Set -5: Modelling data with Zipf's law and a bimodal function

Dhananjay Vora (2020001417)* and Kashish Shroff (202001425)†
Dhirubhai Ambani Institute of Information & Communication Technology, Gandhinagar, Gujarat 382007, India
CS302, Modeling and Simulation

In this report, we studied and modelled the Zipf's law of dependency network of Debian and for the bimodal distribution of a one-dimensional vehicle traffic flow.

I. ZIPF'S LAW IN THE DEPENDENCY NETWORK OF DEBIAN

A. Introduction

The resulting Debian distribution's semantic network of nodes is a simple dependency-based directed network. As a result, two different types of directed networks will develop. In this case, Zipf's law is actually followed by both the networks of incoming and outgoing links.

B. Model

We propose a nonlinear logistic-type equation to offer a continuum model for any power-law feature in this frequency distribution,

$$(x+\lambda)\frac{\mathrm{d}\phi}{\mathrm{d}x} = \alpha\phi\left(1-\eta\phi^{\mu}\right) \tag{1}$$

with ϕ is actual number of software packages that are connected by a particular number of links x. α is a power-law exponent, μ is a nonlinear saturation exponent, η is a "tuning" parameter for non-linearity, and λ parameter in setting a limiting scale for the poorly connected nodes.

$$\phi(x) = \left[\eta + \left(\frac{x+\lambda}{c}\right)^{-\mu\alpha}\right]^{-1/\mu} \tag{2}$$

where c is an integration constant. With $\mu = -1$ and $\alpha = -2$, implying that the power-law is specifically Zipf's law,

$$\phi(x) = \eta + \left(\frac{c}{x+\lambda}\right)^{2} \tag{3}$$

C. Results

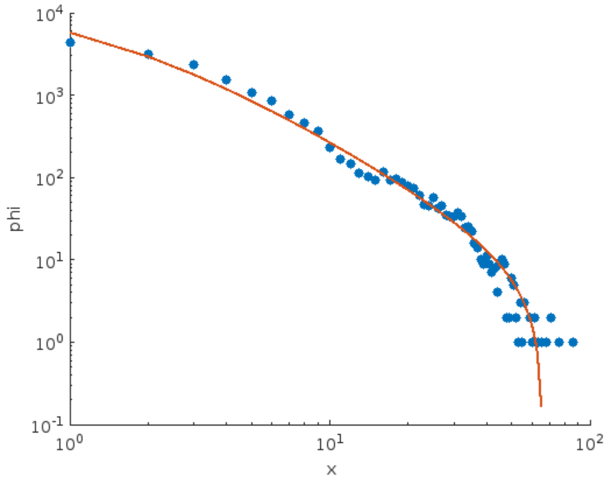

fig. 1. For the network of incoming links in the Etch release

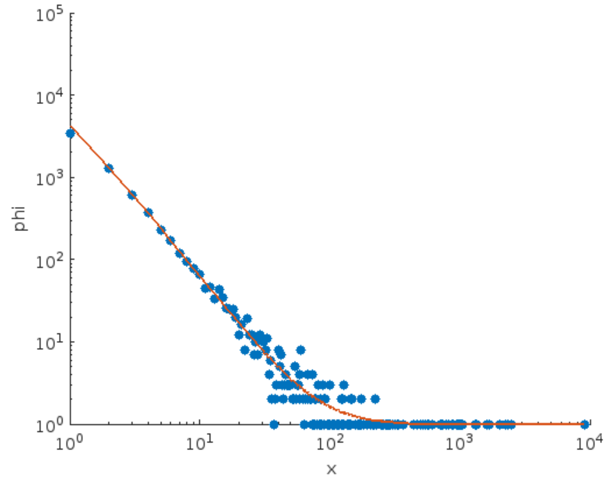

fig. 2. For the network of outgoing links in the Etch release

---

*Electronic address: 202001417@daiict.ac.in
†Electronic address: 202001425@daiict.ac.in

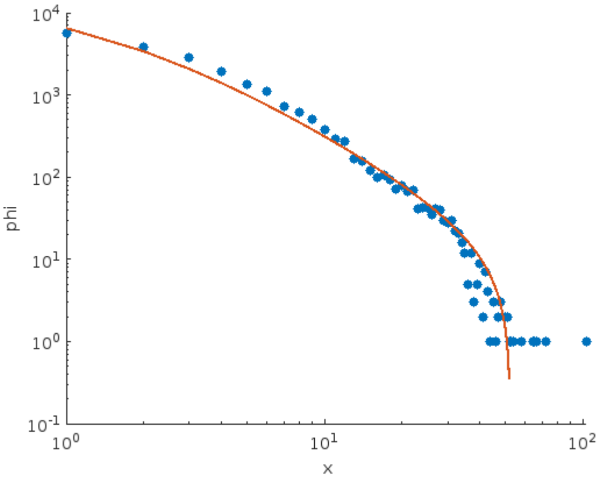

fig. 3. For the network of incoming links in the Lenny release

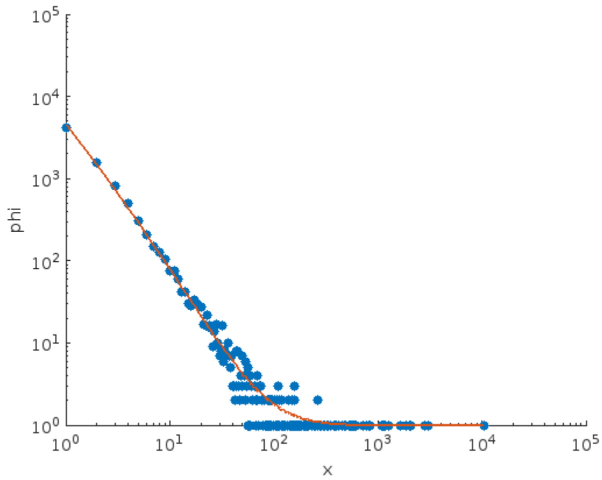

fig. 4. For the network of outgoing links in the Lenny release

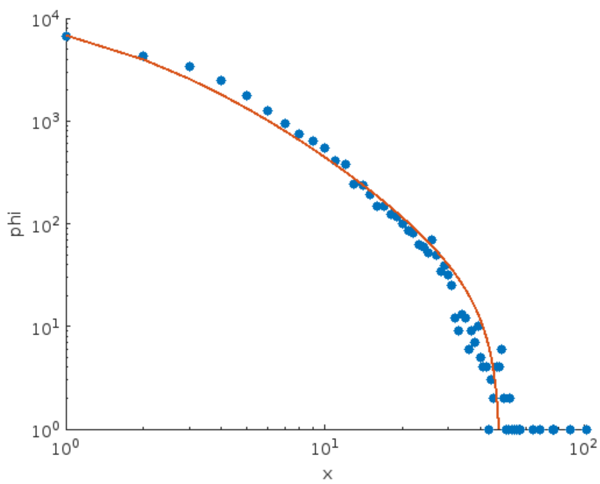

fig. 5. For the network of incoming links in the stable Squeeze release

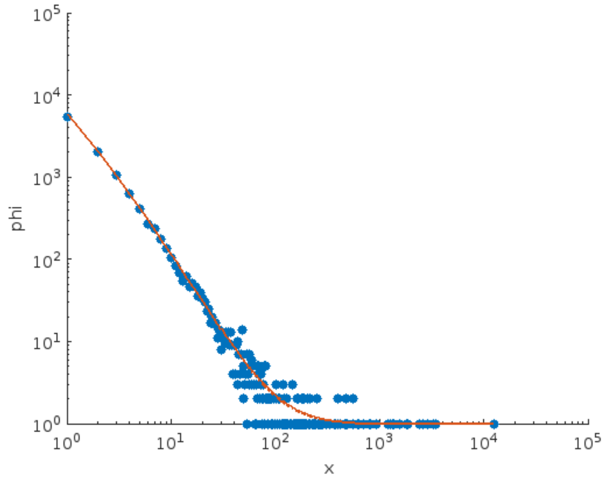

fig. 6. For the network of outgoing links in the stable Squeeze release

Parameter values			
Network Type	η	λ	c
in Etch	-8	1.5	190
out Etch	1	0.25	80
in Lenny	-15	1.6	210
out Lenny	1	0.35	90
in Squeeze	-28	2.2	265
out Squeeze	1	0.45	110

D. Conclusions

- Debian's dependency network adheres to Zipf's law, which implies a power-law relationship between a package's rank and frequency.

- Beyond intermediate scales of x, the power law trend does not remain true when the distribution is finite. Both the densely connected and sparsely linked nodes exhibit a power law trend deviation.

- The distribution of incoming links has a longer tail than the distribution of outgoing links, which indicates that some packages are more frequently relied upon than others, according to the analysis of incoming and outgoing dependency links.

II. BIMODALITY IN TRAFFIC FLOWS

A. Introduction

Here, we create a mathematical framework that provides a unified, comprehensive explanation of bimodal phenomena. The modelling work on the bimodal distribution of data for the case of one-dimensional vehicle traffic flows has consequences that are considerably broader than any one particular situation.

B. Model

$$N(t) = A\left(\mu + t^2\right)\exp\left[-(\lambda t - \beta)^2\right] \tag{4}$$

C. Results

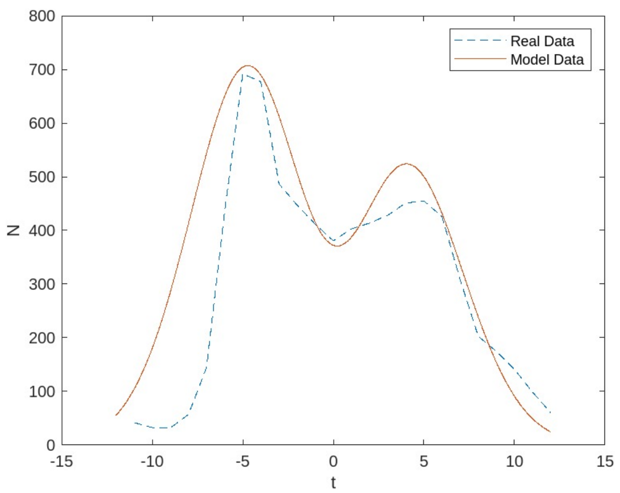

fig. 1. Bimodal distribution of traffic flow due west

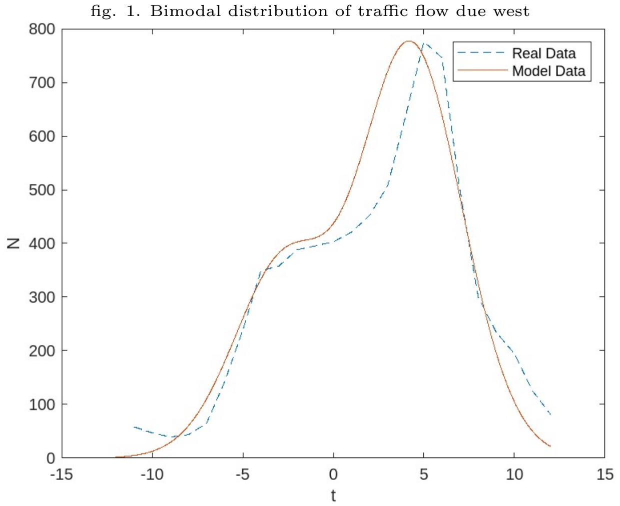

fig. 2. Bimodal distribution of traffic flow due east

Parameter values				
Direction of flow	A	μ	λ	β
West	44	8.53	0.19	-0.09
East	44.1	10.5	0.22	0.24

Statistical Analysis		
Direction of flow	Mean	Standard Deviation
West	-0.0856	0.5395
East	0.7234	2.3252

D. Conclusions

- The key characteristics of bimodality, including asymptotically declining tails, asymmetry of the bimodal peaks, and their oscillatory exchange in a twenty-four-cycle traffic flow, are captured by our model.

- The above- mentioned function effectively encapsulates the key characteristics of bimodality, the first of which is the requirement for three turning points, each with a local minimum and two peaks on either side.

- The western direction has significantly more morning traffic than the eastern. Being unusually heavy in the evening is the eastbound traffic.

[1] Nair, R., Nagarjuna, G. and Ray, A.K., 2009. Finite-size effects in the dependency networks of free and open-source software. arXiv preprint arXiv:0901.4904.

[2] Mullick, A. and Ray, A.K., 2016, October. Nonlinear dynamics of bimodality in vehicular traffic. In Journal of Physics: Conference Series (Vol. 755, No. 1, p. 012009). IOP Publishing.

[3] Ray, A.K. (2010). Modeling Saturation in Industrial Growth. In: Basu, B., Chakravarty, S.R., Chakrabarti, B.K., Gangopadhyay, K. (eds) Econophysics and Economics of Games, Social Choices and Quantitative Techniques. New Economic Windows. Springer, Milano. https://doi.org/10.1007/978-88-470-1501-2$_1$4