Poor–rich demarcation of Matthew effect on scale-free systems and its application*

Yan Dong(闫 栋)[a)†], Dong Ming(董 明)[b)], Abdelaziz Bouras[c)], and Yu Sui-Ran(于随然)[a)]

[a)] *School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai* 200240, *China*

[b)] *College of Economics & Management, Shanghai Jiao Tong University, Shanghai* 200052, *China*

[c)] *IUT Lumière Technology Institute, Université Lumière Lyon 2, Lyon* 69007, *France*

(Received 22 April 2010; revised manuscript received 1 December 2010)

In a scale-free network, only a minority of nodes are connected very often, while the majority of nodes are connected rarely. However, what is the ratio of minority nodes to majority nodes resulting from the Matthew effect? In this paper, based on a simple preferential random model, the poor–rich demarcation points are found to vary in a limited range, and form a poor–rich demarcation interval that approximates to $k/m \in [3, 4]$. As a result, the (cumulative) degree distribution of a scale-free network can be divided into three intervals: the poor interval, the demarcation interval and the rich interval. The inequality of the degree distribution in each interval is measured. Finally, the Matthew effect is applied to the ABC analysis of project management.

Keywords: Matthew effect, scale-free networks, poor–rich demarcation, project management

PACS: 02.50.–r, 05.65.+b, 89.75.Fb **DOI:** 10.1088/1674-1056/20/4/040205

1. Introduction

The Matthew effect,[1] that is, "the rich get richer and the poor get poorer" phenomenon, has been regarded as the inequality phenomenon of the social wealth distribution for a long time. The famous "80–20 rule", Lorenz curve and Gini coefficient have been widely used to describe these inequality phenomena. But, how did the Matthew effect come into being? In 1965, Price proposed the "cumulative advantage" of citation networks, and found that both their in-degree and out-degree obey the power-law distribution.[2] In 1999, Barabási and Albert[3] observed the decaying power-law distribution in the World Wide Web (WWW), i.e., $p(k) \sim k^{-r}$, and proposed the scale-free model (BA model). The Matthew effect of networks shows the heterogeneity. Hu and Wang[4] defined a heterogeneity index to quantify the degree heterogeneity of any given network. Wu *et al.*[5] proposed a normalized entropy of rank distribution for the heterogeneity measure of scale-free networks.

In reality, many technological networks[6−9] often revealed the scale-free behaviour. That said, some modular engineering networks[10−12] exhibited both the power-law and the exponential behaviours. However, one may want to know whether a specified node in a scale-free network belongs to the poor or the rich. This key question will be analysed in this paper. In Section 2, the BA model is reestablished as a simple preferential random model in order to analyse the Matthew effect. In Section 3, the poor–rich demarcation of the "rich-get-richer" phenomenon is proposed. In Section 4, the connectivity (degree) density is introduced to measure the poor–rich degree of each interval. In Section 5, the application of the Matthew effect in project management is presented. In section 6, conclusions are drawn from the present study.

2. Preferential random model

To describe the growing network dominated by the preferential random mechanism, the attachment probability of the preferential random model is defined on the basis of BA model[3] as follows:

$$\Pi(k_i) = r\frac{k_i}{\sum_{j}^{N-1} k_j} + (1-r)\frac{1}{N-1}. \qquad (1)$$

When $r = 0$, it becomes the growing random model (Model A[13]). When $0 < r < 1$, it will be dominated

*Project supported by the "Shu Guang" Project of Shanghai Municipal Education Commission, China (Grant No. 09SG17), and EU ELINK–East–West Link for Innovation, Networking and Knowledge Exchange (Grant No. 149674-EM-1-2008-1-UK-ERAMUNDUS).

†Corresponding author. E-mail: yd_email@yahoo.com.cn

© 2011 Chinese Physical Society and IOP Publishing Ltd

by the preferential random mechanism. When $r = 1$, it becomes the BA model.[3]

According to the Barabási–Albert continuum theory, the degree distribution $p(k)$ and the cumulative degree distribution $P_{\rm c}(k)$ of this model can be obtained respectively as follows:

(i) when $r = 0$,

$$p(k) = \frac{e}{m}{\rm e}^{-k/m}, \quad \text{and}$$
$$P_{\rm c}(k) = \int_k^{\infty} p(k) = {\rm e}^{1-k/m}; \tag{2}$$

(ii) when $0 < r \leq 1$,

$$p(k) = \frac{(2m)(m+A)^{2/r}}{rm}(k+A)^{-(1+2/r)}, \quad \text{and}$$
$$P_{\rm c}(k) = (\frac{k+A}{m+A})^{-2/r}, \tag{3}$$

where $A = 2m(1/r-1)$. The degree of each node in an actual network is always not constant, but varying in a certain range. However, the parameter m in scale-free model is generally regarded as the average degree $\langle m\rangle$ of all nodes in actual networks for simplicity.

3. Interval division of "rich-get-richer" phenomenon

First, the measuring criterion of the "rich-get-richer" phenomenon should be determined. In Ref. [13], the evolution of the growing random model is dominated only by the random mechanism with a uniform probability. However, there still exists the poor–rich difference rather than the "rich-get-richer" phenomenon in the growing random model when $r = 0$, because the older nodes each tend to have a greater connectivity.[14] In the model, the degree of each node depends on its "seniority" (time of life), which is usually considered to be the equity phenomenon. On this basis, the "rich-get-richer" phenomenon of scale-free networks is investigated.

For Eqs. (2) and (3), let $x = k/m(\geq 1)$, then $P_{\rm c}(k)$ can be written as

$$P_{\rm c}(x, r = 0) = {\rm e}^{1-x} \quad \text{and}$$
$$P_{\rm c}(x, r \neq 0) = \left(\frac{x+B}{1+B}\right)^{-2/r}, \tag{4}$$

where $B = A/m = 2/r - 2$.

3.1. Continuity of cumulative degree distribution

In order to obtain the extreme limit of Eq. (4) under the condition of $r \to +0$ and $P_{\rm c}(x, r \to +0) = \lim\limits_{r\to+0}(1+\frac{x-1}{2/r-1})^{-2/r}$, let $S = \frac{2/r-1}{x-1}$, then we will have

$$P_{\rm c}(x, r \to +0) = P_{\rm c}(x, r = 0). \tag{5}$$

Thus $P_{\rm c}(x, r)$ is continuous in the closed interval $r \in [0, 1]$.

3.2. Intersection of cumulative degree distribution

Since there exists at least one point of intersection $(1, P_c(1))$ between the curves $P_{\rm c}(x, r \neq 0)$ and $P_{\rm c}(x, r = 0)$, at the intersection point there exists

$$P_{\rm c}(x, r \neq 0) - P_{\rm c}(x, r = 0) = 0. \tag{6}$$

By Taylor's mean value theorem, at the intersection point, the following expressions can be obtained:

$$x = 1 \quad \text{and} \quad \frac{x-1}{B+1} - \frac{1}{2}\xi^{-2}\left(\frac{x-1}{B+1}\right)^2 = \frac{r(x-1)}{2}, \tag{7}$$

where $1 < \xi < 1 + (x-1)/(B+1)$. When $0 < r \leq 1$, we have

$$x = 1 + 2\xi^2(B+1)\left(1 - \frac{r(1+B)}{2}\right) > 3 - r \geq 2. \tag{8}$$

Thus curves $P_{\rm c}(x, r \neq 0)$ and $P_c(r = 0)$ have at least another intersection point somewhere at $x > 3 - r$.

When $r = 0$, the preferential mechanism does not play a role, but there is still a poor–rich difference in the evolving network, which originates from the "seniority" or the pure "cumulative advantage"[1] rather than the preferential mechanism. Therefore, the intersection points at $x > 3 - r$ can be regarded as the poor–rich demarcation points. Moreover, a set of poor–rich demarcation points generated with the continuous parameter $r > 0$ will form a continuous poor–rich demarcation interval.

3.3. Poor–rich demarcation point

For the cumulative degree distribution curve of a preferential random model, it is difficult to obtain the analytic solution of its poor–rich demarcation point. However, by a geometric method (Monte Carlo method or numerical analysis), the poor–rich demarcation point ($x \approx 3.51$, $P_{\rm c}(x, r = 1) \approx 0.081$) of BA model can be obtained according to $P_{\rm c}(x, r = 1) = x^{-2}$ and $P_{\rm c}(x, r = 0) = {\rm e}^{1-x}$. In Fig. 1, $P_{\rm c}(x, r \neq 0)$ intersects with $P_{\rm c}(x, r = 0)$ at the marked points, as r

values equal 0.2, 0.5, 0.8 and 1, respectively. According to Eq. (8), $x > 3^- \geq 3$ when $r \to 0$.

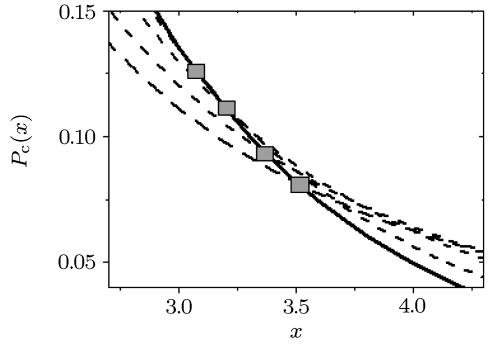


Fig. 1. Points of intersection between $P_c(x, r \neq 0)$ and $P_c(x, r = 0)$.

3.4. Poor–rich demarcation interval

Since $P_c(x, r \neq 0)$ intersects with $P_c(x, r = 0)$, there must be $F(x, r) = P_c(x, r \neq 0) - P_c(x, r = 0) = 0$ and

$$\ln\left(\frac{x+B}{1+B}\right) = \frac{r}{2}(x-1)$$

at the intersection points, where $r \in [0, 1]$. The first partial derivatives $\partial F/\partial r$ and $\partial F/\partial x$ of $F(x, r)$ with respect to r and x can be obtained respectively as follows:

$$\frac{\partial F}{\partial r} = \left(\frac{x+B}{1+B}\right)^{-2/r} \frac{2}{r^2} \times \left[\ln\left(\frac{x+B}{1+B}\right) + \frac{2}{r}\frac{1-x}{(x+\mathrm{B})(1+\mathrm{B})}\right], \quad (9)$$

and

$$\frac{\partial F}{\partial x} = \mathrm{e}^{1-x} - \left(\frac{x+B}{1+B}\right)^{-2/r} \frac{2}{r(x+B)}. \quad (10)$$

For the values of $\partial x/\partial r$ at the intersection points, there exist only three possibilities: $\partial x/\partial r$=0, $\partial x/\partial r < 0$ and $\partial x/\partial r > 0$.

i) If

$$\frac{\partial x}{\partial r} = -\frac{\partial F/\partial r}{\partial F/\partial x} = 0,$$

then $x = 2 + 2/(2 - r)$ (or $x = 1$) under the known condition of

$$\ln\left(\frac{x+B}{1+B}\right) = \frac{r}{2}(x-1).$$

Thus, for any $r \in [0, 1]$, there is always $x \in [3, 4]$. Although actual intersection points in waists are located in the closed interval $x \in [3, 4]$, the monotonic relationship between $P_c(x, r)$ or x and r cannot be determined according to $\partial x/\partial r = 0$.

ii) If $\partial x/\partial r < 0$, then $x > 2 + 2/(2 - r) > 4$ for any $r \in [0, 1]$ under the condition of $\partial F/\partial x > 0$ and $\partial F/\partial r > 0$, but $\partial x/\partial r < 0$ is not true under the condition of $\partial F/\partial x < 0$ and $\partial F/\partial r < 0$.

iii) If $\partial x/\partial r > 0$, then $3 - r < x < 3 + r/(2 - r)$, i.e., $x \in [3, 4]$ for any $r \in [0, 1]$ under the condition of $\partial F/\partial x > 0$, $\partial F/\partial r < 0$ and $x > 3 - r$, but $\partial x/\partial r > 0$ is not true under the condition of $\partial F/\partial x < 0$ and $\partial F/\partial r > 0$. According to Eq. (4), $P_c(x, r)$ monotonically decreases with x, so $y(x, r) \in [\mathrm{e}^{-3}, \mathrm{e}^{-2}]$ for $x \in [3, 4]$.

Since the coordinate variable x of actual intersection point located in the interval $x \in [3, 4]$ monotonically increases with random variable r according to iii) and Subsection 3.3, $x \in [3, 4]$ can be regarded as a poor–rich demarcation interval. In Fig. 2, the whole defined interval $x \in [1, \infty)$ of Eq. (4) is divided into three intervals: the poor interval $x \in [1, 3)$, the (poor–rich) demarcation interval $x \in [3, 4]$, and the rich interval $x \in (4, \infty)$, where the demarcation interval is the fluctuant interval of the poor–rich demarcation points.

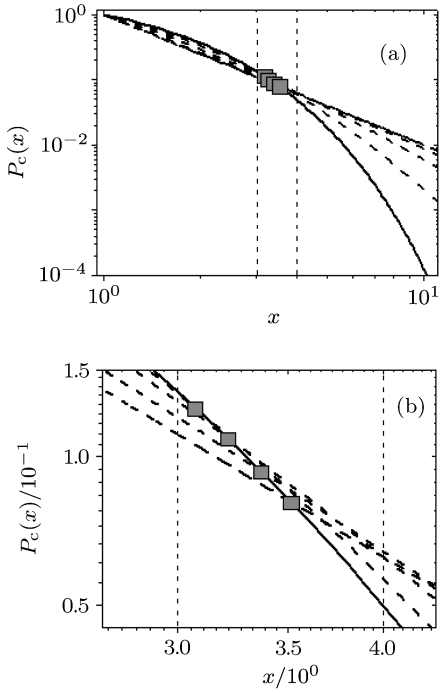


Fig. 2. Poor–rich demarcation intervals of the preferential random model, shown in plots (a) and (b) on log–log scales, respectively.

4. Measure of inequality of the degree distribution

To evaluate the poor–rich degree of each interval, degree density, as an indicator, is defined, which

means the multiple of the average degree, and it equals 1 in the whole defined interval. The inequality phenomenon of the degree distribution of the preferential random model under two extreme conditions of $r = 0$ and $r = 1$ can be measured through three intervals. The same measurement can also be applied under the condition of $0 < r < 1$.

(I) Node ratio: To measure the relative number of individuals in the interval $k \in (am, bm)$, the node ratio is defined as follows:

$$F(am \leq k < bm) = P_{\rm c}(k = am) - P_{\rm c}(k = bm). \quad (11)$$

(II) Degree ratio: The mathematical expectation of the degree of a node in the interval $k \in (am, bm)$ is generally written as

$$E(am < k < bm) = \int_{am}^{bm} kp(k)\mathrm{d}k, \quad (12)$$

and then the degree ratio can be defined as the normalized degree mathematical expectation

$$p_E(am < k < bm) = \frac{E(am < k < bm)}{E(-\infty < k < \infty)}. \quad (13)$$

(III) Degree density: In order to measure the relative poor–rich degree of individuals in a certain interval, the degree density is defined as follows:

$$P_E(am < k < bm) = \frac{p_E(am < k < bm)}{F(am \leq k < bm)}. \quad (14)$$

According to Eqs. (11), (13) and (14), Table 1 shows the poor–rich degrees of the growing random model ($r = 0$) and the BA model ($r = 1$) in three intervals. The degree densities of the growing random model resulting from the "seniority" can be obtained to be 0.84, 1.7 and 2.48 respectively, while these of the BA model from the "seniority" plus the preferential mechanism are 0.75, 1.69 and 3.97. There exists obviously a stronger inequality phenomenon between the three intervals for the BA model, which is the so-called "the poor aid the rich" phenomenon.

Table 1. Node ratios and poor–rich degrees of the growing random and BA models.

		defined interval	poor interval	demarcation interval	rich interval
		$k \in [m, \infty)$	$k \in [m, 3m)$	$k \in [3m, 4m]$	$k \in (4m, \infty)$
growing random model	node ratio	1	$1 - \mathrm{e}^{-2}$	$\mathrm{e}^{-2} - \mathrm{e}^{-3}$	e^{-3}
	degree ratio	1	$1 - 2\mathrm{e}^{-2}$	$2\mathrm{e}^{-2} - 2.5\mathrm{e}^{-3}$	$2.5\mathrm{e}^{-3}$
	degree density	1	0.84	1.7	2.48
BA model	node ratio	1	8/9	7/144	1/16
	degree ratio	1	2/3	1/12	1/4
	degree density	1	0.75	1.69	3.97

5. Application of the Matthew effect in project management

5.1. ABC analysis of project management

The importance of all modules in an engineering system can be classified as A, B and C levels, which is consistent with the ABC classification of human resources (HR): the elementary (class A), the intermediate (class B) and the senior (class C). As a result, the team leader of a project can assign the development tasks with different importances to engineers at different levels as shown in Fig. 3. The number 1 represents an employee at a certain level, who is capable of performing the corresponding task, while the number 0 is on the contrary. Obviously, the optimal scheme lies on the diagonal.

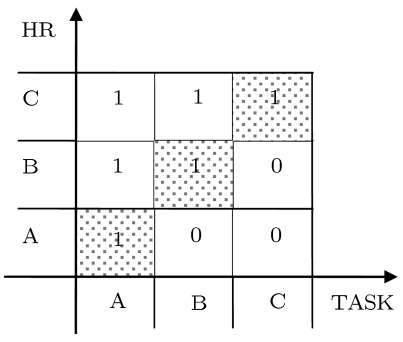


Fig. 3. ABC analysis of HR and task distributions.

5.2. Application examples

The "rich-get-richer" phenomenon and intersections of the out-degree ($k_{\rm out}$) and in-degree ($k_{\rm in}$) distributions of the directed configuration network[15] of an industrial steam turbine are shown in Fig. 4, where the out-degree and the in-degree distributions reveal the exponential and power-law behaviours respectively. In engineering areas, the exponential behaviour of the out-degrees of a system originates from the modularization design technology in nature, while the power-law one of its in-degrees results from reusability.

In this paper, the weights of a system are neglected, or hypothetically proportional to the corresponding degrees. Table 2 shows the poor–rich difference of the system in Fig. 4. The out-degree ratio $E_{\rm out}$ in each interval is approximately proportional to node ratio N, and the average out-degree $\langle k_{\rm out}\rangle$ varies in a small interval, which is conducible to the workload equilibrium. However, the in-degree ratio $E_{\rm in}$ and the average in-degree $\langle k_{\rm in}\rangle$ show a stronger Matthew effect due to reusability. The difference between out-degree and in-degree behaviours can be verified by their Pearson correlation coefficient which is obtained to be –0.014539793. In Table 2, the HR assignment is controlled by Fig. 3, where three HR schemes are proposed, and the quality and quantity of HR are proportional to $\langle k_{\rm in}\rangle$ and $E_{\rm out}$ respectively.

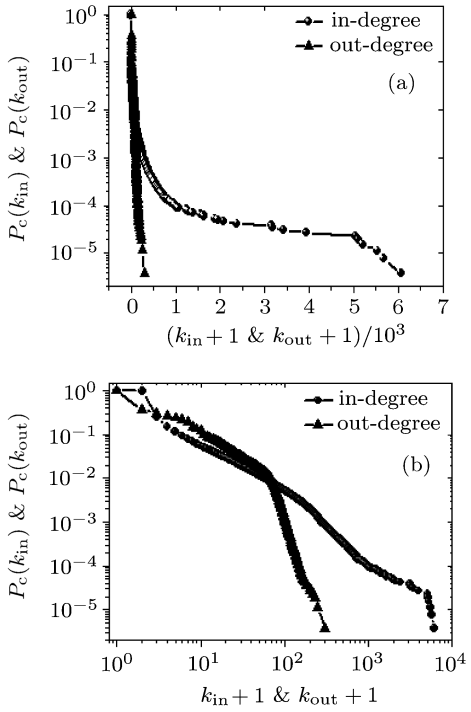


Fig. 4. Cumulative in-degree and out-degree distributions of the turbine directed configuration network on log–linear (a) and log–log (b) scales.

Table 2. Application of the Matthew effect to the ABC analysis of modular systems.

		$[0,+\infty)$	$[0,3m)$	$[3m,4m]$	$[4m,+\infty)$	remark
turbine JGSJ	N	1	0.954	0.012	0.034	
	$E_{\rm out}$	1	0.966	0.011	0.023	$\propto$ workload
	$\langle k_{\rm out}\rangle$	3.5	3.55	3.16	2.38	$\propto \langle$workload$\rangle$
	$E_{\rm in}$	1	0.422	0.041	0.537	$\propto$ reusability
	$\langle k_{\rm in}\rangle$	3.5	1.55	12.29	55.17	$\propto \langle$reusability$\rangle$
HR	scheme 1	A+B+C	A	B	C	HR quantity$\propto E_{\rm out}$ HR quality$\propto \langle k_{\rm in}\rangle$
	scheme 2		A+B		C	
	scheme 3		A	B+C		

6. Conclusions

In this paper, it is found that the cumulative out-degree and in-degree distribution curves of some actual modular systems often intersect at the poor–rich demarcation point. All points of intersection between the cumulative degree distribution curves $P_{\rm c}(x, r \neq 0)$ and $P_{\rm c}(x, r = 0)$ of the preferential random model form a poor–rich demarcation interval. The poor–rich demarcation points and intervals are closely related to parameter r. From the micro perspective, any node in the poor interval has chances to enter into the rich interval in the evolving process of networks. From a macro perspective, each interval does not change with time (network size) due to $x = k/m$. As a result, the quantitative relationship between the minority and the majority in a scale-free network can be determined. Furthermore, the Matthew effect can be applied to the project management of engineering systems.

References

[1] Merton R K 1968 *Science* **159** 56

[2] Price D J de S 1965 *Science* **149** 510

[3] Barabási A L and Albert R 1999 *Science* **286** 509

[4] Hu H B and Wang X F 2008 *Physica* A **387** 3769

[5] Wu J, Tan Y J, Deng H Z and Zhu D Z 2007 *Chin. Phys.* **16** 1576

[6] Newman M E J 2003 *SIAM Rev.* **45** 167

[7] Myers C R 2003 *Phys. Rev.* E **68** 046116

[8] Zhao W, He H S, Lin Z C and Yang K Q 2006 *Acta Phys. Sin.* **55** 3906 (in Chinese)

[9] Wang G Z, Cao Y J, Bao Z J and Han Z X 2009 *Acta Phys. Sin.* **58** 3597 (in Chinese)

[10] Concas G, Marchesi M, Pinna S and Serra N 2007 *IEEE Trans. Software Eng.* **33** 687

[11] Yan D, Qi G N and Gu X J 2006 *Chin. Phys.* **15** 2489

[12] Valverde S, Cancho1 R F and Solé R V 2002 *Europhys. Lett.* **60** 512

[13] Barabási A L, Albert R and Jeong H 1999 *Physica* A **272** 173

[14] Santos F C, Rodrigues J F and Pacheco J M 2006 *Proc. R. Soc.* B **273** 51

[15] Qi G N, Schottner J (Germany), Gu X J and Han Y S 2005 *The Graphic Product Data Management* (Beijing: China Machine Press) p. 133 (in Chinese)