

A bi-logistic growth model for conference registration with an early bird deadline

Research Article

Trevor Fenner, Mark Levene*, George Loizou

Department of Computer Science and Information Systems, Birkbeck, University of London
London WC1E 7HX, U.K.

Received 24 May 2013; accepted 26 June 2013

Abstract: The recent interest in human dynamics has led researchers to investigate the processes that explain human behaviour within different contexts. Here we are concerned in modelling the human response to a deadline, and in particular we look at the process of conference registration with an early bird deadline. We provide empirical evidence from a six-year conference registration data set that the bi-logistic growth function, with the interpretation as registration with an early bird deadline, can be viewed as a social mechanism.

PACS (2008): 89.75.-k

Keywords: data analysis • human dynamics • social mechanism • bi-logistic growth function
© Versita sp. z o.o.

1. Introduction

Recent interest in complex systems, such as social networks, the world-wide-web, email networks and mobile phone networks [1], has led researchers to investigate the processes that may explain the dynamics of human behaviour within these networks. For example, Barabási [2] suggested that the bursty nature of human behaviour, for example when measuring the inter-event response time of email communication, is a result of a decision-based queueing process. In particular, humans tend to prioritise actions, for example when deciding which email to respond to, and therefore a priority queue model was proposed in [2], leading to a heavy-tailed, power-law distribution of inter-event times.

Of particular interest to us has been the formulation of a *generative model*, in the form of a stochastic process, by which a complex system evolves and gives rise to a power-law or other distribution [3, 4]. This builds on the early work of Simon [5]. In the context of human dynamics, the priority queue model [2] mentioned above is a generative model characterised by a heavy-tailed distribution.

In the bigger picture, one can view the goal of such research as being similar to that of *social mechanisms* [6, 7], i.e. looking into the processes, or mechanisms, that can explain observed social phenomena. More precisely, a social mechanism provides an action-based explanation of a social phenomenon. It consists of entities with their properties and the activities they engage in that bring about an observed outcome that is to be explained. According to Schelling [8], a social mechanism can be viewed as an interpretation, in terms of social behaviour (in our case human dynamics), of a mathematical model that reproduces

*E-mail: mlevene@dcs.bbk.ac.uk

the phenomenon that we are studying. The model on its own is not a social mechanism, what it needs in addition is the interpretation or application of the model.

Schelling [8] gave an example of a social mechanism based on the logistic growth model [9, 10], which we shall describe in Section 2. He suggests one such interpretation of the model in the form of the growth of the sales of a book. Here the sales follow an initially exponential growth curve, followed by a decline in sales until they more or less cease. Of course, empirical evidence is needed to validate or invalidate the proposed mathematical model underlying the social mechanism under consideration.

Here we study the problem of modelling human response to a deadline, which has been investigated in the context of conference registration [11, 12] or submission [13]. It is not surprising that conference registration or submission may follow a logistic growth pattern, and indeed Flandrin [13] used a variation of the logistic function to model the process of submissions to conferences. But there is a twist in the tale, which has not yet been investigated, when considering the effect of an *early bird registration deadline* (EBD), which encourages early registration with a discounted fee. Our starting point in the case of an EBD is the logistic model, which gives rise to the ubiquitous S-shaped growth function. However, the effect of an EBD is that it induces a second growth phase just before the early bird deadline, and thus suggests a model based on a bi-logistic growth function [14]. We provide empirical evidence, using a data set spanning six years of an annual conference, that an interpretation of the bi-logistic growth model for conference registration with an EBD can be viewed as a social mechanism for the underlying conference registration process.

The rest of the paper is organised as follows. In Section 2 we introduce the logistic and bi-logistic growth functions, and in Section 3 we present empirical evidence for modelling conference registration with an EBD as a bi-logistic growth function. We give our concluding remarks in Section 4.

2. A model based on logistic growth functions

The logistic growth model describes population growth in terms of an S-shaped sigmoidal curve [9, 10]. The growth rate begins exponentially, but at the inflection point the growth starts slowing down and decreases towards zero as the population size approaches its upper limit. Logistic growth was initially studied in the context of population growth, but has since been applied in many situations; for

example, it is regularly employed in technological forecasting [15].

Logistic growth can be formulated as a first-order differential equation as follows. Let $N(t)$ be the population at time t , κ be the limiting size (or *capacity*) of the population, and α be a constant positive growth rate. Then, the *logistic equation* is the differential equation [10]

$$\frac{dN(t)}{dt} = \alpha N(t) \left(1 - \frac{N(t)}{\kappa} \right), \quad (1)$$

where $N(0) > 0$ is the population size at time $t = 0$.

The solution to the differential equation in (1) is given by

$$N(t) = \frac{\kappa}{1 + \exp(-\alpha(t - \beta))}. \quad (2)$$

We note that the inflection point is at $t = \beta$, the time at which the population reaches $\kappa/2$. The parameter β is the mean of the logistic distribution given by (2) with $\kappa = 1$, and α is inversely proportional to its standard deviation [16].

In some cases, there may be multiple growth phases, each described by the logistic growth function (2). When there are two phases, we have *bi-logistic growth* [14], which is described by the sum of two separate logistic growth functions as follows:

$$N(t) = \frac{\kappa_1}{1 + \exp(-\alpha_1(t - \beta_1))} + \frac{\kappa_2}{1 + \exp(-\alpha_2(t - \beta_2))}, \quad (3)$$

where α_i , β_i and κ_i are the respective parameters of the two logistic phases, for $i = 1, 2$.

The bi-logistic growth model seems to appear in nature, as was shown by Sheehy et al. [17], who described how rice crops can be modelled by two phases of logistic growth: vegetative growth, followed by reproductive growth. The two phases can be readily modelled as superimposed-converging bi-logistic curves. (See Section 3 below for some explanation of the taxonomy of bi-logistic curves [14].) The model can also be used in a social context, as was demonstrated by Ausloos [18], who analysed the evolution of the number of temples built by the Antionist community in Belgium during two non-overlapping periods. Although Ausloos modelled the data in terms of two separate logistic phases, it could readily be modelled in terms of a sequential-diverging bi-logistic curve.

3. Application to conference registration

Modelling the dynamic behaviour when there is a deadline for a response can be applied in the context of conference registration and submission. The proposal in [11, 12] is based on the reasonable assumption that the probability of registration is inversely related to the time remaining before the deadline. Flandrin [13] remarked that this assumption, made by Alfi et al. [11, 12], is somewhat arbitrary, and proposed a new model for submissions based on a variant of the logistic growth function, which proved to be effective on the data set he used for his analysis.

Here we take into account a feature of conference registration not considered in [11, 12] (nor in [13], who only considers submissions), which is the effect of an *early bird registration deadline* (EBD) on the registration process. As in [13], our starting point is the logistic model. However, an EBD induces a second growth phase just before the early bird deadline, and thus suggests a bi-logistic growth function, with the inflection point β_2 of the second phase occurring just before the EBD, as our model for conference registration with an EBD. A simple explanation of the model as the sum of two logistic functions is that there are two types of registrants, “poor” for whom registration before the EBD is critical, and “rich” registrants for whom the EBD is irrelevant.

To test the bi-logistic hypothesis, we made use of weekly registration data for the USENIX LISA (Large Installation System Administration) conference from 2007 to 2012, obtained from www.usenix.org/sites/default/files/lisa12_weeklyregcomp_final.pdf. In this conference, an early bird registration discount is offered, with a deadline of three weeks prior to the start of the conference.

Table 1 shows the percentage of attendees registering *Before* and *During* the week prior to the EBD, the total percentage (*End*) registered by the EBD (i.e. the sum of *Before* and *During*), and κ , the total number of conference participants registered before the first day of the conference. It can be seen that in all years, apart from 2012, the percentage registering before the week prior to the EBD was less than 50% and that about a third of the registrations occurred during the week prior to the EBD. We observe that the total percentage of registrations before the EBD was relatively consistent, being in the range 73–85%, with a standard deviation of 4.6%. Table 2 shows the number of weeks β until registration reached half of the total number as measured before the first day of the conference, the length of the registration period and the number of weeks prior to the EBD. We note that the lowest

number of registrations was in 2009, which had both the highest percentage of registrations during the week prior to the EBD and the lowest percentage of registrations before that.

Table 1. Percentage of attendees registering before and during the week prior to the EBD

Year	Before	During	End	κ
2007	48.27	33.80	82.07	781.0
2008	46.35	34.20	80.55	766.0
2009	33.33	41.67	75.00	612.0
2010	37.47	39.37	76.84	734.0
2011	45.09	28.20	73.29	805.0
2012	53.56	31.77	85.33	702.0
Average	44.01	34.84	78.85	733.3

Table 2. Number of weeks β until registration reached half capacity

Year	β	Weeks	EBD
2007	6.05	10	7
2008	7.11	11	8
2009	7.40	11	8
2010	7.32	11	8
2011	7.17	11	8
2012	6.66	11	8
Average	6.95	10.83	7.83

Table 3 shows the fitted parameters to the bi-logistic curve (3) for the LISA conference data from 2007 to 2012, where $\hat{\kappa} = \kappa_1 + \kappa_2$ is the total estimated capacity, and R^2 is the coefficient of determination [19], which on average was about 0.99. The curve fitting was done in Matlab using non-linear least squares regression; the fitted curves are shown in Figure 1.

We can characterise the fitted curves based on the taxonomy of bi-logistic curves proposed in [14]. Firstly, a bi-logistic curve is said to be *sequential* if the second phase only starts after the first one ends, otherwise it is said to be *superimposed*. Secondly, the curve is said to be *converging* if $\alpha_1 < \alpha_2$, otherwise it is said to be *diverging*. We make the reasonable assumptions that the first phase starts at the beginning of the registration period and the second one ends at the end of the registration period. It can then be readily deduced from Table 3 that for the LISA conferences, apart from 2011, the curves could be categorised as superimposed-converging, while for 2011 as sequential-diverging (we discuss the statistics for 2011 in more detail below).

For all years (with the exception of 2010), the mid-point β_1 of the first logistic growth phase occurred well before the EBD, and the mid-point β_2 of the second phase occurred around half a week before the EBD, i.e. about three and a half weeks before the start of the conference. By symmetry, it follows that the second phase started around the end of week four, except in 2007 when registration started a week later than in subsequent years. We note that $\alpha_1 < \alpha_2$ in all years except 2011. The statistics for that year also show the highest value for α_1 and the lowest for α_2 , indicating particularly fast growth in the first phase. It is also worth noting that in 2011, although the percentage registering during the week prior to the EBD was the smallest, yet the final number of registrations was still the highest (see Table 1).

The worst year during the period in terms of registrations was 2009, when there were at least 90 registrations fewer than in any other year. We note that in this year the percentage registering before the week prior to the EBD was the lowest (see Table 1), and κ_1 was also the lowest (see Table 3). So, for some reason, in 2009 people were slower to register than in other years. Moreover, despite the percentage registering during the week prior to the EBD being the largest (see Table 1), this did not make up for the loss of registration during the first phase. The years 2010 and 2012 stand out in that the growth rate of the second phase is about six times that of the first phase, having the smallest values for α_1 and the highest for α_2 . We note that the fitting procedure does not take into account the fact that the registration process finished after 11 weeks (10 weeks in 2007); this essentially truncates the logistic curve even though this might be showing an upward trend at the end of the registration period. However, Table 3 does indicate that $\hat{\kappa}$, the overall capacity of the fitted bi-logistic growth function, is consistent with κ , the total number of registrations, as shown in Table 1.

Table 3. Bi-logistic fitted parameters for LISA registration data

Year	α_1	β_1	κ_1	α_2	β_2	κ_2	R^2	$\hat{\kappa}$
2007	0.99	2.47	221.1	1.73	6.44	552.6	0.9969	773.7
2008	0.85	2.68	182.9	1.56	7.44	578.0	0.9961	760.9
2009	0.85	3.58	107.6	1.77	7.64	496.2	0.9939	603.8
2010	0.57	6.31	410.0	3.77	7.59	322.2	0.9889	732.2
2011	1.79	1.44	121.8	0.82	7.53	686.9	0.9890	808.7
2012	0.52	4.07	400.0	3.05	7.54	302.6	0.9960	702.6
Average	0.93	3.43	240.6	2.12	7.36	489.8	0.9935	730.3

Looking at the average values of the fitted parameters of the bi-logistic curve (the last row in Table 3), we see that, on average, the first phase finishes in the middle

of the seventh week, while, as noted earlier, the second phase starts around the end of the fourth week; so the two phases overlap for nearly three weeks. From the data and the high R^2 values, it is clearly a plausible hypothesis that the EBD induces bi-logistic growth.

4. Concluding remarks

We have provided empirical evidence that conference registration with an EBD can be modelled with a bi-logistic growth function in order to capture the two growth phases in registration. Thus the bi-logistic growth function, with the interpretation as conference registration with an EBD, can be viewed as a social mechanism. The process has two growth phases one, whose inflection point generally occurs well before the EBD, and a second, whose inflection point occurs about half a week before the EBD. This is, of course, a consequence of predictable human behaviour, viz. a high number of registrations just before the EBD. In particular, in the data set we analysed, about a third of the registrations occurred just before the EBD.

It would be useful to further test the bi-logistic growth model on other conference data sets, and also to examine the effect of normalising the data using the total number of submissions, as was done in [13]. An interesting application of the model would be for forecasting registration [20], which might be utilised for conference planning purposes.

Acknowledgements

The authors would like to thank the referees for their constructive comments, which helped us to improve the paper.

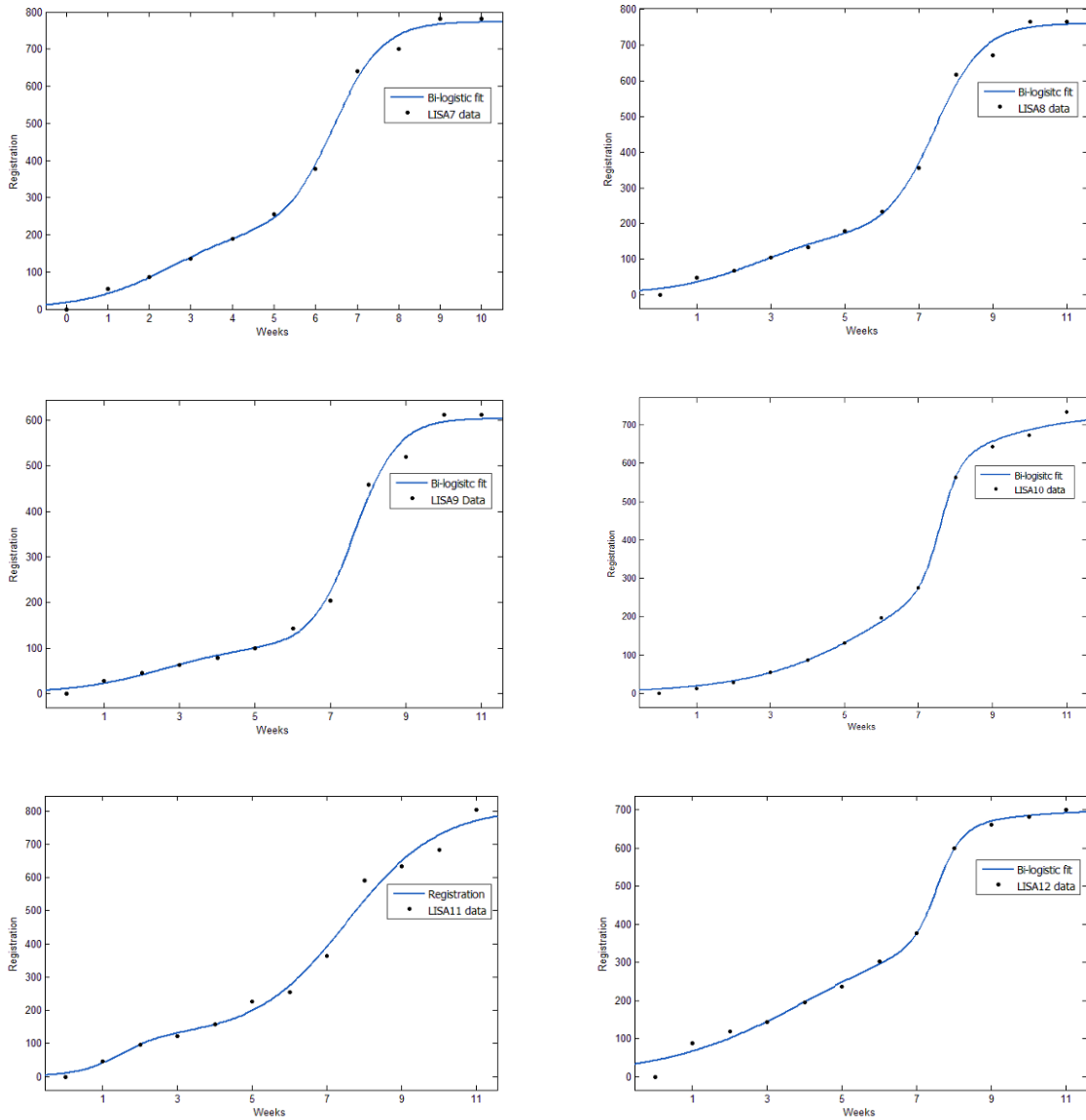


Figure 1. Bi-logistic curve fitting for LISA registration data from 2007 to 2012

References

- [1] A.-L. Barabási, *IEEE Contr. Syst. Mag.* 27, 33 (2007)
- [2] A.-L. Barabási, *Nat.* 435, 207 (2005)
- [3] T.I. Fenner, M. Levene, G. Loizou, G. Roussos, *Comput. Net.* 51, 4586 (2007)
- [4] T.I. Fenner, M. Levene, G. Loizou, *IEEE T. Comput. Intell. AI Games* 4, 84 (2012)
- [5] H.A. Simon, *Biometrika* 42, 425 (2004)
- [6] P. Hedström, R. Swedberg, *Social mechanisms: An introductory essay*, In: P. Hedström, R. Swedberg (Ed.), *Social Mechanisms: An Analytical Approach to Social Theory* (Cambridge University Press, Cambridge, U.K., 1998) 1
- [7] P. Hedström, P. Bearman, *What is analytical sociology all about: An introductory essay*, In: P. Hedström, P. Bearman (Ed.), *Handbook of Analytical Sociology* (Oxford University Press, Oxford, 2009) 3
- [8] T.C. Schelling, *Social mechanisms and social dynamics*, In: P. Hedström, R. Swedberg (Ed.), *Social Mechanisms: An Analytical Approach to Social Theory* (Cambridge University Press, Cambridge, U.K., 1998) 32
- [9] C. Marchetti, P.S. Meyer, J.H. Ausubel, *Technol. Forecast. Soc.* 52, 1 (1996)
- [10] P.S. Meyer, J.W. Yung, J.H. Ausubel, *Technol. Forecast. Soc.* 61, 247 (1999)
- [11] V. Alfi, A. Gabrielli, L. Pietronero, *Nat. Phys.* 3, 746 (2007)
- [12] V. Alfi, A. Gabrielli, L. Pietronero, *Cent. Eur. J. Phys.* 7, 483 (2009)
- [13] P. Flandrin, *Adv. Complex Syst.* 7, 439 (2010)
- [14] P. Meyer, *Technol. Forecast. Soc.* 47, 89 (1994)
- [15] D. Kucharavy, R. De Guio, *Procedia Eng.*, 9, 559 (2011)
- [16] C. Forbes, M. Evans, N. Hastings, B. Peacock, *Statistical Distributions*, 4th edition (John Wiley & Sons, New York, NY, 2011)
- [17] J.E. Sheehy, P.L. Mitchell, A.B. Ferrer, *Ann Bot-London* 94, 811 (2004)
- [18] M. Ausloos, *Adv. Complex Syst.* 16, 1250088 (2012)
- [19] H. Motulsky, *Intuitive Biostatistics* (Oxford University Press, Oxford, 1995)
- [20] N. Mead, T. Islam, *Int. J. Forecast.* 11, 199 (1995)