How the Internet Works ☺

How the Internet Works ☺: Inspired by Per Bak

Lawrence Chou and Caitlin Keane
University of California, Los Angeles
Human Complex Systems 130

Overview

In this paper we discuss the importance of power law distributions as related to eight specific areas of the internet. After reviewing the basics of power laws, we look at the different number of users for social networking websites, instant messaging applications, video sharing websites, news websites, search engines, and blogs. These distributions were produced on a double-log scale of rank vs. size and discussed in terms of how tightly they fit a power law line. We conclude the paper with implications for business and a suggestion of what to research in the future.

Introduction to Power Laws

Generation upon generation has spent hours in the classroom, learning complicated statistical methods in an effort to better explain and predict useful data. In this widely used mathematical science, the focus is usually on the most common phenomena—the "average Joe" of data points. While these procedures may prove useful in predicting the most frequent or likely happenings, the most extreme and unique events are all-too-often overlooked. For example, while Gaussian statistical methods suggest that any extreme events be considered outliers and thus discounted, these data points may be immensely important. In order to better account for these occurrences (which can be anything from devastatingly large earthquakes to very successful companies), it is beneficial to study power law distributions. These distributions are characterized by the existence of "a relationship between two variables, such that one is proportional to the power of the other" (Abbreviated, 2001). This relationship is expressed by the function y =a(b^-k). Surprisingly, there are many groups of data that resemble power law distributions. In fact, it has even been postulated that most of the world is power law distributed, rather than normally distributed (McKelvey, 2009).

The first well-known application of a power law distribution was by Vilfredo Pareto in 1897, when he plotted Italian income on a graph (McKelvey, 2009). Rather than forming a normal distribution, the result was a visual representation showing that as the amount of wealth increased, the likelihood of someone attaining it decreased proportionately. When a power law distribution exists, and it is plotted on a double-log graph, the result is a straight line (McKelvey, 2009). Another major characteristic of these distributions is that they are scale-free, meaning that the "phenomena appear the same no matter what scale the measure is" (McKelvey, 2009). Though it is unknown for sure what causes such power law distributions, it is obvious how they differ from normal bell-shaped distributions. While normal distributions assume that every data point is completely independent of one another, this is the opposite for power laws. In fact, as connectivity or interdependence between the points increases, so does the possibility of having a power law distribution. For this reason, it makes sense that power law distributions would apply to more than normal distributions. It is extremely improbable that a living thing or an agent in a constantly changing, close-knit environment would exist completely independent of its surroundings.

Given that the internet is a system dependent on its connections and networks, we hypothesize that it is an optimal environment in which to study power laws. Furthermore, we believe at least three scale-free theories are evident in the continuously growing world of the internet. The first of these is preferential attachment, a theory which states that larger nodes or hubs seem to attract more connections the larger they become (Mitzenmacher, 239). In other words, there is positive feedback that continues and encourages growth, just as visitors to websites provide advertising revenue to continue building the site. Secondly, the least effort theory (first applied by Zipf to the distribution of word use) states that when there is a difference in the ease of use between subjects, there is a possibility of a power law distribution (McKelvey, 2009). When one site is much easier to navigate than others, more people will tend to use it. Lastly, niche proliferation suggests that "when production, distribution, and search become cheap and easily available, markets develop a long tail of proliferating niches containing fewer and fewer customers" (McKelvey, 2009). As the internet is virtually free to users, and incredibly simple to use, companies are able to supply rare and unpopular to the select group of prospective buyers. Because we believe the internet exhibits these scale-free features, relies on interdependence of sites, and has an appropriate amount of tension between its connections, power law distributions should be evident in different branches of the internet.

The Internet

In discussing present day interconnectivity, the internet and its derivative applications are increasingly being examined due to their proclivity for synergistic networking effects. The escalating prevalence of the internet has allowed unprecedented opportunities for individual heterogeneous agents to connect and spawn emergent order. Among the benefits afforded by the internet are the reductions of the traditional barriers to entry, increased opportunities for interconnectivity with a wide variety of agents, as well as the ability of the internet to act as an equalizer between the individual agents. In addition, the internet has helped to bring about an interesting change in human behavior in the sense that people have an increased belief in themselves, and their power and individuality. Because of the equalization brought about by the internet, a single individual now has a more weighted voice than before, leading to widespread increases in the trust an individual has for his or her own abilities and opinions.

The importance of the reduction of traditional barriers to entry is its direct relationship in increasing the potential for interconnectivity between agents. The internet allows a varied group of people, all from different creeds and backgrounds, to connect and pursue an assortment of relationships online. It is also important to note that there are less meaningful financial restrictions that are in place on the internet. A large majority of portals and internet based applications are available free of explicit financial pressures, which further reduces the inability for heterogeneous agents to connect.

The increase in interconnectivity opportunities is exemplified by the rise of the Web 2.0. A common definition of Web 2.0 is "a perceived second generation of web development and design, that facilitates communication, secure information sharing, interoperability, and collaboration on the World Wide Web" (Web 2.0, 2009). Web 2.0 incorporates innovative new concepts that have helped to spawn new web based applications that promote interconnectivity through the reinforcement and development of web-based communities such as the previously mentioned social networking websites and blogs (Boyd, n.d.).

Equalization brought about by the internet, as well as the behavioral changes that resulted, is an invaluable element of the increase in interconnectivity of the internet. To further elaborate, the behavioral changes that are most apparent are the ones that have contributed to the lessening of "shyness" on the internet. This lessening of shyness has helped agents reduce their inhibitions about making connections (Kittinger, n.d.).

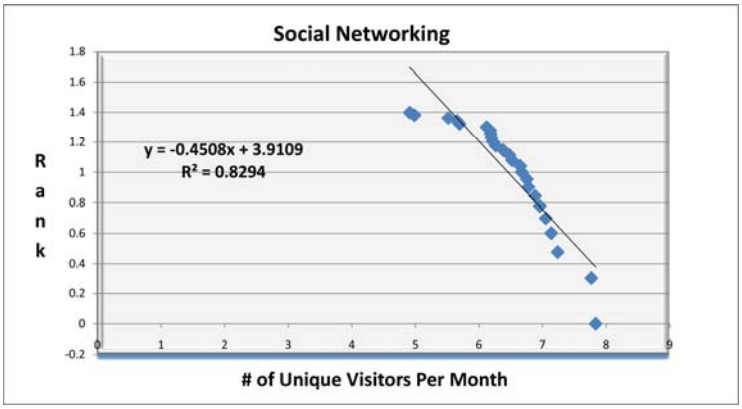
Perhaps one of the most important elements and aspects brought about by the internet is the ability for the internet to disseminate information quickly and parsimoniously. This dissemination of information creates an environment that allows preferential attachment to arise and thrive. Additionally, information that is quick and efficient will serve to support the "least effort theory" in that users should face less uncertainty about what services best fit their idea of "least effort".

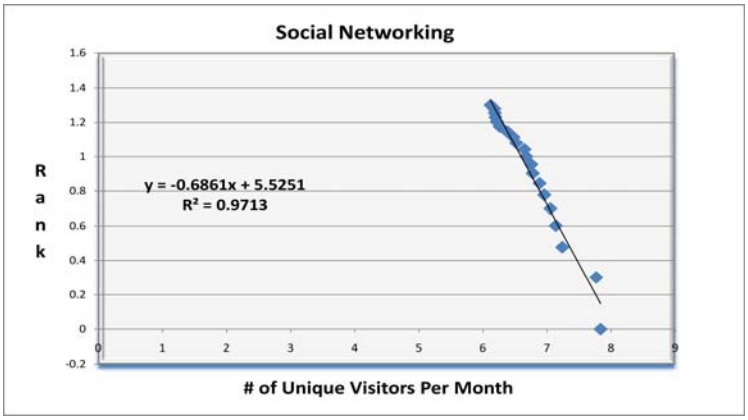
The internet landscape is dominated by various power law distributed industries. These industries include fairly influential and substantial portals such as landing pages, blogs, video entertainment sites, news sites, as well as search engines. In addition, internet based applications exist such as instant messaging, social networking, email, and web browsers.

Data Analysis

Social Networking

Social networking has been a rapidly growing and developing field, both in terms of popularity and prevalence (Krishnamurthy, n.d.). Social networking sites have benefitted significantly from positive feedback, as the strength and attractiveness of the community increases as more users join the service (Boyd, n.d.).





In our research, we examined the number of unique monthly visitors of the top 25 social networking websites. Notably, Myspace.com and Facebook.com were the two social networking websites with a significant unique monthly visitor advantage over the rest of the social networking websites that we examined. The combined unique monthly visitor total for Myspace.com and Facebook.com totaled 127 million, while the rest of the 23 social networking websites had a combined unique monthly visitor total of 99 Million (Social, 2009). These results are important in highlighting the broad differences in popularity between Myspace.com, Facebook.com, and the remainder of the social networking field.

When we plotted rank versus the number of unique visitors per month on a double log scale for the top 25 social networking websites, we were able to generate a linear relationship approximated by the equation "y=-0.4508x+39109" with an R^2 coefficient value of 0.8294. This R^2 coefficient is decently high, and presents us with a modest predictor of the criticality of the industry.

However, if we return to the idea that power law distributions emerge when interconnectivity and preferential attachment are most present, we can apply the inverse of that idea, which is that in a situation where there is a lack of interconnectivity and a lack of preferential attachment, we should see a lack of emergence of power law distributions, and a greater influence of Gaussian distributions due to the independence of the agents. By removing the smallest five social networking websites in our data set, we are in effect removing the social networking websites that have potentially made the least inroads in generating interconnectivity and preferential attachment. Once the five smallest data points were removed, we were able to generate a brand new linear approximation of "y=-0.6861x+5.5251", with a markedly improved R^2 coefficient of 0.9713.

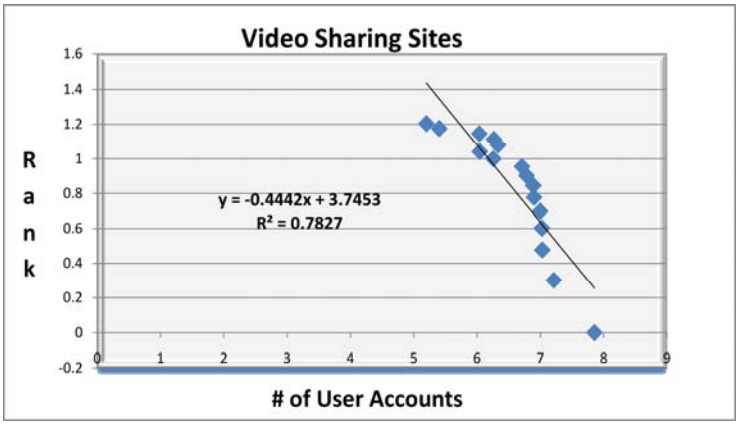
An R^2 value of 0.9713 implies a very strong linear relationship between the points and the linear approximation, leading us to theorize that we have a good predictor of the criticality of the industry for the top 20 social networking websites.

An exploration into the data points that deviate from the linear approximation shows that the three largest values represent Facebook.com, Myspace.com, and Classmates.com. Not surprisingly, these three social networking websites are the also the three largest social networking websites with the positions of Classmates.com, Myspace.com and Facebook.com as residing below the line, above the line, and below the line, respectively.

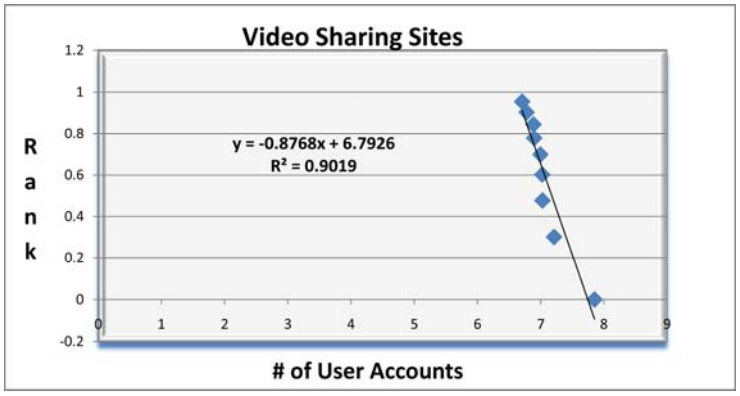
A simple, superficial analysis of the positions of the three largest social networking websites in relation to the generated linear approximation would lead us to assume that Classmates.com and Facebook.com are underperforming, while Myspace.com is outperforming the expectations of the industry based on their rankings. In an effort to propose why we believe Facebook.com signals underperformance, we must keep in mind that there exist two factors that determine the position of each data point; rank and the number of unique visitors. Facebook.com has been experiencing rapid growth and has just recently overtaken Myspace.com as the largest social networking website in the world (Schonfeld, 2009). With such rapid growth, we would expect that Facebook.com would show itself to be outperforming the linear approximation of the criticality of the social networking industry, but that is not the case.

Our hypothesis as to why Facebook.com is signaling underperformance is due to the current overpowering influence of the requirements of the top ranking. This negative, downward pull applies extreme pressure on Facebook.com's data point, which we believe will eventually be overcome with its continual positive growth patterns. If Facebook.com does continue with its current growth rate it will move its respective data point further to the right, without any downward pull at all, as it is already positioned with a top ranking. This movement solely in the rightward direction will allow Facebook.com to eventually outperform the linear approximation for the criticality of the industry, simply by maintaining its current growth rates.

Video Sharing Sites



Recently, with faster and higher definition connections, watching streaming video online has increased in popularity. Whether it be bored college students on Youtube.com, or parents watching the favorite television show they missed, millions of Americans log on to watch videos every month. In plotting a selection of fifteen popular video sharing sites, we expected to find a power law distribution, where the points fall on a straight line. The *x* axis contains the log of the U.S. visitors to the site in millions per month. The *y* axis is the log of the rank of the video sharing site. When every data point (every video site) is plotted on the log-log graph, the relationship between the data distribution and the power law line is not incredibly strong. (The R^2 value is 0.7827.) However, when we plotted only the first eight sites, the strength of the relationship grew to an R^2 value of 0.9019.

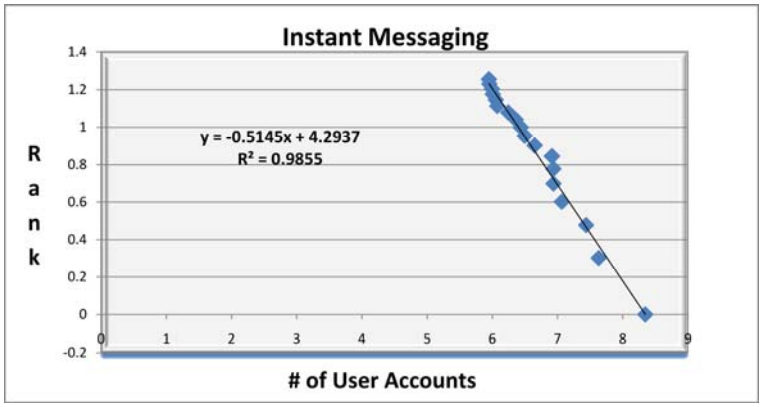


The rationale behind removing the smallest seven sites is that perhaps they are not as connected or developed yet. As the preferential attachment theory prescribes, as the sites increase in size, they grow more connections to other sides, thus increasing their possibility of become power law distributed. In looking at the data on the graph, this seems to be the case. The largest sites follow the line a lot more closely than the smaller sites. Preferential attachment makes sense for video sharing sites due to the fact that the videos on the site are uploaded by those that use it. The more videos that are on the site, the more people will want to visit it. More visitors equates to more videos, which creates the positive feedback that continues the growth of the site. A large site such as Youtube, which, at 71.3 million has nearly five times the users of second-ranked Hulu, probably benefits from preferential attachment (Quantcast, 2009). On the other hand, Hulu's content comes only from the companies that own and operate it, namely Fox and NBC. Therefore, users of the site cannot share their own videos. This could be the reason that YouTube is above the power law line while Hulu and the third and fourth ranked Netflix and BBC are below. YouTube is free to grow by the number of videos users ad while other sites are limited.

Despite this difference, the fact that Hulu is the second-ranked video site means that it has high connectivity regardless. Video sites are ideal for looking at connectivity because they in effect advertise for each other. If one site has certain episodes of a series but is missing the specific one sought by the user, the user may be inclined to search other sites for the episode in question. Sites may also draw attention due to the least effort theory, in that the site is extremely simple to navigate. Even more probable is that the increase in recent television commercials or other advertisements for a site has drawn more users to it. The niche proliferation scale-free theory may also apply to this industry. For example, while there are video sharing sites that only provide the most popular content to a wide audience, other sites such as YouTube contain rare and less popular footage. In this way, a long-tailed distribution forms where fewer and fewer site visitors are searching out less popular videos. The ability for customers to find this content may also contribute to the power law distribution, as it makes the distribution more skewed. Regardless of the specific reason, connectivity is high among video sites and coincides with our theory of power law distributions.

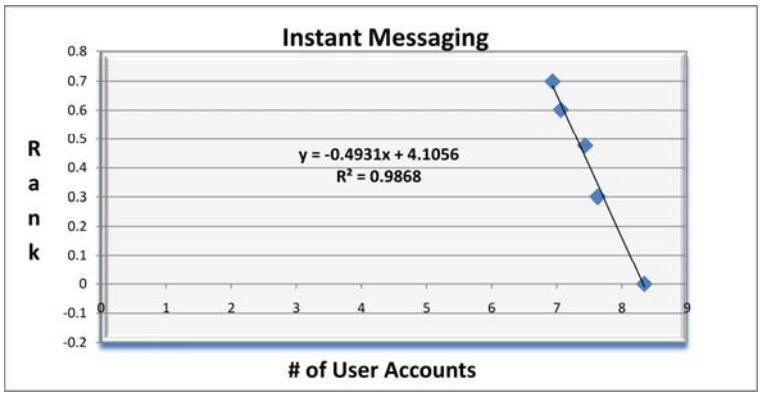
Instant Messaging

Instant messaging is a highly popular method of communicating, used both in social life and business settings. It involves the contemporaneous sending and receiving of messages through the internet from one device to another regardless of location. Instant messaging services have grown in both size and popularity over the last five years. Nardi et. al (2000) described the four primary functions of instant messaging as: "quick questions and clarifications, coordinating and scheduling tasks, coordinating impromptu social meetings, keeping in touch with friends and family" (82). Evidently, instant messaging is used in both social and corporate settings to maintain connectivity within partners or groups. Corporations are now turning to applications such as Sametime Connect to provide messaging services within their departments. Such programs facilitate instant and efficient communication without having to waist time on the phone or walking to another part of the department. Furthermore, instant messenger is advantageous over other internet forms of communication, such as email, for the reason that it is



instantaneous. One does not have to check repeatedly to see if they have received a message, because it automatically notifies the recipient.

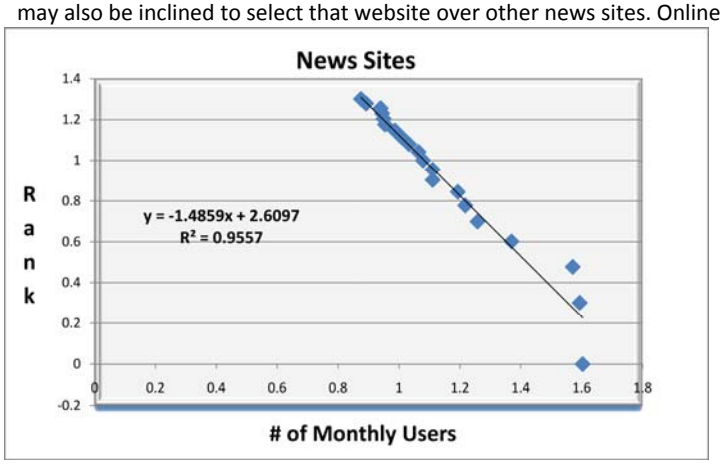
When we plotted eighteen instant messenger applications on the double-log scale (number of total users on *x* and rank on *y*), the strength of the relationship to the power law line had an R^2 of 0.9855. This relationship is very strong. When we remove all but five of the data points on the graph, the R^2 increases to 0.9868. Though this is a stronger power law distribution in relation to the line, five data points is not substantial enough to come to any conclusions. Again, these top-ranked instant messenger providers are probably the most connected to other web pages via embedded links or advertisement, increasing their power law characteristics.



According to McKelvey (2009), companies above or on the power law line of the double-log graph are performing better than those below the line. One would expect MSN messenger to be on or above the line, as it is the first ranked program with 221,095,000 total users, and doing somewhat well (Current, 2007). With this logic, however, it would be assumed that as the amount of users decreases, companies should automatically place farther and farther below the line. The second-ranked instant messaging service would be further below the line than the first, and the third would be further below the line than the second, etcetera. In reality, this is not the case. Looking at the graph, second-ranked AIM is below the line but third-ranked Yahoo! is on the line. There are multiple factors that go into determining where a company falls in relation to the power law line. For example, if AIM were to gain even a substantial amount of users, it could move across the line to the right but may not move down as far in terms of rank. This would be possible because the users would have to increase by many millions in order to get close to the first-ranked MSN. Additionally, managerial strategies and marketing strategies can also influence how well the company performs, which may or may not have a direct influence on the number of users.

News Sites

With the pervasiveness and ease of technology today, it is not surprising that there has been a large decrease in revenue for the newspaper industry. As individuals are able to find the same content on the internet for less money and less effort, online news sites steal business away from print news. In addition to ease and convenience, many news sites have the benefit of being perpetually advertised. Television news programs such as CNN or MSNBC have created online counterparts. As a result, those people that choose a specific television program may also be inclined to select that website over other news sites. Online news sources allow visitors to personally navigate through the site and choose what news stories to read or watch. This may be an example of niche proliferation, as visitors can choose to read obscure stories in unpopular topics rather than being forced to watch or read what is in the highest demand.

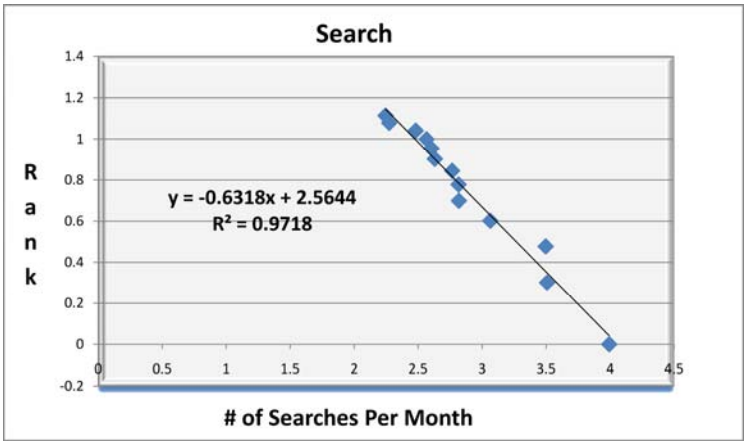


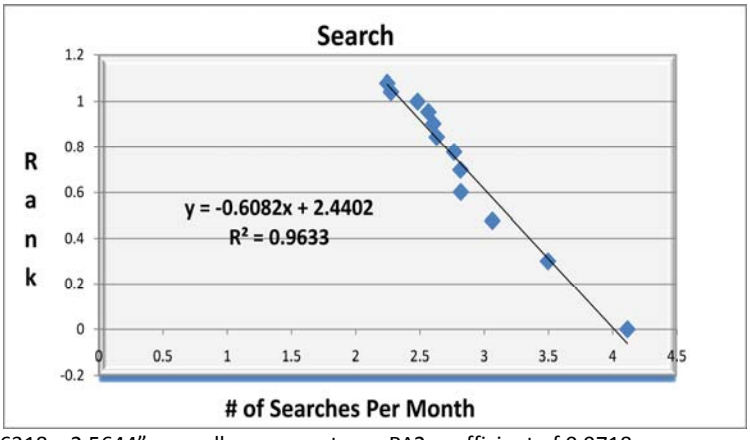
What is interesting about the distribution here is that the top three websites are the ones furthest from the power law line. Whereas other graphs show more of a power law distribution towards as one moves to the highest ranking data, here MSNBC, Yahoo! News, and CNN appear to be almost in a vertical line. They are all within two million viewers, and then the fourth data point has four million less (Exclusive, 2009). It should also be noted that MSNBC is significantly below the power law line, while Yahoo! News is slightly above, and CNN is even further above. This suggests that MSNBC is under-performing while the other two are over-performing. Because they are so close in number of monthly users, it is likely that Yahoo! News will soon overtake MSNBC as the first-ranked news site. Though it is unlikely that the number of monthly users of MSNBC will drastically decrease since the number

of internet users is continuously increasing, it is possible that the number will remain stagnant. One postulation for why this may happen is that MSNBC is the standard homepage programmed on Internet Explorer, a very popular internet browser that comes with Windows. This could account for a large proportion of the users due to the least effort theory. It takes more effort to change the home page to a different news site than it does to keep MSNBC. However, as Macs become increasingly popular (as they are with college students), people may use MSNBC less and less. Macs come standard with Safari as the internet browser, and therefore do not automatically direct people to MSNBC. In other words, it would reverse the least effort theory as it would take people more effort, not less, to reprogram their homepage to MSNBC. In this case, MSNBC retains its 40.1 million monthly users while the other two companies may continue to grow (Exclusive, 2009). Another possible explanation for this could be that Yahoo! and CNN may have increased their connectivity with other sites. To do so they could increase advertising by placing headlines in other sites with links to the full news site. By using the most attention-grabbing stories, perhaps they could draw more users. Regardless of the variance in the largest companies to the line, news sites generally adhere to the power law line.

Search

Perhaps one of the most visible industries on the internet, search has had widespread acceptance and proliferation in terms of usage. A web search engine is a "tool designed to search for information on the World Wide Web" and the search engine acts as the interface between which a user will be connected to the information that he or she is seeking (Web search, 2009). Searches can encompass a variety of different topics and mediums, with no restriction on the qualitative measures on the type of information. Certain search engines now have begun to structure themselves to serve special types of searches, such as image searches, video searches, news feed searches, on all top of the traditional baseline web search. Notable companies such as Google, Yahoo, and Microsoft have competed fiercely for market share by promoting their own selling points, such as ease of use and accuracy of results. Due to the widespread proliferation and acceptance of web searches, we expect to find a very strong power law relationship in existence in the search industry.





We examined the number of searches per month for the top 13 search websites, and plotted the data on a double log scale of rank versus number of searches per month. Google was the corporate entity with the largest volume of searches per month, with 9.8 billion searches. Google significantly outpaced the rest of the industry, as YouTube and Yahoo had monthly searches of 3.2 billion and 3.1 billion, respectively (comScore, 2009). In examining the data, we were able to generate a linear approximation of "y=-0.6318x+2.5644", as well as a very strong R^2 coefficient of 0.9718.

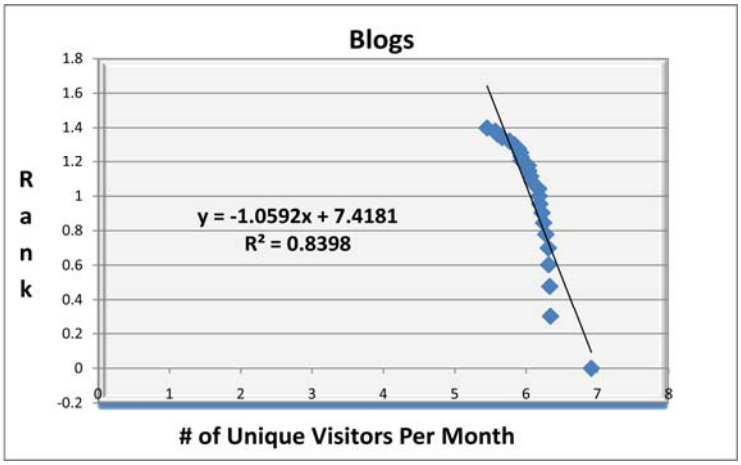
The strength of the R^2 coefficient has led us to believe that search is a very interconnected industry, with bountiful reinforcing of the preferential attachment behavior, as well as seemingly continuous tension to improve the number of searches per month due to the potential profits afforded by the industry. Such a high R^2 coefficient implied a very strong linear relationship between the points and the linear approximation, leading us to theorize that we were able to generate a good predictor of the criticality of the industry.

It was widely publicized in 2006 that Google purchased YouTube for $1.65 billion in an effort to promote synergies between the two companies as well as serve as a platform to further strength Google's search platform. If

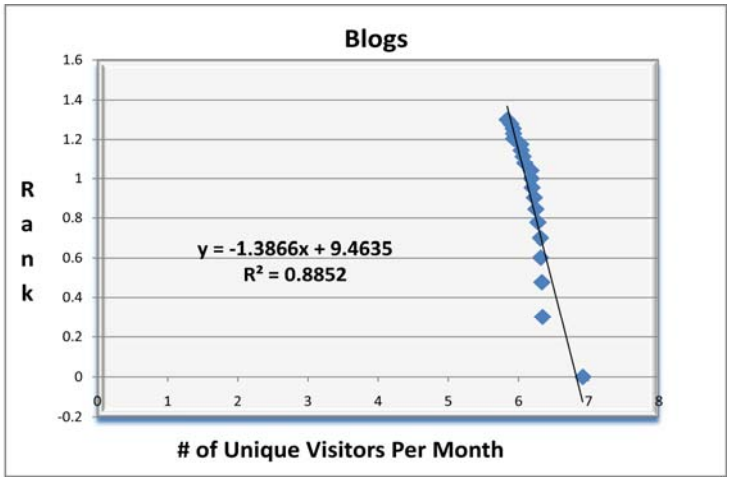
we were to combine YouTube's searches with Google's searches, we would come up with a stunning 13 billion monthly web searches, compared to second place Yahoo at 3.1 billion. By combining the search totals for YouTube and Google, and re-plotting the data on a double log scale of rank versus number of searches per month, we find that the Google corporate entity outperforms the expectation of a leading search engine company in the realm of search volume. We also find that Yahoo no longer outperforms, but rather underperforms in terms of searches for a second ranked company. These findings help to highlight what seems to reflect the current business climate; Google is highly profitable, an industry leader, and dominant in search, while Yahoo holds considerable search, but seems to continue to underperform expectations.

Blogs

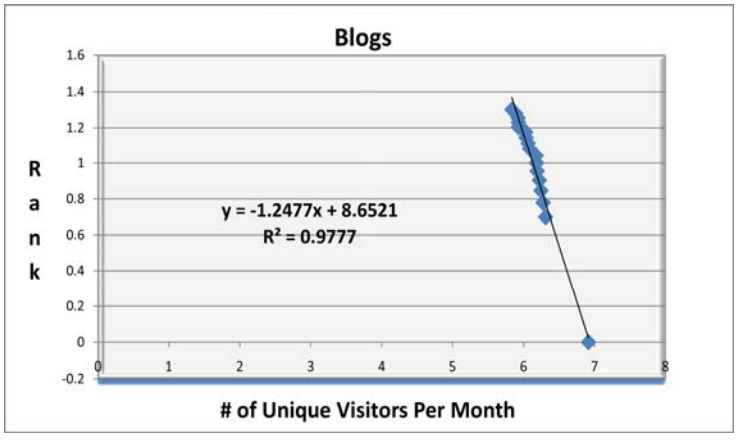
A popular definition of a blog is "a type of website, usually maintained by an individual with regular entries of commentary, description of events, or other material such as graphics or video" (Blog, 2009). Blogs essentially are written by individuals, and read about and commented about by all the blog's readers. It is because of this interesting dynamic that we suspect that the blogging network will display the characteristics of a power law distribution.



In our research, we examined the top 25 blogs on the internet. Celebrity news sites and technology sites dominated the list of most visited blogs, which may have been a reflection of the audience makeup of bloggers, as well as society's growing fascination with celebrities. The most visited blog on the internet was TMZ.com, with 8.2 million visitors per month, followed by gizmodo.com with 2.2 million visitors per month, followed by Perezhilton.com at 2.1 million visitors per month (Top 25, 2009). Rounding out the top five most visited blogs were another set of celebrity gossip websites and technology based blogs, Gawker.com and BoingBoing.net. Gawker.com and BoingBoing.net received web traffic of a little less than 2.1 million visitors per month and 2.0 million visitors per month, respectively.



It is interesting to note that when we examined the data using our standard double log analysis, we received a decent R^2 coefficient of 0.8398, along with a linear approximation of "y=-1.0592x+7.4181". The data also presented some very interesting findings in terms of the distribution, in that we initially found TMZ.com to be underperforming according to our approximation, as well as the immediately following top 4 blogs. However, after examining the data further, we found that the last 5 data points seemed to be following more



of a Gaussian distribution, than a power law distribution. After removing these points from our data set, our new linear approximation became "y=-1.3866x+9.4635", and our R^2 jumped to 0.8852.

With the new data set that adjusted for potentially heavier Gaussian influenced points, we found that TMZ.com was actually outperforming the theoretical criticality of the blogging industry, and the top 5 points all had improved results compared to the initial double log treatment.

From an investment perspective, as we are able to tell from the data, the blogging industry is heavily dominated by celebrity gossip and technology based blogs. It appears as if Gizmodo.com, as the second most visited blog is underperforming by the greatest margin compared to the rest of the industry, while Perezhilton.com seems to also exhibit some underperformance tendencies. If we were to remove Gizmodo.com, Perezhilton.com, and Gawker.com from our data set, we would find that the new linear approximation becomes "y=-1.2477x+8.6521" with a significant increase in the R^2 coefficient; R^2 becomes 0.9777.

By removing the 3 companies that are underperforming the most according to our linear approximation, we find that the rest of the industry appears to act and operate in a healthy, efficient manner according to the linear criticality that we generated for the blogging industry. This marked shift in the strength of our linear approximation leads us to believe that Gizmodo.com, Perezhilton.com, and Gawker.com may potentially serve as a good investment, in the sense that a better management team may be able to attain growth rates that will outperform the blogging industry. Perhaps another possibility is for a consolidation of websites such as Perezhilton.com and Gawker.com under the same company in order to produce synergistic effects and help build each other's website visitor statistics. On the internet, this type of growth may be as simple as reciprocal links between two popular websites, or in this case, reciprocal links between two popular blogs that share the same type of customer base.

Conclusion

As the internet is still considered to be a maturing medium, there is still much left up to hypothesis and theories. While power law distributions have been observed in many online industries, we still cannot present a definitive reason as to why this phenomenon emerges. However, while we cannot definitively say, we can hypothesize.

Our hypothesis as to why these power law distributions occur in different sectors of the internet lies in the increase of interconnectivity on the internet. In addition, the ability of the internet to efficiently facilitate positive feedback through preferential attachment contributes to the existence and proliferation of power law distributions online. As interconnectivity increases, it appears that our respective data trends closer and closer towards a straight line while on a double log scale. Though we chose only a handful of industries to analyze, the structure of the internet appears to increase the possibility, if not probability, of producing power law distributions.

It is important to note that although power law distributions may provide us a linear approximation of the criticality of an industry, any deviations from this linear approximation may not be an authoritative expression as to whether a company is underperforming or outperforming its industry. As we explored with the Facebook.com and Myspace.com example, rapid acceleration through the ranks provides heavy downward pressure, which may signal that a company is underperforming, even if they are experiencing very fast growth rates.

Perhaps a more informative measure of whether a company is outperforming or underperforming relative to its industry is to individually examine each industry to discover whether or not the industry is power law distributed, and then to discover how well the industry is power law distributed without the companies that are created the most deviations from the linear approximations. In our blogging example, when we removed the potential Gaussian points as well as the 3 companies that were significantly underperforming in the industry, we saw that our R^2 coefficient jumped from .8398 to .9777. If we are able to imagine that the blogging industry would exist with a R^2 coefficient in regards to our linear approximation, then it may be more realistic for us to hypothesize that those 3 companies may be underperforming the industry.

Finally, it should be examined why certain industries have stronger linear approximations than others. We propose that that in the industries that are built upon interconnectivity, we see strong evidence of power law distributions, as well as very high R^2 coefficients. Our social networking data, search data, and our instant messaging data supports our assumptions. Conversely, our video sharing data showed a relatively low R^2 coefficient value, which leads us to theorize that as a newer industry, video sharing sites have yet to settle into a power law distribution.

The internet provides a great medium for interconnectivity, positive feedback opportunities, and opportunities for power law distributions. Our linear approximations of double log relationships seem to provide us with strong correlations, but, as with any type of indicators, further examination usually provides an even clearer picture.

References

Adamic, L. A., & Bernardo, A. H. (2002). Zipf's law and the internet. *Glottometrics*, *3,* 143-150. Retrieved June 8, 2009 from http://www.hpl.hp.com/research/idl/papers/ranking/adamicglottometrics.pdf

An abbreviated glossary of system terminology. (2001). Retrieved June 10, 2009 from http://www.esse.ou.edu/glossary_st.html#P

Blog. (2009) Wikipedia: The Free Encyclopedia. Retrieved June 10, 2009 from http://www.en.wikipedia.org/wiki/Blog

Boyd, S. (n.d.). Web 2.0: A social revolution, 1-35.

comScore releasesApril 2009 U.S. search engine rankings. (2009). ComScore. Retrieved June 10, 2009 from http://ir.comscore.com/releasedetail.cfm?releaseid=384605

Current stats and commentary on the instant messenger market: eBuddy and Meebo lead the pack. (2007). Center Networks. Retrieved June 10, 2009 from http://www.centernetworks.com/current-stats-and-commentary-on-the-instant-messenger-market

Exclusive: Top 30 news/current event sites for April- Huff Post and Fox both gain. (2009). Editor & Publisher. Retrieved May 29th, 2009 from http://www.editorandpublisher.com/eandp/news/article_display.jsp?vnu_content_id=1003975048

Instant messenger. (2009). Wikipedia: The Free Encyclopedia. http://en.wikipedia.org/wiki/Instant_messaging

Kill your tv: 15 internet video sites to beat the writers' strike. (2008). PCWorld. Retrieved June 10, 2009 from http://www.pcworld.com/article/141106-5/kill_your_tv_15_internet_video_sites_to_beat_the_writers_strike.html

Kittinger, R. (n.d.) Social behavior of internet users, 1-25.

Krishnamurthy, B. (n.d.) A measure of online social networks, 1-18.

McKelvey, Bill. (2009, May).Power laws. Lecture presented at the University of California, Los Angeles

Mitzenmacher, M. (n.d.) A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, *1*(2), 226-251. Retrieved June 10, 2009 from http://www.internetmathematics.org/volumes/1/2/pp226_251.pdf

Nardi, B., Whittaker, S., & Erin, B. (2000). Interaction and outeraction: Instant

messaging in action, *Computer Supported Cooperative Work*, 79-88.

Quantcast. (2009). Retrieved June 10, 2009 from http://www.quantcast.com

Scale-free networks. (2009) Wikipedia: The Free Encyclopedia. Retreived May 29, 2009 from http://en.wikipedia.org/wiki/Scale-free_network

Sconfeld, E. (2008). Facebook is not only the world's largest social network it is also the fastest growing. Retrieved June 9, 2009 from http://www.techcrunch.com/2008/08/12/facebook-is-not-only-the-worlds-largest-social-network-it-is-also-the-fastest-growing/

Social networks: Facebook takes over top spot, Twitter climbs. (2009). Compete. Retrieved on June 10, 2009 from http://blog.compete.com/2009/02/09/facebook-myspace-twitter-social-network/

Top 25 most popular blogs. (2009). eBizMBA. Retrieved June 10, 2009 from http://www.ebizmba.com/articles/blogs

Web 2.0 (2009). Wikipedia: The Free Encyclopedia. Retrieved June 10, 2009 from http://www.en.wikipedia.org/wiki/web_2.0

Web search engine. (2009). Wikipedia: The Free Encyclopedia. Retrieved June 10, 2009 from : http://en.wikipedia.org/wiki/Web_search_engine