Unified centrality measure of complex networks

Soon-Hyung Yook, Sungmin Lee, Yup Kim
Kyung Hee University

Overview

- Introduction
  - centrality measure
    - shortest path betweenness centrality
    - random walk betweenness centrality
- biased random walk betweenness centrality
  - analytic results
  - numerical simulations
- First systematic study on the edge centrality
- summary and discussion

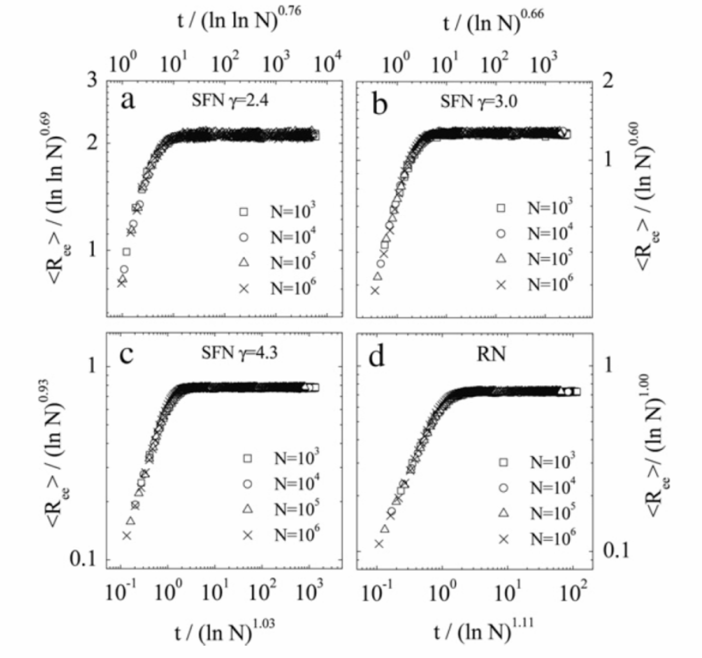
Random walk

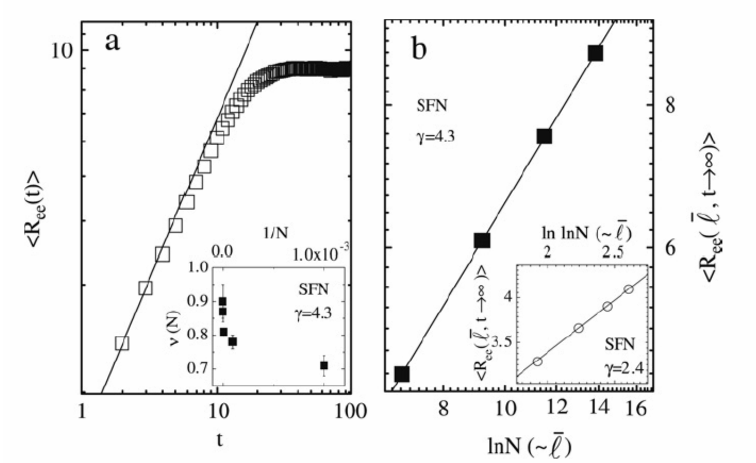
- Many properties of dynamical systems on complex networks are different from those expected by simple mean-field theory
  - Due to the heterogeneity of the underlying topology.
- The dynamical properties of random walk provide some efficient methods to uncover the topological properties of underlying networks

$$\langle R_{ee}(\overline{\ell}, t)\rangle = \overline{\ell}^{\alpha}(\gamma, N)\, g(t/\overline{\ell}^{z})$$

$$g(x) \sim \begin{cases} x^{\nu}, & x \ll 1 \\ \text{const.}, & x \gg 1. \end{cases}$$

Using the finite-size scaling of $<R_{ee}>$
One can estimate the scaling behavior of diameter

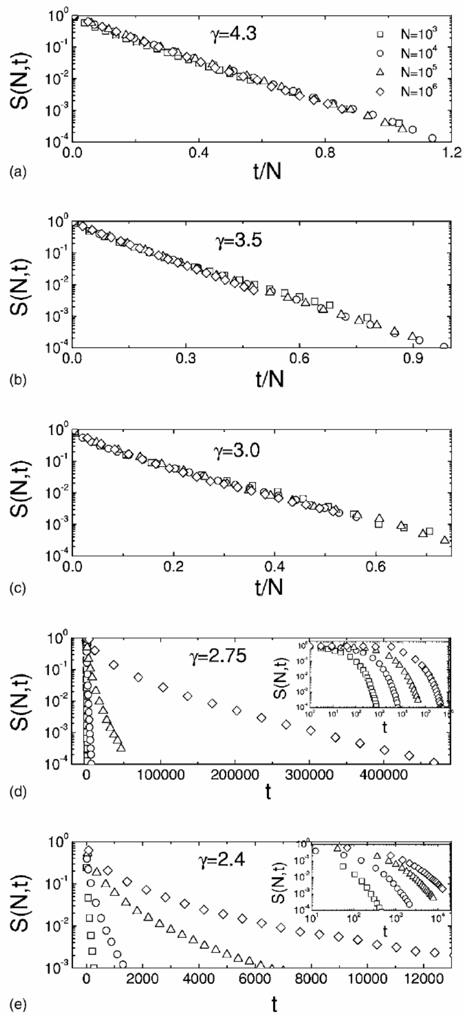


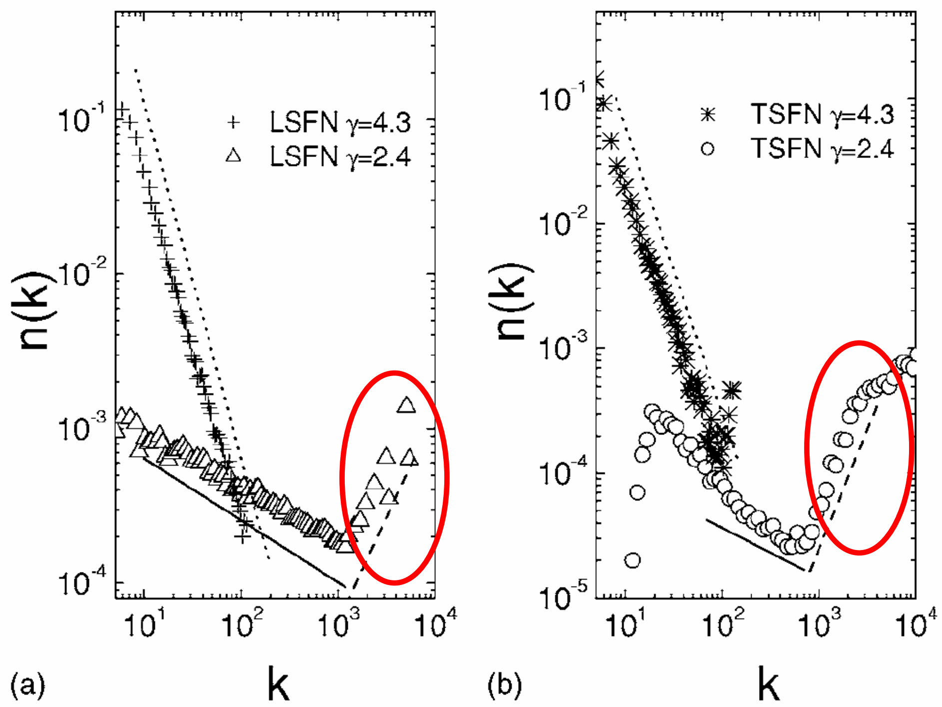


Lee et al, Physica A **387**, 3033 (2008)

Random Walks

- Diffusive capture process
  - Related to the first passage properties of random walker





Nodes of large degrees plays a important role.→ exists some important components [Lee et al. PRE **74** 046118 (2006)]

Centrality

- Centrality: importance of a vertex and an edge
  - The simplest one: degree, k_i
  - Node and edge importance based on adjacency matrix eigenvalue
    - Restrepo, Ott, Hund PRL **97**, 094102
  - Shortest path betweenness centrality (SPBC)
    - b_i: fraction of shortest path between pairs of vertices in a network that pass through vertex i.

$$b_i = \sum_{h<j} \frac{L_{h,i,j}}{L_{h,j}} \frac{2}{N(N-1)}$$

    - h (j): starting (targeting) vertex
    - Total amount of traffic that pass through a vertex
  - Random walk betweenness centrality (RWBC)
  - Information flow between the molecules of a biological networks
    - Provides insight for both the network structure and functions performed by the networks
    - Ex. Diffusion distance in protein-protein interaction networks can be used to predict possible interactions between proteins
  - In food webs: energy transfer between different level of the web is crucial for the organism survival.
  - Other applications:
    - transport of information in the Internet
    - Spreading of disease and/or rumors in social networks

Shortest Path Betweenness Centrality for a vertex

- SPBC distribution: $P(b_v) \sim b_v^{-\delta_v}$

  $\delta_v \simeq 2.2$ on scale-free network with degree exponent $2 < \gamma \leq 3$

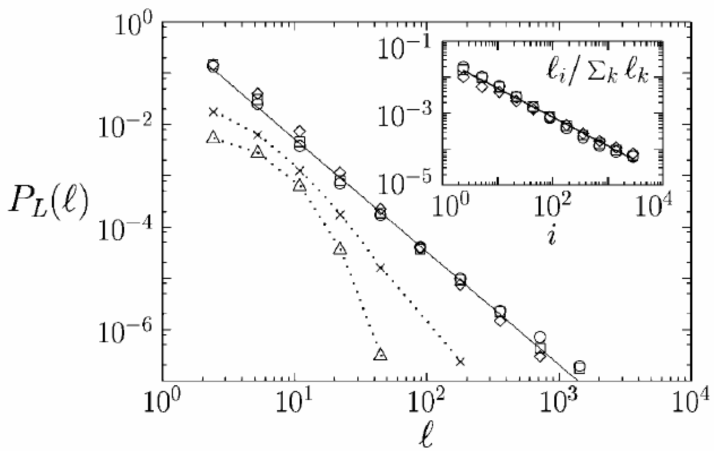


FIG. 1. Plot of the load distribution $P_L(\ell)$ versus ℓ for various $\gamma = 2.2$ ($\circ$), 2.5 ($\square$), 3.0 ($\diamond$), 4.0 ($\times$), and ∞ ($\triangle$) in double logarithmic scales. The linear fit (solid line) has a slope -2.2. Data for $\gamma > 3.0$ are shifted vertically for clearance. Dotted lines are guides to the eye. Simulations are performed for $N = 10\,000$ and $m = 2$ and all data points are log-binned, averaged over ten configurations. Inset: Plot of the normalized load $\ell_i / \sum_k \ell_k$ versus vertex index i in double logarithmic scales for various $\gamma = 2.2$ ($\circ$), 2.5 ($\square$), and 3.0 ($\diamond$).

[Goh et al. PRL **87**, 278701 (2001)]

and $b \sim k^{(\gamma-1)/(\delta_v-1)}$

Betweenness Centrality

- SPBC and RWBC

[Newman, Social Networks **27**, 39 (2005)]

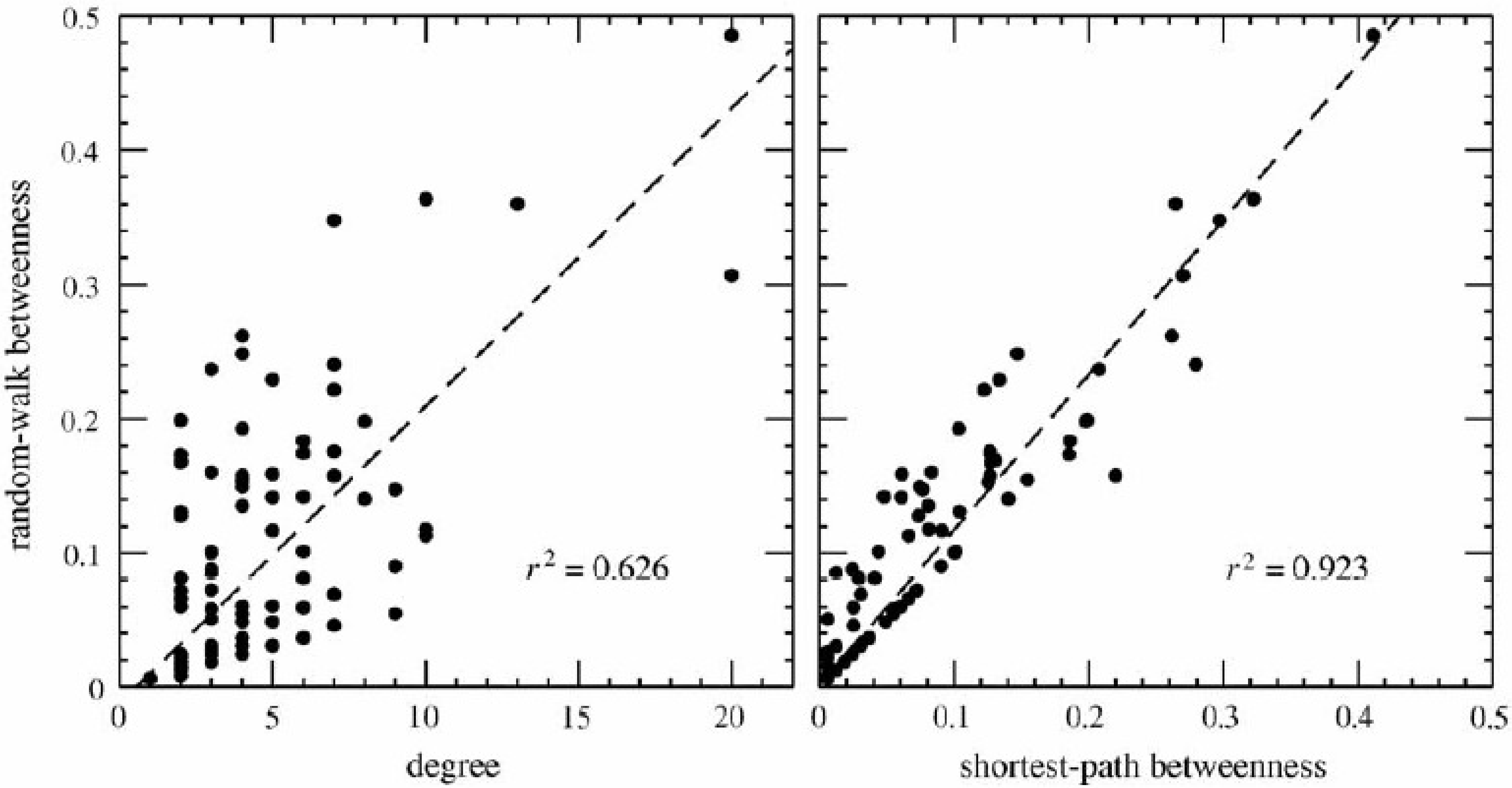


Fig. 4. Scatter plots of the random-walk betweenness of vertices in the sexual contact network of Fig. 5, against vertex degree (left) and standard shortest-path betweenness (right). The dotted lines indicate the best linear fits in each case, which have the correlation coefficients indicated.

Random Walk Betweenness Centrality

- RWBC can find some vertices which do not lie on many shortest paths [Newman, Social Networks **27**, 39 (2005)]

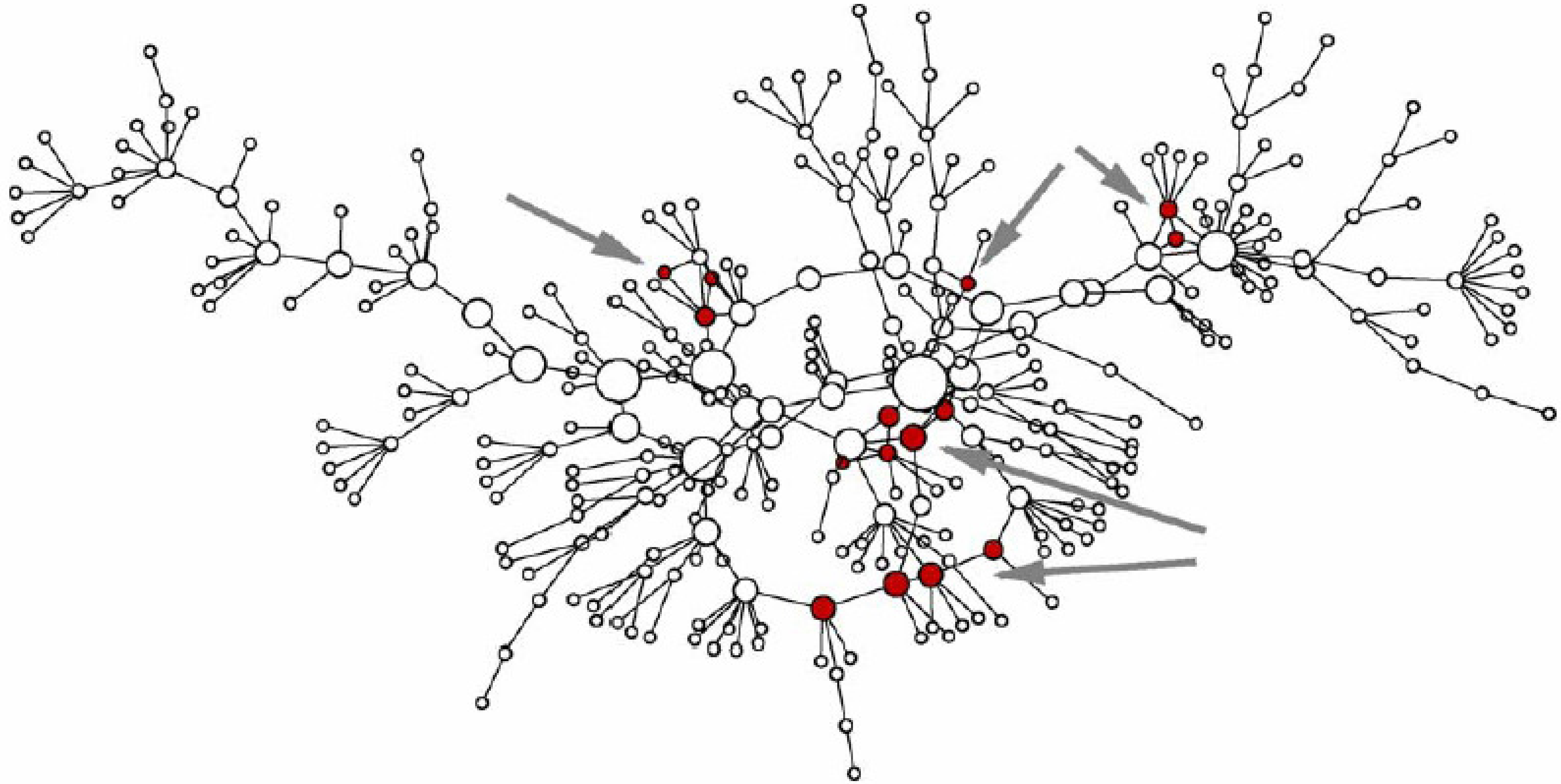


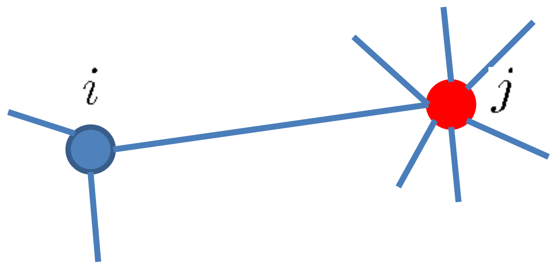
Fig. 5. The largest component of a network of sexual contacts between high-risk actors in the city of Colorado Springs, CO, as reconstructed by Potterat et al. (2002). The size of the vertices increases linearly with their random-walk betweenness, as defined in this paper. The shaded vertices (also indicated by the arrows) are those for which the random-walk betweenness is substantially greater than shortest-path betweenness (a factor of two or more).

Biased Random Walk Betweenness Centrality (BRWBC)

- Generalize the RWBC by biased random walker

jumping probability from a vertex i to another one j

$$P_{i\to j} \propto k_j^{\alpha}$$



- Count the number of traverse, N_T, of a vertex or an edge
- N_T: the measure of BRWBC
- Note that both RWBC and SPBC depend on k

Relationship between BRWBC and SPBC for vertices

- Kwon et al. PRE **77**, 066105 (2008)
  - Mapping to the weight network with weight

$$w_{ij} = (k_i k_j)^\alpha$$

$$P_k^\infty \sim \sum_{i=1}^{\infty} P_i^\infty \delta_{k_i,k} = \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ji} (k_i k_j)^\alpha \delta_{k_i,k}$$

$$P_k^\infty \sim k^{\alpha+1-\gamma}$$

- Therefore, $N_T(k)$ also scales as

$$N_T(k) \sim k^{\alpha+1-\gamma}$$

- Average number of traverse of a vertex having degree k

$$\boxed{n_v(k) \equiv \frac{N_T(k)}{N_v(k)} \sim \frac{P^\infty(k)}{P(k)} \sim k^{\alpha+1} \sim k^{\nu_v}}$$

- $N_v(k)$: number of vertices having degree k

$$N_v(k) \sim k^{-\gamma}$$

Relationship between BRWBC and SPBC for vertices

- SPBC; $b_v(k)$

$$\boxed{b_v(k) \sim k^{\eta_v}}$$

$$\eta_v = (\gamma - 1)/(\delta_v - 1) \text{ for } 2 < \gamma \leq 3$$

thus,

$$\boxed{n_v(k) \sim b_v(k)^{\beta_v}}$$

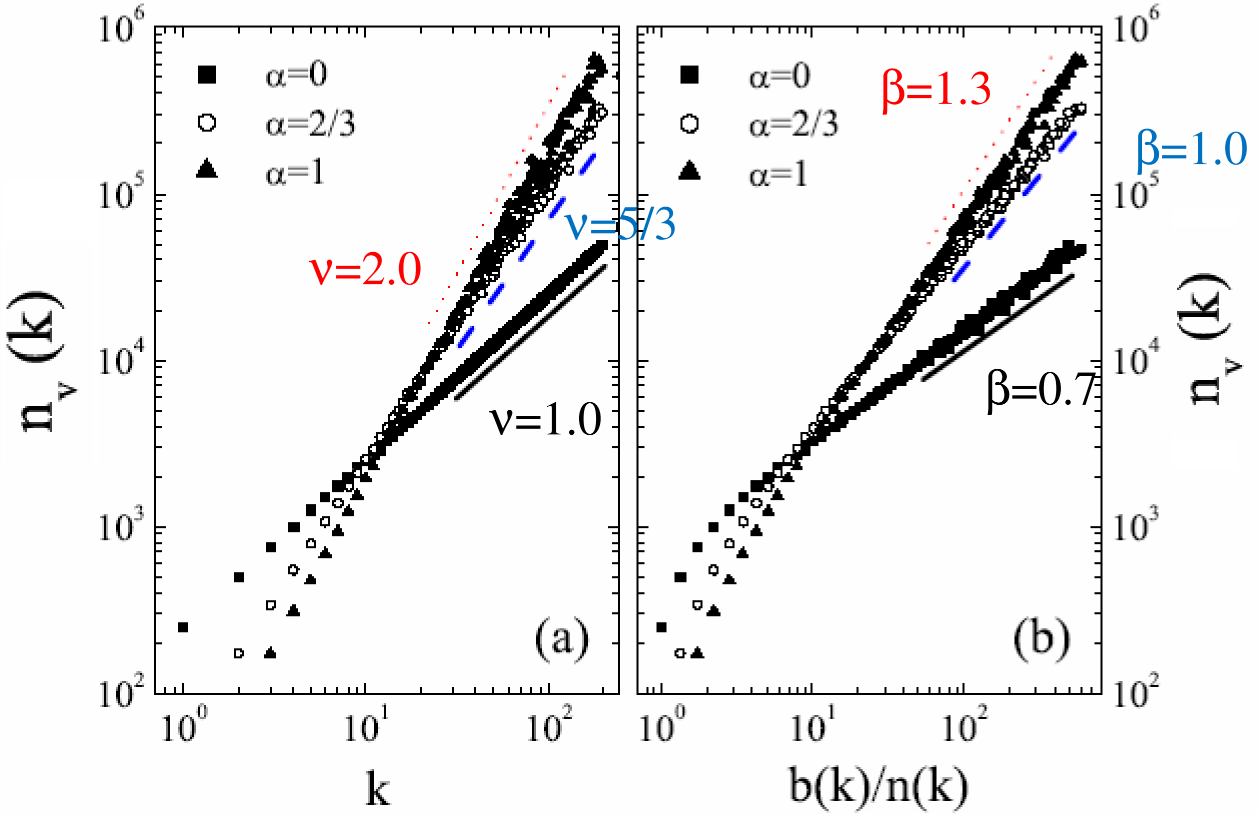
$$\boxed{\beta_v = \nu_v / \eta_v = \frac{(\alpha + 1)(\delta_v - 1)}{(\gamma - 1)}} \quad \text{when } 2 < \gamma \leq 3$$

But in the numerical simulations, we find that this relation holds for $\gamma > 3$

Relationship between BRWBC and SPBC for vertices

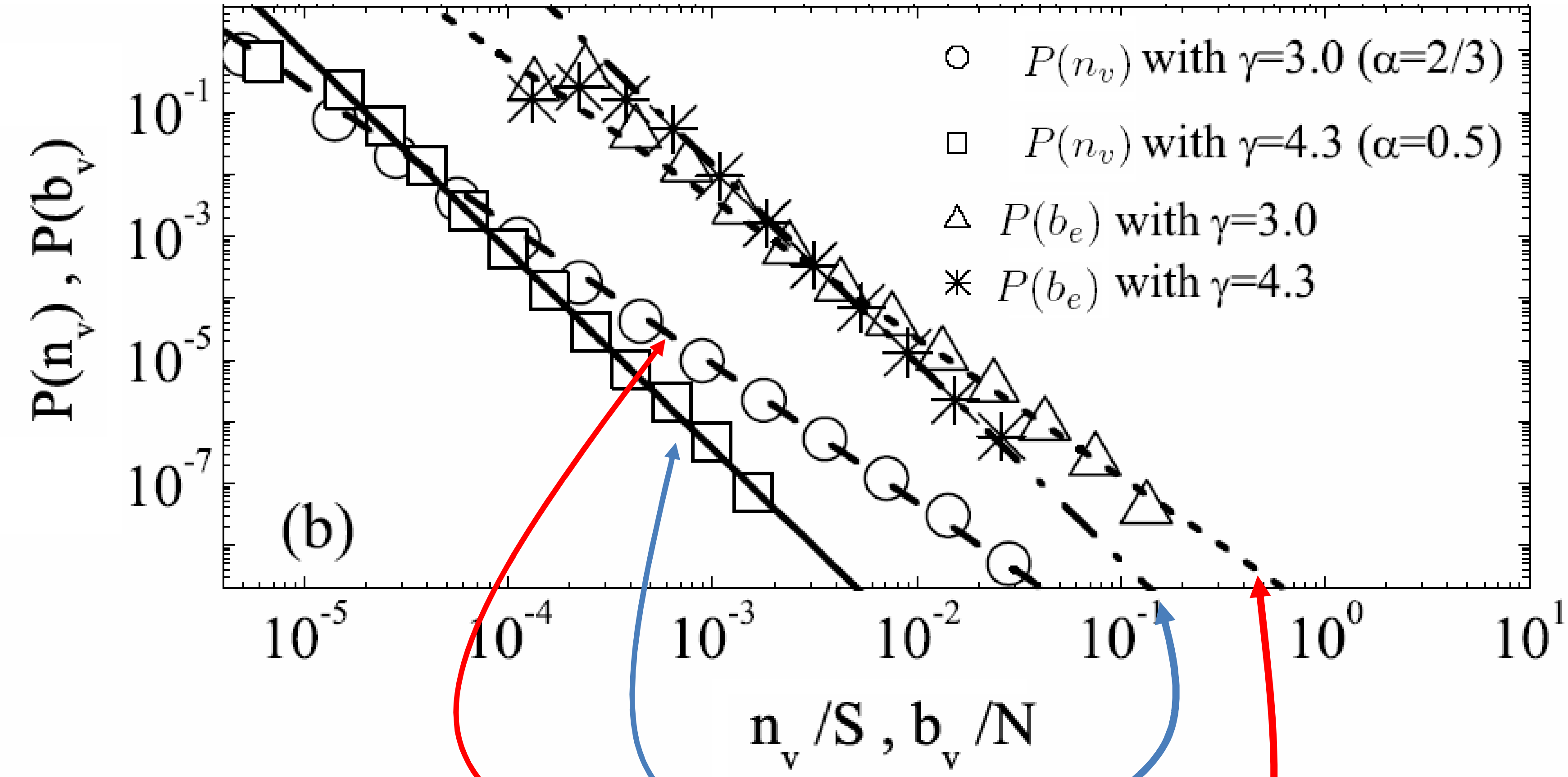
$$\gamma = 3.0$$

$$n_v(k) \equiv \frac{N_T(k)}{N_v(k)} \sim \frac{P^{\infty}(k)}{P(k)} \sim k^{\alpha+1} \sim k^{\nu_v} \qquad n_v(k) \sim b_v(k)^{\beta_v}$$



Relationship between BRWBC and SPBC for vertices

$P(n_v) \sim n_v{}^{-\sigma_v}$ $\qquad$ $P(b_v) \sim b_v{}^{-\delta_v}$



$$\sigma_v = \nu_v \simeq 3.2$$

$$\sigma_v = \nu_v \simeq 2.2$$

Relationship between BRWBC and SPBC for edges

- for uncorrelated network

$$N_T(kk') \sim N_T(k)N_T(k') \sim (kk')^{\alpha+1-\gamma}$$

and the number of edges having *kk*' $N_e(kk') \sim \dfrac{kP(k)k'P(k')}{N^2} \sim (kk')^{1-\gamma}$

thus

$$\boxed{n_e(kk') \equiv \frac{N_T(kk')}{N_e(kk')} \sim (kk')^{\nu_e} \sim (kk')^{\alpha}}$$

- By assuming that

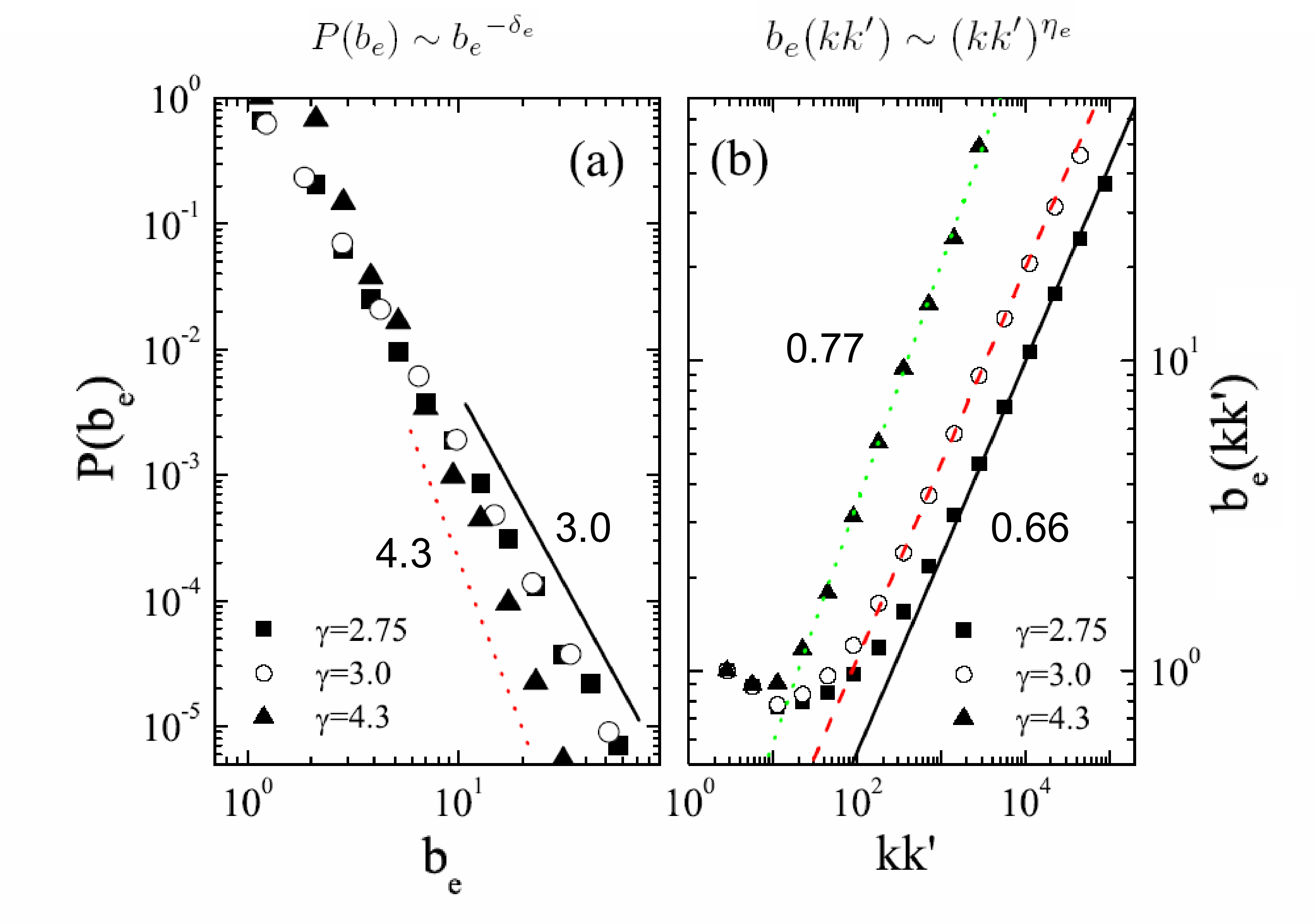
$$b_e(kk') \sim (kk')^{\eta_e}$$

and

$$n_e(kk') \sim b_e(kk')^{\beta_e}$$

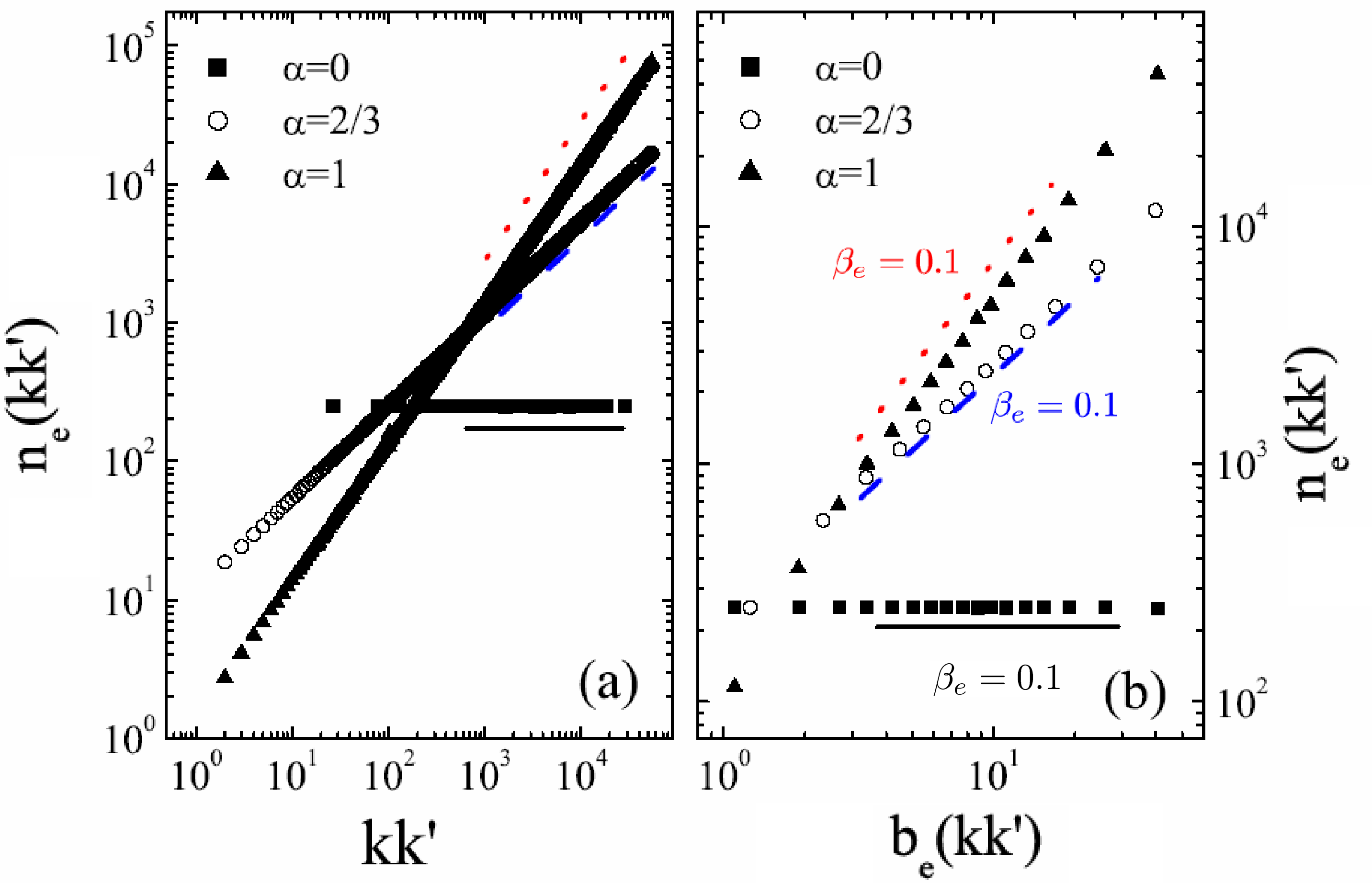
$$\Rightarrow \boxed{\beta_e = \nu_e/\eta_e = \alpha/\eta_e}$$

Relationship between BRWBC and SPBC for edges



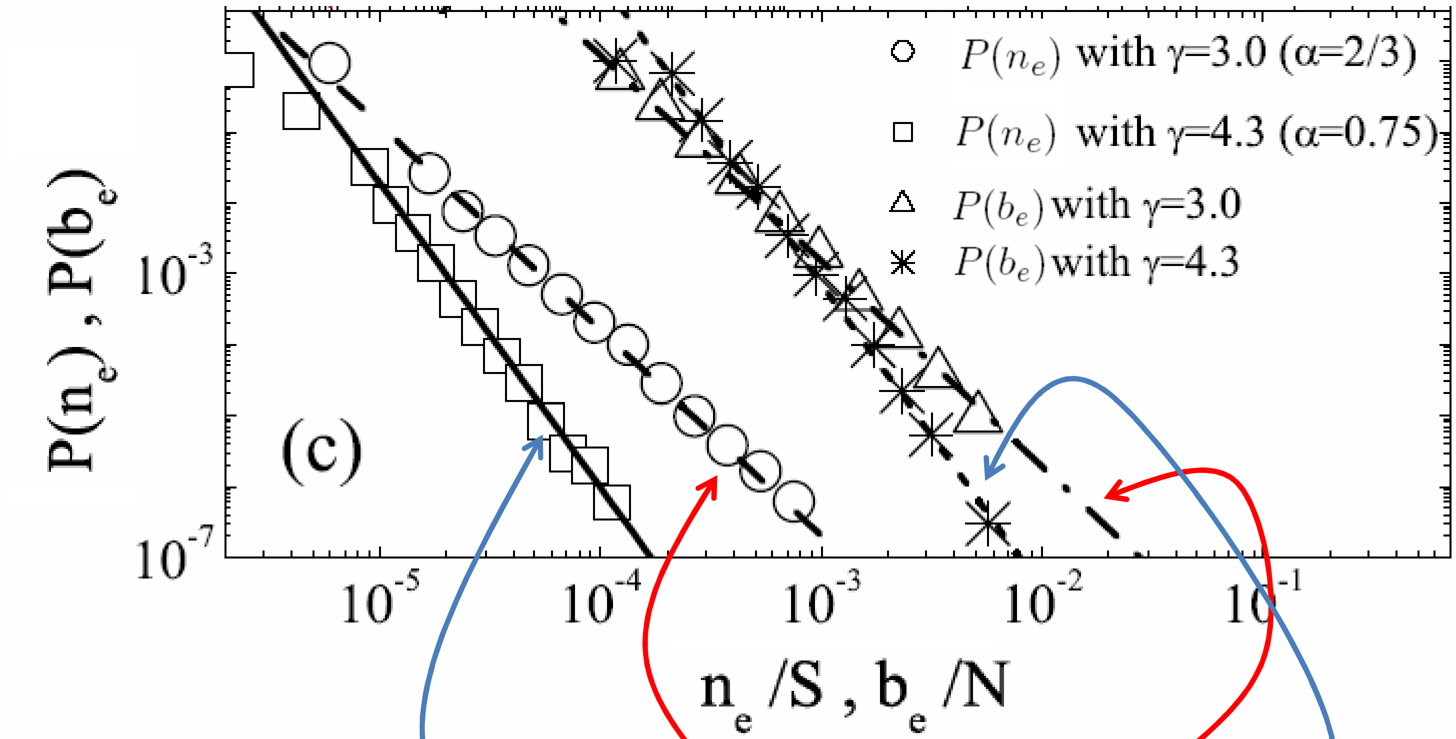
Relationship between BRWBC and SPBC for edges

$$n_e(kk') \equiv \frac{N_T(kk')}{N_e(kk')} \sim (kk')^{\nu_e} \sim (kk')^{\alpha} \qquad n_e(kk') \sim b_e(kk')^{\beta_e}$$



Relationship between BRWBC and SPBC for edges

$$P(n_e) \sim n_e^{-\sigma_e} \qquad P(b_e) \sim b_e^{-\delta_e}$$



$\sigma_e = \delta_e \simeq 3.0$

$\sigma_e = \delta_e \simeq 4.3$

Protein-Protein Interaction Network

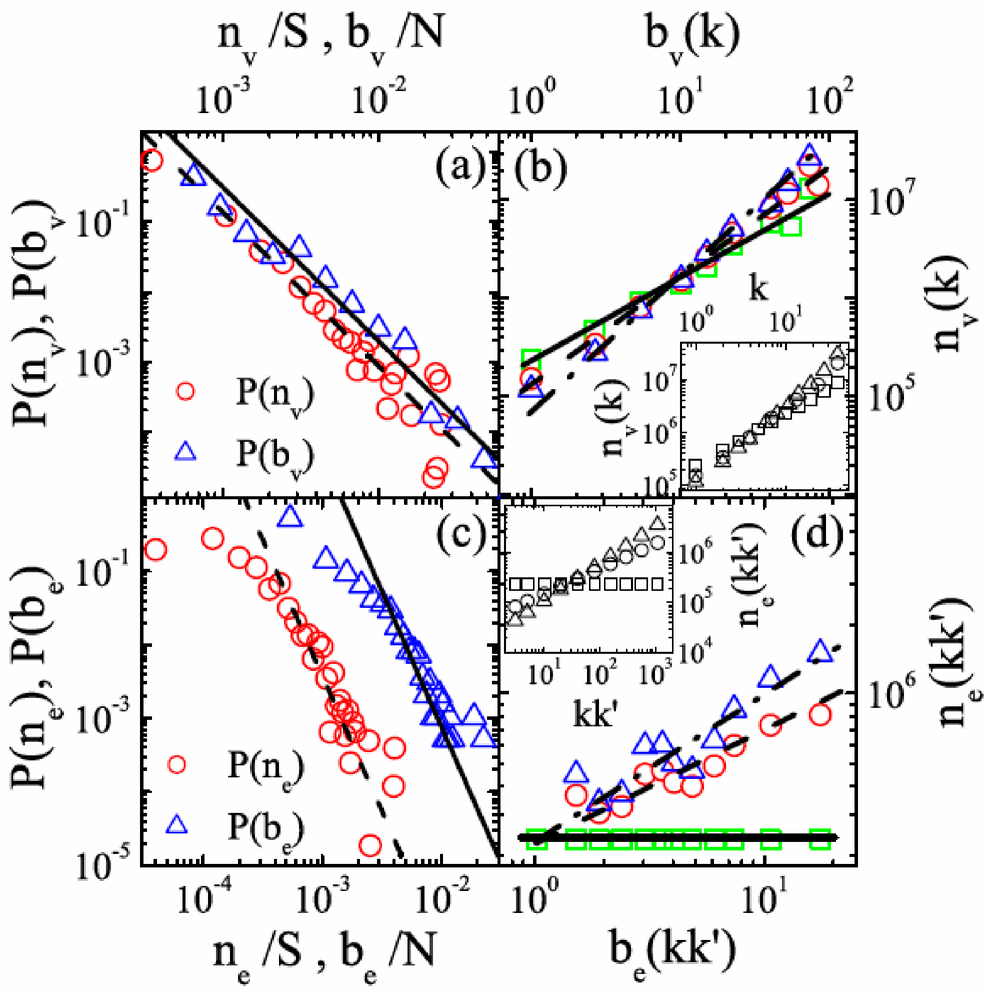


FIG. 3: (Color online) (a) Plots of $P(n_v)$ and $P(b_v)$ in PIN. S is the total number of steps of BRW. We use $\alpha = 0.5$. The dashed line and solid line represents the power-law $P(n_v) \sim n_v{}^{-2.2}$ and $P(b_v) \sim b_v{}^{-2.2}$, respectively. (b) Plot of $n_v(k)$ against $b_v(k)$ with $\alpha = 0$ (□), 0.5 (◯), and 0.75 (△). The lines indicate $\beta_v = 0.8$ (solid line), 1.0 (dashed line), and 1.2 (dashed-dotted line). The inset shows the plot of $n_v(k)$ against k. (c) Plot of $P(n_e)$ and $P(b_e)$ against n_e and b_e. We use $\alpha = 0.75$. The obtained exponents are $\sigma_e = \delta_e \simeq 3.7$ (dashed and solid lines). (d) Plot of $n_e(kk')$ against $b_e(kk')$ with $\alpha = 0$ (□), 0.5 (◯), and 0.75 (△). The lines indicate $\beta_e = 0$ (solid line), 0.65 (dashed line), and 1.0 (dashed-dotted line). In the inset we verified the relation $n_e(kk') \sim (kk')^{\nu_e}$.

Slight deviation of $\alpha+1=\nu$ and $\beta=\nu/\eta=\alpha/\eta$

Summary and Discussion

- We introduce a biased random walk centrality.
- We show that the edge centrality satisfies a power-law.
- In uncorrelated networks, the analytic expectations agree very well with the numerical results.

$$\beta_v = \nu_v/\eta_v = \frac{(\alpha+1)(\delta_v-1)}{(\gamma-1)}, \quad \beta_e = \nu_e/\eta_e = \alpha/\eta_e$$

- In real networks, numerical simulations show slight deviations from the analytic expectations.
  - This might come from the fact that the centrality affected by the other topological properties of a network, such as degree-degree correlation.
- The results are reminiscent of multifractal.
  - D(q): generalized dimension
    - q=0: box counting dimension
    - q=1: information dimension
    - q=2: correlation dimension …
  - In our BC measure
    - for α=0: simple RWBC is recovered
    - If $\alpha \rightarrow$ □; hubs have large BC
    - If $\alpha \rightarrow$ -□ ; dangling ends have large BC