New activity pattern in human interactive dynamics

Marco Formentin$^{1,2}$, Alberto Lovison$^2$, Amos Maritan$^1$ and Giovanni Zanzotto$^3$

1 Dipartimento di Fisica e Astronomia “G. Galilei”, Università di Padova, INFN and CNISM, via Marzolo 8, I-35151 Padova, Italy
2 Dipartimento di Matematica, Università di Padova, via Trieste 63, I-35121 Padova, Italy
3 Dipartimento di Psicologia Generale, Università di Padova, via Venezia 12, I-35131 Padova, Italy
E-mail: marco.formentin@rub.de, maritan@pd.infn.it, lovison@math.unipd.it and zanzotto@dmsa.unipd.it

Received 17 April 2015
Accepted for publication 10 July 2015
Published 8 September 2015

Abstract. We investigate the response function of human agents as demonstrated by written correspondence, uncovering a new pattern for how the reactive dynamics of individuals is distributed across the set of each agent’s contacts. In long-term empirical data on email, we find that the set of response times considered separately for the messages to each different correspondent of a given writer, generate a family of heavy-tailed distributions, which have largely the same features for all agents, and whose characteristic times grow exponentially with the rank of each correspondent. We furthermore show that this new behavioral pattern emerges robustly by considering weighted moving averages of the priority-conditioned response-time probabilities generated by a basic prioritization model. Our findings clarify how the range of priorities in the inputs from one’s environment underpin and shape the dynamics of agents embedded in a net of reactive relations. These newly revealed activity patterns might be universal, being present in other general interactive environments, and constrain future models of communication and interaction networks, affecting their architecture and evolution.

Keywords: socio-economic networks, stochastic processes
1. Introduction

The interaction dynamics of animal and human agents is of interest in many theoretical and applied domains of science, from ecology to sociology to economics. Especially interesting is the clarification of the response function of humans, which has been investigated in a variety of contexts [1–8], a paradigmatic case being written correspondence, especially through email.

When each person is viewed as the node of a graph, written communication generates an evolving weighted and directed network whose large-scale structure and dynamics are still virtually unknown. Many interesting facts have emerged from the investigation of a number of email or paper mail databases collecting basic empirical information on written correspondence spanning from a few months [9, 10] to several decades of writers’ activity [6, 11, 12, 14, 15]. Intermittency was observed in the dynamics of correspondence writers, with bursts of events separated by long pauses, with non-Poissonian, heavy-tailed statistics in both the agents’ inter-event times (IETs) and response times (RTs), see the definitions below and figure 1. This also relates to the heavy-tailed temporal distributions observed in human and animal behavior and locomotion [1–8, 16–21].

A number of approaches have been used to characterize the features of the empirical time statistics of written communication [6, 13, 22–30], with debated indications of scaling behavior for the waiting times, and for their possible modeling through priority queueing. A new method for the analysis of these human reactive phenomena has been recently proposed [15], through which it was shown that, in particular, the mechanisms underpinning the response-time (RT) statistics of written correspondence are best understood, rather than in terms of standard time $t$, in terms of an agent’s
activity, i.e. by a ‘proper time’ parameter $s \in \mathbb{N}^+$ counting an agent’s outbound messages. This approach [15] disentangles from the overall time dynamics of writers the contributions due to their spontaneous pauses between messages, and helped uncover universal power-law features in the RT statistics on written correspondence when the $s$-clocking is utilized, rather than the usual $t$-clocking (figure 1).

Despite the insight given by such earlier enquiries, information of primary importance about the basic features of human interaction is still lacking, both in the data analysis and modeling. First and foremost, solely the total RT distribution $P(\sigma)$ of correspondence writers (with $\sigma = \Delta s$) has so far been considered in the literature, and it is presently unknown in which way the overall interaction of a given agent $A$ is distributed among all of her distinct targets. For instance, the response statistics of a writer $A$ separately with each one of her correspondents have so far never been obtained. This lack of empirical analysis parallels the fact that some main aspects of priority modeling have also remained unexplored in the above context.

In the present study we go beyond the analysis of the total RT distribution $P(\sigma)$ of correspondence writers done in [15], and investigate how the writers’ activity depends on the identity of their distinct contacts. As was done in [31], where the voice-call inter-event times of cell-phone users with their distinct contacts have been considered, such an analysis is the first, natural step in a more in-depth investigation of the activity patterns in interaction networks, and sheds light on their structure that cannot be obtained from the sole total distribution $P(\sigma)$ of the involved agents. We utilize for our inquiry the database presented in [15] and briefly described below, which is to our knowledge the most complete long-term email dataset currently available in the literature. Our findings reveal a new behavioral patterns as well as new modeling effects, evidencing hitherto unknown, possibly universal, aspects of human interaction. The analysis also illustrates how priority, which in the model is a hidden variable not immediately linked to real data, operates in the generation of a dynamics in accord with empirical observations.

**Figure 1.** Activity clock for a node in an interaction network. Representation of the node’s temporal activity, in this case written communication, along the axis of time $t$ for an agent $A$, typically measured in seconds for email data. Arrows pointing into the $t$ axis mark incoming messages from the indicated correspondents $C_1, C_2, \ldots$, of $A$. Arrows pointing out of the $t$ axis mark response messages from $A$ to the same correspondents. The intervals between the outgoing arrows define the inter-event times of $A$. The response times (RTs) of $A$ pertaining to each correspondent $C_i$ can either be clocked through time $t$ (in brown), or (in green) through the activity parameter $s$ which counts the number of outgoing messages from $A$. See the text for definitions.
2. Database and definitions. Proper time

Our written communication data concern the full server-recorded activity of all the email accounts belonging to a Department of a large EU university during two years (see also [15]). The collected data are in the form \{sender, receiver, timestamp\}, with senders and receivers conventionally numbered for identification, and timestamps given in seconds.

Referring to an agent \(A\), the response times (RTs) are defined as the time intervals \(\tau = \Delta t\) (in seconds) separating the arrival of any message \(M\) from any agent \(B\) to \(A\), and the first ensuing message \(M'\) going from \(A\) to \(B\), independently of the subject or contents of \(M\) or \(M'\). Following [15], to extricate from the time dynamics of \(A\) the contributions due to \(A\)’s pauses between messages (given by the individual inter-event time distribution \(P_\tau(\tau)\) of \(A\)), we introduce the activity parameter (proper time) \(N \in \mathbb{N}^+\) of \(A\), which clocks the number of outgoing messages from \(A\). The RTs of \(A\) are thus defined by counting the number \(\sigma = \Delta s\) of outgoing messages from \(A\) intervening between the same messages \(M\) and \(M'\) as above, as represented in figure 1. We remark that inter-event times (IETs) and response times (RTs) constitute altogether different notions of waiting times in the dynamics of individually interacting agents as in written communication. The IETs are the time intervals between two consecutive activity events of \(A\), i.e. the time intervals between two consecutive messages issued by \(A\), while the RTs clock how long a received message sits unreplied-to in an agent’s inbox. IETs and RTs have different statistical properties, for instance when measured through activity, IETs by definition have always \(\sigma = 1\), i.e. the corresponding s-clocked IET probability distribution \(P_\sigma(\sigma)\) is always concentrated at 1, while this is not the case for the RT distributions \(P(\sigma)\) [15] (see also figure 1).

Out of all the nominal monitored accounts in our dataset, we have analyzed the 300 most active agents, whose activity comprises from a minimum of 390 to \(\sim 10^4\) total RTs. A large percentage of these 300 writers have in the order of a few thousand RTs, distributed over a number \(n\) of distinct correspondents ranging from less than 100 to almost 1000. The supplementary figure shows explicitly the RT statistics pertaining to the 84 most active, and the 12 least active, among such 300 agents.

3. Empirical results: new behavioral pattern

It is known [15] that the RT distributions of writers when measured through standard time \(t\) (given in seconds, days,...) are non-universal, as they are observed to depend on the communication medium and on the specific agent considered. In particular, earlier works proposed different exponents, \(-1\) and \(-\frac{3}{2}\), respectively for written correspondence through emails and letters [6, 9, 11, 12, 22]. As mentioned earlier, to shed light on these aspects of written communication, in [15], we have proposed a method to disentangle from the overall response dynamics of a given agent \(A\) the contributions due to the inter-event pauses of \(A\), by clocking her RT statistics through activity (figure 1), i.e. with time steps given by \(A\)’s outgoing messages. The main result in [15] is that
New activity pattern in human interactive dynamics

doi:10.1088/1742-5468/2015/09/P09006

J. Stat. Mech. (2015) P09006

such new clocking unveils the same RT statistics to exist in written communication across all the examined media (letters, email, sms) and agents, with RTs displaying truncated power-law behavior with average exponents near $-\frac{3}{2}$. Furthermore, the RTs of $A$ clocked through standard time $t$ were shown to exhibit, as a consequence, a double-scaling behavior, a prediction verified in the empirical $t$-clocked RTs deduced from new long-term written-communication data. In [15] it was finally shown that these observations regarding the three examined communication media can be well described by means of a suitable queuing model for $A$’s behavior, in which the highest priority messages are replied first (see also section 4 below). However, in [15] we have only analyzed the total RT distribution $P^>(\sigma)$ of correspondence writers (figure 2), which pertains to all messages answered by an agent $A$, neglecting the identity of her correspondents.

For the purposes of this new investigation, we break down the activity of $A$ by considering the RTs of the writer $A$ separately for each one of her correspondents $C_i$ (see figure 1). We then rank $A$’s correspondents $C_1, C_2, \ldots$, through their growing characteristic ($s$-)times $\sigma^c_i = \frac{<\sigma^2>}{<\sigma>}$, the latter being computed from the set of $s$-clocked RTs that $A$ generates with each distinct $C_i$. Then, from the empirical data regarding $A$, we obtain the ordered family of $s$-clocked distributions

$$\mathcal{F}_A = \{P^>(\sigma), \quad i = 1, 2, \ldots, n\},$$

Figure 2. Empirical behavior of a human agent. Log–log plots of the empirical response-time (RT) inverse cumulative distributions $P^>(\sigma)$ in the family $\mathcal{F}_A$ of (1), for a typical email writer $A$ in the database. For clarity, only some of the $n$ curves $P^>(\sigma)$ are plotted, for geometrically growing values of the normalized correspondent rank $r = \frac{i}{n} \in [0,1]$ from left to right. The distribution $P^>(\sigma)$ is the lowest curve in red, while the total RT distribution $P^>(\sigma) = P^>(\sigma)$ is the upper-most curve in black. The inset, whose horizontal axes report the normalized rank $r$, shows the linear-log plot of the values $\sigma^c_i = \frac{<\sigma^2>}{<\sigma>}$ for the distributions $P^>(\sigma)$. This dictates how the curves $P^>(\sigma)$ progressively extend to the right over the $\sigma$-axis for growing $i$, reaching the black total cumulative distribution $P^>(\sigma) = P^>(\sigma)$ of $A$ for $i = n$. The characteristic times $\sigma^c_i$ in the inset grow roughly exponentially as $r \uparrow 1$, see also figures 3 and 4. More statistics are given in the supplementary figure.
where \( n \) is the total number of \( A \)'s correspondents.

For any given \( i \), in (1) we denote by \( P_i^>(\sigma) \) the (inverse) cumulative RT distribution associated to the activity of \( A \) with all her correspondents \( C_j \) with \( j < i \), i.e. \( P_i^>(\sigma) \) gives the probability of finding RTs longer than \( \sigma \) when considering the responses of \( A \) to any \( C_j \) with \( j < i \). In particular: \( P_i^>(\sigma) \) is the distribution of RTs of \( A \) with her correspondent \( C_i \); \( P_i^>(\sigma) \) is the distribution of the aggregated RTs of \( A \) with her correspondents \( C_1 \) and \( C_2 \); and so on, with \( P_i^>(\sigma) \equiv P^>(\sigma) \) giving the total cumulative RT distribution of agent \( A \) with all her correspondents. Figure 2 shows the empirical RT distributions \( P_i^>(\sigma) \) belonging to the family \( \mathcal{F}_A \) of a typical agent \( A \) in the database. A quantitative description of the overall features of a family \( \mathcal{F}_A \) is obtained by computing the characteristic time \( \sigma_i^c \) pertaining to each \( P_i^>(\sigma) \in \mathcal{F}_A \), i.e. the values \( \sigma_i^c = \left\langle \frac{\sigma^2}{\sigma^\sigma} \right\rangle \) computed for the RTs of \( A \) with all her correspondents \( C_j \) with \( j < i \). The \( \sigma_i^c \) indicate how the \( \sigma \)-clocked RTs associated to the correspondents of \( A \) up to the \( i \)-th rank, grow longer as a whole with \( i \). These values, which grow monotonically with \( i \), measure how rapidly the curves \( P_i^>(\sigma) \) progressively spread apart from each other on the plane \( (\sigma, P) \) for growing \( \sigma \), as they approach, for \( i \geq n \), the total cumulative distribution \( P^>(\sigma) \equiv P^>(\sigma) \) of \( A \) (this is the upper-most curve, obtained for \( i = n \), shown in black in each panel of figure 3).

We find in figure 2 that in the family \( \mathcal{F}_A \) the individual distributions \( P_i^>(\sigma) \) are heavy-tailed, but for \( i < n \) do not warrant any simple fitting form, nor collapse property (for instance, fits of the \( P_i^>(\sigma) \) with \( i < n \) through power laws or Weibull distributions are both rejected in about 80% of cases by Kolmogorov-Smirnov and Cramer-von Mises tests applied to \( \sim 10^2 \) randomly selected curves belonging to the agents in figures 2 and 3). Nonetheless, we observe in the inset of figure 2 that the characteristic times \( \sigma_i^c \) computed for each distribution \( P_i^>(\sigma) \), appear to grow roughly exponentially with rank \( i \) (the inset of figure 2 shows \( \sigma_i^c \) as a function of the normalized correspondent rank \( r = \frac{i}{n} \in [0, 1] \)).

As mentioned, the analysis in [15] has revealed a distinctive form of universality in written communication, showing that the activity-clocked total RT probability densities \( P(\sigma) \) of correspondence writers are exponentially truncated power laws, with empirical exponents near \(-\frac{3}{2}\) across all correspondence media. As we see, now an even stronger form of behavioral universality emerges from the analysis of the response patterns of humans described by the empirical distribution families \( \mathcal{F}_A = \{ P_i^>(\sigma) \} \) in (1) for email. This is evidenced by figure 3, which shows the distribution families \( \mathcal{F}_A \) relative to six typical active writers \( A \), evidencing how the \( \mathcal{F}_A \) pertaining to different agents exhibit largely the same features as those in figure 2, clearly pointing to a common pattern in these agents’ reactive dynamics. More statistics of this type are given in the supplementary figure, wherein the same pattern as the families \( \mathcal{F}_A \) highlighted in figures 2 and 3 is clearly present in the great majority of agents, not being recognizable only for the least active agents in the database.

To establish more precisely the statistical commonality indicating this new behavioral pattern of correspondence writers, we show in figure 4(a) the behavior with rank \( r \) of the normalized characteristic times \( \sigma^c \in [0, 1] \) pertaining to a random sample of writers in the database. The F-test [32] confirms that a fit of these curves with exponentials \( \exp(ar) \) is, for about 90% of the users, statistically more significant than a fit with
New activity pattern in human interactive dynamics

An ordinary least squares fit gives a mean value $a = 0.94 \pm 0.01$. We thus consider a fit for the dependence of the empirical characteristic times $\sigma^c_i$ on $r$ by means of the same normalized exponential $\exp(r)$ for all the 300 agents in the database. The corresponding frequency plot for the coefficient of determination $R^2$ derived from this is shown in figure 4(b). We see the histogram is

exp($ar + b$) or with linear functions. Furthermore, an ordinary least squares fit gives a mean value $a = 0.94 \pm 0.01$. We thus consider a fit for the dependence of the empirical characteristic times $\sigma^c_i$ on $r$ by means of the same normalized exponential $\exp(r)$ for all the 300 agents in the database. The corresponding frequency plot for the coefficient of determination $R^2$ derived from this is shown in figure 4(b). We see the histogram is

Figure 3. Same behavioral pattern exhibited by different human agents. Log–log plots of the empirical response-time (RT) cumulative distributions $P_i^c(\sigma)$ in (1), for six typical email writers in the database. The distributions $P_i^c(\sigma)$ are plotted following the same criteria as in figure 2. We see that for all agents the characteristic times $\sigma^c_i$ (shown in the insets) grow roughly exponentially with the normalized rank $r_1$, highlighting the commonality of the newly revealed behavioral pattern, see also figure 4. See the supplementary figure for more statistics.
New activity pattern in human interactive dynamics

strongly peaked near 1, confirming that the $\sigma^c$ grow roughly exponentially with rank for the great majority of agents and giving a quantitative confirmation of the existence of a common character in the activity pattern of email writers, as suggested by figures 2 and 3 and by the supplementary figure.

4. Model: prioritization

Previous work [6, 11, 15, 22–25, 33–35] has analyzed various aspects of priority queuing in relation to written correspondence. Here we show how a simple model based on prioritization as in [15] which accurately describes the power-law behavior of the total $s$-clocked RT distribution $P(\sigma) = P_n(\sigma)$ observed in correspondence writers, also accounts robustly for the RT patterns described by the family $F_A$ of empirical $s$-clocked distributions $P_i(\sigma)$ as in figures 2 and 3.

We describe the model in its simplest form, suitable for agents whose $P(\sigma)$ exponent is (close to) $-1.5$. For different individual exponents see [15]. We consider for an agent $A$ an initial list of $L$ tasks, with assigned priorities $y$ sampled from the uniform distribution on $[0,1]$. At each time step (which corresponds to a unit increment of the activity parameter $s$) the task with highest priority in the list is executed (a message replied), and $m > 1$ new tasks are added on average to the list, each one with priority $y$ sampled as above. It was analytically proven [33, 34, 36] that this queueing mechanism produces an RT probability density $P(\sigma)$ which for $s \to \infty$ decays as a power law with exponent $-\frac{3}{2}$. When finite values of $s$ are considered as in numerical simulations, a truncated $(-\frac{3}{2})$-power-law $P(\sigma)$ is obtained for the RTs. Interestingly, we compute that also such finite-size effect obtained

doi:10.1088/1742-5468/2015/09/P09006
New activity pattern in human interactive dynamics

Figure 5. Behavior of the priority model. (a) Histogram giving the priority $y$ of the replied-to messages in the model (for $3 \cdot 10^4$ activity cycles, with $\rho = m^{-1} \approx 0.64$), showing a threshold effect at the value $y \approx 1 - \rho$, marked by the vertical red line. (b) Linear-log plot for the corresponding characteristic times $\tilde{\sigma}_y^c$ generated by the model for $y \in [1 - \rho, 1]$.

in the model agrees with the cut-off observed in the scaling statistics from the empirical data, because for $s \sim 10^4$ activity cycles, both the model and data give a characteristic time $\sigma^c \sim 10^3$ for the total cumulative RT distribution $P^\sigma()$.

Now, for the purpose of relating the model to the empirical features highlighted above regarding $\mathcal{F}_A$, it is natural to consider the priority-conditioned distributions that are generated by priority queueing. Specifically, let us consider the family of distributions $\mathcal{F} = \{P^\sigma_y()\}$, where $P^\sigma_y()$ is the probability of observing an RT larger than $\sigma$ given that the priority of the replied-to messages has values greater than $y$. When plotted, these $y$-conditioned distributions in $\mathcal{F}$ exhibit heavy tails, and, for decreasing $y$, fan out in the plane $(\sigma, P)$ in a way that is reminiscent of the empirical curves $P^\sigma_i()$ in figure 3 for growing $i \uparrow n$. However, the distribution family $\mathcal{F}$ does not provide a good description for the families $\mathcal{F}_A$ obtained from the empirical data. This is because prioritization induces, in the distribution of $y$-values for the generated RTs, a threshold effect, whose existence can be proven in our context by adapting the arguments in [37]. Accordingly, only the entering messages whose priority is $y \gtrsim 1 - \rho$ are replied to in the model (see figure 5(a)). The corresponding characteristic times $\sigma^c_y = \frac{\langle \sigma^2 \rangle}{\langle \sigma \rangle}$ for the conditional $P^\sigma_y()$, i.e. computed by considering the aggregated RTs given by the model for all priorities greater than $y$, grow supra-exponentially as $y \downarrow 1 - \rho$ (figure 5(b)). Figure 7 below also shows the same effect regarding the normalized $\tilde{\sigma}^c \in [0, 1]$, whose supra-exponential theoretical values do not satisfactorily match the behavior of their normalized empirical counterparts $\tilde{\sigma}^c_i$ for growing rank $i$, see figure 4(a).

5. Gaussian averaging

To actually connect the model to the empirical families $\mathcal{F}_A$, we need to better understand the operation of the priority $y$ in generating the RTs in the queueing process.
As mentioned above, prioritization generates a threshold effect, with the characteristic times $\tilde{\sigma}_y$ exhibiting supra-exponential growth as $\rho \downarrow 1 - \rho$, and likewise for the $\sigma_y^c$. The failure of the family $\mathcal{F}$ to represent correctly the features of $\mathcal{F}_\text{A}$ is not surprising because, while certainly present, the correlation between the identity of correspondents and their messages’ priority cannot be too strict, as each correspondent $C_i$ of $A$ should be associated, rather than to a single value of the priority $y$, to some individual distribution of $y$-values. To describe this, we consider, for a generic agent with $n$ correspondents, a suitable family of kernels $\kappa(y; \tilde{g}_i, d_i)$ which, for each $i = 1, 2, \ldots, n$, describe in the model the distribution (with suitable standard deviation $d_i$, and with mean $\tilde{g}_i$ which is decreasing with growing $i$) of priorities for the messages from the $i$-th correspondent $C_i$ which $A$ has replied to. Given such $\kappa$’s, we compute the distributions

$$P^\gamma_y(\sigma) = \int P^\gamma_y(\sigma) \kappa(y; \tilde{g}_i, d_i) \, dy,$$

where $P^\gamma_y(\sigma)$ is the probability of observing, in the model, an RT in a small neighborhood of $\sigma$ given that the priority of the replied-to messages has values in a small neighborhood of $y$. Then, the behavior exhibited by the empirical families $\mathcal{F}_\text{A}$ in figure 3 should be better captured by a new family of distributions

$$\mathcal{F} = \{P^\gamma_y(\sigma), \ i = 1, 2, \ldots, n\},$$

Figure 6. Behavior of the priority model with averaging. Log–log plots of the response-time (RT) inverse cumulative distributions $P^\gamma_y(\sigma)$ belonging to family $\mathcal{F}$ in (3), obtained through moving Gaussian averages of the $y$-conditioned probabilities computed from the model, for $3 \cdot 10^4$ cycles, with $\rho \simeq 0.64$, $d = 0.2$, and $n = 100$. Twenty curves are plotted, for geometrically growing values of the normalized rank $r = \frac{i}{n} \in [0, 1]$ from left to right (see text). The total RT distribution is the upper-most curve, shown in black, and $P^\gamma_y(\sigma)$ is the lowest curve, in red. The inset shows the linear-log plot for the values $\sigma_y^c$ which increase roughly exponentially with rank (the horizontal axis in the inset reports the normalized rank $r$). See also figure 7. We observe the very good agreement with the behavior of the empirical families $\mathcal{F}_\text{A}$ in (1) which are shown in figures 2 and 3 and in the supplementary figure.
where the $P_y^\sigma(\sigma)$ are the (cumulative) distributions associated to the $P_y(\sigma)$ in (2), i.e. they give the probability of observing RTs greater than $\sigma$ when considering in the model the aggregated replied-to messages pertaining to all the $\kappa$-samples with average priorities $y_J$ greater than $y_i$.

The simplest hypothesis in this context considers, for $i = 1, \ldots, n$, Gaussian kernels

$$\kappa(y; \bar{y}_i, d) = \frac{1}{d \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{y - \bar{y}_i}{d}\right)^2\right) \text{ in (2)},$$

with a common standard deviation $d$, and values of the mean $\bar{y}_i$ which are homogeneously distributed between $1 - \rho$ and 1 (these bounds derive from the threshold effect in the model, see figure 5). The distribution of $y$-values over the RTs in the model remains largely homogeneous after sampling by means of these kernels, as in figure 5(a). For $d \downarrow 0$ we recover the model with no averaging.

Figure 6 shows the family of distributions $F$ in (3) computed through the above Gaussian kernels, for $d = 0.2$ and $n = 100$. We see that the curves in the averaged distribution family $F$ do reproduce qualitatively very well the behavior of the empirical curves in the families $F_A$ in figure 3 and in the supplementary figure. In particular, the inset in figure 6 shows that the values $\sigma_i^c = \langle \sigma_i^2 \rangle / \langle \sigma_i \rangle$, computed by aggregating the RTs given by the model referring to all the Gaussian samples with average priorities greater than $\bar{y}_i$, grow roughly exponentially with rank $i \uparrow n$ (i.e. as $\bar{y}_i \downarrow 1 - \rho$), as was the case for the empirical $\sigma_i^c$ shown in the insets of figures 2 and 3. This was checked for the computed normalized characteristic times $\tilde{\sigma}_i^c$ through analogous goodness-of-fit tests as performed for their empirical counterparts in figure 4. The analysis in figure 7 corroborates the accord of $F$ with the empirical families $F_A$, as we see for a range of values $d$ of the order $10^{-1}$ the computed characteristic times of $F$ (i.e. the colored curves) in

Figure 7. Behavior of the characteristic times in the model. (a) Linear-log plot of the normalized characteristic times $\tilde{\sigma}_i^c \in [0,1]$ (colored curves) computed for values of the standard deviation $d$ of the Gaussian kernels given by the red dots in panel figure 5(b). The dotted black curve gives the characteristic times for the case $d \downarrow 0$, i.e. for the model with no averaging. (b) Relation between the $d$ value in the kernels and the $R^2$ values obtained by fitting the curves in panel figure 5(a) with the normalized exponential function $\exp(r)$. These high $R^2$ indicate the characteristic times obtained through Gaussian kernels with $d \sim 10^{-1}$ are well described by the exponential fit, unlike what happens in the model with no averaging for $d \downarrow 0$, as shown by the dotted curve in panel figure 5(a).
figure 7(a) behave similarly to the empirical ones in figure 4(a). This is further confirmed by figure 4(b) showing that high \( R^2 \) values, above 0.9, are obtained for the exponential fit \( \exp(\tau) \) of the computed characteristic times for \( \mathcal{F} \) for \( d \sim 10^{-1} \). This indicates that the characteristic times produced by the averaged model robustly reproduce the roughly exponential growth exhibited by the set of empirical characteristic times in figures 3–4(b). This effect does not need fine tuning in the model, and is rather rooted in the prioritization process and the weighted averaging used to account for the priority distribution of the messages from each correspondent.

6. Discussion

We have achieved in this study a two-fold result. (a) Firstly, we have uncovered a new activity pattern in the interactive dynamics of correspondence writers, highlighted through the examination of long-term empirical data on written correspondence via email. We find that agents all distribute in the same way their interactions separately with each one of their distinct contacts, generating families of heavy-tailed RT distributions which have largely the same features across writers, with characteristic times which universally exhibit roughly exponential growth with correspondent rank. This analysis considerably extends the scrutiny of the sole total distribution of correspondence writers, on which the literature has focused so far. (b) We have furthermore shown that these previously undetected behavioral structures emerges robustly by considering Gaussian moving averages on the priority-conditioned RT probabilities derived from a basic priority model.

Our findings clarify how priority-queueing contributes to generate the observed activity statistics of human response, and suggest that the evidenced shared pattern, with its reported variability, may result from fundamental constraints imposed by prioritization and by averaging mechanisms on the outcome of any complex underlying individual choice processes. The effects revealed here should affect both the architecture and the evolution of communication and interaction (social) networks, imposing explicit constraints on their future exploration and modeling. They may also contribute to better estimate the possible value of such networks in relation to size \([38, 39]\), which is an important question in computer science, business management, and sociology. Natural extensions of the present study regard the possibility of identifying, within the individual variations of the empirical \( \sigma \)-curves, the existence of core communities \([40, 41]\) within each agent’s ensemble of correspondents.

Another point of interest is the adoption of less schematic averaging kernels than used in Section 5. This would not affect the basic behavioral commonality highlighted here, but may help capture other effects occurring in written communication, and possibly in general reactive dynamics. Indeed, we expect the stylized facts \([42]\) and new activity patterns presently uncovered for email correspondence may occur universally in environments involving directed interactions, and could be successfully investigated through our methods. This should promote our understanding of the dynamics of reciprocal activity in diverse agent-driven domains, as in economics or sociology. For instance, our approach may enhance queueing-based models \([37, 43, 44]\) as valuable tools in finance for investigating order-book dynamics, or for studying behavior in
social media [7, 8, 45], where prioritization effects and scaling of total RTs have been investigated [7, 37]. Also theories for preference formation and extraction, for competing-opinion dynamics, and for information spreading [46–49], may benefit from the knowledge and analysis of reciprocal-action patterns such as we have obtained here on emailing, because decision making at the personal and collective levels, or the shift of sentiments and preferences, are largely based on how individuals communicate and interact with each other. In general, the present analysis should help inform future empirical and theoretical work on the interplay among distinct agents of any kind, animate or inanimate, embedded in networks of reactive relations.

Acknowledgments

We thank Dr M Gravino for providing to us the email data analyzed in this study, and an anonymous referee for suggestions. MF also thanks Prof J M Swart for many conversations. AM acknowledges the Cariparo Foundation (Padova) for financial support. MF acknowledges financial support of GAČR grant P201/12/2613.

Author contributions statement

Authors with initials M F, A L, A M and G Z equally contributed to the manuscript.

Additional information

Competing financial interests: the authors declare no conflict of interest.

References

[1] Henderson T and Bhatti S 2001 Modelling user behaviour in networked games MULTIMEDIA ’01: Proc. of the 9th ACM Int. Conf. on Multimedia pp 90–4
[2] Wang Q and Guo J-L 2010 Human dynamics scaling characteristics for aerial inbound logistics operation Physica A: Stat. Mech. Appl. 389 2127
[3] Dezso Z et al 2006 Dynamics of information access on the web Phys. Rev. E 73 066132
[4] Gonçalves B and Ramasco J J 2008 Human dynamics revealed through Web analytics Phys. Rev. E 78 026123
[5] Gao L, Guo J-L, Fan C and Liu X-J 2010 Individual and group dynamics in purchasing activity arXiv.org 1010
[6] Vazquez A et al 2006 Modeling bursts and heavy tails in human dynamics Phys. Rev. E 73 036127
[7] Mitrovic M and Tadic B 2010 Emergent communities in blog space Eur. Phys. J. B 73 293–301
[8] Garcia D, Weber I and Garimella V 2014 Fiction: the bechdel test of social media Proc. of the 8th Int. AAAI Conf. on Weblogs and Social Media
[9] Eckmann J, Moses E and Sergi D 2004 Entropy of dialogues creates coherent structures in e-mail traffic Proc. Natl Acad. Sci. USA 101 14333
[10] Wu Y, Zhou C, Xiao J, Kurth J and Schellnhuber H J 2010 Evidence for a bimodal distribution in human communication Proc. Natl Acad. Sci. USA 107 18803
[11] Oliveira J G and Barabasi A-L 2005 Human dynamics: darwin and einstein correspondence patterns Nature 437 1251
[12] Qu S, Wang Q and Wang L 2011 The statistical research of human dynamics on correspondence Proc. of the 2011 Int. Conf. on Modelling, Identification and Control pp 282–4
[13] Malmgren R D, Stouffer D B, Motter A E and Amaral L A N 2008 A Poissonian explanation for heavy tails in e-mail communication Proc. Natl Acad. Sci. USA 105 18153–8
[14] Malmgren R D, Stouffer D B, Campanharo A S and Amaral L A 2009 On universality in human correspondence activity Science 325 1696

doi:10.1088/1742-5468/2015/09/P09006
New activity pattern in human interactive dynamics

[15] Formentin M, Lovison A, Maritan A and Zanzotto G 2014 Hidden scaling patterns and universality in written communication Phys. Rev. E 90 012817
[16] Hanai K, Ozaki M, Yamashita D and Nakatomi Y 2006 Scale free dynamics involved in the ant locomotion Proc. of the Int. Conf. on Cellular and Molecular Biology, Biophysics and Bioengineering
[17] Crane R, Schweitzer F and Sornette D 2010 Power law signature of media exposure in human response waiting time distributions Phys. Rev. E 81 056101
[18] Proekt A, Banavar J R, Maritan A and Pfaff D W 2012 Scale invariance in the dynamics of spontaneous behavior Proc. Natl Acad. Sci. USA 109 10564
[19] Jung K, Jiang H, Krallik J D and Jeong J 2014 Bursts and heavy tails in temporal and sequential dynamics of foraging decisions PLoS Comput. Biol. 10 e1003759
[20] Mryglod O, Fuchs B, Szelle M, Holovatch Y and Thurner S 2015 Interevent time distributions of human multi-level activity in a virtual world Physica A 419 681
[21] Chialvo D R, Torrado A M, Gudowska-Nowak E, Ochab J K, Montoya P, Nowak M A and Tagliazucchi E 2015 How we move is universal: scaling in the average shape of human activity arXiv:1506.06717
[22] Barabasi A 2005 The origin of burst and heavy tails in human dynamics Nature 435 207
[23] Grinstein G and Linsker R 2008 Power-law and exponential tails in a stochastic priority-based model queue Phys. Rev. E 77 012101
[24] Masuda N, Kim J S and Kuhl B 2009 Priority queues with bursty arrivals of incoming tasks Phys. Rev. E 79 036106
[25] Walraevens J, Demoor T, Maertens T and Brueel H 2012 Stochastic queueing-theory approach to human dynamics Phys. Rev. E 85 021139
[26] Blanchard P and Hougler M-O 2007 Modeling human activity in the spirit of Barabasi’s queueing systems Phys. Rev. E 75 026102
[27] Min B, Goh K-I and Kim I-M 2009 Waiting time dynamics of priority-queue networks Phys. Rev. E 79 056110
[28] Cho W-K, Min B and Goh K-I 2010 Generalized priority-queue network dynamics: impact of team and hierarchy Phys. Rev. E 81 066109
[29] Kim K and Chae K C 2010 Discrete-time queues with discretionary priorities Eur. J. Operat. Res. 200 473
[30] Jo H-H, Pan R K and Kaski K 2012 Time-varying priority queuing models for human dynamics Phys. Rev. E 85 066101
[31] Jiang Z-Q et al 2013 Calling patterns in human communication dynamics Proc. Natl Acad. Sci. USA 110 1600
[32] Everitt B S 2002 Cambridge Dictionary of Statistics (Cambridge: Cambridge University Press)
[33] Gabrielli A and Caldarelli G 2009 Invasion percolation and the time scaling behavior of a queueing model of human dynamics J. Stat. Mech. P02046
[34] Gabrielli A and Caldarelli G 2009 Invasion percolation on a tree and queueing models Phys. Rev. E 79 041133
[35] Grinstein G and Linsker R 2006 Biased diffusion and universality in model queues Phys. Rev. Lett. 97 130201
[36] Abate J and Whitt W 1996 Asymptotics for M/G/1 low-priority waiting-time tail probabilities Queueing Syst. 25 173
[37] Luckock H 2003 A steady-state model of the continuous double auction Quant. Finance 3 385–404
[38] Briscoe B and Odyzyko A 2006 Metcalfe’s law is wrong IEEE Spectr. 43 34
[39] Tongia R and Wilson E J 2011 The flip side of Metcalfe’s law is wrong IEEE Spectr. 48 34
[40] Dunbar R I 1992 Neocortex size as a constraint on group size in primates J. Hum. Evol. 22 469–93
[41] Gonçalves B, Perra N and Vespignani A 2011 Modeling users’ activity on Twitter networks: validation of Dunbar’s number PLoS ONE 6 e22656
[42] Buchanan M 2012 It’s a (stylized) fact! Nat. Phys. 8 3
[43] Gode D K and Sunder S 1993 Allocative efficiency of markets with zero-intelligence traders: market as a partial substitute for individual rationality J. Polit. Econ. 101 119
[44] Farmer J D, Patelli P and Zovko I I 2005 The predictive power of zero intelligence in financial markets Proc. Natl Acad. Sci. USA 102 2254–9
[45] Tadic B and Suvakov M 2013 J. Stat. Mech. P10014
[46] Hodas N O and Lerman K 2014 The simple rules of social contagion Sci. Rep. 4 4343
[47] Burghardt K, Rand W and Girvan M 2014 Competing opinions and stubborness: connecting models to data arXiv:1411.7415
[48] Xie J et al 2012 Evolution of opinions on social networks in the presence of competing committed groups PLoS ONE 7 e033215
[49] Das S, Gollapudi S and Munagala K 2012 Modeling opinion dynamics in social networks WSDM ’14 Proc. of the 7th ACM Int. Conf. on Web Search and Data Mining pp 403–12

doi:10.1088/1742-5468/2015/09/P09006