



## **Filthy linking rich and getting richer!**

As a kid, I once asked my father: How do you become a multi-millionaire? He looked at me and said: "Easy! First you make a million..." and then he had a good laugh.

Come the latter part of the sixties, my father was a very wealthy man. As a serial entrepreneur his interests in show business had netted him a fortune. He had opened his first nightclub in 1962, which was home to a local group called The Animals, who he helped on their way to international stardom (the bass player of the band and his former business partner would go on to manage Jimi Hendrix).

He drove a Jensen Interceptor (very cool car at the time), wore the most "fab gear" (Beatle speak for hip clothes) and hung around Las Vegas with his pals quite a lot (a little too much for my mother's liking unfortunately!). He was at the peak of his career and a major influencer within his social group. Not bad, for a guy who had to catapult himself from his early beginnings as a typewriter salesman.

It was some time later in life when my father and I were talking about wealth creation that he used the expression: The rich get richer.

I'll come back to my dear departed Dad again, but stay with me right now, as this is going somewhere. And I'm afraid some of it may be pretty bleak reading if you have newly created, lowly indexed web pages and you're desperately waiting for someone to link to them so that they stand a chance of ranking in a search engine with a static link based algorithm.

I have to tell you, for the past 18 months, I've been absorbed in an entirely new world of research. Network science can be regarded as a branch of complexity theory. Complexity itself describes any number of different sciences, theories, and world views such as chaos theory, emergence and network science. And it's fascinating. In fact, I'll go as far as saying enthralling.

By trying to further increase my understanding of the real power of linkage data in search engine algorithms (to be shared with you in the upcoming third edition of my book), I've become even more aware of how "the rich get richer" power law affects search engine results and also the ecology of the web itself. The richer you are with links pointing back to your site, the richer you are likely to become in search marketing terms.

Are search engines giving a fair representation of what's actually available on the web? Not really. If pages were judged on the quality and the relevance for ranking, then there would be less search engine bias towards pages which are simply popular by "linkage voting". Unfortunately, quality is subjective so finding a universally acceptable measurement or metric is not going to be easy.

If you're involved in the search marketing industry, particularly on the link building side, then you'll know better than most, that getting links for a large and more visible web site is easier than that of getting links for a start-up or a mom-and-pop type outfit.

Now you may feel that you're about to read something obvious and decide to skip the rest here. But please stay with me a little longer.

I believe you may be very interested to know that the scale of the problem is rapidly getting greater with the bias of a static based "link popularity" algorithm such as PageRank, largely the cause of the problem.

"Now wait a minute Grehan," I hear you say. "Aren't you the biggest sceptic about PageRank and its role in Google search results?" And the answer is "you bet I am."



However, what I want to do with this feature article is to try and highlight just how great the bias is for high ranking pages which are fundamentally ordered on link based algorithms, to attract more links. And why I believe (along with many in the research community) it's becoming necessary for search engines to seek a new paradigm. I think it is of benefit to the search marketing community as whole to understand the implications of such concerns by search engines to move away from current methodology as it will certainly have its impact on our industry.

When speaking on the subject of linking at conferences, seminars and workshops, I always attempt to explain the different manner in which search engine marketers look at links on the web compared to that of the way search engines view the same data. Search engine marketers are concerned, basically, with a hyperlink from another page back to theirs. And generally speaking, the more the merrier!

However, search engines take a much more mathematical, philosophical and analytical impression of the entire web (or more to the point, the fraction of it that they have captured). To search engines, web pages which are linked together are nodes in the web graph. By applying random graph theory to the web, they have viewed it as a type of static, equilibrium network with a classic Poisson type distribution of connections.

Even though graph theory has made great progress and been an important factor in the way that search engines have been able to plot crawling of the web and ranking of documents: We now know that the most important natural and artificial networks have a specific architecture based on a fat-tailed distribution of the number of connections of vertices that differs crucially from the classical random graphs studied by mathematicians. Because, as a rule, these networks are not static but evolving objects.

It's really only been the last five years or so that physicists have started extensive empirical and theoretical research into networks which are organised this way. The main focus prior to this research was in neural and Boolean networks where the arrangements of connections was secondary.

So now I hear you say: "Hey Mike... Whoa! Stick with search engines and optimisation and suchlike. I'm a search engine marketer, not a physicist! Random graph theory, equilibrium network with a classic Poisson type distribution, fat-tailed distribution, neural networks... I'm brain zonked already and this is only the second page of your article."

Yup, I understand that, I've scrambled my own brains a bit a few times recently. But it's important that you do know a little more about what's really going on in the research field to provide a better analysis of what currently makes one web page more important than another and how that is likely to change.

Your business may depend greatly on being able to optimise for search engines. And that's only going to become harder and harder.

Tell you what I'll do - I'll back up a bit here and try and do a brief history to what I've been kind of "glossing over".

Let's have a slightly (and I do mean slightly) more in-depth look at network science and how it is mathematically, philosophically and otherwise, applied to search. As Russian physicist and genius Sergei Dorogovstev put it in his excellent text, *Evolution of Networks*, from last year, I feel it's "more important to be understood than to be perfectly rigorous". Although I have tried to eliminate the math as best as I can and stick to the principles, there are bound to be sections which do reference formula. Therefore brave reader, as I frequently do myself, don't be afraid to skim over some parts you don't understand to reach those that you do.

Remember: Like you, I'm a search engine marketer - not a scientist.



The history behind social network concepts and graph theory applied to ranking algorithms is something which can help you understand a lot more about their complexities and why some of your best SEO endeavours may, already, not be working.

It may seem as if the web grows in a very unorganised and haphazard way. But that's not really the case. It's beginning to show powerful underlying regularities from the way in which web pages link together to the patterns found in the way users surf.

And it's interesting to note that, these regularities have been predicted on the basis of theoretical models from a field of physics, statistical mechanics, that few would have thought would apply to the web.

Among the chaos of activity and information on the web, scientists have analysed data which has been collected by the internet archive and other sources which has helped to uncover hidden patterns which hold many clues to what's really happening in cyberspace.

These patterns are being discovered all the more by the many researchers worldwide who are intrigued by the new science of networks. And the discoveries they make are both surprising and very, very interesting.

What is of most interest to us in our little search marketing community, is how quickly researchers have established that the distribution of pages and links per web site follows a universal and lawful behaviour. The simple truth of the matter is, few sites have enormous numbers of pages and many have few. And it follows that few sites have many links pointing to them whereas many have few.

How is it that the web in its distribution follows some kind of known patterns, when there is no central planner of the web?

There is no central body to suggest how it should grow and who should have links and who should not.

You need to look at the origins of network theory which throw a light on a number of social mechanisms which operate beyond the world wide web. These theories help to explain why the web has become a huge informational ecosystem that can be used to quantitatively measure and test theories of human behaviour and social interaction.

Phrases such as "it's a small world" and "the rich get richer" and "well connected" have worked their way into everyday vocabulary. It's interesting that such phrases have been the by-product of a mixture of research in social network analysis, physics, mathematics and computer science. All of which can (and do) apply to the algorithms used by the major search engines.

### **It's a small world:**

In the 1960s, American psychologist Stanley Milgram, was intrigued by the composition of the web of interpersonal connections that link people into a community. To inform himself more about this, he sent letters to a random selection of people living in Nebraska and Kansas, asking them, in turn, to forward the letters to a stockbroker in Boston. But he didn't give them the address of the stockbroker. Instead he asked them to forward the letter only to someone they knew personally and whom they thought may be 'socially' closer to the stockbroker.

Most of the letters did, in fact, eventually make it to the stockbroker. But the much more startling fact was how quickly they did so. It wasn't a case of hundreds of mailings to reach the final target, but typically, just six or so.

This true experiment has passed into folklore and is now famously known as "six degrees of separation". Although, it wasn't Milgram who named it so. That was from the 1993 play of the same name:

"I read somewhere that everybody on this planet is separated by only six other people. Six degrees of separation between us and everyone else on this planet."

Ouisa Kitteridge. From John Guare's play, Six degrees of Separation.

Popular culture plays its part again, when in 1997, a new game called "The Kevin Bacon Game" arrived on the scene. The game was invented by a couple of movie buffs who (for some reason of their own) had come to the conclusion that Kevin Bacon was the true centre of the movie universe (it has been proven that he is actually NOT the most connected actor in Hollywood circles, but nevertheless...).

If you haven't heard of the game, here's how it works. The movie network consists of actors who are connected by virtue of the fact that they have acted together in one or more feature films.

And this is not just Hollywood. This is any movie made anywhere. According to the Internet Movie Database (IMDB), between the years 1898 and 2000, roughly half a million people have acted in over two hundred thousand feature films.

So, here we go... If you have acted in a movie with Kevin Bacon, then you have a Bacon number of one (Bacon himself has a Bacon number of 0). As Kevin has acted in more than fifty movies, he has acted with more than 1,150 other actors. It follows, therefore, that over 1,150 actors have a Bacon number of one. Moving outward from Bacon, if you ever acted with an actor who had appeared with Bacon, then you have a Bacon number of two. And so on and so forth...

But the Kevin Bacon game is not the only one in town.

It's actually based on "Erdos numbers" these being applied to the distance between mathematicians who authored a paper with the great Paul Erdos (we're coming to him in more detail in just a few paragraphs) and those who authored a paper with a person who authored a paper with Erdos, and so on and so forth.

Much like the Kevin Bacon game, the smaller your "Erdos number" the higher the prestige you have within the mathematician community.

The existence of these short chains of acquaintance have actually been observed and documented by social network scientists for years.

Milgram's experiment with the Boston Stockbroker raised a couple of interesting issues. One regarding the properties that networks must have to become small worlds. If you were to draw a network of people (nodes) and links between those nodes relating to who knows whom, it wouldn't be at all obvious that any two nodes would be separated by six links. This is because there is something peculiar about a social network that is reflected in its link structure.

The second issue concerns what the best strategies are for navigating such small-world graphs in a short number of steps. Think about the people in Milgram's experiment. They did not have detailed knowledge of the social network in which they were embedded, but they still managed to pass the messages in a fairly short number of links.

These issues have since been addressed by Duncan Watts and Steven Strogatz at Cornell University and also Jon Kleinberg (in a small world way) also at Cornell. However, the conclusions and findings are vastly beyond the scope of this paper as an introduction and covered in more detail in the third edition of my book.



Does the same small world phenomenon exist between web sites and web pages? A few years ago, Lada Adamic, of Xerox, Palo Alto Research Centre, undertook a study of the average number of links you would need to traverse to get from one site on the web to another. She discovered that, just as in the social sphere, one could pick two sites at random and get from one to the other within four clicks.

This phenomenon was again shown to exist for the number of links between any two pages on the web. Albert Lazlo Barabasi, at Notre Dame University (more about Barabasi coming) discovered that, in the case of pages, the number is nineteen.

### **Getting connected:**

Hungarian mathematician and genius, Paul Erdos, was the first to address the fundamental question pertaining to our understanding of an interconnected universe: How do networks form?

His solutions laid the foundations of the theory of random networks. To explain: suppose we take a collection of dots on a page and then just haphazardly wire them together - the result is what mathematicians refer to as a random graph.

Okay, now imagine that you've been given the task of building roads to connect up the towns of an undeveloped country. At this time there are no roads at all, just fifty isolated towns scattered across the map. Because the construction guys are likely to misunderstand your plans and build roads linking the wrong towns and, of course, the country has so little money, you need to build as few roads as possible. The question is then: how many will be enough?

Mathematician and author Mark Buchanan, in his excellent book, *Nexus*, explained it this way:

If finance wasn't a problem then you'd simply order the construction guys to keep building until every last pair of towns were linked together. To link each of the fifty towns to all forty nine-others would take 1,225 roads. But what is the smallest number of roads you need to build to be reasonably sure that drivers can go between any two towns without ever leaving a road?

It's one of the most famous problems in graph theory and could be expressed in any number of ways; houses and telephone links; the power grid etc. It's a very difficult problem to solve and took the considerable mind power of Erdos back in 1959.

In this particular problem, it turns out that the random placement of 98 roads is adequate to make sure that the towns are connected. Even if that seems like a lot of roads, it actually only represents 8% of the original figure of 1,225 roads in total.

Erdos discovered that, no matter how many points there might be, a small percentage of randomly placed links is always enough to tie the network together into a more or less completely connected whole. To put this into the internet perspective, the percentage required dwindles as the network gets bigger. For a network of 300 points, there are nearly 50,000 possible links that could run between them.

But if no more than 2% of these are in place, the network will be completely connected. For 1,000 points, the crucial factor is less than 1%. And for 10 million points, it is only 0.00000166.

So, does that mean that if people were linked more or less at random, the typical person would have to know only about one out of every 250 million for the entire population of the world to be linked into a social web?



Let me just make a note here; one of the primary features of a random graph is that its degree distribution always has a particular mathematical form known as *Poisson distribution* (named in honour of the French mathematician).

### **The rich get richer:**

I want to just skip forward very quickly here for a moment. By 1999, Hungarian physicist Albert-Lazlo Barabasi had become completely engrossed in network theory, in particular, its application to the World Wide Web.

He himself had been schooled in the Hungarian tradition of graph theory, including the Erdos model of random graphs.

His excellent work in the field has given great insight into networks as diverse as those which begin as cocktail parties right up to the growth of the national power grid. And his innovative work has shown that many networks in the real world have degree distributions that don't look anything like a Poisson distribution. Instead, they follow what is known as a *power law*.

Barabasi's body of work has transformed the study of links and nodes. He has discovered that all networks have a deep underlying order and operate according to simple but powerful rules.

Duncan J Watts, one of the principle architects of network theory, has argued that, the origin of the Poisson degree in a random graph, and its corresponding cut-off, lies with its most basic premise: that links between nodes come into existence entirely independently of one another.

This means that, in an egalitarian system, things average out over time. An individual node can be unlucky for a while, but eventually, it has to be on the receiving end of a new connection. And in the same way, no run of luck can go on forever, so if one node gets picked up more frequently than average for some period of time, eventually others will catch up.

But you know, real life is not that fair unfortunately. Particularly when it comes to matters of wealth and success. Let's just think about the growth of a social web, as posed earlier, from the mathematicians viewpoint to begin with.

Duncan Watts puts it this way. Imagine you have a hundred friends. And each one of those also has a hundred friends. This means that at one degree of separation you can connect to one hundred people and within two degrees you can reach one hundred times one hundred which is ten thousand people. By three degrees you are up to almost one million; by four, nearly a hundred million; and in five degrees about nine billion people. What this would mean is, that if everyone in the world had one hundred friends, then within six steps, you can easily connect yourself to the population of the entire planet.

But as he also points out, if you're at all socially inclined, you'll already have spotted the fatal flaw in the reasoning.

A hundred friends is a lot to think about. So think about your ten best friends and then ask yourself who their ten best friends are. And the chances are that you'll come up with many of the same people. Go to Orkut now (if you can remember what Orkut is!) and check on your ten best pals in the search marketing network to get a real life understanding of this.

It's what known as *clustering*. We tend not so much to have friends as we do groups of friends, based on shared interests, experience and location, all of which overlap with other groups. And this is an almost universal feature, not just in social networks, but of networks in general.



It's this social network phenomenon which is the underlying cause of the rich getting richer. And this phenomenon has been with us for a long, long time. The great twentieth century sociologist Robert Merton dubbed it the "Mathew effect" as a reference to a passage in the Bible, in which Mathew observes, "For unto everyone that hath shall be given, and he shall have abundance; but from him that hath not shall be taken away even that which he hath."

[Once the vision of Michael Palin, in Monty Python's *Life of Brian* dissipates from your mind, I'll continue here...]

The Mathew effect, when applied to networks, basically equates to well connected nodes being more likely to attract new links, while poorly connected nodes are disproportionately likely to remain poor.

In fact, it has been proposed that the rich-get-richer effect drives the evolution of real networks. If one node has twice as many links as another node, then it is precisely twice as likely to receive a new link.

Let's return, for a moment, to Barabasi's introduction of power laws to bring us into a real world example. The distribution of wealth in the Unites States, for instance, resembles a power law. The nineteenth century Parisian engineer Vilifredo Pareto was the first to notice this phenomenon which subsequently became known as *Pareto's law* and demonstrated that it held true in every European country for which the relevant statistics existed.

The law shows that very many people possess very little wealth, while a very small minority are extremely wealthy. We tend to refer to *Pareto's law* more generally as the 80/20 principle.

Interestingly, a similar process tends to underlie the growth of social networks. A study by sociologists Fredrick Liljeros and Christopher Edling of Stockholm University, working with a team of physicists from Boston University, looked at the links of the sexual contact between 2,810 randomly selected individuals in Sweden. If acquaintance is a fairly loosely defined relationship, the existence of or non existence of a sexual link is not.

In the sexual context, these are the people who Malcolm Gladwell, in his book, *The Tipping Point*, referred to as "the connectors" a socially prolific few who tie an entire social network together.

You might put the prolific performance of the connectors in a sexual contact network down to special skills given at birth, or in early childhood...

But in the experiment carried out by Liljeros, another plausible explanation for the structure of the sexual contact network includes the increased skill in acquiring new partners, as the number of previous partners grows, and the motivation to have many new partners simply to sustain self-image... As you can see, it's evident in the sexual contact network that, the rich get richer here too.

This is a scale free network described by Albert-Lazlo Barabasi as a power-law or fat-tail distribution for network elements according to the number of links they have.

The physicists themselves believe that their approach is the best for understanding the evolution of networks. It's a direct generalisation of the usual physics of growth, percolation phenomena, diffusion, self-organised criticality, mesoscopic systems etc.

The physicists' approach has brought with it enough mathematical and computational data that you can fry your brains just trying to get your head around the basic concepts. But that's not what I'm trying to achieve here. I'm hoping that I'm able to give some basic but useful background to the way that social network concepts are applied and identified in the connectivity graphs used by each search engine when analysing linkage data.



You should also note, of course, that "connectionism" is very much a descriptive word applied to the field of AI.

Perhaps the greatest discovery of the laws of network organisation focuses on the idea of "hubs" and how they form. These are the centrepieces of networks, around which many links form.

Before we even apply it to the web and search, a strong case has been made by Barabasi (and subsequently by others) that the best way to combat AIDs, for instance, would be to concentrate on identifying and treating the hubs in sexual contact networks.

Information about the structure of the web is of great importance to search engines. The common observation is that, one good web document tends to link to other good documents of similar content. So therefore, there will be groups of pages of similar content (and similar quality) which refer to each other. The quality of the pages is presumed to be guaranteed by the recommendations implicit in the links between them. However, as we shall discover, this is not necessarily a good metric overall for quality as many had first thought.

Lada Adamic (Xerox) tested her theories (mentioned earlier) built around an application to examine a repository of web pages crawled by Google. For any given search word, she brought back results to the queries which provided PageRank, text match and link information for each page.

She then identified all the connected clusters and selected the largest one, as it would most likely contain links across sites other than just the common ones. What she discovered was, connected clusters spanning several sites tend to contain the main relevant pages and are rich in "hubs" (pages which contain links to many other good pages). It is then possible to find the centre of the cluster by computing the number of links among all the members of the cluster.

This shows that, rather than presenting a list of documents that contain many sequential entries from the same site, a search engine, using the phenomenon of the "small world" can present just the centre from each cluster. Users can then explore the rest of the cluster on their own.

### **Hyperlink based "popularity" algorithms.**

Maybe it was a small world event in a scale free network (pun), or simply a quirk of fate that, Jon Kleinberg, foremost computer scientist, found himself as a professor at Cornell University at just the same time as foremost physicist and sociologist Duncan Watts. Whatever it was, the information exchange between them in the study of networks has helped to transform the way that search engines relied almost solely on methods such as the vector space model which had pages "standing in isolation" to the two major hyperlink analysis algorithms: HITS and PageRank.

The application of network analysis and physics has given search engines fundamental principles to base ranking mechanisms on, among other things, clustering, interconnectivity and popularity.

I don't need to tell you that, the majority of web page accesses are referred by search engines, you already know this. Given the sheer quantity of information on the web, it's no wonder that search engines have become an indispensable tool.

An individual could never sift through the billions of pages online trying to find the ten best. So, that becomes the job of the search engine: To narrow it down to a smaller number of pages worth looking at.





This method of "topic distillation" to tackle the issue of "the "abundance problem" i.e. too many relevant pages being returned for a query, with little indication as to which are the most important, or authoritative, is centred around PageRank and HITS.

These algorithms applied to the link structure of the web fundamentally suggest the higher number of quality links you have pointing back at you, the higher you should rank in the results. It's a popularity metric.

Of course, the fact that strong hubs form in networks such as the web, the utopian dream of a free and equally democratic internet which many have dreamed of, becomes somewhat nonsense. Having covered the basic ideas of how networks form and strengthen and dominate topological degrees gives us an indication of what is bound to follow with a static hyperlink based ranking algorithm.

I write a lot about HITS/CLEVER which is a query specific algorithm. Using this approach, HITS builds a subgraph of the web which is relevant to the query and then uses link analysis to rank the pages of the subgraph. But for this particular article I just want to stick with PageRank as it is this approach which causes the accelerating "rich get richer" problem which many search marketers struggle with.

PageRank, is the most visible of the link based algorithms due to its association with Google and can be referred to more as a static ranking scheme. Using this method, all pages to be indexed are ordered once-and-for-all in a best-to-worst rank, regardless of any query. When the query does arrive, the index returns the best ten pages that satisfy the query at the top of the pile. Best, here, being determined by the static ranking.

But this is also the creator of a very worrisome problem which affects new web pages with low linkage data, regardless of the quality of those pages. Quality and relevance are sometimes at odds with each other. And the ecology of the web may be suffering because of the way search engines are biased towards a page's popularity more than its quality. In short, "currently popular" pages are repeatedly being returned at the top of the results at the major search engines.

So, the "filthy linking rich" get richer and currently popular pages continue to hit the top spots. The law of "preferential attachment" as it is also known, wherein new links on the web are more likely to go to sites that already have many links, proves that the scheme is inherently biased against new and unknown pages. When search engines constantly return popular pages at the top of the pile, more web users discover those pages and more web users are likely to link to them. This therefore means that currently unpopular pages (as such) are not returned by search engines (regardless of quality) so they are discovered by very few web users. And this, of course, is unfortunate for both the publishers of web pages and the seekers of their information. (Not to mention web marketers!)

This has been a lengthy journey already and we're still only scratching the surface. I want to finish by making you aware of an experiment which took place by scientists in America earlier this year.

First of all, they suggest that by 2002, around 70% of all web searches on line were being handled by Google. They also suggest that, while Google takes into account more than 100 factors in its ranking algorithm, the core of it is based on PageRank. This is a "static" link popularity metric to represent importance or authority for ranking purposes.

Now it's important to understand that there is a distinction between the importance or quality of page to that of the relevance of page following a user query.

The scientists suggest that the relevance is a quantity which relies heavily on the particular search issued by the user. But the importance or quality of a document could actually be computed at crawl time and could be seen as intrinsic to the document itself.



And the reason they are looking at this intrinsic quality is based on the desire to find a new paradigm for ranking web pages which is not so heavily based on link popularity.

The problem being that Google repeatedly returns "currently popular" pages at the top of the results and ignores newer pages which are not so densely connected. Therefore it is inherently biased against "unknown" pages.

So are the "rich getting richer" insofar as linkage is concerned at search engines? Yes and it's a rapidly worsening factor. The experiment carried out covered data collected over a seven month period. And from that experimental data, they observed that the top 20% of the pages with the highest number of incoming links obtained 70% of the new links after seven months, while the bottom 60% of the pages obtained virtually no incoming links at all during that period.

So where's the good news Mike? Well, there is a little consolation in that, the "rich get richer" behaviour varies in different categories. A new model has been developed which can be used to predict and analyse competition and diversity in different communities on the web.

However, that is covered in more detail again in the third edition of Search Engine Marketing: The essential best practice guide.

I'll round up here where I started with my dear departed Dad. Just as he became a social bright light earning (and burning) lots of cash, so he attracted lots of new friends (links). But when the Gaming and Lotteries Act in the late sixties forced him to close many of his venues (something he hadn't seen coming) the cash reserves slipped away... and so did the friends.

Still he left me with one excellent piece of advice. I said to him, it's alright saying you become a multi millionaire by becoming a millionaire first, but how do you do that?

He looked and smiled and said, in my experience I've discovered that looking for the million dollar deal is very difficult. Getting a million dollars from one person is hard. However, getting one dollar from a million people is really not so difficult.

Like myself, my father was much more of an optimist than a physicist!

You can find the research paper covering the "rich get richer" problem at Google here:

Impact of search engines on page popularity.  
<http://oak.cs.ucla.edu/~cho/papers/cho-bias.pdf>

Another interesting paper is:  
A new paradigm for ranking web pages on the world wide web.  
<http://www2003.org/cdrom/papers/refereed/p042/paper42.html/p42-tomlin.htm>

And if you really want to get into the real substance of network science, then I recommend:

Evolution of networks.  
[http://www.amazon.com/exec/obidos/ASIN/0198515901/qid=1097073023/sr=2-1/ref=pd\\_ka\\_2\\_1/104-3808226-6152737](http://www.amazon.com/exec/obidos/ASIN/0198515901/qid=1097073023/sr=2-1/ref=pd_ka_2_1/104-3808226-6152737)