Correlations between power law parameters in complex networks

Nelly Litvak

joint work with
Yana Volkovich, Werner Scheinhardt, Bert Zwart

University of Twente

COST meeting, 19-05-2011

Outline

- Power laws in complex networks
- Dependence between power law graph parameters
- Angular measure
- Example: in-degree and PageRank

Power laws: formal description

- Power laws: Internet, WWW, social networks, biological networks, etc...

Power laws: formal description

- Power laws: Internet, WWW, social networks, biological networks, etc...
- degree of the node = # (it-/out-) links
- [fraction nodes degree at least k] = p_k,
- **Power law:** $p_k \approx const \cdot k^{-\alpha}$, $\alpha > 0$.

Power laws: formal description

- Power laws: Internet, WWW, social networks, biological networks, etc...
- degree of the node = # (it-/out-) links
- [fraction nodes degree at least k] $= p_k$,
- **Power law:** $p_k \approx const \cdot k^{-\alpha}$, $\alpha > 0$.
- **Regular variation:**
  - X is regularly varying random variable with index α
  - $P(X > x) \sim L(x)x^{-\alpha}$ as $x \to \infty$
  - $L(x)$ is slowly varying: for every $t > 0$, $L(tx)/L(x) \to 1$ as $x \to \infty$

Power laws: formal description

- Power laws: Internet, WWW, social networks, biological networks, etc...
- degree of the node = # (it-/out-) links
- [fraction nodes degree at least k] $= p_k$,
- **Power law:** $p_k \approx const \cdot k^{-\alpha}$, $\alpha > 0$.
- **Regular variation:**
    - X is regularly varying random variable with index α
    - $P(X > x) \sim L(x)x^{-\alpha}$ as $x \to \infty$
    - $L(x)$ is slowly varying: for every $t > 0$, $L(tx)/L(x) \to 1$ as $x \to \infty$
- Power law is the model for high variability

Power laws: formal description

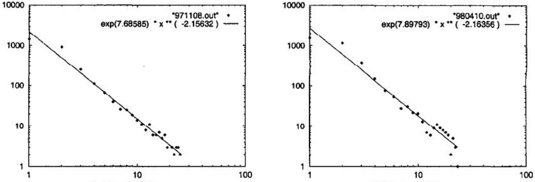
- Power laws: Internet, WWW, social networks, biological networks, etc...
- degree of the node = # (it-/out-) links
- [fraction nodes degree at least k] = p_k,
- **Power law:** $p_k \approx const \cdot k^{-\alpha}$, $\alpha > 0$.
- **Regular variation:**
  - X is regularly varying random variable with index α
  - $P(X > x) \sim L(x)x^{-\alpha}$ as $x \to \infty$
  - $L(x)$ is slowly varying: for every $t > 0$, $L(tx)/L(x) \to 1$ as $x \to \infty$
- Power law is the model for high variability
- $\log p_k = log(const) - \alpha \log k$

Power laws: formal description

- Power laws: Internet, WWW, social networks, biological networks, etc...
- degree of the node = # (it-/out-) links
- [fraction nodes degree at least k] = p_k,
- **Power law:** $p_k \approx const \cdot k^{-\alpha}$, $\alpha > 0$.
- **Regular variation:**
  - X is regularly varying random variable with index α
  - $P(X > x) \sim L(x)x^{-\alpha}$ as $x \to \infty$
  - $L(x)$ is slowly varying: for every $t > 0$, $L(tx)/L(x) \to 1$ as $x \to \infty$
- Power law is the model for high variability
- $\log p_k = log(const) - \alpha \log k$
- Straight line on the log-log scale

Power laws in Internet graphs

- Network of routers (physical)
- Routers are grouped in autonomous systems (AS), or domains
- Faloutsos, Faloutsos, Faloutsos (1999):
  - The degree of the nodes follow power laws, exponent 2.5
  - The degree of the network of domains also follow power laws, exponent 2.1



(a) Int-11-97 (b) Int-04-98

Figure 5: The outdegree plots: Log-log plot of frequency f_d versus the outdegree d.

But Power Law is not everything!

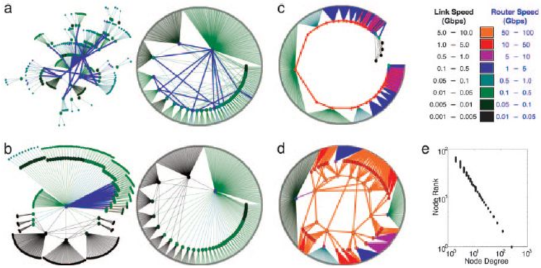
Example: Robustness of the Internet.

- Albert, Jeong and Barabasi (2000): Achille's heel of Internet: Internet is sensitive to targeted attack

But Power Law is not everything!

Example: Robustness of the Internet.

- Albert, Jeong and Barabasi (2000): Achille's heel of Internet: Internet is sensitive to targeted attack
- Doyle et al. (2005): Robust yet fragile nature of Internet: Internet is not a random graph, it is designed to be robust



But Power Law is not everything! (cont.)

Example: Spread of infections

- Classical epidemiology, e.g. Adnerson and May (1991): a virus turns into an epidemic if infection rate exceeds a certain critical value

But Power Law is not everything! (cont.)

Example: Spread of infections

- Classical epidemiology, e.g. Adnerson and May (1991): a virus turns into an epidemic if infection rate exceeds a certain critical value
- Vespignani et al. (2001): power law networks have a zero critical infection rate!

But Power Law is not everything! (cont.)

Example: Spread of infections

- Classical epidemiology, e.g. Adnerson and May (1991): a virus turns into an epidemic if infection rate exceeds a certain critical value
- Vespignani et al. (2001): power law networks have a zero critical infection rate!
- Eguiluz et al. (2002): a specially wired highly clustered network is resistant up to a certain critical infection rate.

How do we measure dependencies in networks

- Newman (2002): assortativity measure r

How do we measure dependencies in networks

- Newman (2002): assortativity measure r
- r is a statistical estimation for the coefficient of variation

$$\rho = \frac{E(XY) - [E(X)]^2}{Var(X)},$$

- X and Y are the degrees of the nodes on the two ends of a randomly chosen edge

How do we measure dependencies in networks

- Newman (2002): assortativity measure r
- r is a statistical estimation for the coefficient of variation

$$\rho = \frac{E(XY) - [E(X)]^2}{Var(X)},$$

- X and Y are the degrees of the nodes on the two ends of a randomly chosen edge
- Measurements by Newman (2002):
  - Social networks: assortative, $r > 0$
  - Engineering networks (Internet, WWW): disassortative, $r < 0$

How do we measure dependencies in networks

- Newman (2002): assortativity measure r
- r is a statistical estimation for the coefficient of variation

$$\rho = \frac{E(XY) - [E(X)]^2}{Var(X)},$$

- X and Y are the degrees of the nodes on the two ends of a randomly chosen edge
- Measurements by Newman (2002):
  - Social networks: assortative, $r > 0$
  - Engineering networks (Internet, WWW): disassortative, $r < 0$
- Problems?

How do we measure dependencies in networks

- Newman (2002): assortativity measure r
- r is a statistical estimation for the coefficient of variation

$$\rho = \frac{E(XY) - [E(X)]^2}{Var(X)},$$

- X and Y are the degrees of the nodes on the two ends of a randomly chosen edge
- Measurements by Newman (2002):
  - Social networks: assortative, $r > 0$
  - Engineering networks (Internet, WWW): disassortative, $r < 0$
- Problems? YES!!!

How do we measure dependencies in networks

- Newman (2002): assortativity measure r
- r is a statistical estimation for the coefficient of variation

$$\rho = \frac{E(XY) - [E(X)]^2}{Var(X)},$$

- X and Y are the degrees of the nodes on the two ends of a randomly chosen edge
- Measurements by Newman (2002):
  - Social networks: assortative, $r > 0$
  - Engineering networks (Internet, WWW): disassortative, $r < 0$
- Problems? YES!!!
  - X and Y are power law r.v.'s, exponent $\alpha - 1$

How do we measure dependencies in networks

- Newman (2002): assortativity measure r
- r is a statistical estimation for the coefficient of variation

$$\rho = \frac{E(XY) - [E(X)]^2}{Var(X)},$$

- X and Y are the degrees of the nodes on the two ends of a randomly chosen edge
- Measurements by Newman (2002):
  - Social networks: assortative, $r > 0$
  - Engineering networks (Internet, WWW): disassortative, $r < 0$
- Problems? YES!!!
  - X and Y are power law r.v.'s, exponent $\alpha - 1$
  - In real networks, $1 < \alpha < 2$, so $E(X) = \sum_k k \frac{kp_k}{E(\text{degree})} = \infty$

How do we measure dependencies in networks

- Newman (2002): assortativity measure r
- r is a statistical estimation for the coefficient of variation

$$\rho = \frac{E(XY) - [E(X)]^2}{Var(X)},$$

- X and Y are the degrees of the nodes on the two ends of a randomly chosen edge
- Measurements by Newman (2002):
  - Social networks: assortative, $r > 0$
  - Engineering networks (Internet, WWW): disassortative, $r < 0$
- Problems? YES!!!
  - X and Y are power law r.v.'s, exponent $\alpha - 1$
  - In real networks, $1 < \alpha < 2$, so $E(X) = \sum_k k \frac{kp_k}{E(\text{degree})} = \infty$
  - ρ is not defined in the power law model!

How do we measure dependencies in networks

- Newman (2002): assortativity measure r
- r is a statistical estimation for the coefficient of variation

$$\rho = \frac{E(XY) - [E(X)]^2}{Var(X)},$$

- X and Y are the degrees of the nodes on the two ends of a randomly chosen edge
- Measurements by Newman (2002):
  - Social networks: assortative, $r > 0$
  - Engineering networks (Internet, WWW): disassortative, $r < 0$
- Problems? YES!!!
  - X and Y are power law r.v.'s, exponent $\alpha - 1$
  - In real networks, $1 < \alpha < 2$, so $E(X) = \sum_k k \frac{kp_k}{E(\text{degree})} = \infty$
  - ρ is not defined in the power law model!
  - Then: what are we measuring?

Google PageRank

- S. Brin, L. Page, The anatomy of a large-scale hypertextual Web search engine (1998)
- Google ranking: use the information contained in the links!

Google PageRank

- S. Brin, L. Page, The anatomy of a large-scale hypertextual Web search engine (1998)
- Google ranking: use the information contained in the links!
- PageRank r_i of page $i = 1, \ldots, n$ is defined as a stationary distribution of a random walk with jumps:

$$R_i = \sum_{j \to i} \frac{c}{d_j} R_j + (1 - c) b_i, \quad i = 1, \ldots, n$$

$d_j = \#$ out-links of page j;
$c \in (0, 1)$, originally 0.85, probability of a random jump;
b_i probability to jump to page i.

Google PageRank

- S. Brin, L. Page, The anatomy of a large-scale hypertextual Web search engine (1998)
- Google ranking: use the information contained in the links!
- PageRank r_i of page $i = 1, \ldots, n$ is defined as a stationary distribution of a random walk with jumps:

$$R_i = \sum_{j \to i} \frac{c}{d_j} R_j + (1-c) b_i, \quad i = 1, \ldots, n$$

  $d_j = \#$ out-links of page j;
  $c \in (0,1)$, originally 0.85, probability of a random jump;
  b_i probability to jump to page i.

- A page is important when many important pages link to it

Google PageRank

- S. Brin, L. Page, The anatomy of a large-scale hypertextual Web search engine (1998)
- Google ranking: use the information contained in the links!
- PageRank r_i of page $i = 1, \ldots, n$ is defined as a stationary distribution of a random walk with jumps:

$$R_i = \sum_{j \to i} \frac{c}{d_j} R_j + (1-c) b_i, \quad i = 1, \ldots, n$$

  $d_j = \#$ out-links of page j;
  $c \in (0,1)$, originally 0.85, probability of a random jump;
  b_i probability to jump to page i.

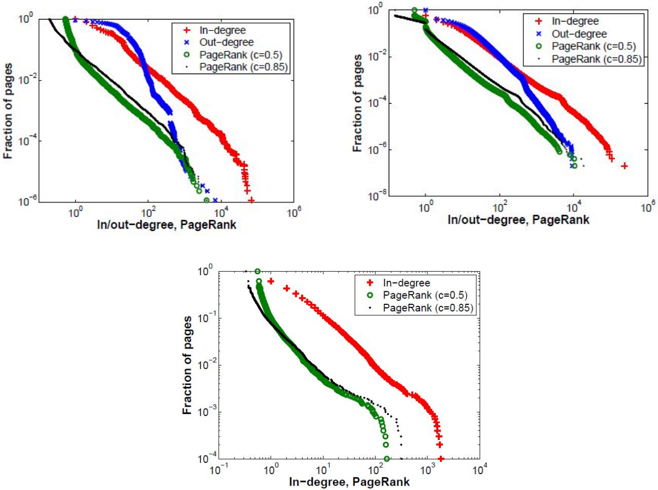
- A page is important when many important pages link to it
- Modifications of PageRank are used in search. Many other applications: clustering, spam detection, measuring node distances, citation analysis, etc.

Power Law behaviour of PageRank

- Pandurangan et al. (2002) PageRank, scaled with the number of pages, $R_i \to nR_i$, has a power law distribution

Power Law behaviour of PageRank

- Pandurangan et al. (2002) PageRank, scaled with the number of pages, $R_i \to nR_i$, has a power law distribution
- Data for Web, Wikipedia and Preferential Attachment graph



Modelling the Power Law behaviour of PageRank

Stochastic equation for PageRank

$$R \stackrel{d}{=} c \sum_{j=1}^{N} \frac{1}{D_j} R_j + c p_0 + (1-c) B$$

- N is the in-degree of the randomly chosen page
- D is the out-degree of page that links to the randomly chosen page
- p_0 is the fraction of pages with out-degree zero
- R_j is distributed as R; N, D, R_j are independent; N and B can be dependent

Modelling the Power Law behaviour of PageRank

Stochastic equation for PageRank

$$R \stackrel{d}{=} c \sum_{j=1}^{N} \frac{1}{D_j} R_j + cp_0 + (1-c)B$$

- N is the in-degree of the randomly chosen page
- D is the out-degree of page that links to the randomly chosen page
- p_0 is the fraction of pages with out-degree zero
- R_j is distributed as R; N, D, R_j are independent; N and B can be dependent

Theorem (V&L 2010)

If $P(B > x) = o(P(N > x))$, *then the following are equivalent:*

- $P(N > x) \sim x^{-\alpha_N} L_N(x)$ *as* $x \to \infty$,
- $P(R > x) \sim C_N x^{-\alpha_N} L_N(x)$ *as* $x \to \infty$,
  where $C_N = (E(c/D))^{\alpha_N}[1 - \mathbb{E}(N)\mathbb{E}((c/D)^{\alpha_N})]^{-1}$

Extremal dependencies

- X, Y are r.v's; F_X, F_Y are distribution functions

Extremal dependencies

- X, Y are r.v's; F_X, F_Y are distribution functions
- $1 - F_X(X) =$ fraction of occurrences of the value $> X$ (rank)

Extremal dependencies

- X, Y are r.v's; F_X, F_Y are distribution functions
- $1 - F_X(X) =$ fraction of occurrences of the value $> X$ (rank)
- Define $(R, \Theta) = \text{POLAR}\left(\frac{1}{1 - F_X(X)}, \frac{1}{1 - F_Y(Y)}\right)$

Extremal dependencies

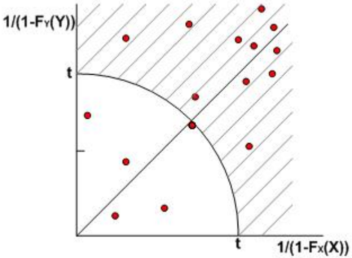
- X, Y are r.v's; F_X, F_Y are distribution functions
- $1 - F_X(X) =$ fraction of occurrences of the value $> X$ (rank)
- Define $(R, \Theta) = \text{POLAR}\left(\frac{1}{1 - F_X(X)}, \frac{1}{1 - F_Y(Y)}\right)$
- Angular measure: $S(A) = \lim_{t \to \infty} tP(R > t, \Theta \in A)$

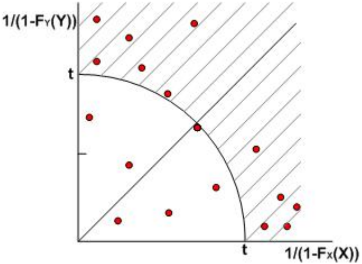
Extremal dependencies

- X, Y are r.v's; F_X, F_Y are distribution functions
- $1 - F_X(X)$ = fraction of occurrences of the value $> X$ (rank)
- Define $(R, \Theta) = \text{POLAR}\left(\frac{1}{1-F_X(X)}, \frac{1}{1-F_Y(Y)}\right)$
- Angular measure: $S(A) = \lim_{t\to\infty} tP(R > t, \Theta \in A)$

Dependence
S concentrated around $\pi/4$

Independence
S concentrated around 0 and $\pi/2$





Statistical procedure

- graph's parameters: $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$
- node j: (X_j, Y_j)

Statistical procedure

- graph's parameters: $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$
- node j: (X_j, Y_j)
- rank transform:
  $\{(X_j, Y_j), j = 1, \ldots, n\} \to \{(r_j^x, r_j^y), j = 1, \ldots, n\}$,

Statistical procedure

- graph's parameters: $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$
- node j: (X_j, Y_j)
- rank transform:
  $\{(X_j, Y_j), j = 1, \ldots, n\} \to \{(r_j^x, r_j^y), j = 1, \ldots, n\}$,
- r_j^x is the descending rank of X_j in $(X_1, \ldots, X_n)$
  r_j^y is the descending rank of Y_j in $(Y_1, \ldots, Y_n)$

Statistical procedure

- graph's parameters: $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$
- node j: (X_j, Y_j)
- rank transform:
  $\{(X_j, Y_j), j = 1, \ldots, n\} \rightarrow \{(r_j^x, r_j^y), j = 1, \ldots, n\}$,
- r_j^x is the descending rank of X_j in $(X_1, \ldots, X_n)$
  r_j^y is the descending rank of Y_j in $(Y_1, \ldots, Y_n)$
- $(R_{j,k}, \Theta_{j,k}) = \text{POLAR}\left(\frac{k}{r_j^x}, \frac{k}{r_j^y}\right)$
- $k = 1, \ldots, n$ 'upper' order statistics

Statistical procedure

- graph's parameters: $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$
- node j: (X_j, Y_j)
- rank transform:
  $\{(X_j, Y_j), j = 1, \ldots, n\} \to \{(r_j^x, r_j^y), j = 1, \ldots, n\}$,
- r_j^x is the descending rank of X_j in $(X_1, \ldots, X_n)$
  r_j^y is the descending rank of Y_j in $(Y_1, \ldots, Y_n)$
- $(R_{j,k}, \Theta_{j,k}) = \mathrm{POLAR}\left(\frac{k}{r_j^x}, \frac{k}{r_j^y}\right)$
- $k = 1, \ldots, n$ 'upper' order statistics
- empirical distribution of Θ for k largest values of R
- cumulative distribution function $\{\Theta_{j,k} : R_{j,k} > 1\}$

Statistical procedure

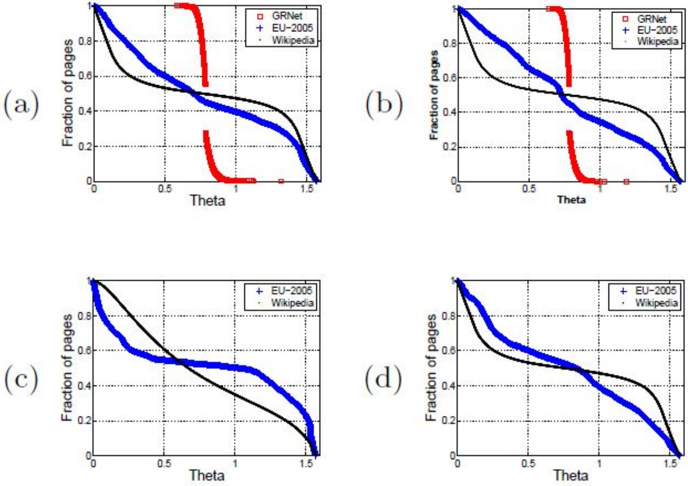
- graph's parameters: $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$
- node j: (X_j, Y_j)
- rank transform:
  $\{(X_j, Y_j), j = 1, \ldots, n\} \to \{(r_j^x, r_j^y), j = 1, \ldots, n\}$,
- r_j^x is the descending rank of X_j in $(X_1, \ldots, X_n)$
  r_j^y is the descending rank of Y_j in $(Y_1, \ldots, Y_n)$
- $(R_{j,k}, \Theta_{j,k}) = \text{POLAR}\left(\frac{k}{r_j^x}, \frac{k}{r_j^y}\right)$
- $k = 1, \ldots, n$ 'upper' order statistics
- empirical distribution of Θ for k largest values of R
- cumulative distribution function $\{\Theta_{j,k} : R_{j,k} > 1\}$
- We measure correlations on fraction k/n of the data.

Statistical procedure

- graph's parameters: $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$
- node j: (X_j, Y_j)
- rank transform:
  $\{(X_j, Y_j), j = 1, \ldots, n\} \to \{(r_j^x, r_j^y), j = 1, \ldots, n\}$,
- r_j^x is the descending rank of X_j in $(X_1, \ldots, X_n)$
  r_j^y is the descending rank of Y_j in $(Y_1, \ldots, Y_n)$
- $(R_{j,k}, \Theta_{j,k}) = \text{POLAR}\left(\frac{k}{r_j^x}, \frac{k}{r_j^y}\right)$
- $k = 1, \ldots, n$ 'upper' order statistics
- empirical distribution of Θ for k largest values of R
- cumulative distribution function $\{\Theta_{j,k} : R_{j,k} > 1\}$
- We measure correlations on fraction k/n of the data.
  But same applies to power laws!

Dependencies between parameters of a node

- Measurements (Volkovich et al., 2008)



- (a) In-PR ($c = 0.85$), (b) In-PR ($c = 0.5$), (c) In-Out, (d) Out-PR

And how is our branching model doing?

$$R \stackrel{d}{=} \sum_{i=1}^{N} A_i R_i + B$$

And how is our branching model doing?

$R \stackrel{d}{=} \sum_{i=1}^{N} A_i R_i + B$

Lemma

As $u \to \infty$, for any constant $C > 0$,

$$P(N > u, R > Cu) \sim \min\{\bar{F}_N(u), (\mathbb{E}A/C)^{\alpha}\bar{F}_N(u)\}.$$

And how is our branching model doing?

$R \overset{d}{=} \sum_{i=1}^{N} A_i R_i + B$

Lemma

As $u \to \infty$, for any constant $C > 0$,

$$P(N > u, R > Cu) \sim \min\{\bar{F}_N(u), (\mathbb{E}A/C)^{\alpha} \bar{F}_N(u)\}.$$

'Proof':
Recall $R \overset{d}{=} \sum_{i=1}^{N} A_i R_i + B$.
By the SLLN we have $R \approx \mathbb{E}A \cdot N$ when N is large.
Hence:

- When $\mathbb{E}A > C$, the event $\{R > Cu\}$ is 'implied' by $\{N > u\}$, leading to $\mathrm{P}(N > u)$
- When $\mathbb{E}A < C$, N needs to be larger for $R > Cu$ to hold, leading to $\mathrm{P}(N > Cu/\mathbb{E}A)$.

Obtaining tail dependence

$\bar{F}_N(x) = P(N > x)$, $\bar{F}_R(x) = P(R > x) \sim C_N \bar{F}_N(x)$

Obtaining tail dependence

$$\bar{F}_N(x) = P(N > x),\ \bar{F}_R(x) = P(R > x) \sim C_N \bar{F}_N(x)$$

Theorem (L et al. 2009)

The function $r(x, y)$ for N and R is given by

$$r(x,y) := \lim_{t \to 0} t^{-1} P(\bar{F}_N(N) \le tx, \bar{F}_R(R) \le ty) = \min\left\{x, \frac{y(\mathbb{E}A)^{\alpha}}{C_N}\right\}$$

Obtaining tail dependence

$$\bar{F}_N(x) = P(N > x),\ \bar{F}_R(x) = P(R > x) \sim C_N \bar{F}_N(x)$$

Theorem (L et al. 2009)

The function $r(x,y)$ for N and R is given by

$$r(x,y) := \lim_{t\to 0} t^{-1} P(\bar{F}_N(N) \le tx, \bar{F}_R(R) \le ty) = \min\left\{x, \frac{y(\mathbb{E}A)^\alpha}{C_N}\right\}$$

'Proof': For fixed $x, y > 0$,

$$\begin{aligned}
& P(\bar{F}_N(N) \le tx, \bar{F}_R(R) \le ty) \\
&\quad = \ P(N \ge \bar{F}_N^{-1}(tx), R \ge \bar{F}_R^{-1}(ty)) \\
&\quad = P\left(N \ge \bar{F}_N^{-1}(tx), R \ge \left(\frac{y}{C_N x} \frac{L(\bar{F}_1^{-1}(tx))}{L(\bar{F}_R^{-1}(ty))}\right)^{-1/\alpha} \bar{F}_N^{-1}(tx)\right) \\
&\quad \sim P\left(N \ge \bar{F}_N^{-1}(tx), R \ge \left(\frac{y}{C_N x}\right)^{-1/\alpha} \bar{F}_N^{-1}(tx)\right)
\end{aligned}$$

Analytical derivation of the angular measure

- Extreme value theory: A unique (nonnegative) measure $S(\cdot)$ exists on $\Xi = \{\omega \in \mathbb{R}^2_+ \ : \ ||\omega||_1 = 1\}$ s.t.

$$r(x, y) = \int_0^1 \min\{wx, (1-w)y\} S(dw),$$

$$\int_0^1 wS(dw) = \int_0^1 (1-w)S(dw) = 1.$$

Analytical derivation of the angular measure

- Extreme value theory: A unique (nonnegative) measure $S(\cdot)$ exists on $\Xi = \{\omega \in \mathbb{R}^2_+ \ : \ ||\omega||_1 = 1\}$ s.t.

$$r(x,y) = \int_0^1 \min\{wx, (1-w)y\} S(dw),$$

$$\int_0^1 w S(dw) = \int_0^1 (1-w) S(dw) = 1.$$

- From this and the shape of $r(x,y)$ we get a two-point measure

Analytical derivation of the angular measure

- Extreme value theory: A unique (nonnegative) measure $S(\cdot)$ exists on $\Xi = \{\omega \in \mathbb{R}^2_+ \ : \ ||\omega||_1 = 1\}$ s.t.

$$r(x,y) = \int_0^1 \min\{wx, (1-w)y\} S(dw),$$

$$\int_0^1 wS(dw) = \int_0^1 (1-w)S(dw) = 1.$$

- From this and the shape of $r(x,y)$ we get a two-point measure

Theorem

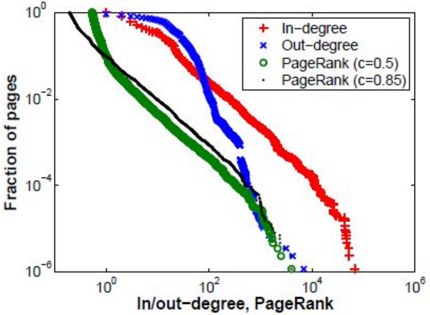
The angular measure with respect to the $||\cdot||_1$ norm of N and R is a two-point measure, with masses

$$S(0) = 1 - \frac{(\mathbb{E}A)^\alpha}{C_N} \qquad \text{in } 0,$$

$$S(a) = 1 + \frac{(\mathbb{E}A)^\alpha}{C_N} \qquad \text{in } a = \frac{C_N}{C_N + (\mathbb{E}A)^\alpha}.$$

Numerical results: Web

- EU-2005 data set due to the Laboratory for Web Algorithmics (LAW) of the Universit‘a degli studi di Milano, Boldi and Vigna (2004)
- Total of 862,664 nodes and 19,235,140 links
- Fitting gives $\alpha = 1.1$, both for In-degree and PageRanks, see log-log plots, with $c = 0.85$ and $c = 0.5$



Numerical examples: Web data

With the graph parameters, $d = 22.2974$, our results give the following angular measure:

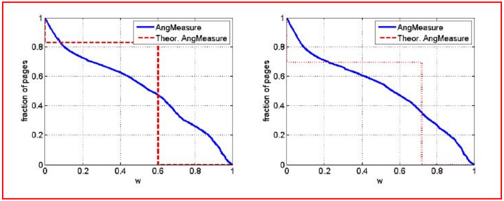
c	a_c	$S(a_c)/2$
0.5	0.6031	0.8290
0.85	0.7210	0.6934

Numerical examples: Web data

With the graph parameters, $d = 22.2974$, our results give the following angular measure:

c	a_c	$S(a_c)/2$
0.5	0.6031	0.8290
0.85	0.7210	0.6934

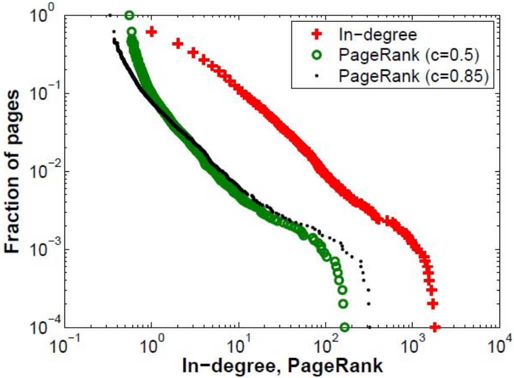
Comparison for $c = 0.5$ and for $c = 0.85$, respectively:



Interpretation of S is that high PR is due to high in-degree or a high PR of the neighbors. But reality is more complex...

Numerical results: PA graph

- Network of 10.000 nodes
- Constant out-degree $d = 8$
- With prob. 0.1, new node links to random page, with prob. 0.9, new node follows the preferential attachment rule



Numerical examples: Growing network

- Assuming $P(R_i > u) = o(P(N > u))$ we derive a one-point measure:

$$a = 1/2, \quad S(a) = 2, \quad S(0) = 0$$

Numerical examples: Growing network

- Assuming $P(R_i > u) = o(P(N > u))$ we derive a one-point measure:

$$a = 1/2, \quad S(a) = 2, \quad S(0) = 0$$

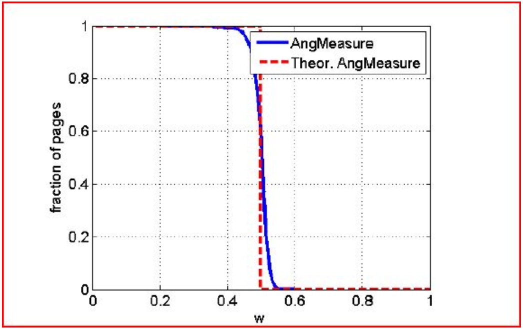
- Large PageRank always due to large In-degree

Numerical examples: Growing network

- Assuming $P(R_i > u) = o(P(N > u))$ we derive a one-point measure:

$$a = 1/2, \quad S(a) = 2, \quad S(0) = 0$$

- Large PageRank always due to large In-degree
- Empirical and theoretical measures:



Conclusions

- Measuring dependencies in complex networks is a largely unresolved problem

Conclusions

- Measuring dependencies in complex networks is a largely unresolved problem
- Solution is needed, because mixing patterns play important role in network processes

Conclusions

- Measuring dependencies in complex networks is a largely unresolved problem
- Solution is needed, because mixing patterns play important role in network processes
- The models such as PA do not reflect the dependencies properly

Conclusions

- Measuring dependencies in complex networks is a largely unresolved problem
- Solution is needed, because mixing patterns play important role in network processes
- The models such as PA do not reflect the dependencies properly
- Extremal dependencies is a promising start for rigorous modelling and analysis