TrustRank

Trust Rank

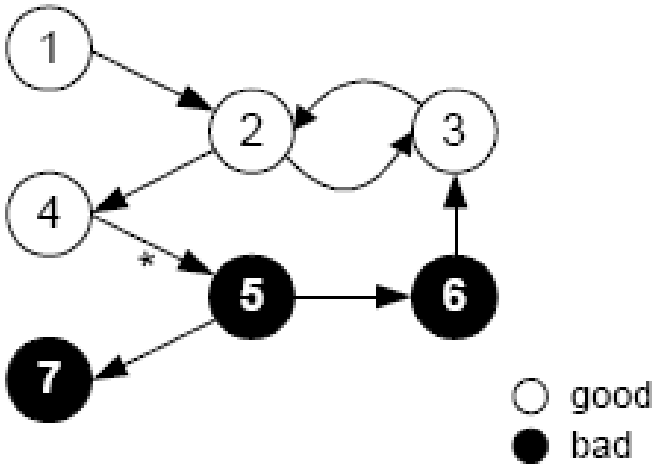
- Observation
  - Good pages tend to link good pages.
  - Human is the best spam detector
- Algorithm
  - Select a small subset of pages and let a human classify them
  - Propagate goodness of pages

Propagation

- Trust function T
  - *T(p)* returns the propability that *p* is a good page
- Initial values
  - *T(p) = 1*, if *p* was found to be a good page
  - *T(p) = 0*, if *p* was found to be a spam page
- Iterations:
  - propagate Trust following out-links
  - only a fixed number of iteration *M*.

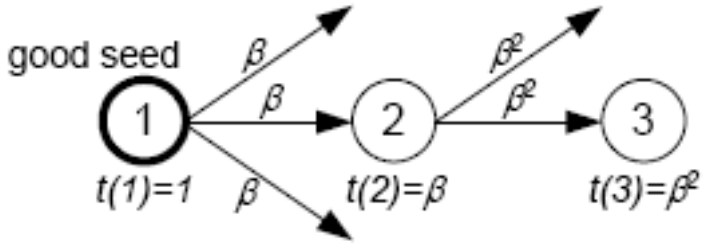
Propagation (2)

- Problem with propagation
  - Pages reachable from good seeds might not be good
  - the further away we are from good seed pages, the less certain we are that a page is good.



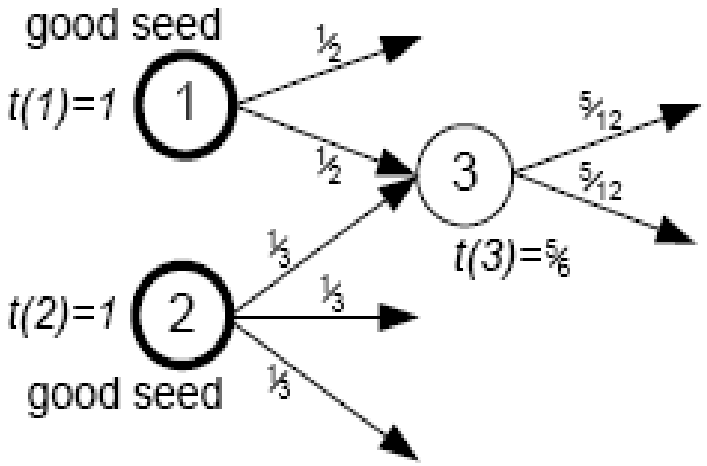
  - **solution**: reduce trust as we move further away from the good seed pages (trust attenuation).

Trust attenuation – dampening



- Propagate a dumpened trust score $\beta < 1$ at first step
- At n-th step propagate a trust of β^n
- How to deal with multiple in-links? (max, mean, etc.)

Trust attenuation – splitting



- Parent trust value is splittet among child nodes
- Observation: the more the links the less the care in choosing them
- Mix damp and split? *β^n*(splitted trust)

Selection – Inverse PageRank

- The seed set S should:
  - be as small as possible
  - cover a large part of the Web
- Covering is related to out-links in the very same way PageRank is related to in-link
  - ***Inverse PageRank !***
- Perform PageRank on a graph with inverted links
  - G' = (V, E') where $(p,q) \in$ E' $\Leftrightarrow (q, p) \in$ E.
- Alternatively, using *High PageRank* showed slighly worse performance

Algorithm

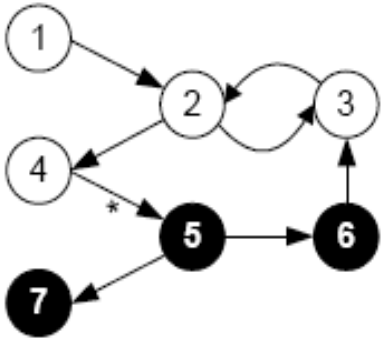
1. Select seeds (s) and order by preference
2. Invoke oracle (human) on the first L seeds,
3. Initialize and normalize oracle response d
4. Compute TrustRank score (as in PageRank formula):

$$\boldsymbol{t}^* = \beta \cdot \boldsymbol{T} \cdot \boldsymbol{t}^* + (1-\beta) \cdot \boldsymbol{d}$$

T is the adjacency matrix of the Web Graph.

β is the dampening factor. (usually .85)

Algorithm - example



- **s** = [0.08, 0.13, 0.08, 0.10, 0.09, 0.06, 0.02]
- Ordering = [2, 4, 5, 1, 3, 6, 7]
- L=3 {2, 4, 5} d=[0, 0.5, 0, 0.5, 0, 0, 0]
- β=0.85 M=20
- t* = [0, 0.18, 0.12, 0.15, 0.13, 0.05, 0.05]

- NB. max=0.18
- Issues with page 1 and 5

Evaluation metrics

- Pairwise orderness
  - fraction of pairs without mistakes
- Precision
  - fraction of good pages among those with trust above threshold
- Recall

$$\text{pairord}(\mathsf{T},\mathsf{O},\mathcal{P}) = \frac{|\mathcal{P}| - \sum_{(p,q)\in\mathcal{P}} \mathsf{I}(\mathsf{T},\mathsf{O},p,q)}{|\mathcal{P}|}$$

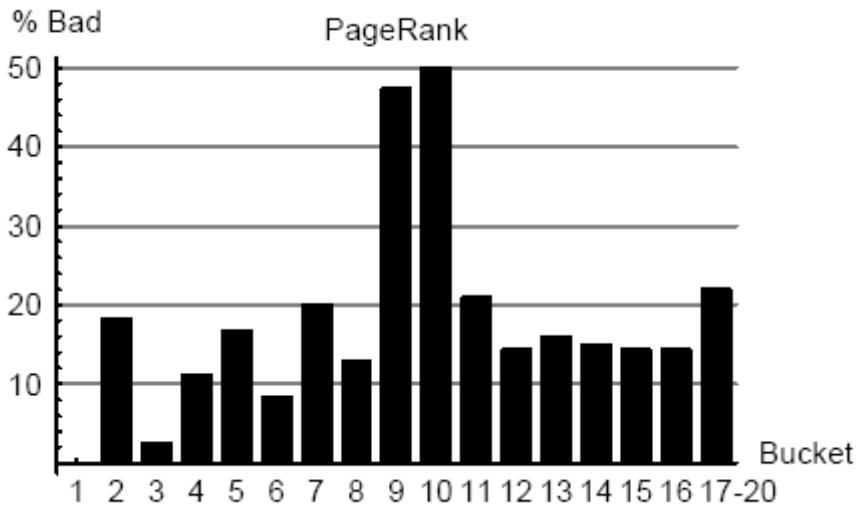
$$\text{prec}(\mathsf{T},\mathsf{O}) = \frac{|\{p \in \mathcal{X} | \mathsf{T}(p) > \delta \text{ and } \mathsf{O}(p) = 1\}|}{|\{q \in \mathcal{X} | \mathsf{T}(q) > \delta\}|}$$

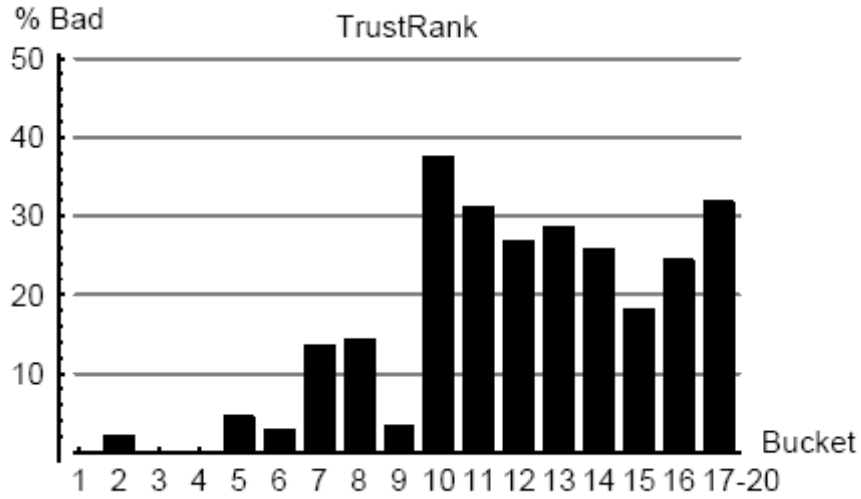
$$\text{rec}(\mathsf{T},\mathsf{O}) = \frac{|\{p \in \mathcal{X} | \mathsf{T}(p) > \delta \text{ and } \mathsf{O}(p) = 1\}|}{|\{q \in \mathcal{X} | \mathsf{O}(q) = 1\}|}$$

Results – evaluation data

- August 2003 dataset
- Approximation to websites instead of page
- 31 million websites
- 1 third (13 million) were unreferenced
- 178 seeds were choosed among those the oracle evaluated as good seeds
- 748 sample sites used to evaluate TrustRank

Results – compare with PageRank

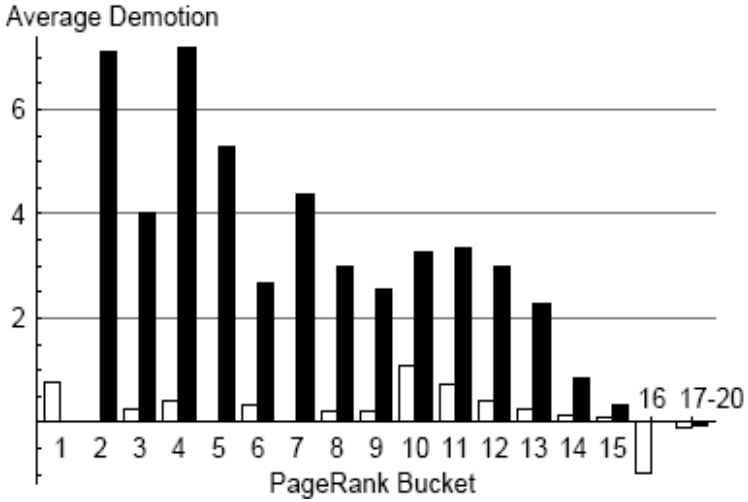




- Almost no spam in the first 5 buckets of TrustRank

Results – compare with PageRank

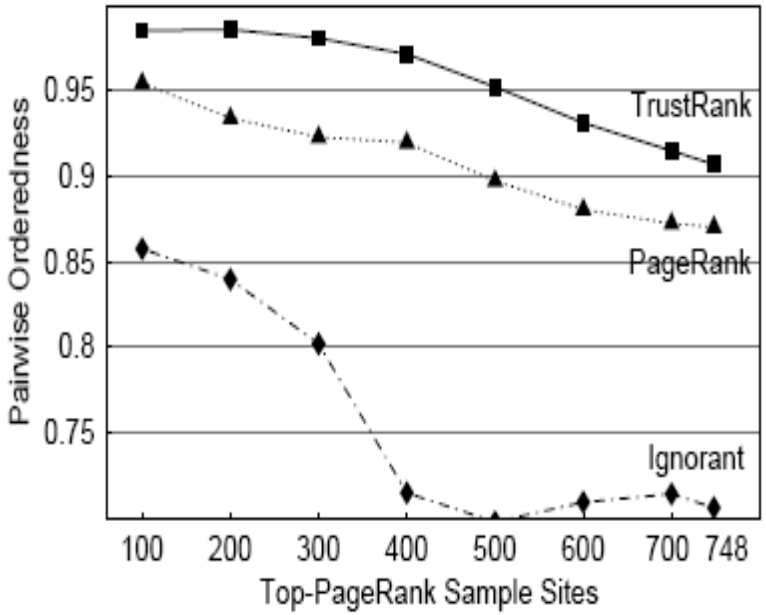
- The vertical axis shows the number of buckets by which sites from a specific PageRank bucket got demoted in TrustRank on average.
- White bars represent the reputable sites, while black ones denote spam.
- Example: spam sites in PageRank bucket 2 got demoted seven buckets on average (around bucket 9)

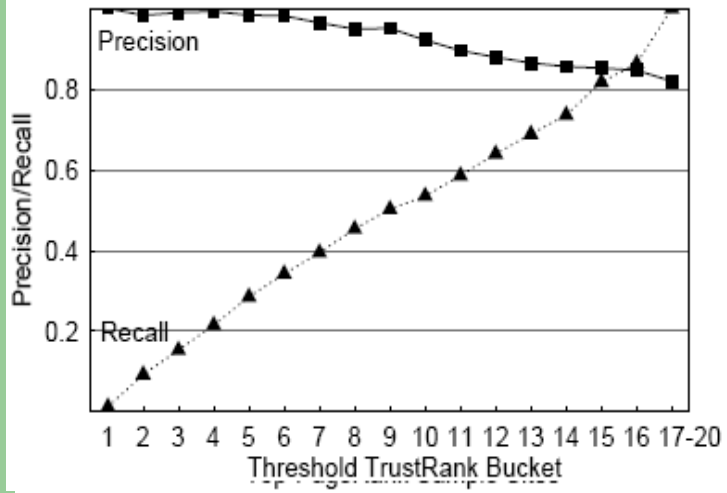


- Promotion exaple: in PageRank bucket 16, good sites appear on average one bucket higher in the TrustRank ordering.

Results – evaluation metrics

- Pairwise orderness in TrustRank, PageRank and the ignorant trust funtion.





- Precision and recall. Threshold choosed according to buckets.

Further refinements

- further explore the interplay between dampening and splitting for trust propagation.
- iterative process: after the oracle has evaluated some pages, we could reconsider what pages it should evaluate next, based on the previous outcome.

fine.

PageRank

- *PageRank in one equation:*
  - $PR(p) = \alpha M + (1-\alpha) V_p$
  - M is the adjacency matrix of the Web Graph.

  α is the *damping* factor. *(usually .85)*
  - in case of fairness V_p=1/N (N = # of pages in the Web).
  - V is the personalization vector.