How to rank nodes in a network? (i.e., how does Google sort the web?)
Outline

* The ranking problem
  * Path metrics
  * Iterative metrics
  * Spectral analysis
The ranking problem

* Assume a graph and a subset of “relevant” nodes
* Can we decide
  - Which nodes are the most important? (i.e. more precisely, their relative rank)
  - Based on topological properties only?

* Applications:
  - job hunt, search engine, find best articles, etc.
Outline

* The ranking problem
* Path metrics
* Iterative metrics
* Spectral analysis
Path measure 1:

- Degree of nodes
  Larger is more central

- Pretty accurate when each link is costly to maintain or especially meaningful

- Somewhat short-sighted

- Cheating with cheap links
Path measure 2:

- **Closeness centrality**: Average distance from u to all other nodes. Smaller is more central.

- **Betweenness centrality**
  - not-very discriminative
  - all scores are similar
  - Very sensitive
Path measure 3

- Betweenness centrality: 
  # of shortest paths going through node $u$
  Larger is more central

- Intuitively, relate to damage of removing $u$
Path measure 4

k-shell decomposition

1. Remove all nodes with a single link (assign $k_s=1$)

2. In the remaining graph $G'$ all nodes have 2 links. Remove now nodes with 2 links in $G'$ (assign $k_s=2$)

... Until all nodes removed.

1. All have counterexamples
2. Computationally costly?
3. What if nodes places link (or make friends) strategically?

- Can we define more robust / flexible metrics?

The Dirty Little Secrets of Search, NYTimes (Feb. 11th 2011).
The ranking problem
Path metrics
Iterative metrics
Spectral analysis
1. Count Votes
2. See who voted judiciously
3. Reweight votes accordingly

Authoritative sources in a hyperlinked environment.
J. Kleinberg JACM (1999)
We wish to find **Authorities**
- Nodes receiving votes from judicious people
- This also boils down to identify Hubs

- Does this process **converge** (after renormalization)?
- Where and how **fast**?
- Does it depend on our initial vote **counts**?
* The ranking problem
* Path metrics
* Iterative metrics
* Spectral analysis
An iteration follows \( h = A \cdot a \) and \( a = A^T \cdot h \)

- Hence \( a(k+1) = A^T A \cdot a(k) = (A^T A)^{k+1} \cdot a(0) \)

Thm: if \( a_j(0) > 0 \ \forall \ j \), and \( A^T A \) positive spectral gap

- (i.e. largest eigenvalues of \( A^T A \) satisfies \( \lambda_1 > \lambda_2 \))
- Then \( a(k)/\lambda_1^k \) converges to \( x_1 \) as \( k \) gets large
- Where \( x_1 \) is eigenvector with norm 1 for \( \lambda_1 \)

Bottom line: under weak conditions

- Fast convergence to a global property of the graph