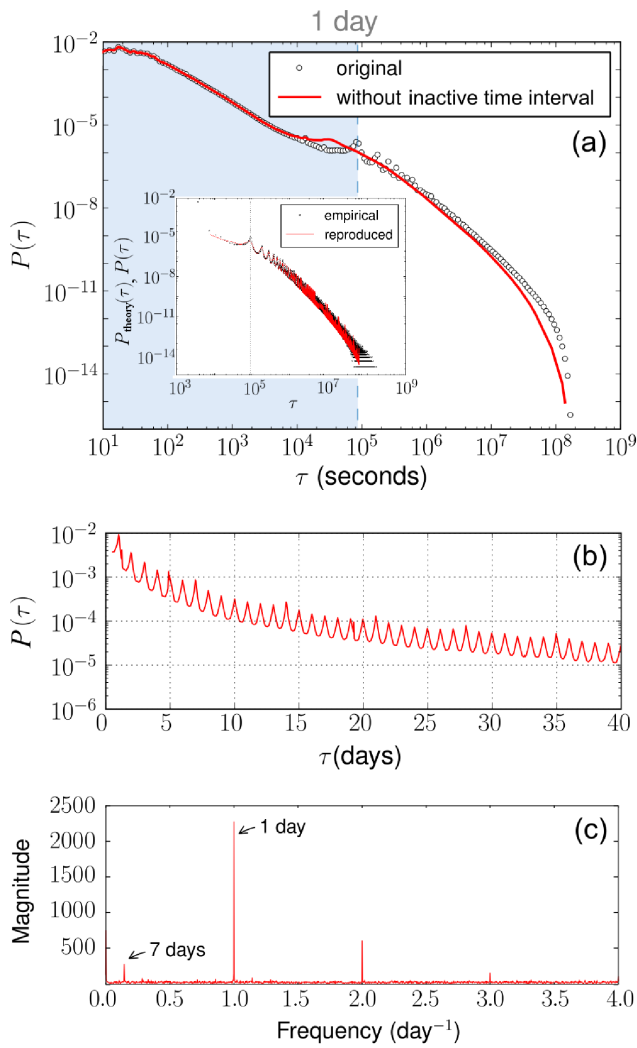


Figure 1. Empirical IET distribution. (a) Plot of the IET distribution $P(\tau)$ based on the empirical data (\circ). The IET distribution $P_{ad}(\tau)$ after the removal of the inactive time interval is also shown (solid curve). Inset: Comparison of the IET distribution obtained from the empirical data (\circ) with that from the theory $P_{theory}(\tau)$ (solid curve). (b) Enlarged representation of the IET distribution $P(\tau)$, in which clear periodic peaks are observed. (c) The Fourier transform of the IET distribution. Peaks are located at frequencies $\omega = (one\ week)^{-1}$ and $(one\ day)^{-1}$. Other peaks at multiples of $(one\ day)^{-1}$ are redundant. doi:10.1371/journal.pone.0058292.g001

by bloggers that had authored more than 100 articles and worked for more than one month. This selection aims to exclude those bloggers who had posted suspicious spam content. After this data filtering, the number of remaining articles is 379,627,193, contributed by 908,409 users.

From this dataset, we obtained the following empirical results: (i) The IET distribution decays following a power law with the exponent $\alpha \approx 1.5$ in a time regime shorter than one day. (ii) The IET distribution exhibits a heavy-tailed decay behavior in the long-time regime, which is nonuniversal depending on individual activities. (iii) An oscillating pattern appears with a period of one day; this pattern persists over the entire long-time regime. However, the amplitude of the oscillation pattern decreases with time. Details regarding these results are presented below.

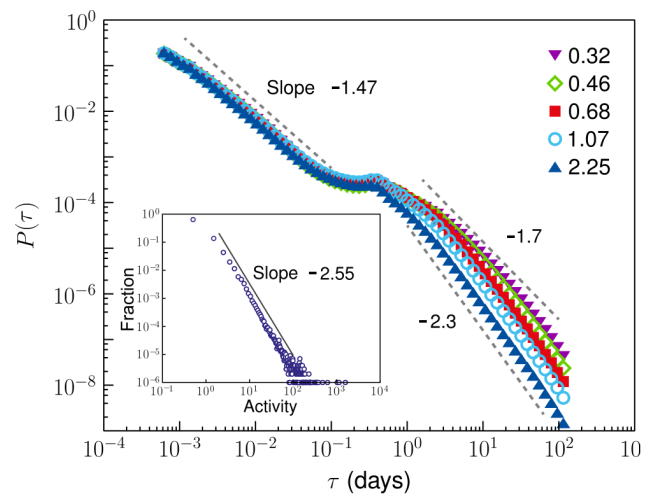


Figure 2. Dependence of the IET distributions of individual users on their activities. Inset: Population as a function of activity, showing decay as a power law with exponent $\alpha \approx 2.55$. doi:10.1371/journal.pone.0058292.g002

We measured the IETs defined as the interval between two consecutive time stamps for each user. Then the distribution $P_i(\tau)$ of the IETs of user i is obtained as $P_i(\tau) \equiv N_i(\tau) / \sum_{\tau} N_i(\tau)$, where $N_i(\tau)$ is the number of events having an IET of τ . The total number of articles, N_i , written by user i is given as $N_i = \sum_{\tau} N_i(\tau) + 1$. To determine the collective behavior of all the users, we calculate

$$P(\tau) = \frac{\sum_i N_i(\tau)}{\sum_{i,\tau} N_i(\tau)}. \quad (1)$$

$P(\tau)$ is shown in Fig. 1(a). When $\tau < 1$ day, $P(\tau)$ behaves as $\sim \tau^{-1.5}$. When $\tau > 1$ day, $P(\tau)$ follows a skew distribution. Interestingly, there exists an oscillating pattern in $P(\tau)$, which can be seen more clearly in the finer scale shown in Fig. 1(b). Moreover, peak heights periodically change with a period of one week [6,13,18,27]. To check the periodicity of the oscillating pattern, we perform a Fourier transformation, $F(\omega) = \sum_{\tau} P(\tau) e^{i\omega\tau}$. Figure 1(c) shows that there indeed exist

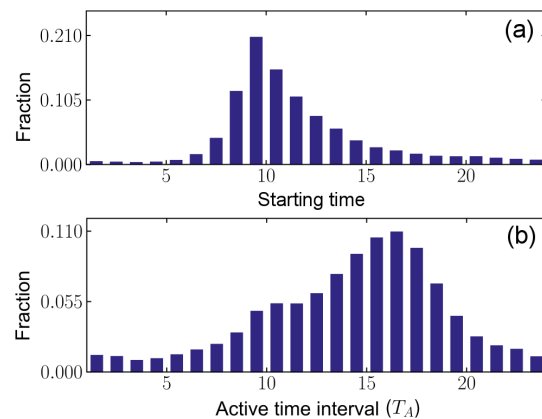


Figure 3. Distribution of active time intervals. (a) Distribution of the starting time of the active time interval. A peak is located between 9 and 10 am. (b) Distribution of active time intervals. The model is located at 16 h. doi:10.1371/journal.pone.0058292.g003

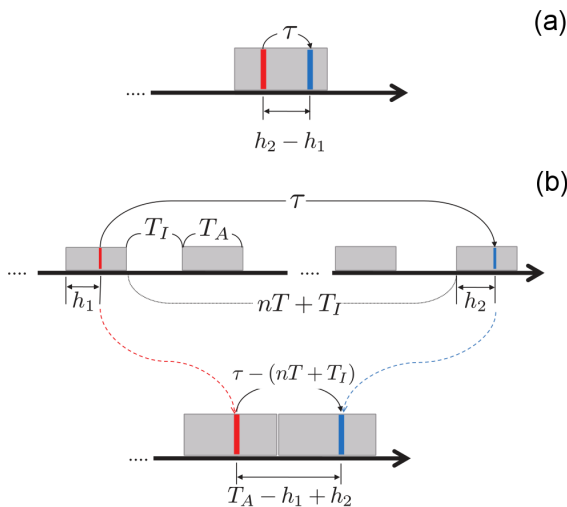


Figure 4. Schematic illustration of the model with circadian periodicity. It is assumed that an individual essentially lives a well-regulated daily life consisting of active and inactive time intervals. To reproduce the oscillating behavior of IET distribution within the framework of the queueing model, we construct an *ad hoc* time domain in which separated active time intervals are connected by removing inactive time intervals between them. See text for details. doi:10.1371/journal.pone.0058292.g004

two distinct meaningful peaks in $P(\tau)$ at the frequencies corresponding to one day and one week, respectively. Other peaks correspond to multiples of one day. We study how such an oscillating pattern can be reproduced within the framework of the priority-based model later.

Next, we examine the dependence of the IET distribution on the activity of individuals. The activity A_i of user i is the number of articles written per unit time interval. Thus, when user i writes N_i articles during the time interval $T_{tot}^{(i)}$ [13,18], where $T_{tot}^{(i)}$ is the time interval between the first and the last time stamp of user i , the activity of user i is given as $A_i = N_i / T_{tot}^{(i)}$. To determine the heterogeneity of individual activities, we measured the distribution

of individual activities as shown in the inset of Fig. 2. Indeed, the distribution decays, following a power law with the exponent ≈ 2.6 , indicating that individual activities are considerably heterogeneous. Thus, it is worth investigating how the heterogeneity of activities affects the IET distribution [28,29]. In Fig. 2, we can see that as one’s activity level becomes higher, the IET distribution decays faster in the long time regime. This behavior is rather natural in the sense that a user with higher activity has a shorter mean IET. Accordingly, it would be interesting to introduce a new model to illustrate this activity-dependent behavior, and such a model is presented later.

Results and Discussion

Modelling Oscillating Pattern

In previous studies, the heavy-tailed behavior of the IET distribution was investigated by using the priority-based queueing model. In this approach, time was considered as continuous without any intermission. However, humans do not work continuously, and hence, intermission, for example, those that account for sleeping, must be considered. Moreover, the pattern of daily life during weekdays is almost regular, but it differs from that during weekends. Thus, it is natural to assume that each person can have a regular time interval during which the person is away from on-line world. This time interval is called the inactive time interval, and the remaining time of a day is called the active time interval. Moreover, the duration and starting time of the active time interval depend on the individual (see Fig. 3).

We suppose the situation that two events occur in the active period of one day (see Fig. 4a) at times h_1 and h_2 , where $h_1 < h_2$ and h_1 and h_2 belong to the same active time interval. Then, the inter-event time is defined as $\tau = h_2 - h_1$. More generally, when two events are executed in different active intervals separated by $nT + T_I$, where n is an integer $n \geq 0$ (see Fig. 4b), we can obtain the following relation,

$$\tau - (n + 1)T_I = nT_A + (T_A - h_1 + h_2), \tag{2}$$

where $\tau - (n + 1)T_I$ is the IET after removing the inactive time intervals. This quantity is defined as the IET in the *ad hoc* time

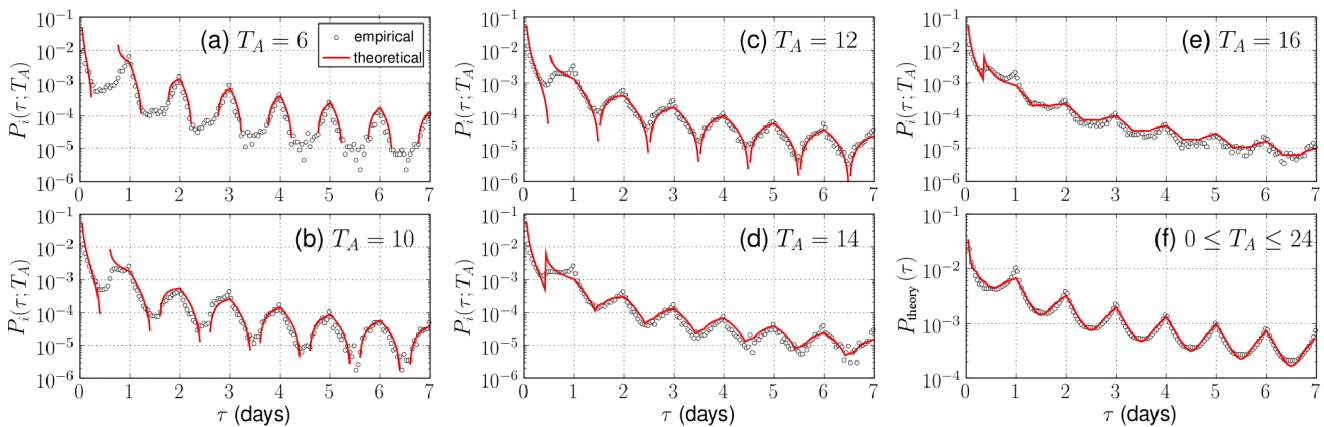


Figure 5. Comparison of empirical (open circles) and theoretical (solid lines) inter-event time distributions with the circadian active-inactive pattern for different T_A . Empirical distributions $P_i(\tau; T_A)$ are obtained by aggregating the top 100 users who have a clear periodicity with an active time interval T_A , and the distributions suitably show the change in the peak height and width. The weighted average of $P_i(\tau; T_A)$ is also displayed in (f), and we can observe the characteristic peaks. doi:10.1371/journal.pone.0058292.g005

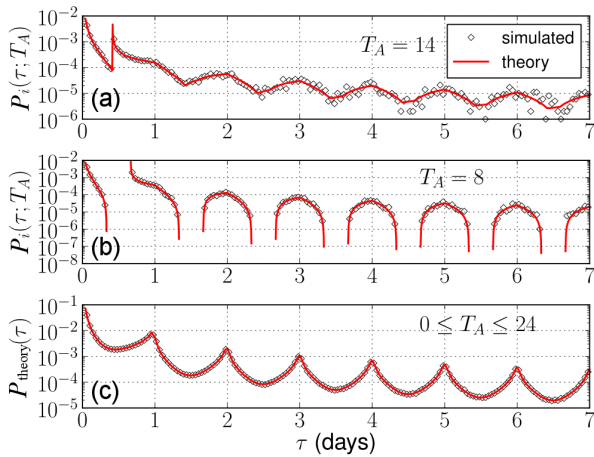


Figure 6. Comparison between simulated and theoretical IET distributions with the circadian pattern. To calculate $P(\tau; T_A)$ in Eq.(8), we assume that $P_{ad}(\tau) \sim \tau^{-1.5}$. We consider the two cases (a) $T_A=14$ and (b) $T_A=8$, as examples. (c) The distribution in Eq.(10) collected over the flat distribution of T_A . The resulting theoretical IET distribution is consistent with the one obtained from the simulated data.

doi:10.1371/journal.pone.0058292.g006

domain, and is denoted as τ' . Then the *ad hoc* time domain is continuous. We find that any inter-event time τ belongs to one of the two sets of intervals T_1 and T_2 , defined as

$$T_1 = \{\tau | 0 < \tau - (n+1)T < T_A\} \text{ for } h_1 < h_2, \quad (3)$$

and

$$T_2 = \{\tau | T_I < \tau - nT < T\} \text{ for } h_1 > h_2. \quad (4)$$

The fraction of each category is given as

$$q_1(\tau; T_A, n) = \frac{(n+1)T + T_A - \tau}{T_A} = \frac{h_2 - h_1}{T_A}, \quad (5)$$

and

$$q_2(\tau; T_A, n) = \frac{\tau - nT - T_I}{T_A} = \frac{T_A - h_1 + h_2}{T_A}, \quad (6)$$

respectively.

Let $P_{ad}^{(i)}(\tau')$ be the IET distribution of user i in the *ad hoc* time domain, and let $P_{ad}(\tau')$ be the collective one from individuals, defined as

$$P_{ad}(\tau') \equiv \frac{1}{N} \sum_i P_{ad}^{(i)}(\tau'). \quad (7)$$

Here, τ' is the IET defined in the *ad hoc* time domain, which is related to τ in the original time domain as $\tau' = \tau - mT_I^{(i)}$, where m is the largest non-negative integer satisfying $\tau' > 0$, which implies that there exist m inactive time intervals during τ . $P_{ad}^{(i)}(\tau')$ is obtained

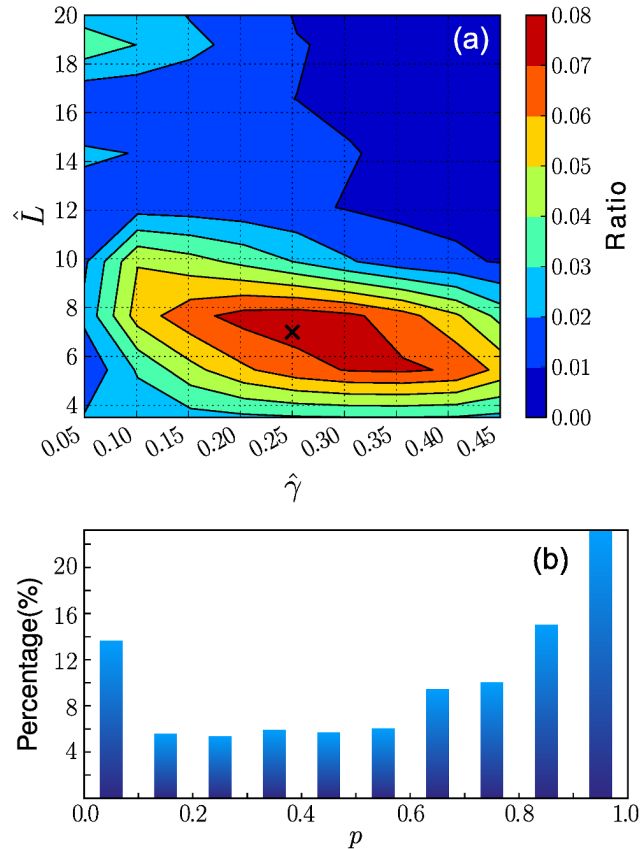


Figure 7. Modelling activity dependence. (a) Distribution of the best-estimate parameters \hat{L} and $\hat{\gamma}$ of individuals. Contour lines are obtained by interpolation between each nearest point. The most dense point is described by $\hat{L}=7$ and $\hat{\gamma}=0.25$, and a large portion of cases settle around the peak point. (b) A fraction of the p values in the KS test between synthetic and empirical probability distribution functions. Over 86% of cases have p values that are larger than 0.1, and hence, the null hypothesis cannot be rejected for those cases.

doi:10.1371/journal.pone.0058292.g007

from the queueing model [30], which is discussed later. Collecting all individuals' $P_{ad}^{(i)}(\tau')$, i.e., using the formula (7), we obtain $P_{ad}(\tau')$, which exhibits a heavy-tailed distribution shown in Fig. 1.

We consider how to reproduce the oscillating behavior. For this purpose, we assume that an IET distribution is given, for example, the previously obtained $P_{ad}(\tau')$ from the empirical data, or $P_{model}(\tau') \sim \tau'^{-1.5}$ from the queueing model [30]. Then, we can obtain the IET distribution of user i with the active time interval $T_A^{(i)}$ as follows:

$$P_i(\tau; T_A^{(i)}) = \sum_{m=0}^{\infty} q_1(\tau; T_A^{(i)}, m) P_x(\tau - mT_I^{(i)}) \mathbf{II}(\tau; mT, mT + T_A^{(i)}) + \sum_{m=1}^{\infty} q_2(\tau; T_A^{(i)}, m-1) P_x(\tau - mT_I^{(i)}) \mathbf{II}(\tau; mT - T_A^{(i)}, mT), \quad (8)$$

where P_x represents either P_{ad} or P_{model} . \mathbf{II} is a rectangle function defined as

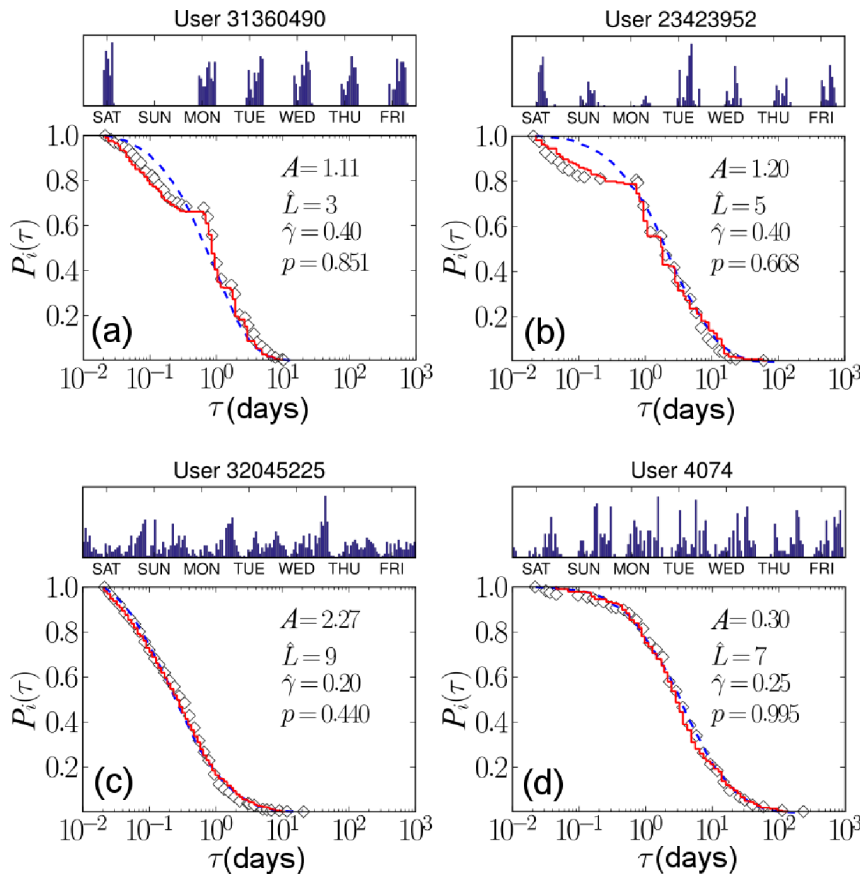


Figure 8. Comparison between empirically observed individual inter-event time probability distributions (open diamonds) and model predictions that are fit to the data. Model predictions are calculated by two methods by using $\mu(t)$ and $\lambda(t)$ (red solid lines), and by using time-averaged rates of $\langle \mu \rangle$ and $\langle \lambda \rangle$ (blue dotted lines). The histograms in the upper panel of each plot represent the relative ratio of blog posts written during a certain hour of the day. In cases with clear periodicity (a) and (b), red solid and blue dotted lines show apparent differences. Otherwise (c) and (d), they exhibit very similar patterns, and the periodicity assumption seems to be irrelevant to them. On the other hand, we only consider data points on scales larger than 30 min, because the resolution of $\mu(t)$ and $\lambda(t)$ is 1 h. doi:10.1371/journal.pone.0058292.g008

$$\mathbf{II}(x; a, b) \equiv \begin{cases} 1 & \text{for } a < x < b, \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

which represents the intervals defined in \mathcal{T}_1 and \mathcal{T}_2 . Next, we obtain the average $P_i(\tau; T_A^{(i)})$ over all users and obtain

$$\begin{aligned} P_{\text{theory}}(\tau) &= \frac{1}{N} \sum_i P_i(\tau; T_A^{(i)}) \\ &= \int_0^T r(T_A) P_i(\tau; T_A) dT_A, \end{aligned} \quad (10)$$

where $r(T_A)$ is the fraction of users whose active time interval is T_A . The distribution of $r(T_A)$ exhibits a peak at $T_A = 16$ h as shown in Fig. 3(b). By plugging the empirical distribution P_{ad} into P_x in Eq. (8), we successfully reproduce the oscillating pattern of the IET distribution $P_{\text{theory}}(\tau)$ in the inset of Fig. 1(a) and Fig. 5. When P_x is replaced by the theoretical formula $P_{\text{model}}(\tau') \sim \tau'^{-1.5}$ [30], the obtained result for $P_{\text{theory}}(\tau)$ is consistent with the simulated one, as shown in Fig. 6. It is noteworthy that the functional form of $r(T_A)$ does not play an important role in determining the oscillating behavior of the IET distribution. For

example, even for the flat distribution of $r(T_A)$, the oscillating pattern of $P_{\text{theory}}(\tau)$ can be obtained.

Modelling Activity Dependence

As discussed in the previous section, we have shown that the activities of individuals are heterogeneous and that their distribution follows a power law: $P_a(k) \sim k^{-\kappa}$ with $\kappa \approx 2.6 \pm 0.1$ as shown in the inset of Fig. 2. That is, a few people post many articles and many others post only a few articles in a given interval. Moreover, individuals have their own active time intervals. Thus, it would be interesting to study how such heterogeneities affect the IET distribution. We categorize users into groups according to their activities, and we measure the IET distributions of each group as shown in Fig. 2. It is interesting to notice that the IET distribution appears to be independent of activities in the short-time regime within one day, but it depends on activities in the long-time regime.

In the priority-based queueing model introduced in Ref. [30], packets arrive at a queue with the rate λ and are executed with the rate μ , where the rates λ and μ are regarded as constants, independent of time and individuals. Here, however, since the activity and the period of the active time interval are different, we assign user index i to the rates as λ_i and μ_i , and those quantities are assumed to depend on time. We consider $\mu_i(t)$ as proportional

to the frequency of blog postings at time t by user i . Next, we use the following relation between the execution rate $\mu_i(t)$ and the activity A_i ,

$$\frac{\gamma_i}{T_{\text{tot}}^{(i)}} \int_0^{T_{\text{tot}}^{(i)}} \mu_i(t) dt = A_i, \quad (11)$$

where γ_i is a proportionality constant. For the arriving rate $\lambda_i(t)$, since we do not have any information of when a new task is arriving, we assume $\lambda_i(t)$ to be the same as $\mu_i(t)$.

Based on this idea, for each user i , we perform numerical simulations as follows:

- i) We numerically generate both arrival and execution time sequences $\{t_k\} \in (0, T_{\text{tot}}^{(i)})$ through the Poisson process with the rates $\mu_i(t)$ and $\lambda_i(t)$ [31].
- ii) Subsequent these time sequences, we input a task into the queue when it is not full of L_i tasks, where the queue size L_i is determined at a later stage. Upon arrival, the task is given a priority $x \in [0, 1]$. At the same time, a task with the highest priority is executed and removed from the queue. The waiting time of the task is also recorded.
- iii) We repeat this procedure until N_i waiting times are obtained. N_i is regarded as the number of blog posts uploaded by user i .

In this model, the activity is determined to be $A_i = N_i / T_{\text{tot}}^{(i)}$, whereas the queue size L_i and the proportionality constant γ_i remain to be determined.

To determine L_i and γ_i , i.e., to generate a synthetic probability distribution function fit to the empirical data, we use the Kolmogorov-Smirnov (KS) statistical test [32]. We obtain a set of \hat{L}_i and $\hat{\gamma}_i$ for each user i by minimizing the KS statistic between the empirical data and simulated data. They are distributed as shown in Fig. 7. The closeness between the empirical data and the simulated data is tested (see Fig. 8): the obtained p value is shown in the legend. It is known that if the p -value is higher than a preassigned value ($p = 0.05$), then one can accept the null hypothesis that the probability distribution functions are identical. As we can see in the p -value histogram of Fig. 7(b), most cases show good agreement between synthetic and empirical data with high p values: The fraction of users is 23.2% for $p > 0.9$, and

86.3% for $p > 0.1$. Thus, it can be said that our theoretical result reasonably reproduces the empirical pattern.

Moreover, we simulate the queuing process by using the average rates of $\langle \mu \rangle$ and $\langle \lambda \rangle$ instead of the time-dependent form of $\mu(t)$ and $\lambda(t)$ for each user. In most cases, there is only a slight difference between the two simulated results with different types of parameters as shown in Fig. 8. However, there are apparent different cases for the two results; these occur when periodic time intervals appear in the activity of writing blog posts. In this case, the time-dependent forms $\mu(t)$ and $\lambda(t)$ are better for fitting to the empirical data.

Conclusions

In this work, we have studied the inter-event time statistics of human dynamics based on a large scale of on-line records of blog writings at a Korean portal site. We observed that the IET distributions of each user exhibit a universal pattern in the short-time regime, but they exhibit different decay patterns in the long-time regime, which depends on the activities of individual users. Moreover, we observed a clear periodic pattern with a period of one day, which reflects the circadian pattern of human behavior. We explained these patterns within the framework of the queuing model. First, we identified active and inactive time intervals of individual behaviors and then removed inactive time interval and constructed an *ad-hoc* time domain. Next, we applied the priority-based queuing model in the *ad-hoc* time domain by adjusting the arrival and execution rates of tasks to the empirical data. Following this, we returned to the real time domain and found our theoretical results to be in agreement with the empirical results including the positions of circadian peaks [6,13,18,27]. The microscopic studies performed in this paper enable us to understand these empirical results from a theoretical perspective.

Acknowledgments

We would like to thank Mr. Youn Sik Lee, Director of Data Information Center, for allowing for using the data after deleting user names, and Mr. Sukwon Kang for helpful discussion.

Author Contributions

Conceived and designed the experiments: BK JK. Analyzed the data: JK DL. Contributed reagents/materials/analysis tools: JK. Wrote the paper: BK JK.

References

1. Lazer D, Pentland A, Adamic L, Aral S, Barabási AL, et al. (2009) SOCIAL SCIENCE: computational social science. *Science* 323: 721–723.
2. Castellano C, Fortunato S, Loreto V (2009) Statistical physics of social dynamics. *Rev Mod Phys* 81: 591646.
3. Song C, Qu Z, Blumm N, Barabási AL (2010) Limits of predictability in human mobility. *Science* 327: 1018–1021.
4. Barabási AL (2005) The origin of bursts and heavy tails in human dynamics. *Nature* 435: 207–211.
5. Johansen A (2004) Probing human response times. *Physica A* 338: 286–291.
6. Eckmann JP (2004) Entropy of dialogues creates coherent structures in e-mail traffic. *Proc Natl Acad Sci U S A* 101: 14333–14337.
7. Vázquez A, Oliveira JG, Dezső Z, Goh KI, Kondor I, et al. (2006) Modeling bursts and heavy tails in human dynamics. *Phys Rev E* 73: 036127.
8. Vázquez A (2007) Impact of memory on human dynamics. *Physica A* 373: 747–752.
9. Malmgren RD, Stouffer DB, Motter AE, Amaral LAN (2008) A poissonian explanation for heavy tails in e-mail communication. *Proc Natl Acad Sci U S A* 105: 18153–18158.
10. Oliveira JG, Barabási AL (2005) Human dynamics: Darwin and einstein correspondence patterns. *Nature* 437: 1251.
11. Dezső Z, Almaas E, Lukács A, Rácz B, Szakadát I, et al. (2006) Dynamics of information access on the web. *Phys Rev E* 73: 066132.
12. Scalas E, Kaizoji T, Kirchner M, Huber J, Tedeschi A (2006) Waiting times between orders and trades in double-auction markets. *Physica A* 366: 463–471.
13. Zhou T, Kiet HAT, Kim BJ, Wang BH, Holme P (2008) Role of activity in human dynamics. *Europhys Lett* 82: 28002.
14. Johansen A, Sornette D (2000) Download relaxation dynamics on the WWW following newspaper publication of URL. *Physica A* 276: 338–345.
15. Johansen A (2001) Response time of internetauts. *Physica A* 296: 539–546.
16. Chessa AG, Murre JM (2004) A memory model for internet hits after media exposure. *Physica A* 333: 541–552.
17. Harder U, Paczusi M (2006) Correlated dynamics in human printing behavior. *Physica A* 361: 329–336.
18. Radicchi F (2009) Human activity in the web. *Phys Rev E* 80: 026118.
19. Barabási AL (2011) Bursts: the hidden patterns behind everything we do, from your e-mail to bloody crusades. New York: Plume.
20. Karsai M, Kaski K, Barabási AL, Kertész J (2012) Universal features of correlated bursty behaviour. *Sci Rep* 2: 397.
21. Wu Y, Zhou C, Xiao J, Kurths J, Schellnhuber HJ (2010) Evidence for a bimodal distribution in human communication. *Proc Natl Acad Sci U S A* 107: 18803–18808.
22. Jo HH, Pan RK, Kaski K (2012) Time-varying priority queuing models for human dynamics. *Phys Rev E* 85: 066102.
23. Hidalgo R CA (2006) Conditions for the emergence of scaling in the inter-event time of uncorrelated and seasonal systems. *Physica A* 369: 877–883.
24. Malmgren D, Stouffer D, Campanharo A, Nunes Amaral L (2009) On universality in human correspondence activity. *Science* 325: 1696–1700.
25. Vajna S, Tóth B, Kertész J (2012) Modelling power-law distributed interevent times. arXiv: 1211.1175.

26. Jo HH, Karsai M, Kertész J, Kaski K (2012) Circadian pattern and burstiness in mobile phone communication. *New J Phys* 14: 013055.
27. Holme P (2003) Network dynamics of ongoing social relationships. *Europhys Lett* 64: 427–433.
28. Goh KI, Barabási AL (2008) Burstiness and memory in complex systems. *Europhys Lett* 81: 48002.
29. Kivela M, Pan RK, Kaski K, Kertész J, Saramäki J, et al. (2011) Multiscale analysis of spreading in a large communication network. *J Stat Mech*: P03005.
30. Grinstein G, Linsker R (2006) Biased diffusion and universality in model queues. *Phys Rev Lett* 97: 130201.
31. Lewis PAW, Shedler GS (1979) Simulation of nonhomogeneous poisson processes by thinning. *Naval Research Logistics Quarterly* 26: 403–413.
32. Conover WJ (1999) *Practical nonparametric statistics*. New York: Wiley.