Social and Information Networks

University of Toronto CSC303
Winter/Spring 2019

Week 6: February 11,13 (2019)

Announcements and agenda

Announcements

- TA office hour today following class in SF 4302B
- Assignment 1 is due this Saturday (Feb 16) at 11:59.
- Next week is reading week
- The term test is Friday, March 1

Agenda

- We continue to discuss Chapter 18 of the text. The focus is on a *preferential attachment* model. This model is propsed as a model explaining the power law distribution for the number of Web page in-links.

- We will then discuss Chapter 14 which provides two ranking algorithms, *Hubs and Authorities* and *Page Rank*, that utilize links to rank Web pages relating to a search query.

- We return to the ongoing issue of influence in decision making; in particular, we start the discussion of the cascade model in Chapter 16.

A power law distribution and network dynamics

We repeat the definition from last week:

A power law distribution for an event E satisfes $Prob[E|\text{parameter}\, k] \approx \frac{a}{k^c}$ for some constants a and c. (We often just focus on on the exponent c and say that the probability is proportional to k^{-c}.) Such distributions are called "scale-free" in the sense that the stated probability is independent of the size of the network.

Having observed that many events in social and information networks have a power law distribution, the big question is how this happens.

Chapter 14 considers the observed power law for the number of in-links in the Web graph. We understand that this could not evolve from independent decisions (that have averaged out) but rather results from the feedback coming from correlated decisions.

In an influential article, Kumar et al [2000] proposed a preferential attachment model that can explain the power law distribution. Recall, the observed distribution of in-links is that Prob[a site has k in-links] is proportional to $k^{-(2+\epsilon)}$ for a small $\epsilon > 0$.

A "rich get richer model" for in-links on the Web

Here is the model proposed in Kumar et al article (which has been uploaded to the course Web page).

Aside: I found this link from a paper by Bollobas which was cited in Chapter 18.

1. Web pages are created sequentially, and named $1, 2, \ldots N$. (Of course, N keeps growing but we are looking at the web at some point in time.)
2. With probability p, the j^{th} page chooses a page $i < j$ uniformly at random and links to page i.
3. With probability $q = 1 - p$, page j chooses a page $i < j$ uniformly at random and then creates a link to the page (say $k < i$) to which i has a link.

Note: The model is more general in that multiple links from page j are created in this stochastic model. Chapter 18 simplifies the model and only creates one link. However, **this does not change the power law exponent**. As will be seen, the key parameter is p.

The linking model continued

There is an equivalent way to state the indirect linking that takes place that makes clear the "rich get richer" preferential attachment phenomena.

- [3'] With probability $q = 1 - p$, page j chooses a page ℓ with probability proportional to ℓ's current number of in-links and creates a link to ℓ.

This is, of course, the idea behind *popularity*. For example, the more people that are reading a current novel, the more likely that you might want to read it. And for various social and economic reasons why some large cities continue to grow.

The linking model continued

There is an equivalent way to state the indirect linking that takes place that makes clear the "rich get richer" preferential attachment phenomena.

- [3'] With probability $q = 1 - p$, page j chooses a page ℓ with probability proportional to ℓ's current number of in-links and creates a link to ℓ.

This is, of course, the idea behind *popularity*. For example, the more people that are reading a current novel, the more likely that you might want to read it. And for various social and economic reasons why some large cities continue to grow.

Note: As $p \to 0$ (and $q \to 1$), pages are more likely to copy the same previous pages and the more likely that the process is creating some popular pages.

The linking model continued

There is an equivalent way to state the indirect linking that takes place that makes clear the “rich get richer” preferential attachment phenomena.

- [3’] With probability $q = 1 - p$, page j chooses a page ℓ with probability proportional to ℓ’s current number of in-links and creates a link to ℓ.

This is, of course, the idea behind *popularity*. For example, the more people that are reading a current novel, the more likely that you might want to read it. And for various social and economic reasons why some large cities continue to grow.

Note: As $p \to 0$ (and $q \to 1$), pages are more likely to copy the same previous pages and the more likely that the process is creating some popular pages.

Hedge: As the text states clearly, the goal of this model is not to capture all the reasons why people create links on the Web (or links in other networks) but rather to explain why it is reasonable to expect such popularity effects.

Sensitivity to unpredictable initial stages in network dynamics

As we are all are familiar, it is never clear why say some "pop" singers become so popular while other (perhaps of equal talent) never "make it". Clearly, the initial stages of a dynamic process are critical and that is why advertising, promotions, etc. are so important.

How can we better understand the impact of the randomness in the initial stages of a dynamic process? What if we could replay history many times? We would, of course, expect the resulting distribution to be the same. But would the same books, the same movies, the same pop stars, the same web pages, etc continue to be the most popular?

Sensitivity to unpredictable initial stages in network dynamics

As we are all are familiar, it is never clear why say some “pop” singers become so popular while other (perhaps of equal talent) never “make it”. Clearly, the initial stages of a dynamic process are critical and that is why advertising, promotions, etc. are so important.

How can we better understand the impact of the randomness in the initial stages of a dynamic process? What if we could replay history many times? We would, of course, expect the resulting distribution to be the same. But would the same books, the same movies, the same pop stars, the same web pages, etc continue to be the most popular?

Our intuition (and experience) suggets that there is often considerable “luck” in exactly who or what becomes popular, On the flip side, we also believe that “quality” is also important.

But how do we “rewind history”?

An experiment to “rewind history”

While we do not have the ability to really rewind history, it is something that I think we often think about.
Aside: There are some classic movies (e.g. Sliding Doors, Run Lola Run, Blind Chance) that explore this theme about initial random effects that lead to very different alternative outcomes. These are very interesting movies but, as you might expect, do not consider distributions. See https://www.denofgeek.com/us/movies/run-lola-run/256336/7-movies-and-tv-shows-that-master-the-multiple-reality-narrative.

Salganik et al perform an interesting experiment (in fact, two experiments at different times with different participants) to observe the impact of the initial random stages in a dynamic process. (I have downloaded the article and the supporting material.)

The Salganik et al experiment

Here is their experiment:

- They created 9 copies of 48 "obscure" (as determined by some experts) songs of varying "quality"
- In the experiment, approximately 7200 young participants were recruited to listen to the music. At the start of the experiment, all that is known is the name of the band and the name of the song.
- In each of the copies, participants sequentially listened to some music selections, rated the music and then were given the opportunity to download copies of songs they liked.
- In each of 8 copies of the music, 10% of the participants were also given the number of times each song had been previously downloaded.
- In the 9^{th} version, this previous history of downloads was not provided to the remaining 20% of the particpants. The average of the ratings (from $1 =$ "I hated it" to $5 =$ "I loved it") in this "no influence" version determined the song "quality".

The findings in the Salganik experiement

The experiment was designed to measure the extent that social influence leads to different outcomes in the “success” (i.e. the number of downloads) of a particular song.
Simply stated, the results show that:

- Increasing the strength of social influence increased both the inequality (i.e. degrees of popularity) and unpredictability (i.e., relation to quality) of success.
- However, quality was also a factor: the best rated songs rarely did poorly and the the worst songs rarely did well.

As I said, this is an interesting study and one where the authors carefully try to eliminate sources of bias. The article is worth reading.

As the text points out in section 18.6, how recommendation systems are designed can impact how people make choices, leading to increased “rich can richer” phenomena, or alternatively exposing people to less popular items.

Visualizing the long tail of a power law distribution

Once we accept a power law nature of popularity, it is instructive to consider the consequences for a given industry. Namely, the nature of the sales curve that would be dictated by a power law distribution.

The shape of the long tail in a power distribution raises the question as to how many sales can be obtained from less popular (e.g. niche items).

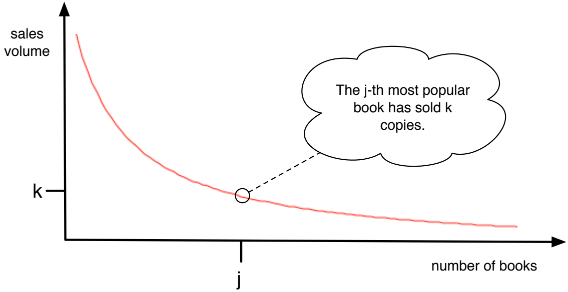


Figure: [Fig 18-4 in E&K] text; how many copies of the j^{th} most popular items have been sold.

An informal analysis for the simplied preferential attachment model proposed for Web in-links

A precise analysis of even the simple one link per page preferential attachment model is technical. In section 18.7, the text provides a heuristic argument as to how the power law exponent is determined by the probability p (of the j^{th} page linking uniformly at random to some page $i < j$ vs linking indirectly with probability $q = 1 - p$ to a page ℓ based on the popularity of page ℓ).

While we often gain insight by viewing a continuous process as a long sequence of discrete events, it is often adventageous to model a sequence of discrete events as a continuous process.

More specifically, the approximate analysis considers a continuous deterministic variable $x_\ell(t)$, which is an approximation of the discrete random variable $X_\ell(t)$, the number of in-links to a page ℓ at time $t \geq 0$.
Aside: We often do this in the analysis of algorothms. For example, we consider the continuous extension of a submodular function and what is called a continuous greedy algorithm.

The deterministic continuous model of the rendom discrete process

Initially, $X_\ell(0) = x_\ell(0) = 0$.
The discrete propbability that the number of links to a page ℓ increases is :

$$\frac{p}{t} + \frac{q \cdot X_\ell(t)}{t}$$

and the corresponding continuous rate of growth is modeled by the differential equation:

$$\frac{dx_\ell}{dt} = \frac{p}{t} + \frac{q \cdot x_\ell(t)}{t}$$

The rest of the section uses some basic calculus to show that this leads to a power law distribution proportional to k^{-c} with $c = 1 + 1/q$. This makes sense as the closer p gets to 0 (and $q = 1 - p$ goes to 1), the exponent $c = 1 + 1/q$ limits to the observed exponent $c = 2 + \epsilon$ for the observed in-link power law distribution. The closer p goes to 1, the exponent limits to ∞ making a large number of in-links very unlikely.

Search and ranking on the Web

Our next topic is to undertstand how the popularity of a web page is determined and how that impacts its rank in the responses to a query.

But first, how do search engines find and rank responses to a query?

Search and ranking on the Web

Our next topic is to undertstand how the popularity of a web page is determined and how that impacts its rank in the responses to a query.

But first, how do search engines find and rank responses to a query?

The specific algorithms used by search engines such as Bing and Google is a trade secret. To some extent this has to be kept secret as there is always a "war" between a search engine and companies that create web sites to enhance the ranking of a site.

However, we do have a basic idea as to how these search engines rank sites given a query. In fact, at the most elementary level, the main idea is an old one, but one that was not well accepted for many years. Here is my sense of things.

Aside: In the 1960s and 70s, there was a basic argument as to whether online search and ranking was a more or less normal algorithmic search and optimization problem or one that required "intellegience" (i.e. the ability to understand natural language). Who won this argument?

Search and ranking of Web documents; the role of link popularity

The most basic approach is to treat a document as a bag of words and then use "normalized" word counts (and pairs,triplets of words) to identify and rank docuemnts relating to the query. This became enhanced by more sophisticated contextual aspects of word occurrences, etc and today machine learning algorithms are also used in classifiying a search query.

But early in the development of popular search engines, a popularity aspect was added where the ranking of a docuemnt also depended on the link structure and the popularity of a Web page in the Web network (or at least that part that seems relevant to the query).

Two algorithms were independently proposed for determinining the popularity of a Web page, namely Hubs and Authorities developed at IBM and used in their never released search engine, and Page Rank, developed and integrated into Google's search engine.

End of Monday, February 11 lecture

Monday's lecture ended at slide 14.
Today's agenda:

- Continue discussion of search engines and the role of link structure in helping to determine the ranking of a document.
- Hubs and Authorities
- Page Rank
- Why these algorithms converge
- Begin Chapter 16 and discussion of Information Cascades

Link analysis and page popularity

Neither Hubs and Authorities nor Page Rank use link in-degree as the popularity measure but link analysis is (or at least was) used to determine page popularity. Currently, it seems clear that popularity also depends on recent behaviour of users to related queries.

We will not try to infer more precisely how say Google (or any search engine) precisely determines the ranking of a document in resposne to a query. In particular, we do not know how much page ranking depends on content vs link analysis. But we do know that this ranking is essential in determining how often a page will be downloaded. The quality of the ranking algorithm leads to user activity and thus the resulting advertising.

We will begin with the Hubs and Authorites ranking algorithm and then the Page Rank algorithm.

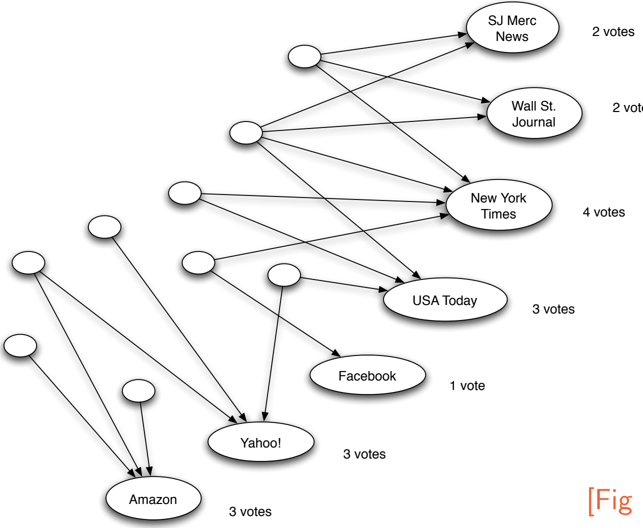
Hubs and Authorities

- A simple way to utilize links to rank web pages would be to think of each link from A to B as an endorsement or vote by A for B.
- And then use the number (or weight) of endorsements as a key feature determining the rank. Of course, one would have to adjust such scores coming from say the same domain name.
- Even after adjusting for such "vote fixing", if Auston Mathews or John Tavares has a web site and a link suggesting where he buys his hockey equipment you might think that is more meaningful than if say where I recommend you buy your hockey equipment.
  Spoiler alert: I don't play hockey.

Reinforcement of Hubs and Authorities.

- This then becomes the motivation (and seemingly circular reasoning) behind hubs and authorities.
- The best "authorities" on a subject (places to buy equipment) are being endorsed by the best hubs (people who know where to buy equipment).
- Similarly, the best hubs are those sites that recommend the best authorities. Conceptually the link structure induces a bipartite graph. The same web page can be both a hub and an authority.
- Comment: The word "authority" is not generally an accurate way to describe high ranking documents. These might better be referred to (barring other information) as the most relied upon sites. This is also different from "the most popular" sites which might better be measured in terms of the number of clicks being received. Hubs then are the most reliable endorsers.

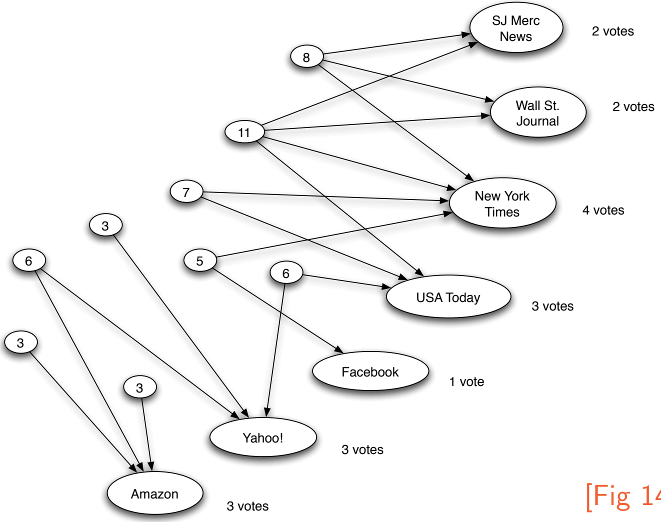
- The result of applying the authority update rule: for each page p, $\text{auth}(p)$ is the sum of hub values (initially just the number) of hubs pointing to p.



[Fig 14.1, E&K]

Figure: Counting in-links to pages for the query "newspapers."

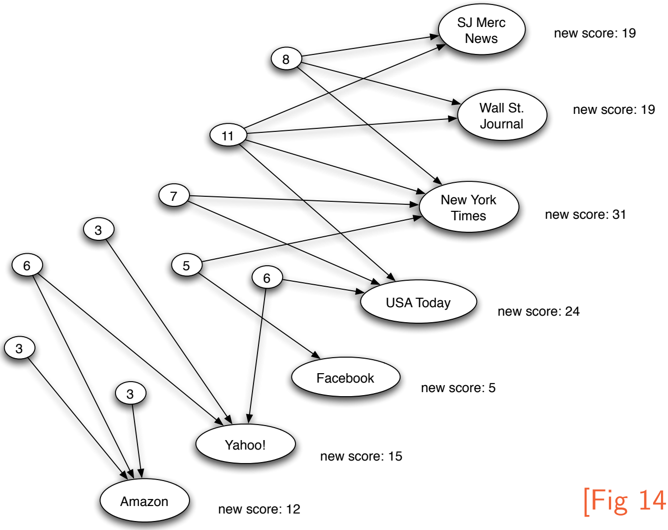
- Then to recalibrate hub values, we use the hub update rule: for each page p, hub(p) is the sum of values of all authorities that p points to.



[Fig 14.2, E&K]

Figure: Finding good lists for the query "newspapers": each page's value as a list is written as a number inside it.

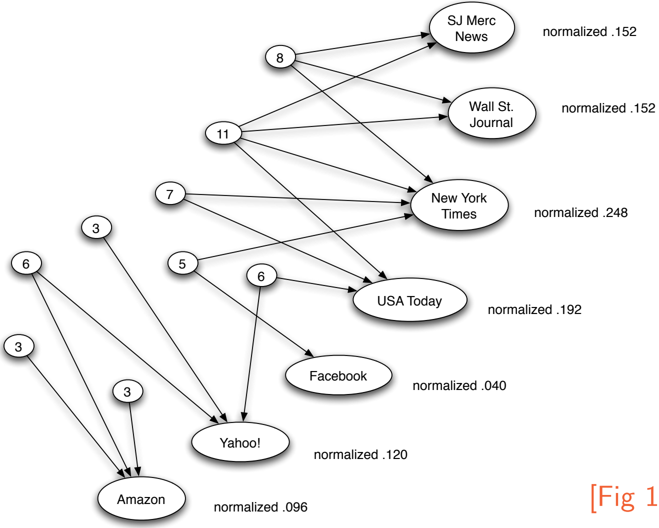
- Applying the authority update rule again we get figure 14.3.



[Fig 14.3, E&K]

Figure: Re-weighting votes for the query "newspapers": each of the labeled pages new score is equal to the sum of the values of all lists that point to it.

- Since we only care about the relative values of these numbers, both authority and hub scores can be normalized to sum to 1 (to allow convergence and avoid dealing with large numbers).

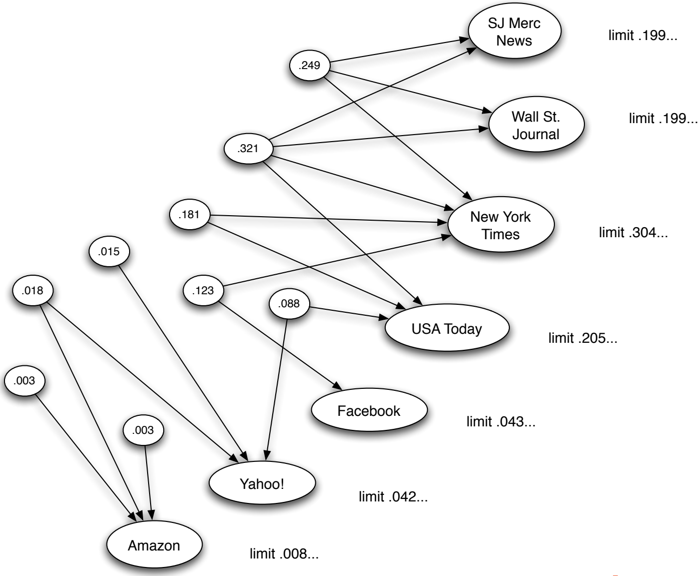


[Fig 14.4, E&K]

Figure: Re-weighting votes after normalizing for the query "newspapers".

Keep repeating a good idea

- Now having recalibrated and normalized both the authority and hub scores, we can continue this process to continue to refine these scores.
- That is, the hubs and authorities procedure is as follows:
  - Initialize all hub values (say to some positive vector perhaps depending on usage or content)
  - For sufficiently large k, perform the following k times
    - Apply authority update rule to each page
    - Apply hub update rule to each page
    - Normalize so that sum of A and H weights $= 1$.
- Using linear algebra, it can be shown (in Section 14.6) that these A and H normalized values will converge to a limit as $k \to \infty$ (which can be approximated by some sufficiently large k)!
- Hubs and Authorities can be extended to work for weighted edges (e,g. weighting links in achor text, or near a section heading, etc.)



[Fig 14.5, E&K]

Figure: Limiting hub and authority values for the query "newspapers".

Page Rank

- The motivation behind page rank is a somewhat different view of how authority is conferred.
  - Endorsement of authority is conveyed by other authorities
  - That is, no hub concept
  - This is how peer review works in the academic and scholarly world.

- Authorities themselves convey authority on those they link to. This naturally leads to a forumation in terms of two equivalent views of page rank:
  1. Authorities directly conveying authority (without hubs)
  2. Authority values resulting from long term behaviour of a random walk on a graph.

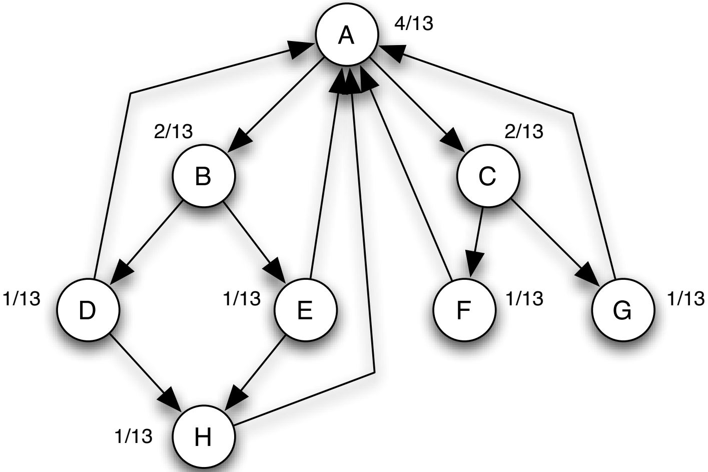
How does Page rank spread authority?

- Suppose at any point of time we have relevant authority scores.
  - A page spreads its authority equally amongst all of its out links.
  - If a page has no outlinks then all authority stays there.
- This redistributes the authority scores. (We are not creating or losing any authority, we are just redistributing it.)
- We can initially start with every relevant page having authority $1/n$ where there are n pages. Then we repeat this process k times for some sufficiently large k .
- With the exception of some "degenerate cases" (e.g. the process is periodic) it can be proven (again using linear algebra) that this process has a limiting behavior as $k \to \infty$.
- The resulting limit values will form an equilibrium.
- If the network is strongly connected then there is a unique equilibrium,

Remark

In many cases this won't reflect the desired authority. Namely, if the network has any sinks (or strongly connect components that are sinks) which it will surely have, then all of the authority will pass to such sinks.

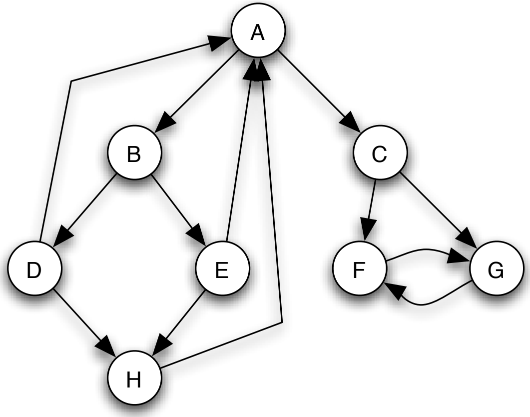
Page rank equilibrium for a network



[Fig 14.7, E&K]

Figure: Equilibrium PageRank values for the network of eight Web page.

Where has all the authority gone when we redirect (F, A) and (G, A) edges?



[Fig 14.8, E&K]

Figure: The same collection of eight pages, but F and G have changed their links to point to each other instead of to A. Without "scaling", all the PageRank would go to F and G.

Scaled page rank

- The way around this sink hole of authority is to have a scaled version of page rank where
  - only a fraction s of the authority of a page is distributed to its out links
  - the remaining $(1 - s)$ fraction is distributed equally amongst all relevant pages.
- For any value of $s < 1$ (which effectively makes the graph strongly connected), we get convergence to a unique set of scores for each page and that is its page rank (for that particular value of s). It is reported that Google uses $0.8 \leq s \leq 0.9$.
- (See the footnote on page 410 of E&K as to why in the previous example, nodes F and G will still get most of the authority but that for realistically large networks, the process works well.)

Some additional remarks

- The limiting scores for both the authority and hubs approach and the page rank approach are equilibrium points for an appropriate algebraic process.
- That is, if we actually were in the limiting state, we would be in the equilibrium state. In practice, we can stop the process when the change in each iteration is sufficiently small.
- We can weight the network edges (say according to some concept of link importance) and apply the same authority and hubs or page rank approach distributing authority in proportion to these weights.

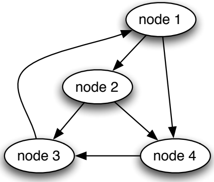
Advanced material (section 14.6): Handwaving argument why these processes converge

We have already suggested that both the page rank and hubs and authorities processes can be understood in terms of an algebraic process, namely, a linear transformation.

- Suppose we are considering a web network of n pages. We can represent the hub, authority or page rank values at any time k of the process by an n-dimensional (column) vector, denoted (respectively) by $\mathbf{h}^{(k)}, \mathbf{a}^{(k)}, \mathbf{r}^{(k)}$.
- Here we are using boldface $\mathbf{v} =< v_1, \ldots, v_n >$ to represent a vector whose components are the v_j so that (for example), $r_j^{(k)}$ repesents the page rank of the j^{th} web page after k steps of the page rank process.
- Let $\mathbf{v}$ be any of the hub, authority or page rank vectors. In each case it is not difficult to see that the process can be viewed as a linear transformation $\mathbf{v}^{(k+1)} = M\mathbf{v}^{(k)}$ for some appropriate $n \times n$ matrix M whose entries are non negative real numbers.

Advanced material continued: page rank convergence

- Section 14.6 tells us how to define the appropriate matrices and gives the conditions that will guarantee the convergence of the process; that is, when there exists $\mathbf{v}^{(*)} = \lim_{k \to \infty} \mathbf{v}^{(k)}$ and when this limit vector $\mathbf{v}^{(*)}$ is unique and independent of the starting vector $\mathbf{v}^{(0)}$.
- Figure 14.3 of the text illustrates a simple directed graph and the matrix N that defines the unscaled page rank update process. That is, $< r_1^{k+1}, \ldots, r_n^{k+1} >= N^{tr} < r_1^k, \ldots, r_n^k >$ where N^{tr} is the transpose of matrix N.



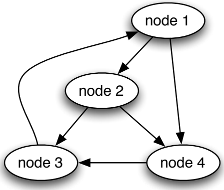
$$\begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

[Fig 14.13, E&K]

Figure: A toy web graph and the associated matrix N describing the unscaled update process.

Page rank analysis for the scaled update

Similarly Figure 14.4 illustrates the same graph and the matrix $\tilde{N}$ that defines the scaled page rank update process with scaling factor $s = .8$.



$$\begin{bmatrix} .05 & .45 & .05 & .45 \\ .05 & .05 & .45 & .45 \\ .85 & .05 & .05 & .05 \\ .05 & .05 & .85 & .05 \end{bmatrix}$$

[Fig 14.14, E&K]

Figure: The same toy web graph and the associated matrix $\tilde{N}$ describing the scaled update process with $s = 0.8$.

- It follows that $\mathbf{r}^k = (\tilde{N}^{tr})^k \mathbf{r}^0$
- If the process is converging then it would be converging to some $\mathbf{r}^*$ satisfying $\mathbf{r}^* = N^{tr}\mathbf{r}^*$

Now comes the necessary linear algebra

So far we have mainly used matrices as a convenient way to represent the process. But to understand convergence we need to mention some more essential aspects of linear algebra.

- Let $M_{n \times n}$ be a full rank matrix. Recall that the matrix-vector multiplication $M\mathbf{v}$ can rotate and expand/shrink the vector $\mathbf{v}$.
- Since it is hard to "visualize" an n-dimensional vectore space, we can simply think about the meaning of such a linear transformation in 2-space or 3-space.
- A vector $\mathbf{v}$ is an eigenvector of M with associated eigenvalue λ if $M\mathbf{v} = \lambda\mathbf{v}$. It follows that $\mathbf{v}$ is also an eigenvector of M^k with eigenvalue λ^k.
- When $\lambda = 1$, the eigenvector then becomes an equilibrium of the process!

More linear algebra

- For each full rank matrix there is a set of n eigenvectors with (not necessarily distinct) associated eigenvalues $\lambda_1, \ldots, \lambda_n$; these eigenvectors span the n-dimensional Euclidean space so that any vector can be expressed as a linear combination of the eigenvectors.
- An important result from linear algebra (Perron's Theorem) states that any matrix which has all positive entries has a unique eigenvector $\mathbf{y}$ corresponding to the largest positive eigenvalue λ_1 and furthermore $\lambda_1 > |\lambda_i|$ for $i > 1$.
- Since $\lambda_1 > |\lambda_i|$ for $i > 1$, and since every vector is a linear combination of the eigenvectors, it follows that as $k \to \infty$, the transformation M^k is being dominated by the largest eigenvalue acting on its associated eigenvector.
- For the scaled matrix $\tilde{N}^{tr}$, all entries are positive and the largest eigenvalue is 1. It follows that as $k \to \infty$, $(\tilde{N}^{tr})^k\mathbf{v}$ will converge to the eignevector $\mathbf{y}$ associated with the largest eignevalue 1.

Similar analysis for hubs and authorites

- If M is the adjacency matrix of the web graph, then the process can be described by $\mathbf{h} = M\mathbf{a}$ and $\mathbf{a} = M^{tr}\mathbf{h}$.

- Then
  1. $\mathbf{a}^{(1)} = M^{tr}\mathbf{h}^{(0)}$
  2. $\mathbf{h}^{(1)} = M\mathbf{a}^{(1)} = MM^{tr}\mathbf{h}^{(0)}$

- It follows that
  1. $\mathbf{a}^{(k)} = (MM^{tr})^{k-1}M^{tr}\mathbf{h}^{(0)}$
  2. $\mathbf{h}^{(k)} = (MM^{tr})^{k}\mathbf{h}^{(0)}$

Hubs and authorities analysis continued

- The matrices (MM^{tr}) and $(M^{tr}M)$ are symmetric and have non-negative entries.

- Any $n \times n$ symmetric matrix S with non negative entries has an orthonormal set of n eigenvectors all of whose associated eigenvalues are real. By normalizing the scores, we can assume that the largest eigenvalue $\lambda_1 = 1$.

- If the largest eigenvalue is unique (which is what would happen in a real web graph), then the same analysis for page rank applies (assuming that the starting hub scores are all positive).

Returning to the issue of influence

In some sense or another we are often talking about social influence in this course. Even in Chapter 14, we can view, for example, hubs as influencing which Web pages will be ranked highly.

In chapter 18, we observed two sequential processes where previous individual decisions had a significant impact on
1) The evolution of links on the Web, and
2) The evolution of opinions in evaluating music.

The music evaluation experiment is closer to reality in the sense that it explicitly integrates a changing measure of quality into the decision making process. (We could augment the link generation process to use a measure of similarity between web pages to enhance the process by which Web pages are generated, but the goal of that discussion was to illustrate how power law distributions can arise.

The spread of influence

- This will be the beginning of a several week discussion of
  - influence/technology/disease spread
    i.e. "contagion" in a very general sense in social networks;

- We will first be discussing Chapter 16 (information cascades) where (as we have seen before) sequential decisions are influenced by previous decisions. The chapter argues that being infuenced by previous decisions is *rational* and not necessarily mindless. Here the benefit is indirect in the sense that the probability of making a better decision can be improved by following othetrs.

- Then in Chapter 17 , influence comes in the form that there is a *direct benefit effect* (i.e. a change in the reward) for following others.

- That is, for the next several weeks we will be studying

various social processes that channel individual behaviour into collective behaviour.

Information cascades (herding) - Chapter 16

- Chapter 16 concerns the phenomena of information cascades
  - whereby individuals observe and then make decisions sequentially based on the behaviour of people having made decisions earlier;
  - e.g. deciding on a restaurant by observing how many people are currently eating there, what clothes you buy, other fashions/fads.
- Chapter 17 discusses decisions based on direct benefit (e.g., using a popular operating system/ laptop because wide use implies more software support).
- Clearly both phenomena can be interacting when people make decisions (e.g. busy restaurants are more able to use fresh ingrediants); the text organization is to try to first isolate and model these phenomena so as to gain insight.

A simple information cascade model

Assumptions for an information cascade:

- Individuals make decisions sequentially and can observe the decision of those who have acted earlier.
- Each individual has some private information that can be used in making their decision.
- Individuals only observe the behavior of earlier people but do not know their private information beyond any inferences that can be made from the previous decisions/actions.
- Note how the musice evaluation experiment fits into this model.