Modified Lomax model: a heavy-tailed distribution for fitting large-scale real-world complex networks

Swarup Chattopadhyay1,3 · Tanujit Chakraborty2 · Kuntal Ghosh1 · Asit K. Das3

Received: 10 August 2020 / Revised: 14 April 2021 / Accepted: 17 April 2021 / Published online: 27 April 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2021

Abstract
Real-world networks are generally claimed to be scale-free. This means that the degree distributions follow the classical power-law, at least asymptotically. However, closer observation shows that the classical power-law distribution is often inadequate to meet the data characteristics due to the existence of an identifiable nonlinearity in the entire degree distribution in the log-log scale. The present paper proposes a new variant of the popular heavy-tailed Lomax distribution which we named as the modified Lomax (MLM) distribution that can efficiently capture the crucial aspect of heavy-tailed behavior of the entire degree distribution of real-world complex networks. The proposed MLM model, derived from a hierarchical family of Lomax distributions, can efficiently fit the entire degree distribution of real-world networks without removing lower degree nodes, as opposed to the classical power-law-based fitting. The MLM distribution belongs to the maximum domain of attraction of the Frechet distribution and is right tail equivalent to Pareto distribution. Various statistical properties including characteristics of the maximum likelihood estimates and asymptotic distributions have also been derived for the proposed MLM model. Finally, the effectiveness of the proposed MLM model is demonstrated through rigorous experiments over fifty real-world complex networks from diverse applied domains.

Keywords Complex networks · Degree distribution · Lomax distribution · Heavy-tailed distribution · Power-law · Statistical properties

1 Introduction
The modeling and structural aspects of large-scale real-world complex networks, including social, information, collaboration, communication, etc., have been well studied during the past decade (Albert and Barabási 2002; Albert et al. 2000; Newman 2001, 2003; Chacoma et al. 2018) by many researchers. The World Wide Web, Twitter, Orkut, Youtube, DBLP, Wiki talk, Facebook, LinkedIn are examples of such large-scale real-world complex networks. These networks are characterized by several important structural, emergent properties like degree distribution, correlation coefficient, average nearest neighbor, average path length, clustering coefficient, community structure, etc. Recently, the modeling and statistical aspects of such emergent structural properties, therefore, remain an important research area in the study of large-scale real-world complex networks (Zarandi and Rafsanjani 2018; Newman 2003; Cui et al. 2014; Nie et al. 2016; Shakibian and Charkari 2018). In this regard, the node degree distribution has been well studied and viewed as an important structural characteristic of real-world networks (Muchnik et al. 2013). In 1999, (Barabási and Albert 1999; Albert et al. 1999) modeled the node degree distribution of the World Wide Web (WWW) using a power-law. Since then, many researchers have also favored the use of heavy tailed power-law in modeling the node degree distribution of real-world networks such as collaboration networks, communication networks, social networks, biological networks, etc. (Clauset et al. 2009; Barabasi 2005). Mathematically, a quantity $x$ follows a power-law if it is drawn from a probability distribution $P(x) \propto x^{-\alpha}$ where the parameter $\alpha$ is a positive constant and is known as exponent or scaling parameter of the distribution. Thus, it is common to encounter the
claim that most of the real-world networks are scale-free, meaning that the degree distributions follow single power-law. Despite this, a closer observation, while fitting, shows that the classical power-law distribution is often inadequate to meet the data characteristics adequately because of the existence of an identifiable nonlinearity (bend) when the entire degree distribution is considered in log-log scale as shown in Fig. 1a, b (elaborated later).

This feature (nonlinearity) of the entire degree distribution, depending on when and where it is considered or ignored, possibly constitutes the reason why the universality vis-a-vis scarcity of scale-free networks has remained controversial ever since its inception (Liljeros et al. 2001; Jones and Handcock 2003). The debate has continued to crop up time and again throughout the last twenty-one years (Newman 2005; Seshadri et al. 2008; Clauset et al. 2009; Barabasi 2005) and in very recent times too whence it has been claimed through an empirical and an extensive study that the power-law distribution does not fit well in most cases and thereby produces a significant fitting error, followed by counter-claims (Brodo and Clauset 2019).

This apart, researchers have also argued differently in favor of scale-free structure while suggesting some softer statistical criteria for scale-freeness (Holme 2019; Voitalov et al. 2019; Stumpf and Porter 2012). Especially significant in this context is the following quote (Stumpf and Porter 2012): “The fact that heavy-tailed distributions occur in complex systems is certainly important (because it implies that extreme events occur more frequently than would otherwise be the case)... However, a statistically sound power-law is no evidence of universality without a concrete underlying theory to support it. Moreover, knowledge of whether or not a distribution is heavy-tailed is far more important than whether it can be fit using a power-law”.

Several other heavy-tailed distributions such as lognormal, Pareto lognormal (PLN), double Pareto lognormal (DPLN), etc., also have been proposed in modeling the degree distribution of real-world networks instead of power-law (Seshadri et al. 2008; Sala et al. 2010). Recent research also recognized the deviations from a pure power-law distribution over various network data sets and recommended some other distributions for better modeling the heavy-tailed node degree distribution (Voitalov et al. 2019; Chattopadhyay et al. 2019, 2014). Thus, identifying the reasons for deviation of single power-law while fitting and looking for alternative models which can efficiently capture the crucial aspect of heavy-tailed and long-tailed behavior of the entire degree distribution of real-world complex networks continue to remain a challenging task of current research in the field of complexity science even as it steadily gravitates toward data science (Holme 2019; Voitalov et al. 2019; Stumpf and Porter 2012).

Motivation: Networks are a powerful way to represent and study the structure of real-world complex systems. Across various applied domains of networks, it is common to encounter the claim that most of the real-world networks are scale-free, meaning that the degree distributions follow single power-law, though the universality of scale-free networks remains controversial as already discussed above.

Figure 1a, b depicts the plot of entire degree distribution in the log-log scale of the Twitter and LiveJournal social networks. The horizontal axis represents the unique degree value \(x\), and the vertical axis represents the corresponding frequency. In these networks, a node represents a single user,
and an edge represents a follower of that user. From these figures, it is clear that the pattern of the degree distribution of these networks does not match with the straight-line representation in the log-log scale through a single power-law. Usually, while fitting the node degree distribution, the single power-law is applied only for values of degree higher than some minimum (say, $x_{\text{min}}$) and the exponent $\alpha$ is estimated from the data using MLE accordingly.

Thus power-law distribution provides better fitting or in other words better inclined to the right tail of the data unless otherwise, some “unimportant” (i.e., lower degree) nodes are left out. Analytically, we can say that this inadequacy of fitting a single power-law occurs because of nonlinear behavior of the degree distribution curve in the log-log scale. This motivates the researchers to use other heavy-tailed probability models with nonnegative exponent for better modeling the entire degree distribution of real-world networks. To capture these nonlinearities in the degree distribution of the real-world complex networks in a log-log scale, previous studies used various heavy-tailed probability distributions (Voitalov et al. 2019; Seshadri et al. 2008; Chattopadhyay et al. 2019, 2021; Golosovsky 2017). In this current research, we study the behavior of the entire degree distributions with a new variant of Lomax distribution that has wide applications in the field of actuarial science, reliability modeling, economics and computer science (Lomax 1954; Ahsanullah 1991; Hassan et al. 2016; Childs et al. 2001). The Lomax distribution is essentially a Pareto Type-II distribution that has been shifted so that its support begins at zero (Ahsanullah 1991; Lomax 1954). Some extension and generalization of the Lomax distribution has been carried out for analyzing reliability and survival data sets in the past (Ahsanullah 1991; Al-Awadhi and Ghitany 2001; Balakrishnan and Ahsanullah 1994). Recent research also focused on a new generalization of Pareto distribution with application to the breaking stress data (Jayakumar et al. 2020). This paper proposes a modified Lomax (MLM) distribution to be derived from a hierarchical family of Lomax distributions where the nonnegative shape parameter is assumed to be expressible as a nonlinear function of the data.

Our contribution: The major contribution here is to develop a modified Lomax (MLM) distribution from a hierarchical family of Lomax distributions for efficient modeling of the entire degree distribution of real-world complex networks (Ahsanullah 1991; Lomax 1954). The reasons for introducing MLM distribution are to provide greater flexibility and better fitting to the entire node degree distribution of complex networks compared to other popularly used heavy-tailed distributions. In other words, the proposed MLM model can be used for effective modeling the degree distribution of complex networks, coming from different disciplines, in the whole range of the data without discarding some of the lower degree nodes. Moreover, some statistical properties including extreme value and asymptotic behavior of the proposed MLM distribution have been studied in this context. We also provide mathematical arguments to explain the behavior of the likelihood surface for this nonlinear variant of the Lomax distribution, i.e., MLM distribution. A sufficient condition for the existence of the global maximum for the likelihood estimates is given using the notion of the coefficient of variations (CV) and discuss the parameter estimation procedures of the proposed MLM distribution. In order to justify the effectiveness of the proposed MLM distribution, we have compared it with the other common power-law-type distributions, viz. power-law, Pareto, lognormal, exponential, power-law with exponential cutoff and Poisson (Sala et al. 2010; Clauset et al. 2009; Newman 2005). The goodness-of-fit of the observed degree distribution is evaluated and compared using a few statistical measures, viz. bootstrap Chi-square, KL-divergence (KLD), mean absolute error (MAE) and root mean square error (RMSE). Several real-world complex networks from diverse fields have been used for experimental evaluation. Empirical results confirm the effectiveness of the proposed MLM distribution compared to other common distributions.

The remainder of the paper is organized as follows. Section 2 provides the details of the hierarchical family of Lomax distributions. We propose and interpret proposed modified Lomax (MLM) distribution in Sect. 3. Section 4 discusses the statistical properties, including extreme value and asymptotic behaviors of the proposed MLM distribution. Section 5 is devoted to the experimental results with a detailed analysis of the results over several real-world complex networks. Finally, Sect. 6 concludes the paper with a brief discussion.

2 Model

In this section, we first introduce a new family of heavy-tailed Lomax (HLM) distributions. Further, we propose a relevant model from this newly introduced family to model the real-world heavy-tailed network data sets in the whole range.

2.1 Genesis

Lomax distribution has been used as an alternative to exponential, power-law, gamma and Weibull distribution for modeling heavy tailed data sets (Atkinson and Harrison 1978; Bryson 1974; Chahkandi and Ganjali 2009; Hassan and Al-Ghamdi 2009). The cumulative distribution function (CDF) and the probability density function (PDF) of the Lomax model are defined as follows:
Definition 1 A random variable $Z$ follows Lomax distribution with parameters $\alpha$ and $\sigma$ if the CDF is of the form:

$$F(z) = 1 - \left(1 + \frac{z}{\sigma}\right)^{-\alpha} ; z \geq 0,$$

where $\alpha > 0$ is the shape parameter (real) and $\sigma > 0$ is the scale parameter (real). The corresponding PDF is defined as follows:

$$f(z) = \frac{\alpha}{\sigma} \left(1 + \frac{z}{\sigma}\right)^{-\alpha-1} ; z \geq 0 (1)$$

Below we introduce a new family of heavy tailed Lomax distributions which is right tail-equivalent to a power-law distribution.

Definition 2 A continuous random variable $X$ follows a family of heavy-tailed Lomax (HLM) distributions if and only if it has the following CDF:

$$F(x) = 1 - (1 + x)^{-m(x)} ; x \geq 0 (2)$$

and $F(x) = 0$ if $x \leq 0$, where $m : (0, \infty) \rightarrow \mathbb{R}^+$ is a real, continuous, positive function which is differentiable on $(0, \infty)$ and satisfies the following conditions:

1. The function $m$ is strictly positive and have finite limit at infinity, i.e., $\lim_{x \to \infty} m(x) = \alpha > 0$.
2. $\lim_{x \to 0^+} (1 + x)^{m(x)} = 1$ and $\lim_{x \to \infty} (1 + x)^{m(x)} = \infty$.

3. $\frac{m'(x)}{m(x)} \geq \frac{1}{(1 + x) \log(1 + x)}, x > 0$.

It can be easily verified that the CDF in (2) satisfying conditions (1), (2) and (3) is a genuine CDF which can also be expressed as follows:

$$F(x) = 1 - \exp \left[-m(x) \log(1 + x)\right], \quad x > 0$$

The PDF of this new family of heavy-tailed Lomax distribution is of the form:

$$f(x) = (1 + x)^{-m(x)} \left[\frac{m(x)}{(1 + x)} + m'(x) \log(1 + x)\right], \quad x > 0 \text{ and } f(x) = 0, \quad x \leq 0.$$  

There can be a wide variety of choices of $m(x)$ satisfying $\lim_{x \to \infty} m(x) = \alpha > 0$. It is noted that the simplest choice of $m(x) = \alpha$ and $x = \frac{z}{\sigma}$ corresponds to the Lomax distribution.

We further represent this newly introduced family of Lomax distributions as a hierarchical family in accordance with Pareto distribution (Arnold 2015).

Definition 3 (HLM Type-I family of distributions) Suppose that a random variable $X$ follows HLM family of distributions as defined in (2). Then with a scale parameter $\sigma \in (0, \infty)$, the CDF of HLM Type-I family of distributions takes the following form:

$$F(x) = 1 - \left[1 + \left(\frac{x}{\sigma} - 1\right)^{\sigma \cdot (x - 1)}\right], \quad x > \sigma$$

By taking $m \left(\frac{z}{\sigma} - 1\right) = \alpha > 0$, we obtain the classical Pareto Type-I distribution.

Definition 4 (HLM Type-II family of distribution) Supposed that a random variable $X$ follows HLM family of distributions as defined in (2). Then with a location parameter $\mu \in \mathbb{R}$ and a scale parameter $\sigma \in (0, \infty)$, the CDF of HLM Type-II family of distributions takes the following form:

$$F(x) = 1 - \left(1 + \frac{x - \mu}{\sigma} \right)^{-\frac{m}{\sigma}}, \quad x > \mu$$

By taking $m \left(\frac{x - \mu}{\sigma} \right) = \alpha > 0$, we obtain the Pareto Type-II distribution. Also, in addition $\mu = 0$ corresponds to the Lomax distribution.

Definition 5 (HLM Type-III family of distribution) Suppose that a random variable $X$ follows HLM family of distributions as defined in (2). Then with a location parameter $\mu \in \mathbb{R}$, scale parameters $\sigma \in (0, \infty)$ and a shape parameter $\gamma > 0$, the CDF of HLM Type-III family of distributions takes the following form:

$$F(x) = 1 - \left[1 + \left(\frac{x - \mu}{\sigma} \right)^{\gamma}\right]^{-m \left(\frac{x}{\alpha} \right)}, \quad x > \mu$$

By taking $m \left(\frac{x - \mu}{\sigma} \right) = \alpha > 0$, we obtain the Pareto Type-III distribution.

Obviously, the choice of $m(\cdot)$ function is subjective and any function $m$ satisfying conditions (1), (2) and (3) will give some known (unknown) heavy-tail Lomax distributions.

### 3 Modified Lomax (MLM) model

The Lomax distribution does not provide great flexibility in modeling heavy-tailed data sets in the whole range similar to the power-law distribution. Due to this, the trend of
parameter(s) induction to the baseline Lomax distribution has received increased attention in the recent years. Several generalized classes of distributions by adding additional parameters such as shape and or scale and or location in the distribution are available such as exponentiated Lomax (EL) (Abdul-Moniem 2012), Beta-Lomax (BL) (Rajab xxx), exponential Lomax (ELomax) (Bassiouny yyy), Gamma-Lomax (GL) (Cordeiro et al. 2015) and Gumbel–Lomax (GuLx) model (Tahir et al. 2016).

This paper provides a new modified version of the Lomax distribution called modified Lomax (MLM) distribution. MLM distribution is shown to be an asymmetric distribution, which provides great fit in modeling large-scale heavy-tailed data sets. The proposed MLM model is derived from the HLM family of distributions (in particular, HLM Type-II model) that can efficiently model the entire degree distribution of real-world complex networks in the whole range without discarding lower degree nodes. We define a relevant model from the newly introduced HLM Type-II family with the location parameter \( \mu = 0 \) and we choose a flexible \( m(\cdot) \) function that depends on two shape parameters \( \alpha \) and \( \beta \) satisfying \( \lim_{x \to \infty} m(x) = \alpha \). The rational behind adding an additional shape parameter in the HLM Type-II family of distribution will make the statistical model more flexible, simple and have physical interpretation. This idea of generalization should suffice the practical needs of working with the nonlinear exponent to address the structural issue (degree distribution) of real-world complex networks.

Now we choose a nonlinear function \( m(\cdot) \) that adds a nonlinear exponent while fitting heavy-tailed HLM Type-II model in the degree distributions is as follows:

\[
m(x) = \alpha \left( \frac{\log(1 + x)}{1 + \log(1 + x)} \right)^{\beta}.
\]

The chosen \( m(x) \) approaches to \( \alpha \) from below if \( -1 < \beta < 0 \) as \( x \to \infty \) and approaches to \( \alpha \) from above for \( \beta > 0 \) as \( x \to \infty \). Note that, the function \( m(x) \) as defined above includes the constant function (in this case \( \alpha \)) as special cases by setting \( \beta = 0 \). The derivative of \( m(x) \) is given by

\[
m'(x) = \frac{\alpha \beta}{x + 1} \left( \frac{\log(1 + x)}{1 + \log(1 + x)} \right)^{\beta-1} \left( \frac{\log(1 + x)}{1 + \log(1 + x)} \right)^{\beta} \left( 1 + \log(1 + x) \right)^{-2}.
\]

Now, we define a relevant model with the above choice of \( m(\cdot) \) in the HLM Type-II model with \( \mu = 0 \) and name it as Modified Lomax Model to be denoted by MLM(\( \alpha, \beta, \sigma \)). This modification to the Lomax distribution provides more flexibility in the data modeling since the nonnegative shape parameter is assumed to be expressed as a nonlinear function of the empirical data. Thus the proposed MLM model with parameters \( \alpha, \beta, \sigma \) could be useful for modeling the heavy-tailed degree distribution of real-world complex network data sets in the whole range.

**Definition 6 (Modified Lomax Distribution)** A continuous random variable \( X \) follows MLM(\( \alpha, \beta, \sigma \)) distribution with \( \alpha (> 0) \) and \( \beta (> -1) \) as the shape parameters and \( \sigma (> 0) \) as the scale parameter if the CDF takes the following form:

\[
F(x) = 1 - \exp \left[ -\alpha \frac{\log^{\beta+1}(1 + x/\sigma)}{[1 + \log(1 + x/\sigma)]^{\beta}} \right], \ x > 0,
\]

and \( F(x) = 0 \) if \( x \leq 0 \). The corresponding PDF is given by,

\[
f(x) = \frac{\alpha [\beta + 1 + \log \left( 1 + \frac{x}{\sigma} \right)] \left[ \log \left( 1 + \frac{x}{\sigma} \right) \right]^{\beta}}{\sigma \left[ 1 + \log \left( 1 + \frac{x}{\sigma} \right) \right]^{\beta+1}} \exp \left[ -\alpha \frac{\log \left( 1 + \frac{x}{\sigma} \right)^{\beta+1}}{\left[ 1 + \log \left( 1 + \frac{x}{\sigma} \right) \right]^{\beta}} \right], \ x > 0
\]

and \( f(x) = 0 \) if \( x \leq 0 \).

This MLM model includes Lomax distribution \( \beta = 0 \) as particular case. In addition, it belongs to the new family of HLM Type-II distribution satisfying the condition: \( \lim_{x \to \infty} m(x) = \alpha (> 0) \). Due to the addition of an additional parameter \( \beta \) in the exponents of the Lomax distribution generates various shapes (unimodal and bimodal) and provides greater flexibility (nonlinearity and heavy-tail) as shown in Fig. 2. We study the monotonicity for the PDF of the proposed MLM model in Theorem 1 below.
Theorem 1 Let \( X \) be the random variable follows MLM(\( \alpha, \beta, \sigma \)) distribution, then the PDF as in (4) is a decreasing function for \(-1 < \beta < 0\).

Proof Differentiating (4) w.r.t. \( x \), we have

\[
f'(x) = -\frac{\alpha^2(1-F(x))\left[\beta + 1 + \log\left(1 + \frac{x}{\sigma}\right)\right]^2\left[\log\left(1 + \frac{x}{\sigma}\right)\right]^{\beta+2}}{\sigma^2\left[1 + \frac{x}{\sigma}\right]^2\left[1 + \log\left(1 + \frac{x}{\sigma}\right)\right]^{\beta+2}}
\]

\[
= \frac{\alpha[1-F(x)]\left\{\beta + 1 + \log\left(1 + \frac{x}{\sigma}\right)\left[\log\left(1 + \frac{x}{\sigma}\right)\right]^{\beta} + (1 + \beta)\left[\log\left(1 + \frac{x}{\sigma}\right)\right]^{\beta-1}\right\}}{\sigma^2\left[1 + \frac{x}{\sigma}\right]^2\left[1 + \log\left(1 + \frac{x}{\sigma}\right)\right]^{\beta+1}}
\]

\[
+ \frac{\alpha[1-F(x)](1 + \beta)\left[\beta + 1 + \log\left(1 + \frac{x}{\sigma}\right)\left[\log\left(1 + \frac{x}{\sigma}\right)\right]^{\beta-1}}{\sigma^2\left[1 + \frac{x}{\sigma}\right]^2\left[1 + \log\left(1 + \frac{x}{\sigma}\right)\right]^{\beta+2}}
\]

Trivially, if \(-1 < \beta < 0\), then \( f'(x) < 0 \). Thus, \( f(x) \) is decreasing function if \( \beta \in (-1,0) \). □

4 Statistical properties of the MLM distribution

4.1 Characterization and existence of the likelihood

Initially, we characterize the maximum likelihood estimates (MLEs) of the parameters \( \alpha \) and \( \sigma \) of a Lomax distribution. Subsequently, we derived a sufficient condition for the existence of MLEs of the MLM distribution using coefficient of variation (CV). Given a set of samples \( \{x_i\} \) of size \( n \), the

\[\text{Fig. 2 Plot of the PDFs of MLM distribution}\]
log-likelihood function for the Lomax distribution, after dividing it by the sample size \( n \), is given by

\[
\ell(a, \sigma) = \log a - \log \sigma - \frac{(a + 1)}{n} \sum_{i=1}^{n} \log \left(1 + \frac{x_i}{\sigma}\right)
\]  

(6)

Differentiating (6) w.r.t. \( a \) and \( \sigma \), respectively, we have:

\[
\frac{d\ell(a, \sigma)}{da} = \frac{1}{a} - \frac{1}{n} \sum_{i=1}^{n} \log \left(1 + \frac{x_i}{\sigma}\right)
\]  

(7)

\[
\frac{d\ell(a, \sigma)}{d\sigma} = -\frac{1}{\sigma} + \frac{(1 + a)}{n\sigma} \sum_{i=1}^{n} \left(\frac{x_i}{\sigma + x_i}\right)
\]  

(8)

Equating to zero the derivative of \( \ell(a, \sigma) \) w.r.t. \( a \) in (7), we obtain \( \hat{a} = a(\sigma) \) as follows:

\[
\hat{a} = a(\sigma) = \frac{n}{\sum_{i=1}^{n} \log \left(1 + \frac{x_i}{\sigma}\right)}
\]  

(9)

Differentiating (9) w.r.t. \( \sigma \) we have,

\[
a'(\sigma) = \frac{\hat{a}^2}{n\sigma} \sum_{i=1}^{n} \frac{x_i}{\sigma + x_i}
\]  

(10)

It is important to note that there is no closed form solution to the likelihood based on (7) and (8), and a suitable numerical algorithm (for example, Newton-Raphson method) can be employed to obtain the maximum likelihood estimates (MLEs) of the \( a \) and \( \sigma \). Different estimation procedures of the MLEs have been discussed in previous literature, for example see (Giles et al. 2013). But for small- or medium-sized samples, anomalous behavior of the likelihood surface can be encountered when sampling from the Lomax distribution. In this paper, we characterize the profile log-likelihood function in terms of the coefficient of variation (CV), defined as follows:

**Definition 7** The CV is the ratio of the standard deviation \((s)\) to the mean \((\mu)\).

\[
CV = \frac{s}{\mu};
\]

where \( \mu = \frac{1}{n} \sum_{i=1}^{n} x_i \) and \( s = \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2 - \mu^2} \).

By using standard notation, the profile log-likelihood function based on equation 6, is given by

\[
\ell_p(\sigma) = \sup_{a(\sigma)} \ell(\hat{a}, \sigma) = \log(a(\sigma)) - \log \sigma - 1 - \frac{1}{a(\sigma)}
\]  

(11)

Differentiating (11) w.r.t. \( \sigma \), we have the following:

\[
\ell'_p(\sigma) = \frac{a'(\sigma)}{a(\sigma)} - \frac{1}{\sigma} + \frac{a'(\sigma)}{[a(\sigma)]^2}
\]  

(12)

Below we present the following lemmas which will be useful to find the sufficient condition for the existence for the global maximum of the profile log-likelihood function (11).

**Lemma 1** The following limit holds:

1. \( \lim_{\sigma \to \infty} \frac{1}{\sigma} = 0; \)
2. \( \lim_{\sigma \to \infty} \frac{a(\sigma)}{a(\sigma) + x} = \frac{1}{\bar{x}}, \) where \( \bar{x} \) is the sample mean;
3. \( \ell_0 \equiv \lim_{\sigma \to \infty} \ell_p(\sigma) = \log \left(\frac{1}{\bar{x}}\right) - 1. \)

**Proof** The proof is elementary and can easily be done using series expansions.

**Lemma 2** The following limit holds:

1. \( \lim_{\sigma \to \infty} \frac{1}{\sigma a(\sigma)} = 0; \)
2. \( \lim_{\sigma \to \infty} \frac{a(\sigma)}{a(\sigma) + x} = \frac{1}{\bar{x}}; \)
3. \( \ell'_0 \equiv \lim_{\sigma \to \infty} \ell'_p(\sigma) = \log \left(\frac{1}{\bar{x}}\right) - 1. \)

**Proof** The proofs are straightforward and can be done using Lemma (1).

A sufficient condition for monotonic increasing (decreasing) for the profile log-likelihood function is presented in Theorem (2) below, for sufficiently large \( \sigma \). Also, we present a sufficient condition for the existence of global maximum corresponding to the likelihood function for the Lomax distribution to be at a finite point in Corollary (1).

**Theorem 2** Let \( X \) follows \( LM(a, \sigma) \) distribution with \( a, \sigma > 0 \). A sufficient condition for \( \ell'_p(\sigma) \) to be monotonically decreasing function is \( CV > 1 \) for \( \sigma \to \infty \), and if \( CV < 1 \), it is monotonically increasing.

**Proof** Using (9) and (10) in Eqn. (12), we can write \( \ell'_p(\sigma) \) as:
Using the limits of Lemma (1) in Eqn. (13), we have

\[
\lim_{\sigma \to 0} \frac{1}{\sigma} \sum_{i=1}^{n} \frac{x_i}{\sigma + x_i} = \frac{1}{n} \sum_{i=1}^{n} \frac{x_i}{\sigma + x_i}
\]

(13)

Finally, we note that \( \lim_{\sigma \to \infty} \sigma^2 \ell'_p(\sigma) > 0 \) when the R.H.S of Eqn. (14) is strictly greater than 0. Alternatively, the likelihood function is monotonic decreasing when \( \frac{1}{\sigma} \sum_{i=1}^{n} x_i^2 - \bar{x}^2 > 0 \), or, equivalently, CV > 1. In a similar way, we can show that if CV < 1, then the \( \ell'_p(\sigma) \) is monotonic increasing function for sufficiently large \( \sigma \).

**Remark 1** As a consequence of Theorem (2), it can be immediately concluded that \( \ell'_p(\sigma) \) tends to \( \ell'_\alpha \) based on Lemma

\[
\ell' \equiv \ell(x; \alpha, \beta, \sigma) = n \log(\alpha) - \sum_{i=1}^{n} \log(\sigma + x_i) + \sum_{i=1}^{n} \log \left[ \beta + 1 + \log \left( 1 + \frac{x_i}{\sigma} \right) \right]
\]

\[
+ \beta \sum_{i=1}^{n} \log \left[ \log \left( 1 + \frac{x_i}{\sigma} \right) \right] - (\beta + 1) \sum_{i=1}^{n} \log \left[ 1 + \log \left( 1 + \frac{x_i}{\sigma} \right) \right]
\]

\[
- \alpha \sum_{i=1}^{n} \frac{\log \left( 1 + \frac{x_i}{\sigma} \right)^{\beta+1}}{\left[ 1 + \log \left( 1 + \frac{x_i}{\sigma} \right) \right]^\alpha},
\]

(17)

(2) and \( \ell'_p(\sigma) \) is a monotonic function for sufficiently large \( \sigma \). The value of CV as a measure that can be useful to determine when \( \ell'_p(\sigma) \) will be monotonic increasing or decreasing function for sufficiently large \( \sigma \).

**Corollary 1** Given a set of samples \( \{x_i\} \) of (+)ve numbers with CV < 1, the profile likelihood function for the LM(\( \alpha, \sigma \)) distribution has a global maximum at a finite point.

**Proof** For small or moderate values of \( \sigma \), using (9), we have

\[
\lim_{\alpha \to 0} \alpha = \lim_{\sigma \to 0} \frac{n}{\sum_{i=1}^{n} \log \left( 1 + \frac{x_i}{\sigma} \right)} = 0.
\]

(15)

Now, using (15) in (11) we have the following:

\[
\lim_{\sigma \to 0} \ell'_p(\sigma) = -\infty.
\]

Since \( \ell'_p(\sigma) \) is a continuous and monotonic decreasing function for sufficiently large \( \sigma \) (as in Theorem 2) and using (16), we can conclude that a global maximum exists at a finite point when CV > 1.

**Remark 2** Corollary 1 shows that the likelihood function for the Lomax distribution has a global maximum for the samples \( \{x_i\} \) with CV > 1 at a finite point. The calculation of CV is completely based on available empirical data and easy to compute. The existence of MLE based on CV for the MLM distribution will also holds as because MLM model reduce to Lomax distribution when \( \lim_{x \to \infty} m(x) = \alpha \). This can be empirically validated in Sects. 5, 5.3 and will be useful from practitioner’s point of view.

### 4.2 MLE of parameters

In this section, the maximum likelihood estimates are derived for parameters \( \alpha, \beta, \) and \( \sigma \) of MLM distribution. Let \( x_1, x_2, ..., x_n \) be a sample of size \( n \) from MLM(\( \alpha, \beta, \sigma \)) distribution. Then the log-likelihood function for the vector of parameters \( \Theta = (\alpha, \beta, \sigma)^T \) is given by

\[
\frac{\partial \ell'}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} \frac{\log \left( 1 + \frac{x_i}{\sigma} \right)^{\beta+1}}{\left[ 1 + \log \left( 1 + \frac{x_i}{\sigma} \right) \right]^\beta},
\]

(18)
\[ \frac{\partial \ell}{\partial \beta} = \sum_{i=1}^{n} \frac{1}{(1 + \beta + w_i)} + \sum_{i=1}^{n} \log \left( \frac{w_i}{1 + w_i} \right) \times \left[ 1 - \frac{\alpha w_i^{\beta+1}}{(1 + w_i)^\beta} \right] \]
\[ \frac{\partial \ell}{\partial \alpha} = -\frac{n}{\alpha^2} \]  
\[ \frac{\partial \ell}{\partial \sigma} = \sum_{i=1}^{n} \frac{x_i}{\sigma(\sigma + x_i)} \left[ \frac{\beta + 1}{(1 + w_i)} \right] - \frac{\beta}{w_i} - \frac{1}{(1 + \beta + w_i)} + \alpha \sum_{i=1}^{n} \frac{x_i}{\sigma(\sigma + x_i)} \left[ \frac{(1 + w_i)w_i^{\beta}}{(1 + w_i)^{\beta+1}} \right], \]
\[ \frac{\partial \ell}{\partial \sigma} = \sum_{i=1}^{n} \frac{x_i}{\sigma(\sigma + x_i)^2} \left[ \frac{w_i^{\beta+1}}{(1 + w_i)^\beta} \right] \]
\[ \frac{\partial \ell}{\partial \beta^2} = -\sum_{i=1}^{n} \frac{1}{(1 + \beta + w_i)^2} - \alpha \sum_{i=1}^{n} \log^2 \left( \frac{w_i}{1 + w_i} \right) \]
\[ \frac{\partial \ell}{\partial \alpha} = \sum_{i=1}^{n} \frac{x_i}{\sigma(\sigma + x_i)} \left[ \frac{(1 + \beta + w_i)w_i^{\beta}}{(1 + w_i)^{\beta+1}} \right] \]

where \( w_i = \log \left( 1 + \frac{w_i}{\sigma} \right) \).

The MLEs of the three parameters of the MLM(\( \alpha, \beta, \sigma \)) distributions are obtained by setting these above equations to zero and solving them simultaneously. Closed forms of the solutions are not available for the equations (18), (19) and (20). So, iterative methods will be applied to solve these equations numerically.

### 4.3 Asymptotic distribution

Fisher information matrix, a measure of the information content of the data relative to the parameters to be estimated, plays an important role in parameter estimation. The Fisher information matrix \( F \) can be obtained by taking the expected values of the second-order and mixed partial derivatives of \( \ell(\alpha, \beta, \sigma) \) w.r.t. \( \alpha, \beta, \) and \( \sigma \). Since, the analytical expression is hard to compute. Thus, it can be approximated by numerically investing the \( F = (F_{ij}) \) matrix. The asymptotic \( F \) matrix can be given as follows:

\[
F = \begin{bmatrix}
\frac{\partial^2 \ell}{\partial \alpha^2} & \frac{\partial^2 \ell}{\partial \alpha \beta} & \frac{\partial^2 \ell}{\partial \alpha \sigma} \\
\frac{\partial^2 \ell}{\partial \beta \alpha} & \frac{\partial^2 \ell}{\partial \beta^2} & \frac{\partial^2 \ell}{\partial \beta \sigma} \\
\frac{\partial^2 \ell}{\partial \sigma \alpha} & \frac{\partial^2 \ell}{\partial \sigma \beta} & \frac{\partial^2 \ell}{\partial \sigma^2}
\end{bmatrix}
\]

The second and mixed partial derivatives of the log likelihood function are obtained as follows:
The asymptotic distribution of \(\hat{\alpha}, \hat{\beta}, \hat{\sigma}\) can be written as

\[
\left(\hat{\alpha} - \alpha, (\hat{\beta} - \beta), (\hat{\sigma} - \sigma)\right) \sim N_3(0, F^{-1}(\hat{\Theta}))
\]

Then the approximate 100(1 - k)% confidence intervals for \(\alpha, \beta, \text{and } \sigma\) are given by \(\hat{\alpha} \pm 2\sqrt{\text{Var}(\hat{\alpha})}, \hat{\beta} \pm 2\sqrt{\text{Var}(\hat{\beta})}, \text{and } \hat{\sigma} \pm 2\sqrt{\text{Var}(\hat{\sigma})}\), where \(\hat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{\sigma})\) and \(2\sqrt{\text{Var}(\hat{\theta})}\) is the upper 100 k-th percentile of the standard normal distribution.

**4.4 Extreme value properties**

Here we study some of the interesting extreme value theoretical properties. The concept of regular variation is an important notion of extreme value theory. Below we show the extreme value results for the MLM distribution that can characterize the asymptotic behavior of extremes along with well grounded statistical theory.

**Definition 8** (Maximum domain of attraction) A function \(F\) is said to be regularly varying at infinity, if for every \(t > 0\),

\[
\lim_{x \to \infty} \frac{1 - F(tx)}{1 - F(x)} = t^{-\alpha}; \quad \alpha > 0.
\]

Then we say that \(F\) is a function with regularly varying tails with \(\alpha > 0\) as the tail index and \(F\) belongs to the maximum domain of attraction (MDA) of the Frechet distribution with index \(\alpha\).

**Theorem 3** The CDF (Eqn. 3) of the MLM distribution is a function with regularly varying tails, and it belongs to MDA of the Frechet distribution with index \(\alpha\).

\[
1 - F(tx) = \exp \left[ -\alpha \frac{\log^\beta (1 + \frac{t\alpha}{\sigma})}{\beta} \left(1 + \log (1 + \frac{t\alpha}{\sigma})\right) \right]; \quad t > 0
\]

**Proof**

Now, we have (using expansions of \(\log(1 - x)\) and \(\exp(x)\)):

\[
\left(\frac{\log (1 + \frac{t\alpha}{\sigma})}{1 + \log (1 + \frac{t\alpha}{\sigma})}\right)^\beta
\]

\[
= \left(1 - \frac{1}{1 + \log (1 + \frac{t\alpha}{\sigma})}\right)^\beta
\]

\[
= \exp \left[ \beta \log \left(1 - \frac{1}{1 + \log (1 + \frac{t\alpha}{\sigma})}\right)\right]
\]

\[
= \exp \left[ \beta \left( - \frac{1}{\log (1 + \frac{t\alpha}{\sigma})} + O\left(\frac{1}{\log^2 (1 + \frac{t\alpha}{\sigma})}\right)\right)\right] \quad \text{ (28)}
\]

\[
= 1 - \frac{\beta}{\log (1 + \frac{t\alpha}{\sigma})}
\]

\[
+ O\left(\frac{\beta}{\log^2 (1 + \frac{t\alpha}{\sigma})}\right)
\]

Using Eqn. (27) and (28) together, we get

\[
1 - F(tx) = \exp \left[ -\alpha \log \left(1 + \frac{tx}{\sigma}\right)\right]
\]

\[
\left\{ 1 - \frac{\beta}{\log (1 + \frac{tx}{\sigma})}
\right.\]

\[
+ O\left(\frac{\beta}{\log^2 (1 + \frac{tx}{\sigma})}\right)\right\} \quad \text{ (29)}
\]

Similarly for \(t = 1\), Eqn. (29) becomes
1 − F(x) = \exp \left[ - \alpha \log \left( 1 + \frac{x}{\sigma} \right) \right]
\begin{align*}
&\left\{ 1 - \frac{\beta}{\log \left( 1 + \frac{x}{\sigma} \right)} \\
&\quad + O\left( \frac{\beta}{\log^2 \left( 1 + \frac{x}{\sigma} \right)} \right) \right\} \tag{30}
\end{align*}

Now,
\begin{align*}
\lim_{x \to \infty} \frac{1 - F(tx)}{1 - F(x)} &= \lim_{x \to \infty} \exp \left[ - \alpha \log \left( 1 + \frac{x}{\sigma} \right) \right] \\
&\quad + O\left( \frac{1}{\log^2 \left( 1 + \frac{x}{\sigma} \right)} \right) \\
&= \exp \left( - \alpha \log t \right) \\
&= t^{-\alpha}.
\end{align*}
Thus, \( F \in MDA(\Phi_u) \).

\begin{align*}
\lim_{x \to \infty} \frac{1 - F(x)}{1 - G(x)} &= \lim_{x \to \infty} \frac{\exp \left[ - \alpha \log \left( 1 + \frac{x}{\sigma} \right) \right] + O\left( \frac{1}{\log \left( 1 + \frac{x}{\sigma} \right)} \right)}{1 - G(x)} \\
&= \lim_{x \to \infty} \exp \left( \frac{\alpha \beta}{\log \left( 1 + \frac{x}{\sigma} \right)} \right) \\
&= c < \infty.
\end{align*}

Now we study the tail-equivalent and heavy-tailed behavior of the proposed MLM distribution as follows:

**Definition 9** (Tail-equivalent) Two distributions \( F \) and \( G \) are said to be tail-equivalent if
\[ \lim_{x \to \infty} \frac{1 - F(x)}{1 - G(x)} = c; \quad 0 < c < \infty. \]

**Theorem 4** The MLM(\( \alpha, \beta, \sigma \)) distribution, defined in Eqn. (3), is right tail-equivalent to the power-law distribution.

**Proof** Let \( G(x) \) be the CDF of the power-law distribution, i.e.,
\[ 1 - G(x) = \left( 1 + \frac{x}{\sigma} \right)^{-\alpha} \]
and \( F(x) \) is the CDF of MLM distribution as given in Eqn. (3). Then,
\[ \lim_{x \to \infty} \exp \{ \lambda x \} \left( 1 - F(x) \right) = \infty, \quad \text{for any} \ \lambda > 0. \]

**Definition 10** (Heavy-tailed distribution) A distribution function \( F \) is heavy-tailed if
\[ \lim_{x \to \infty} \exp \{ \lambda x \} \left( 1 - F(x) \right) = \infty, \quad \text{for any} \ \lambda > 0. \]
Theorem 5 The MLM($\alpha, \beta, \sigma$) distributions, defined in Eqn. (3), are heavy-tailed distributions.

$$\lim_{x \to \infty} \exp \left( \frac{x}{\sigma} \right) \left( 1 - F(x) \right)$$

$$= \lim_{x \to \infty} \exp \left[ \frac{x}{\sigma} - \alpha \log \left( 1 + \frac{x}{\sigma} \right) + \alpha \beta + O \left( \frac{1}{\log \left( 1 + \frac{x}{\sigma} \right)} \right) \right]$$

$$= \infty,$$

since $\log \left( 1 + \frac{x}{\sigma} \right) \approx x^\epsilon$ for any $\epsilon > 0$ and for sufficiently large $x$.

Proof

There are two other important class of distributions (Embrechts et al. 2013), viz. the class $\mathcal{D}$ of dominated-variation distributions and the class $\mathcal{L}$ of long-tailed distributions that are used in the risk theory and queueing theory. The proposed MLM distributions also follows these two properties.

Definition 11 A distribution $F$ belong to the class $\mathcal{D}$ of dominated-variation distributions if

$$\lim_{x \to \infty} \frac{1 - F(x)}{1 - F(2x)} < \infty.$$

Theorem 6 If $\alpha > 0$, then MLM($\alpha, \beta, \sigma$) distribution, defined in Eqn. (3), belongs to the class $\mathcal{D}$ of dominated-variation distributions.

$$\lim_{x \to \infty} \frac{1 - F(x)}{1 - F(2x)}$$

$$= \lim_{x \to \infty} \frac{\exp \left[ - \alpha \log \left( 1 + \frac{x}{\sigma} \right) + \alpha \beta + O \left( \frac{1}{\log \left( 1 + \frac{x}{\sigma} \right)} \right) \right]}{\exp \left[ - \alpha \log \left( 1 + \frac{2x}{\sigma} \right) + \alpha \beta + O \left( \frac{1}{\log \left( 1 + \frac{2x}{\sigma} \right)} \right) \right]}$$

$$= \lim_{x \to \infty} \exp \left[ \alpha \log \left( 1 + \frac{x}{\sigma} \right) + O \left( \frac{1}{\log \left( 1 + \frac{x}{\sigma} \right)} \right) \right]$$

$$+ \frac{1}{\log \left( 1 + \frac{x}{\sigma} \right)} \right) \right]$$

$$= \exp \left( \alpha \log 2 \right)$$

$$= 2^\alpha < \infty.$$

Proof

where $\alpha > 0$.

Definition 12 A distribution $F$ is said to belong to the class $\mathcal{L}$ of long-tailed distributions if $F$ has right unbounded support and, for any fixed $y > 0$,

$$\lim_{x \to \infty} \frac{1 - F(x + y)}{1 - F(x)} = 1.$$
5.2 Performance measures

Here we use some evaluation measures which justify that the degree distribution of a real-world complex network can plausibly been drawn from the proposed MLM distribution. As here the actual distribution is discrete, we can quantify the goodness-of-fit test (i.e., how closely a hypothesized distribution resembles the actual distribution) by calculating the Chi-square statistic value based on bootstrap resampling by generating 50000 synthetic data sets. The Chi-square test will return a $p$ value which quantifies the probability that our data were drawn from the hypothesized distribution. If the $p$ value is small (less than the significance level), we can reject the null hypothesis that the data come from the MLM distribution. We have also computed few other statistical measures such as KL-divergence, RMSE and MAE for quantifying the goodness-of-fit of the proposed MLM distribution model in comparison to the other standard distribution functions related to other heavy-tailed distributions.

5.3 Analysis of results

Table 1 represents some of the statistical measures corresponding to the network data and also provides the statistical evidences of the proposed fitting over the node degree distribution in the whole range using MLM distribution. CV is also calculated corresponding to each of the degree distribution data, and it gives us the sufficient condition for the existence of the global maximum at finite point of the $MLM(\alpha, \beta, \sigma)$ distribution. From Table 1, it is clear that the value of CV is greater than 1 in all the network data sets under consideration. Thus it confirms that the maximum likelihood estimates for the parameters $(\alpha, \beta, \sigma)$ of the proposed MLM distribution attain at the finite points which has been theoretically described in Sects. 4, 4.1. To estimate the parameters $(\alpha, \beta, \sigma)$ of the MLM distribution numerically, we have used “optim” function along with the quasi-Newton L-BFGS-B algorithm in R statistical software by taking the initial parameters value $(\alpha, \beta, \sigma) = (1, 0, 1)$. The estimated values of the parameters for all the data sets satisfied the condition, i.e., $(\alpha > 0, \beta > -1$ and $\sigma > 0$) as clearly seen in Table 1, for the complete characterization of the proposed MLM distribution. Empirically it is observed that in almost all the cases the estimated value of the parameter $\sigma$ attains the higher values as compared to the estimated value of $\alpha$. On the other hand, the estimated value of the parameter $\beta$ lies between $(0, 1)$ lies between $-1$ and $1$ except a few which can be clearly seen from Table 1.

5 Experimental analysis

5.1 Description of data sets

We present here the results of fitting modified Lomax (MLM) distribution over 50 real-world complex networks (Leskovec and Krevl 2014; Rossi and Ahmed 2015) coming from broad variety of different disciplines such as Social Networks, Collaboration Networks, Communication Networks, Citation Networks, Temporal Networks, Web Graphs, Product co-purchasing Networks, Biological Networks, Brain Networks, etc. Please go through the supplementary materials for more details about the data sets under consideration. Some statistical measures of the data sets, and the detailed experimentation of the performances of the proposed MLM distribution compared to the other common power-law related distribution such as Lomax, Pareto, Log-normal, power-law cutoff, Exponential and Poisson are discussed in the following sub sections.
Table 1 Performance of the proposed MLM model over different real-world networks

| Data sets          | No. of nodes | No. of edges | Stat. Prop. | Estimated parameters | Bootstrap Chi-square value (p) |
|--------------------|--------------|--------------|-------------|----------------------|--------------------------------|
| Social Networks    |              |              |             |                      |                                |
| ego-Twitter(In)    | 81,306       | 1,768,149    | 57.965      | 1.9922               | 30.543                         |
| ego-Gplus(In)      | 107,614      | 13,673,453   | 140.48      | 0.7108               | −0.4983                        |
| soc-Slashdot       | 70,068       | 358,647      | 35.069      | 0.8663               | −0.6228                        |
| soc-Delicious(In)  | 536,108      | 1,365,961    | 39.826      | 1.3630               | −0.6819                        |
| soc-Digg(In)       | 770,799      | 5,907,132    | 166.61      | 0.7931               | −0.6928                        |
| soc-Academia       | 200,169      | 1,398,063    | 48.297      | 2.7429               | −0.3737                        |
| LiveJournal(In)    | 4,847,571    | 68,993,773   | 44.969      | 2.6892               | 51.933                         |
| Dogster-Friendship | 426,821      | 8,546,581    | 284.06      | 1.5634               | 0.3108                         |
| Higgs-Twitter(In)  | 456,626      | 14,855,842   | 530.91      | 1.6797               | 36.204                         |
| Artist-Facebook    | 50,615       | 819,307      | 63.427      | 2.0117               | −0.1455                        |
| Athletes-Facebook  | 13,866       | 86,859       | 17.978      | 3.1229               | 21.180                         |
| Citation Networks  |              |              |             |                      |                                |
| cit-HepTh(In)      | 27,770       | 352,807      | 43.139      | 1.8410               | −0.3093                        |
| cit-Patents(In)    | 3,774,768    | 16,518,948   | 6.9125      | 4.8222               | −0.2534                        |
| cit-Citeseer(In)   | 227,320      | 814,134      | 9.8260      | 2.2630               | −0.7288                        |
| Collaboration      |              |              |             |                      |                                |
| ca-CondMat         | 23,133       | 93,497       | 10.671      | 3.0686               | 10.533                         |
| Networks           |              |              |             |                      |                                |
| ca-AstroPh         | 18,772       | 198,110      | 30.568      | 16.434               | 37.276                         |
| ca-GQc             | 5242         | 14,496       | 7.9186      | 2.2624               | 0.7651                         |
| Bio-Mouse-Gene     | 3926         | 7823         | 8.0800      | 2.8780               | 0.6791                         |
| Bio-Dmela          | 45,101       | 14,506,199   | 856.67      | 10.921               | 0.4825                         |
| Bio-WormNet-v3     | 16,347       | 762,822      | 138.17      | 704.71               | 0.9846                         |
| Biological Networks|              |              |             |                      |                                |
| Yeast-PPIN         | 2361         | 7182         | 8.0800      | 10.535               | −0.4527                        |
| Diseasome          | 3926         | 7823         | 9.1009      | 10.9688              | −0.9493                        |
| Bio-Mouse-Gene     | 45,101       | 14,506,199   | 856.67      | 6.3e−08              | 2.1e+00                        |
| Bio-Dmela          | 7393         | 25,569       | 10.782      | 14.979               | −0.0535                        |
| Bio-WormNet-v3     | 16,347       | 762,822      | 138.17      | 704.71               | 0.9846                         |
| Product Networks   |              |              |             |                      |                                |
| amazon0601(In)     | 403,394      | 3,387,388    | 15.279      | 3.8261               | 19.522                         |
| co-purchasing      | 410,236      | 3,356,828    | 15.313      | 3.8367               | 19.984                         |
| networks           |              |              |             |                      |                                |
| Amazon0312(In)     | 400,727      | 3,200,444    | 15.073      | 3.7631               | 18.747                         |
| Temporal           |              |              |             |                      |                                |
| sx-mathoverflow(In)| 24,818       | 506,550      | 31.476      | 1.4452               | 0.8241                         |
Table 1 (continued)

| Data sets          | No. of nodes | No. of edges | Stat. Prop. | Estimated parameters | Bootstrap Chi-square value (p) |
|--------------------|--------------|--------------|-------------|----------------------|------------------------------|
|                    |              |              | s           | μ                    | â     | ß     | θ     |                     |
| Networks sx-stackoverflow(In) | 2,601,977  | 63,497,050  | 186.00      | 27.647               | 6.7278 |       |       | 1.0218 −0.8224 4.4865 | 0.9490 |
| sx-supervisor(In)  | 194,085      | 1,443,339    | 23.782      | 5.8239               | 4.0836 |       |       | 1.7401 2.1405 0.7284 | 0.9780 |
| sx-askubuntu(In)   | 159,316      | 964,437      | 18.404      | 4.3856               | 4.1966 |       |       | 2.1923 2.2069 0.7665 | 0.9300 |
| Communication Email-Enron | 36,692      | 183,830      | 36.100      | 10.021               | 3.6027 |       |       | 1.2417 −0.1275 2.9045 | 0.9641 |
| Networks Wiki-Talk(In) | 2,394,385  | 5,021,410    | 12.259      | 2.1195               | 5.7844 |       |       | 1.5167 −0.2846 0.0016 | 0.9900 |
| Rec-Libimseti(In)  | 220,970      | 17,359,347   | 413.71      | 102.85               | 4.0227 |       |       | 2.5008 −0.8496 33.118 | 0.9670 |
| Ground-truth Wiki-Topcats | 1,791,489  | 28,511,807   | 283.78      | 15.915               | 17.831 |       |       | 1.1811 0.1998 2.6142 | 0.8310 |
| Networks com-Friendster | 65,608,366 | 1,806,067,135 | 137.81      | 55.056               | 2.5031 |       |       | 4.5863 −0.9188 590.01 | 0.9000 |
| com-LiveJournal    | 3,997,962    | 34,681,189   | 42.957      | 17.349               | 2.4759 |       |       | 2.8206 −0.6020 65.638 | 0.7980 |
| com-Orkut          | 3,072,441    | 117,185,083  | 154.78      | 76.281               | 2.0291 |       |       | 3.7049 0.1292 167.93  | 0.9890 |
| com-Youtube        | 1,134,890    | 2,987,624    | 50.754      | 5.2650               | 9.6398 |       |       | 1.6113 8.3355 0.0094  | 0.8410 |
| Brain Human25890-session1 | 177,584   | 15,669,036   | 319.01      | 176.47               | 1.8078 |       |       | 1.6098 −0.2076 168.75 | 0.8710 |
| Networks Human25890-session2 | 723,881   | 158,147,409  | 667.91      | 436.94               | 1.5286 |       |       | 14.423 −0.3466 188.63 | 0.9980 |
| Human25864-session2 | 692,957    | 133,727,516  | 554.48      | 385.96               | 1.4366 |       |       | 16.250 −0.3379 192.178 | 0.9660 |
| Human25913-session2 | 726,197    | 183,978,766  | 446.92      | 258.99               | 1.7256 |       |       | 7.1013 −0.4681 577.98  | 0.9290 |
| Human25886-session1 | 780,185    | 158,184,747  | 558.41      | 405.50               | 1.3771 |       |       | 21.591 −0.3119 269.759 | 0.9768 |
| Data sets                  | MLM          | Pareto          | Power-law        | Lognormal          | Exponential          | Normal          |
|---------------------------|--------------|-----------------|------------------|--------------------|----------------------|-----------------|
| **Table of different statistical measures of different competitive models over real-world networks** |
| **RMSE**                  | **KLD**      | **MAE**         | **RMSE**         | **KLD**            | **MAE**              | **RMSE**        |
| Social Networks          | 16.800       | 0.00819         | 1.3498           | 29.566             | 0.01354              | 2.4701          |
| Social ego-Twitter(In)   | 16.115       | 0.01832         | 1.8213           | 20.45              | 0.03358              | 2.4701          |
| Social ego-Gplus(In)     | 3.157        | 0.001355        | 5.7823           | 91.06              | 0.01007              | 15.067          |
| Social ego-Academia      | 12.864       | 0.001829        | 2.8414           | 24.841             | 0.01291              | 10.074          |
| Social soc-Slashdot      | 16.233       | 0.001351        | 4.5803           | 176.6              | 0.01211              | 247.87          |
| Social soc-Digg(In)      | 3.1203       | 0.010785        | 5.7823           | 91.06              | 0.01007              | 15.067          |
| Social soc-Delicious(In) | 12.921       | 0.001079        | 5.7823           | 91.06              | 0.01007              | 15.067          |
| Social soc-Academia      | 12.864       | 0.001829        | 2.8414           | 24.841             | 0.01291              | 10.074          |
| Social Higgs-Twitter(In)| 11.708       | 0.001079        | 5.7823           | 91.06              | 0.01007              | 15.067          |
| Social Artist-Facebook   | 14.380       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social Athletes-Facebook | 19.821       | 0.001079        | 5.7823           | 91.06              | 0.01007              | 15.067          |
| Social BerkStan(In)      | 71.819       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social Wikipedia2009(In)| 86.510       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social ca-GrQc           | 15.986       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social ca-HepPh          | 36.062       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social ca-HepTh          | 71.819       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social Google2009(In)    | 86.510       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social Yeast-PPIn        | 36.062       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social Yeast2009(In)     | 86.510       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social Hudos(In)         | 71.819       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social Bio-Mouse-Cast     | 7.6919       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social Bio-WarmNet-v3    | 7.6919       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social amazon0312(In)    | 109.983      | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social Temporal-product-networks | 109.983 | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
| Social sx-superserver(In) | 7.6919       | 0.001289        | 4.5788           | 18.07              | 0.01211              | 247.87          |
Furthermore, we leverage one of the popular statistical method, viz. bootstrapping Chi-square test to evaluate the goodness-of-fit test of the proposed MLM distribution. From Table 1, it is clear that the proposed MLM distribution produces higher p values (i.e., closure to 1) in almost all the data sets which suggest that the null hypothesis, i.e., the data drawn from MLM distribution cannot be ruled out at the 0.05 level of significance. This indicates that the observed degree distribution is plausibly drawn from the MLM distribution. Thus from Table 1 it can be concluded that the proposed MLM distribution is effective in modeling the entire degree distribution of real-world complex networks without ignoring some of the lower degree nodes as oppose to the procedure of fitting power law distribution. In addition, we also used some other statistical measures, viz. KLD, RMSE and MAE in order to compare the performance of the proposed MLM distribution with the each of the other common power-law-related distributions as given in the following Tables 2 and 3.

Tables 2 and 3 depict the values of different statistical measures (viz. RMSE, MAE and KLD) which has been used for the measure of performances of the MLM distribution in comparison to the competitive distributions while modeling the data. RMSE and MAE are two different variants, carrying information about the differences between actual and predicted degree frequencies corresponding to a network. Higher similarity between actual and mapped distributions is achieved by generating smaller values of RMSE and MAE. From Tables 2 and 3, it is clear that the proposed MLM distribution provides smaller RMSE and MAE values compared to other competitive distributions in almost all the networks except a few where the power-law cutoff distribution outperforms the others. The worst performance observed for the poisson distribution in minimizing the RMSE and MAE values compared to the other competing distributions over all the real-world networks as clearly seen from Table 3. The Kullback–Leibler divergence (KLD), or relative entropy, is a quantity which measures the dissimilarity between two probability distributions. Thus the smaller value of KLD represents the higher similarity between the actual and the predicted distribution. From Tables 2 and 3, it is clear that the proposed MLM distribution generates smaller KLD values compared to other competitive distributions in almost all the networks except a few where power-law cutoff distribution outperforms the others. This indicates that the observed degree distribution satisfactorily matches the proposed MLM distribution in almost all the networks. Note that, in terms of KLD, the Poisson and exponential distributions always perform worse than the others in all the networks as in the case RMSE and MAE. The performances of power-law and Pareto distributions are alike in terms of KLD, but remain below par compared to the MLM and power-law cutoff distributions. The performance of the proposed MLM distribution is always superior to the competitive in terms of KLD over
| Data sets | Log-normal | Poisson | Exponential | Power-law cutoff |
|-----------|------------|---------|-------------|-----------------|
| Social ego-Twitter(In) | 53.863 | 0.0169 | 2.9494 | 410.93 |
| Networks ego-Gplus(In) | 10.155 | 0.0678 | 0.2523 | 95.967 |
| soc-Slashdot | 69.438 | 0.0552 | 1.9087 | 323.59 |
| soc-Delicious(In) | 237.63 | 0.1058 | 10.549 | 684.36 |
| soc-Digg(In) | 281.34 | 0.0579 | 10.781 | 957.82 |
| soc-Academia | 91.003 | 0.0169 | 2.0921 | 542.38 |
| LiveJournal(In) | 3475.6 | 0.0355 | 70.64 | 13.61K |
| Dogster-Friendship | 42.539 | 0.0272 | 1.1459 | 494.49 |
| Higgs-Twitter(In) | 41.955 | 0.0134 | 0.5995 | 448.91 |
| Artist-Facebook | 44.951 | 0.0189 | 0.8679 | 793.2 |
| Athletes-Facebook | 96.127 | 0.0929 | 2.755 | 752.71 |
| Collaboration | 353.26 | 0.0299 | 2.066 | 219.21 |
| Networks ca-CondMat | 42.665 | 0.0082 | 6.576 | 576.49 |
| ca-AstroPh | 28.399 | 0.0312 | 1.3281 | 289.84 |
| ca-GrQc | 30.184 | 0.0015 | 1.2148 | 93.81 |
| ca-HepTh | 29.993 | 0.0111 | 1.6898 | 88.958 |
| ca-HepPh | 55.618 | 0.0078 | 2.1052 | 370.15 |
| ca-Mat | 1514.5 | 0.0878 | 4.0657 | 2442.6 |
| Collaboration | 2720.9 | 0.0082 | 20.009 | 279.72 |
| Networks Wikipedia2009(In) | 240.18 | 0.0563 | 6.8234 | 672.72 |
| ca-AstroMat | 30.184 | 0.0151 | 1.2148 | 93.81 |
| Collaboration | 353.26 | 0.0299 | 2.066 | 219.21 |
| Networks Yeast(PITP) | 42.665 | 0.0082 | 6.576 | 576.49 |
| ca-GrQc | 28.399 | 0.0312 | 1.3281 | 289.84 |
| ca-HepTh | 30.184 | 0.0015 | 1.2148 | 93.81 |
| ca-HepPh | 55.618 | 0.0078 | 2.1052 | 370.15 |
| Collaboration | 1514.5 | 0.0878 | 4.0657 | 2442.6 |
| Networks Yeast(PITP) | 2720.9 | 0.0082 | 20.009 | 279.72 |
| ca-AstroMat | 240.18 | 0.0563 | 6.8234 | 672.72 |
| Collaboration | 353.26 | 0.0299 | 2.066 | 219.21 |
| Networks Yeast(PITP) | 2720.9 | 0.0082 | 20.009 | 279.72 |
almost all the networks. Thus overall, by considering RMSE, MAE and KLD values, the performance of the proposed MLM distribution for majority of the networks is found to be better than the other competing distributions which suggest that the observed distribution plausibly comes from the proposed MLM distribution. Lomax distribution and power-law with exponential cutoff remained in 2nd and 3rd positions. The best results are presented in bold letters in Tables 2 and 3.

The effectiveness of the proposed MLM distribution can also be verified through the plotting of the fitted results of competitive distributions. For this purpose, the log-log plots of the of the original frequency distribution, the estimated frequency by MLM distribution and the frequency estimated by power-law, pareto, log-normal, power-law cutoff and exponential distributions are drawn for all the networks under consideration. Twenty-four such examples have been provided in Figs. 3, 4, 5, 6, 7, 8, 9 and 10. These are the soc-Academia network, ego-Twitter network, Higgs-Twitter network, ego-Gplus network, cit-HepTh network, cit-Citeseer network, ca-CondMat network, ca-AstroPh network, Web-Google network, web-Hudong network, sx-stackoverflow, sx-mathoverflow, Bio-Dmela network, Bio-Wormnet-V3 network, com-LiveJournal and com-WikiTopcats network. Few more plotted results are also provided in the supplementary section. We have omitted the plot of the poisson distribution due to its poor performances over all the networks. It is visually clear from Figs. 3, 4, 5, 6, 7, 8, 9 and 10 that the proposed MLM distribution provides better fit compared to the other competitive distributions in almost all of the networks since the proposed curve always passes through the middle of the scatter plot of the observed distribution. In a few cases, the power-law cutoff and log-normal provide a better fit than the proposed distribution. It is visually clear from observing the social, biological, brain and citation networks that the entire node degree distribution can be better represented by the MLM distribution compared to other heavy tailed distributions.

Thus the proposed MLM distribution, a modification of the Lomax distribution with non linear exponent in the shape parameter, can be used for effective and efficient modeling of the entire degree distribution of real-world networks without ignoring the lower degree nodes. The proposed MLM distribution provides more flexibility in the degree distribution modeling since the nonnegative shape parameter are assumed to be expressed as a nonlinear function of the data. Empirical results also suggests the effectiveness of the proposed MLM distribution compared to others as depicted through Tables 1, 2 and 3 and Figs. 3, 4, 5, 6, 7, 8, 9 and 10.

6 Conclusion and discussion

In this article, we have proposed a modified Lomax (MLM) distribution derived from a hierarchical family of Lomax distributions for flexible and efficient modeling of the entire
node degree distribution of real-world complex networks. The proposed MLM distribution can be thought of as a generalization of the Lomax distribution with the nonlinear exponent in the shape parameter. We have theoretically established that the MLM distribution is heavy-tailed and right-tailed equivalent to the power-law distribution. Furthermore, we have shown a sufficient condition for the existence of the MLE for the parameters of MLM distribution using the notion of CV. The proposed MLM distribution can find MLE for the parameters at finite points when the value of CV > 1. We also theoretically justified that the MLM distribution is a function with regularly varying tails that belongs to the maximum domain of attraction of the Frechet distribution. We have further studied the asymptotic behaviors of the MLM distribution in this context.

The proposed MLM distribution captures the heavy-tailed and nonlinear behavior of the entire degree distributions of real-world networks in the original and the log-log scale more adroitly. It also enables us to accurately characterize the degree distribution pattern which may have a significant impact on analyzing real-world networks in terms of their social or biological aspects, as the case may be. We have applied the proposed MLM distribution in modeling the entire degree distribution over 50 different real-world empirical data sets taken from diverse fields. Empirical results suggest that as compared to the power-law distribution or any other well-known distribution, our proposed MLM distribution produces a lower fitting error in terms of three statistical tests, viz. RMSE, KL-divergence, and MAE. We also demonstrated the statistical significance of the estimated MLM distribution with the help of the bootstrap Chi-square value. This generalization of the Lomax distribution by adding an additional parameter in the base model results in flexible modeling to the entire degree distribution of a real-world network compared to other heavy-tailed distributions, unlike power-law. The proposed fit distribution sometimes helps us in better characterization of the evolution process of large-scale real-world networks instead of explicitly performing the empirical study at each time step. Thus, by simulating the parameters of a proposed fit MLM distribution, one can easily capture the spatial structure and dynamical pattern of a real-world network as the network evolves over time. The dynamic pattern analysis of such

![Fig. 3 Degree distribution of soc-Academia and ego-Twitter networks in log-log scale](image)

![Fig. 4 Degree distribution of Higgs-Twitter and ego-Gplus networks in log-log scale](image)
Fig. 5 Degree distribution of cit-HepTh and cit-Citeseer networks in log-log scale

Fig. 6 Degree distribution of ca-CondMat and ca-AstroPh networks in log-log scale

Fig. 7 Degree distribution of Web-Google and Web-Hudong networks in log-log scale
Fig. 8 Degree distribution of sx-stackoverflow and sx-mathoverflow networks in log-log scale

Fig. 9 Degree distribution of Bio-Dmela and Bio-Wormnet-V3 networks in log-log scale

Fig. 10 Degree distribution of LiveJournal and Wiki-Topcats networks in log-log scale
structural properties in real-world networks is one of the future scopes of research.

Acknowledgements The authors would like to thank the Editor-in-Chief, the Associate Editor and the reviewers for their valuable comments and detailed suggestions for the improvement of the contents of this article. The authors also gratefully acknowledge the financial assistance received from Indian Statistical Institute (I.S.I.) and Visvesvaraya PhD Scheme awarded by the Government of India.

References

Abdul-Moniem IB (2012) Recurrence relations for moments of lower generalized order statistics from exponentiated Lomax distribution and its characterization. J Math Comput Sci 2(4):999–1011
Ahsanullah M (1991) Record values of the Lomax distribution. Stat Neer 45(1):21–29
Al-Awadhi S, Ghitany M (2001) Statistical properties of poission-Lomax distribution and its application to repeated accidents data. J Appl Stat Sci 10(4):365–372
Albert R, Barabási A-L (2002) Statistical mechanics of complex networks. Rev Mod Phys 74(1):47
Albert R, Jeong H, Barabási A-L (1999) Diameter of the world-wide web. Nature 406(6794):130–131
Albert R, Jeong H, Barabási A-L (2000) Error and attack tolerance of complex networks. Nature 406(6794):378–382
Arnold BC (2015) Pareto distributions. Chapman and Hall/CRC, Boca Raton
Atkinson AB, Harrison AJ (1978) Distribution of personal wealth in Britain, Cambridge Univ Pr
Balakrishnan N, Ahsanullah M (1994) Relations for single and product moments of record values from Lomax distribution. Sankhyā Indian J Stat B 140–146
Barabási A-L (2005) The origin of bursts and heavy tails in human dynamics. Nature 435(7039):207–211
Barabási A-L, Albert R (1999) Emergence of scaling in random networks. Science 286(5439):509–512
Broido AD, Clauset A (2019) Scale-free networks are rare. Nat Commun 10(1):1–10
Bryson MC (1974) Heavy-tailed distributions: properties and tests. Technometrics 16(1):61–68
Chacoma A, Mato G, Kuperman MN (2018) Dynamical and topological aspects of consensus formation in complex networks. Phys A 551:22–38
Chakraborty T, Ghosh K, Das AK (2021) Uncovering patterns in heavy-tailed networks: a journey beyond scale-free. In: Proceedings of the 8th ACM IKDD CODS and 26th COMAD, pp 136–144
Chattopadhyay S, Ghosh K, Das AK (2021) Uncovering patterns in heavy-tailed networks: a journey beyond scale-free. In: Proceedings of the 8th ACM IKDD CODS and 26th COMAD, pp 136–144
Chattopadhyay S, Ghosh K (2019) Finding patterns in the degree distribution of real-world complex networks: going beyond power law. Patt Anal Appl 1–20
Childs A, Balakrishnan N, Moshtreef M (2001) Order statistics from non-identical right-truncated Lomax random variables with applications. Stat Pap 42(2):187–206
Cordeiro GM, Ortega EM, Popović BV (2015) The gamma-Lomax distribution. J Stat Comput Simul 85(2):305–319
Cui Y, Wang X, Eustace J (2014) Detecting community structure via the maximal sub-graphs and belonging degrees in complex networks. Phys A 416:198–207
El-Bassiouny A, Abdo N, Shahen H Exponential Lomax distribution. Int J Comput Appl 121(13)
Embretichs P, Klüppelberg C, Mikosch T (2013) Modelling extremal events: for insurance and finance, vol 33. Springer, New York
Foss S, Korshunov D, Zachary S et al (2011) An introduction to heavy-tailed and subexponential distributions, vol 6. Springer, New York
Giles DE, Feng H, Godwin RT (2013) On the bias of the maximum likelihood estimator for the two-parameter Lomax distribution. Commun Stat Theo Methods 42(11):1934–1950
Golosovsky M (2017) Power-law citation distributions are not scale-free. Phys Rev E 96(3):032306
Hassan AS, Assar SM, Shelbaa A (2016) Optimum step-stress accelerated life test plan for Lomax distribution with an adaptive type-II progressive hybrid censoring. J Adv Math Comput Sci 1–19
Hassan AS, Al-Ghamdi AS (2009) Optimum step stress accelerated life testing for Lomax distribution. J Appl Sci Res 5(12):2153–2164
Holme P (2019) Rare and everywhere: perspectives on scale-free networks. Nat Commun 10(1):1–3
Jayakumar K, Krishnan B, Hamedani G (2020) On a new generalization of pareto distribution and its applications. Commun Stat Simul Comput 49(5):1264–1284
Jones JH, Handcock MS (2003) Sexual contacts and epidemic thresholds. Nature 423(6940):605–606
Klüppelberg C (1988) Subexponential distributions and integrated tails. J Appl Prob 25(1):132–141
Leskovec J, Krevl A (2014) SNAP datasets: stanford large network dataset collection. http://snap.stanford.edu/data
Liljeros F, Edling CR, Amaral LAN, Stanley HE, Åberg Y (2001) The web of human sexual contacts. Nature 411(6840):907–908
Lomax K (1954) Business failures: another example of the analysis of failure data. J Am Stat Assoc 49(268):847–852
Muchnik L, Pei S, Parra LC, Reis SD, Andrade JS Jr, Havlin S, Makse HA (2013) Origins of power-law degree distribution in the heterogeneity of human activity in social networks. Sci Rep 3(1):1–8
Newman ME (2001) The structure of scientific collaboration networks. Proc Nat Acad Sci 98(2):404–409
Newman ME (2003) The structure and function of complex networks. SIAM Rev 45(2):167–256
Newman ME (2005) Power laws, pareto distributions and Zipf’s law. Contemp Phys 46(5):323–351
Nie T, Guo Z, Zhao K, Lu Z-M (2016) The dynamic correlation between degree and betweenness of complex network under attack. Phys A 457:129–137
Rajab M, Aleem M, Nawaz T, Daniyal M On five parameter beta Lomax distribution. J Stat Comput Simul 85(2):305–319
Rossi RA, Ahmed N (2015) The network data repository with interactive graph analytics and visualization. In: Twenty-ninth AAAI conference on artificial intelligence
Sala A, Zheng J, Zhao BY, Gaito S, Rossi GP (2010) Brief announcement: revisiting the power-law degree distribution for social graph analysis. In: Proceedings of the 29th ACM SIGACT-SIGOPS symposium on Principles of distributed computing, pp 400–401
Seshadri M, Machiraju S, Sridharan A, Bolot J, Faloutsos C, Leskovec J (2008) Mobile call graphs: beyond power-law and lognormal distributions. In: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, pp 596–604
Shakibian H, Charkari NM (2018) Statistical similarity measures for link prediction in heterogeneous complex networks. Phys A 501:248–263
Stumpf MP, Porter MA (2012) Critical truths about power laws. Science 335(6069):665–666
Tahir M, Hussain MA, Cordeiro GM, Hamedani G, Mansoor M, Zubair M (2016) The gumbel-Lomax distribution: properties and applications. J Stat Theo Appl 15(1):61–79
Voitalov I, van der Hoorn P, van der Hofstad R, Krioukov D (2019) Scale-free networks well done. Phys Rev Res 1(3):033034
Zarandi FD, Rafsanjani MK (2018) Community detection in complex networks using structural similarity. Phys A 503:882–891