Social Scaling: From scale-free to stretched exponential models for scalar stress, hierarchy, levels and units in human and technological networks and evolution[1]

Douglas R. White © 2003

ISCOM working paper (live http: links in PDF version)
For eventual submission to: *Computer and Mathematical Organization Theory*

Abstract. Johnson's (1982) model of scalar stress deals with how networks are stacked at different levels to reduce information and energy load by substituting relationships among leaders of hierarchically ordered groups for relationships among members of larger groups at a lower level in the hierarchy. The logic and scaling results of this model are important elements in a theory of network and social scaling. They point to the possibility of scale-free modeling of the modularity of networks based on the relative constancy of the basic units at the individual level that give structure to these networks, the flexibility of how particular groups are organized, the fact that network hierarchies are population-filling with scale-free relationships to population size, and the bulking, organization and conservation of energy, information and material in ways that match the constraints on populations of individuals. These characteristics of scale-free modeling have been successful in biology, and social scaling may well follow the same principles. This article suggests the kinds of modifications that made be needed for larger-scale integrative projects in social scaling.

Hierarchical and power law models have been much debated in recent decades and their limitations exposed. While Johnson's work contains important insights, this paper examines new types of models that account for observed attenuations in the finite regimes of scale-free distributions (the stretched exponential model) and broken scale-free regimes. A combination of stretched exponentials and network modeling is found to be a productive approach to social and economic scaling that yields theoretical predictions about basal units, moments of distributions, attenuated and broken power-law regimes.

Studies of scale-free, cutoff, and hierarchical properties of the U.S. airlines network in 1997 and a physics citation network are used to compare Johnson's findings with basal unit and scale-free regimes in a more suitable class of scaling model that uses the stretched exponential. This model estimates hierarchy levels and basal unit characteristics and finds a similar basal unit of 6 for renormalization at a second level (hubs for local neighborhoods) in the airline industry, suggestive of Johnson's results. The citation network suggests three-levels of multiplicative effects and a basal unit of 3 that is well under Johnson's limit of 6 but constitutes a minimum unit of social cohesion.

[1] This paper was developed out of the stimulation and materials provided by the first meeting at the Santa Fe Institute on "Social Scaling" under the auspices of the ISCOM (Information Society as a Complex System) project directed by G. West, D. Lane and S.E. van der Leeuw and funded by the European Union (Project # IST 2001-35505). Special thanks to S. E. van der Leeuw for providing the G. A. Johnson article, to G. West for providing a clear demonstration in the biological realm for scale-free modeling, and to the face-to-face and email discusions in the ISCOM networks project having Luis Bettencourt and Tim Evans as core members.

Johnson (1982) develops an empirically based scaling of human performance stress as a function of network linkages that are rescalable from individuals to groups as a way of renormalizing information load with respect to human attentional capabilities, limited to tracking $\boldsymbol{s} = 6 \pm 2$ simultaneous dimensional judgments (Simon 1973). For sets of individuals performing tasks, experimental data show that individual performance increases when the group size grows from 1 to 6 and then degrades. Just as memory and information capacity can be bulked into chunks, however, Individuals can form groups with leaders and informational capacities can be renormalized around the same optimum of 6 ± 2 ties between for each leader in a clique of leaders. In doing so, information overload (stress) is renormalized. This process can be carried on through many levels and iterations.[2]

As a human population increases to size p, optimal renormalization of stress, in its simplest form, will generate a fractal-like structure of groups in a hierarchy that begins with one central node, branching at each level until a level h is reached that minimally satisfies the constraint $h \geq \log_S p$ or $s^h \geq p$. A fractal hierarchy is one with successive divisions with no overlap among them. This corresponds to what Johnson calls a sequential hierarchy. By fractal-like I mean that the number of splits at each level (and thus the sizes of groups) are not strictly uniform but vary, with new detail appearing at each scale.

As a simplification, consider that there are s^h (renormalized) groups that grow exponentially as h increases. In this or the looser fractal-like form, there is a second form of information overload, that of too many leaders in the network, when a smaller number will do for purposes of social stability or synchronization.

A finding of Johnson's (1982: 413-414) empirical analysis is that the total number of recognized types of leaders,[3] relative to the fractal hierarchy, attenuates as scale increases.

My purpose here is to show that this attenuation can be shown as a scale-free model (power-law) that fits his data and provides a correct intercept and slope of his log-log scaling result (Johnson's Figure 21.5, r=.96) between log population and log of recognized types of leaders. Figure 21.5 is redrawn as my Figure 1, extended to show that extrapolating to a population of 1, the leadership rate ~ 1000/1000 ~ 1.

---

[2] The value s ~ 6 is taken to represent a situation where information on a single dimension or coming through a single link is itself multiples. If links or incoming information streams are themselves simplified, s may increase. Say the norm is that each link has 12 subdimensions, and the true capacity for the product of dimensions is 72, so that s ~ 72/12 = 6 is a normal information load but s ~ 72 is possible if each link or incoming stream of information has but a single subdimension. The normal case of handing 6 x 12 of course, results because recognition patterns are already bulked.

[3] Johnson takes the number of recognized types of leaders as a proxy for the number of recognized leaders.

Variables

I define the following variables in accordance with Johnson to replicate his Figure 21.5. Let:
h = the number of hierarchical levels of units that divide a population (or its subpopulations) into sequential units

b_i = population size of i^{th} level, i=0 (top) to h.
b_0 = p = total population size
$b_h = 1$, $\log_{\boldsymbol{s}} b_h = 0$
b_{h-1} = s ~ 6 in first approximation of scalar capacity, later we see that s can vary where s = c/a and
a = # activity streams watched (e.g., 1-12 ~ limit)
c = capacity ~ 72

Fractal groups hierarchically filling populations

In representing Johnson's model, fractal groupings are assumed to fill a hierarchical structure for a population. Each node has assigned to it the entire population in the groups it subsumes at lower levels. The b_i populations subsume successively smaller populations while the number of nodes at each level grows as s to the power i. Their product is constant. For example, if h=4, s = 10 (not 6!) = b_{h-1}, and p=10000, we have the following stacks of values for the b_i, the powers of s^i, and the product b_0s^0:

$b_0 = 10000$	$s^0 = 1$	$b_0s^0 = 10000$
$b_1 = 1000$	$s^1 = 10$	$b_1s^1 = 10000$
$b_2 = 100$	$s^2 = 100$	$b_2s^2 = 10000$
$b_3 = 10$	$s^3 = 1000$	$b_3s^3 = 10000$
$b_4 = 1$	$s^4 = 10000$	$b_4s^4 = 10000$

for all i, $b_i \cdot s^i = p$ thus $\log_s b_i + i = \log_s p$ (1)
thus $0 + h = \log_s p$ (1')

Leadership and Attenuation

If we assume each fractal level i with s^i units has a leadership position, then

$d = \Sigma_{i=1,h-2}\, s^i = s^{h-1}-1$ sum of leadership positions, discounting single individuals and elevating representatives of lower groups up the chain to fill all positions.

d/p = leadership rate per person ~ s^{h-1}/p, so (2)

$\log_s (d/p) = (h-1) \log_s s - \log_s p$, by (1') = h–1–h = –1, resulting in a constant (3).

Johnson does not get the result in his Figure 21.5 that leadership rate per person is invariant with respect to population size. He hypothesizes attenuation of who is considered a leader according to the number of levels. The literature in ecological psychology finds ample confirmation of this in the higher per capital opportunities for leadership in smaller-scale populations.

A Scale-free attenuation model

Let $d/p \sim (s^{h-1})^{\alpha}/p$ represent the attenuated sum of leadership positions, adding to (2) the attenuation coefficient α. Then:

$\log_s (d/p) = \alpha\,(h–1) \log_s s – h = \alpha\, h – \alpha – h = \alpha – (1–\alpha)\, h$, and, by (3),

$$\mathbf{\log_s (d/p) = – \alpha – (1- \alpha) \log_s p} \qquad (4)$$

It is easier here to use $\beta = 1– \alpha$ as the power coefficient since s is a constant, where by raising both sides of (4) to powers of s,

$$\mathbf{d/p = s^{\beta–1}\, p^{–\beta}} \qquad (5)$$

where the intercept is $s^{\beta–1}$ and the slope is – β, as in Figure 21.5.

Johnson's Figure 21.5 suggests attenuation on the order of $\beta \sim 6/7$. This implies a value for s, but not one that is sufficiently precise to pinpoint in the range $s = 6 \pm 2$, although this range is not inconsistent the observed intercept. In one sense that makes the model robust since there is no great dependence on a special value of s, and a wide range of values near 6 are consistent with the intercept in Johnson's data.

Thus for human populations, Johnson's scale-free model for the hierarchical sociopolitical dimension of human evolution with leadership recognition decaying at a $\beta \sim 6/7$ power provides indirect evidence for uniform modularity in human hierarchical groups, but the estimate of sizes that play a role in modularity are within a range that contains $s \sim 6 \pm 2$. A tighter range is not needed for consistency of the beta decay in recognizing leaders of groups with the general principle that fractal-like rescaling of groups at different levels in political hierarchies is consistent with his result, but not tightly constraining as to the uniformity of those groupings.

The attenuated rate of political leaders (or types of leaders) in a population in this fitted empirical model (r=.96 with agreement between theory and data as to intercept and slope) is a function of three parameters, increasing in $\boldsymbol{p}$ with $\boldsymbol{p}^{-\beta}$ attenuating by the power –β, and with an initial intercept of $s^{\beta–1}$ that depends on only very rough bounds to s within which groups of different sizes, or groups of groups, are brought under common leadership thereby forming new levels in political hierarchies. Fits for a range of values of s are shown in Figure 2. The attenuation to the power $\beta \sim 6/7$ is reasonably close to the type of power-law attenuation observed in biological scaling (West et al, 1997), the network-based construction of this law follows the same structure-filling and energy-optimization principles as the biological scaling laws, and the intercept of $\mathbf{s^{\beta–1}}$ for a single individual converges to a reasonable value ~ 1.

What have we learned by revisiting Johnson's model? First, as to the quality of his original log-log fit, it is significant that the intercept of his power-law scales at 1 leader for a population of 1, which reinforces the plausibility of the finding given the relatively small number of cases (N=23, but with r=.96 along two continuous scales). Second, a mathematical derivation of the power-law from the principles that he advances in his study shows consistency with his argument and findings, but also finds, importantly, that the power-law is relatively indifferent to the bulking coefficient for sequentially stacked groups in the hierarchical model, which may indicate that variations of group size have little effect on how efficiently the slope of the power-law can be estimated. It would then be well worth seeking more data on the attenuation of recognized political leaders (or types of leaders) with population. Third, since we know that group-bulking occurs at different levels of leadership and that there is a basis for Johnson's scalar stress hypothesis, the consistency of the mathematical derivation with s ~ 6 ± 2 is a credit to his theory.

Perhaps most importantly, the level of goodness-of-fit of a scale-free model that incorporates hierarchy in human groups (measured by the leadership variable) is encouraging for further research. Following the logic of Johnson's (1982) model of scalar stress, understanding how networks are stacked at different levels to reduce information and energy load by substituting relationships among leaders of hierarchically ordered groups for relationships among members of larger groups at a lower level in the hierarchy is an important element in a theory of network scaling. It points to the possibility of scale-free modeling of the modularity of networks based on the relative constancy of the basic units at the individual level that give structure to these networks, the flexibility of how particular groups are organized, the fact that network hierarchies are population-filling with scale-free relationships to population size, and the bulking, organization and conservation of energy, information and material in ways that match the constraints on populations of individuals. These characteristics of scale-free modeling have been successful in biology, and social scaling may well follow the same principles.

Critique

Johnson's concept of scalar stress on individuals and the management by leaders of relationships among groups in sequential hierarchies is a valuable one, and well grounded in the way he uses it at the individual level. At that level, experimental evidence suggests that productivity tends to increase exponentially as groups grow from 1 to 6, and then decreases following exponential decay with the same apparent coefficient. When he moves to scalar stress in relationships among groups, however, as managed by leaders, he treats scalar stress as increasing at a near-exponential function of the number n of ties, namely the formula for the number of ties in a clique, $n(n-1)/2$. I believe his model would do better at this level if he used exactly the same curves that he uses for productivity.

To follow this logic, Johnson defines capacity, load and performance in the following ways for individuals in groups of size n:

Capacity=15= $(n^2 - n)/2$ *for n=6*
Load= $(n^2 - n)/2$ *for n=group size*
Performance= $L/15$ when load L is below capacity
Performance= $15/L$ when load L is above capacity

Then instead of defining scalar stress as $(n^2 - n)/2$, network stress can be defined for a set of leaders of n groups as suboptimal performance expected from:

Network stress= $(15-L)/15$ when load L is below capacity
Network stress= $15/(15-L)$ when load L is above capacity
Capacity, however, should also be defined just as Johnson did for the bulking number s = 6 ± 2, by defining s = c/a and
a = # activity streams watched (e.g., 1-12 ~ limit)
c = capacity ~ 72.

We need to adjust possibility that an individual's capacity can handle larger networks when the *effective n* of a network, *e*, is expanded through simplification of relationships.

Effective n, e=c/a (normally 6 but much higher if links are simplified)
Effective capacity= $(e^2 - e)/2$ *(e.g.,* $6 \leq e \leq 72$*)*
Network stress= $2L/(e^2 - e)$ when load L is below capacity
Network stress= $(e^2 - e)/(e^2 - e)-2L)$ when load L is above capacity

Given this renormalized definition of network stress a more precise analysis could be done of what Johnson calls basal units, those groups (and their sizes) that are the effective groups in organizing a given activity, and beneath which there are supportive hierarchies.[4]

Extending Johnson's Theory with Contemporary Data

Power laws for critical phenomena in physics such as the boundary conditions for magnetization/demagnetization or the lambda point for superfluid helium are well established, but are observed only asymptotically through infinite control over measurements that approach criticality. Their applicability to Laws of Nature or a great many social and economic statistics is much debated. Many are unlikely to survive closer scrutiny. 'Fat tails' typically contain a great deal of measurement error and fluctuations that are obscured by the binning of data and the logarithmic contraction of tails (Laherrère and Sornette 1998).

---

[4] Question to ISCOM group members: can we show Johnson's own material be better analyzed with the measure of network stress?

Another direction to go is from a network measure of stress (and effective n) that is dependent not on the concept of a clique of complete ties but on k-components where individuals in the group need have only a minimum of k ties with others. Then

Cohesive n, c= ek/n, i.e., reduced by k/n, where k is the measure of cohesion and k < n (Moody and White 2003).
Effective capacity= $(c^2 - c)/2$ *(e.g.,* $6 \leq c \leq 72$*)*
Network stress= $2L/(c^2 - c)$ when load L is below capacity
Network stress= $(c^2 - c)/(c^2 - c)-2L)$ when load L is above capacity

The concept of cohesive n permits a much greater manageable group size given that a person or leader needs only k links within the group, not n.

Power law descriptions, whether in the physics of criticality in phase transitions or social and economic phenomena, necessarily give way to another regime dominated by finite size effects where the probability distribution functions cross over into exponential decay, the curvature at the tails of many log-log plots. For Laherrère and Sornette (1998), this poses the question "of whether these observed deviations from a power law description result simply from a finite-size effect or does it invade the main body of the distribution, thus calling for a more fundamental understanding and also a completely different quantification of the pdf" generating the distribution.

This problem gives rise to two competing hypotheses: finite-size effects versus an alternate description of the whole range of the distribution. Laherrère and Sornette note that the power law Gutenberg-Richter distribution of earthquake sizes, for example, if extrapolated to sizes predicted from a power-law, would predict an infinite mean release of energy, so that appeal is made to finite-size effects. Similarly for oilfield extrapolation for a finite earth. The fact remains, however, that many such distributions show a curved log-log plot that avoids such divergences by thinner tails than predicted by power laws.

A theoretical explanation for power-law pdfs for orders of magnitude that have exponential cutoffs at the tails was recently demonstrated by Frische and Sornette (1997) using a stretched exponential distribution.[5] The exponent c of in this model is the inverse of the number of generations in a multiplicative process that generates the distribution. The usual exponential distribution is the special case where c=1, while the stretched exponential corresponds to values of $c < 1$, for which there is a relatively large regime of linear behavior in a log-log plot, the more so the smaller the c. In addition there is a cutoff where curvature is evident. This provides a single model for process that are multiplicative over different levels that account for parts of a distribution that are scale-free (power law regimes) and cutoffs for tails that that are exponential. In addition to the exponent c, directly interpretable as the inverse of the number of levels in a multiplicative process, this model has a second directly interpretable parameter which represents the characteristic scale of a basal or smallest unit in these processes and from which moments of the distribution can be calculated knowing the value of c.

A stretched exponential distribution that covers a whole range of discrete events rank-ordered by the variable x (e.g., size, degree, etc.) is expressed for a cumulative probability distribution $P_c(x)$, ranked from largest to smallest, as

$$P_c(x) = \exp[-(x/x_o)^c] \text{, and hence} \qquad (6)$$

$$\ln P_c(x) = -(x/x_o)^c = -x^c/x_o^c,$$

[5] It has long been recognized that power law distributions can result from a wide variety of aggregative and interactive phenomena that are not necessarily scale independent, especially those that result from multiplicative processes. The stretched exponential distribution has the advantage of bringing the modeling of these processes directly into the model to account for both the scale-free regime and the cutoff.

Since $P_c(x) \leq 1$, $\ln P_c(x) \leq 0$, so $- \ln P_c(x) = x^c/x_o^c$, where x_o^c is a constant, x_o is the basal unit of scale.

In experimenting with estimating the stretched exponential for network data, I began with the network for connections among the 332 U.S. airports in 1997 (USAir97.net), and calculated the degree of each node (number of connections to other airports) for each node. The degree follows a power law distribution with a slope ~ 1 (Figure 3). The log-log plot showed considerable fluctuation in the tail.

The stretched exponential $P_c(x) = \exp[-(x/x_o)^c] = \exp[-(x/6)^c]$, shown in Figure 4, in which the largest airline hubs start at the upper left since they are first on the cumulative distribution $P_c(x)$, was a fitted by least squares of at r=.99 for c=.55 and a basal unit constant of $x_o = 6$. Since 1/c is indicative of the number of levels in the multiplicative model, in this case 1/c ~ 2, such as the second level hierarchy of hubs listed in appendix 1 is suggested within sets of on-average six local airports as the basic structure of the airline routing system. The number 6 reflects Johnson's scalar stress regrouping of local clusters of airports into sets in which one operates as a hub. This might be the scale or level at which transport by plane or alternate transport is locally synchronized.

Thus, within a single model, instead of a power-law, the fitted stretched exponential for the airline data is

$$P_c(x) = \exp[-(x/x_o)^{.55}] = \exp[-(x/6)^{.55}],$$

has a scale-free regime that by inspection of Figure 4 includes roughly 99% of the airports, with attenuation at the tail (the upper left of the figure) in that the top three airports have somewhat fewer connections than expected, presumably due to congestion. (I have yet to learn how the cutoff point might be predicted).

Univariate Frequencies versus Network Frequencies

Many power laws of social and economic phenomena are analyses of simple univariate distributions: income and wealth (the Pareto distributions), city size (the urban hierarchy), earthquake frequencies at different energies (the Richter scale) and word frequency (the Zipf distribution), each having a power law structure. What is especially useful from the stretched exponential distributions are the depths of the multiplicative cascades such as are estimated by Laherrère and Sornette (1998):

	Inverse of C	x_o
Light waves from galaxies	20-50	E-34 (intensities)
Radio waves from galaxies	10-20	E-8 (intensities)
Earthquake sizes	10	E-9 (??)
Urban aggregates	6	7-20 (thousands?????
Oilfield formations	3-6	3±1 (millions of barrels)

Top physicist citations	3	~3 (citations)
Country sizes	2.5	7 (millions)
Temperature, S. Pole	2	13 (normalized temperature)
Species extinction times	~1	22-25 MY lifespan (time scale)

What is the mathematical meaning of x_o, as for example $x_o = 2.7 \sim 3$ for the citation network of physics articles and authors? Combined with c=0.3, and the Gamma function $\Gamma(z)$ as of z (extending the factorial z! to real number arguments), the moments (mean, standard deviation, skewness, kurtosis, etc.) can be directly calculated from x_o (Laherrère and Sornette 1998, appendix):
$\langle x \rangle = x_o\ (1/c)\ \Gamma(1/c)$.
$\langle x^2 \rangle = x_o\ (2/c)\ \Gamma(2/c)$.
$\langle x^k \rangle = x_o\ (k/c)\ \Gamma(k/c)$.

So for "top" physics citations,
$\langle x \rangle \sim 2.7 \cdot 3 \cdot 3! \sim 49$.
$\langle x^2 \rangle \sim 2.7^2 \cdot 6 \cdot 6! \sim 5832$.
$\langle x^k \rangle \sim 2.7^k \cdot k3 \cdot (3k!)$.

Degree distributions in networks are more complex than univariate distributions in which magnitudes are ranked because degree determines the rank, and the mean and other moments of the degree distribution reflect network density and topology. They also have a natural dimensionless units in the node and clusters of nodes. If we think of there being three levels in the production of physics articles, we have for example (1) characteristics of the node, such as the competence, skill, expertise, training, knowledge of the individual, (2) the complementarities of the groups within which this individual is embedded, and (3) resonance with the reception by similar groups who evaluate the outcome, which reflects back on the recognition by the individual and working group to recognize a good problem. We might think of $x_o = 2.7 \sim 3$ as telling us that the minimal coauthorship set that has these multilevel resonances is not the individual and not the dyad, but the minimal human group that Simmel identified as the group of three. The fitted model here has $x_o < 6$, so no scalar stress is embodied in this unit. For the airlines problem the values of $x_o \sim 6$ and $c \sim ½$ might be telling us that the airline routing problem is a simpler one in which a two-level solution has been reached with geographic constraints allowing the full efficiencies of six airports in a local group in which there is a complete graph for local transportation linkages have been synchronized.

Ferrer-i-Cancho and Solé (2001a) noted that "Zipf's law states that the frequency of a word is a power function of its rank. The exponent of the power is usually accepted to be close to (-)1. Great deviations between the predicted and real number of different words of a text, disagreements between the predicted and real exponent of the probability density function and statistics on a big corpus, make evident that word frequency as a function of the rank follows two different exponents, ~ (-)1 for the first regime and ~ (-)2 for the second. The implications of the change in exponents for the metrics of texts and for the origins of complex lexicons are analyzed." They decided (Ferrer-i-Cancho and Solé 2001b) to study this bifurcation as a network problem and constructed a network of word co-occurrence at the sentence level in the 70 million word British National Corpus. Their graph of the degree distribution for their word web, reproduced as Figure X, shows two power law regimes, one with a coefficient of 1.5 and the other (the high frequency words), with one of 2.7. They argue that the upper regime is the kernel word lexicon that is available to all speakers (perhaps ~5,000 words) while the lower frequency regime consists of lexicons that are available only to some.

Dorogovtsev and Mendes (2002, 2003) follow up on the study of the word web with a language evolution model with t words at a given time, m ~ 1 preferential attachments for a new word as it enters the lexicon, and for a constant c, c·m·t new attachments that emerge at each time step that are doubly preferential in proportion to the product of degrees of pairs of nodes,. Hence new edges are produced by a combination of linear growth in new nodes and their edges but is increasingly quadratic in the upper region that is doubly preferential. From this model of the doubly preferential network component and parameters c, m, and t they predict the cutoff point for the shift in the distribution. Network phenomena, then, might also have broken regimes from the dynamics of network growth or relinking.

Conclusions

We started with Johnson's findings of how groups are renormalized around leaders at successively higher levels, and his model of power-law attenuation in fractional recognition of potential leaders at higher scales. Groups of six individuals or leaders were the units of renormalization that performed most efficiently, and above which scale stress was found to grow exponentially without renormalization.

[shorten: paragraph] A prediction of Johnson's model, if the number of leadership positions ***d*** derives from the total number of number of groups in a population of size ***p*** divided into ***h*** hierarchical levels with a constant number ***s*** of fractal splits at each level, is that an attenuated rate of recognition of political leaders in a population is a function of three parameters, increasing in ***p,*** and attenuating for $\boldsymbol{p}^{-\beta}$ by the power $-\beta$, with an intercept at $\boldsymbol{s}^{1-\beta}$. I verify that a scale-free law $\boldsymbol{d/p} \approx \boldsymbol{s}^{1-\beta}\boldsymbol{p}^{-\beta}$ is not inconsistent with $\boldsymbol{s} \sim 6 \pm 2$, which is the limit of human attention to complex simultaneous signals (Simon 1973) and sets the framework in G. A. Johnson's (1982) model for optimal and empirically observed hierarchies. The mathematically derived model fits Johnson's empirical distribution for a log-log correlation of r=.96 between the political and population variables, and theory and data agree as to slope and intercept of the linear scatterplot. The ***d/p*** rate intercept of $\boldsymbol{s}^{1-\beta}$ for a single individual converges to ~ 1. An attenuation slope to the 6/7th power-slope is not unlike the type of power-law attenuation observed in biological scaling (West et al., 1997) and the network-based construction of this law follows the same structure-filling and energy-optimization principles.

While power laws of social and economic phenomena are suggestive of fractal, hierarchical, multilevel and multiplicative processes, they do not actually model them at finite scales, and suffer the defect of failing to predict the limits of scale-free invariance and thus overpredicting in the direction of infinite variance.

The stretched exponential captures power law invariance over many orders of magnitude (a large-regime) as an explicit model of multilevel and multiplicative processes, and has the benefit of predicting as well the basal unit size in fractal hierarchies. A first experiment with data from the airline routing system in the U.S. converged to Johnson's result: 6 ± 2 as the unit of renormalization, in this case 6 ± 2 local airports around long distance hubs, and a hierarchy of depth two. Basal unit sizes may prove to be fairly easy to interpret for network data. Comparison with a physics citation network is instructive. Here, the basis unit is of size three, well below the Johnson limit, but estimated to have a small cascade of three levels that interact (individual, groups of collaborators, groups of receptors) to produce multiplicative probabilities of outcomes.

This short set of experiments in social scaling, then, has produced positive results and opens a new methodological path that may well lead to significant and integrated results if carried out in a larger research program.

References

Dorogovtsev, Serguei N., and José F.F. Mendes. 2002. Evolution of networks. *Advances in Physics* 51 : 1079-1187. http://arxiv.org/abs/cond-mat/0106144

——— 2003. *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford : Oxford University Press.

Ferrer-i-Cancho, R. and Solé, R. V. 2001a. Two regimes in the frequency of words and the origins of complex lexicons: Zipf's law revisited. *Journal of Quantitative Linguistics*, 8(3):165-173. http://www.isrl.uiuc.edu/~amag/langev/paper/ferrer01twoRegimes.html

Ferrer-i-Cancho, R. and Solé, R. V. 2001b. The small world of human language. *Proceedings of The Royal Society of London. Series B, Biological Science*s, 268(1482):2261-2265. http://www.isrl.uiuc.edu/~amag/langev/paper/ferrer01theSmall.html

Frisch, U., D. Sornette 1997. Extreme deviations and applications J. *Phys. I France* 7, 1155-1171. http://arxiv.org/abs/cond-mat/9705132

Johnson. Gregory A. 1982. Organizational Structure and Scalar Stress, pp 389-422 in Renfrew, Colin, Michael Rowlands and Barbara A. Segraves-Whallon, *Theory and Explanation in Archaeology: The Southampton Conference*. New York: Academic Press. http://eclectic.ss.uci.edu/~drwhite/ISCOM/GAJ1-17.pdf (live link)

Laherrère, Jean, and Didier Sornette. 1998. Stretched exponential distributions in Nature and Economy: ``Fat tails'' with characteristic scales. *Eur.Phys.J.* B2: 525-539. http://arxiv.org/abs/cond-mat/9801293

Moody, James and Douglas R. White. 2003 Structural Cohesion and Embeddedness: A Hierarchical Conception of Social Groups. *American Sociological Review* 68(1):1-25. http://eclectic.ss.uci.edu/~drwhite/soc_con17.pdf

Simon, H. A., 1973, The organization of complex systems, in Pattee, H. H. (ed.), *Hierarchy Theory: The Challenge of Complex Systems,* New York, George Braziller, 3-27.

West, G.B., J.H. Brown, and B.J. Enquist. 1997. A general model for the origin of allometric scaling laws in biology. *Science* 276:122-126. [See also: "Allometric scaling laws in biology," Letters Section, *Science* 278:371-373] http://eeb37.biosci.arizona.edu/~brian/West_Brown_Enquist_1997.pdf

Check:
Chapple – book
Ecological Psychology – book
Arensberg and Kimball -

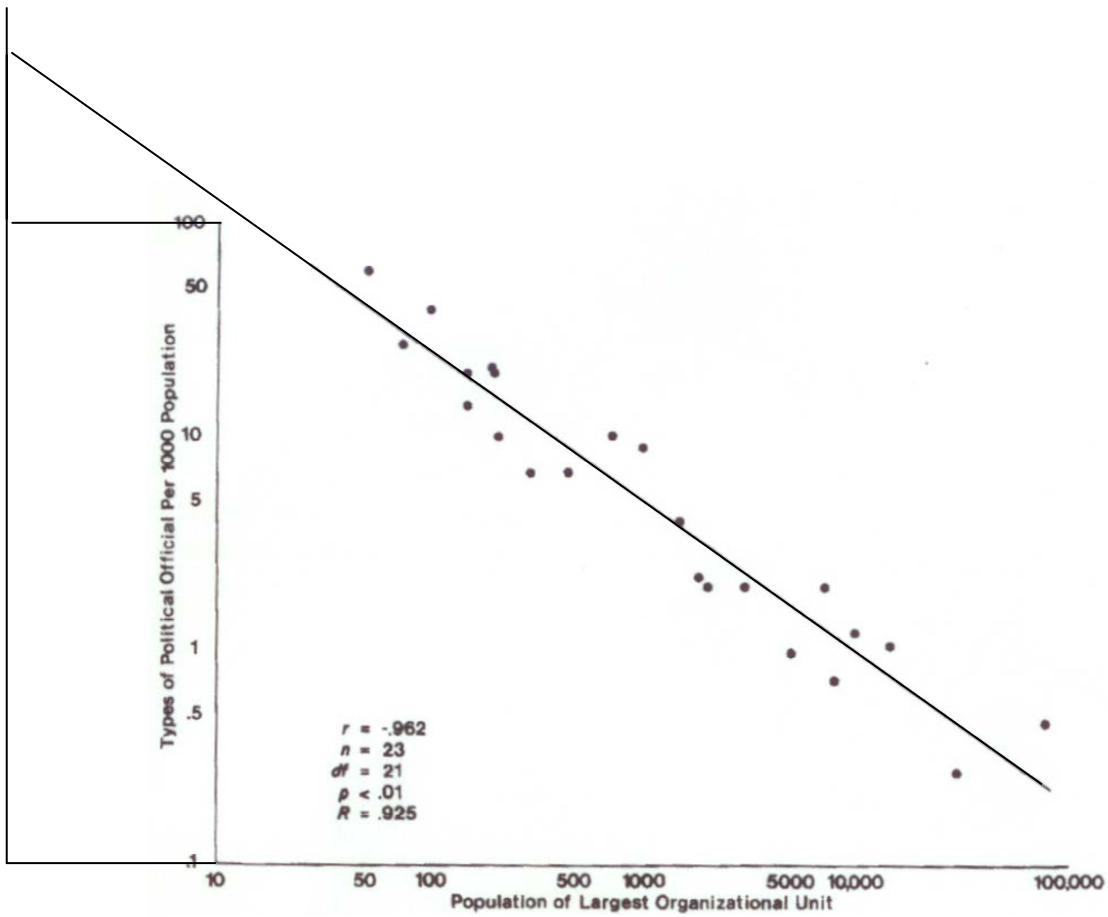


Figure 1. Societal size and relative size of controlling component.
(reprinted with permission from Johnson, 1982, Figure 21.5)

Figure 2. Societal size and relative size of controlling component, derived model, power-law beta = 6/7, goodness-of-fit is relatively indifferent to the group bulking number s

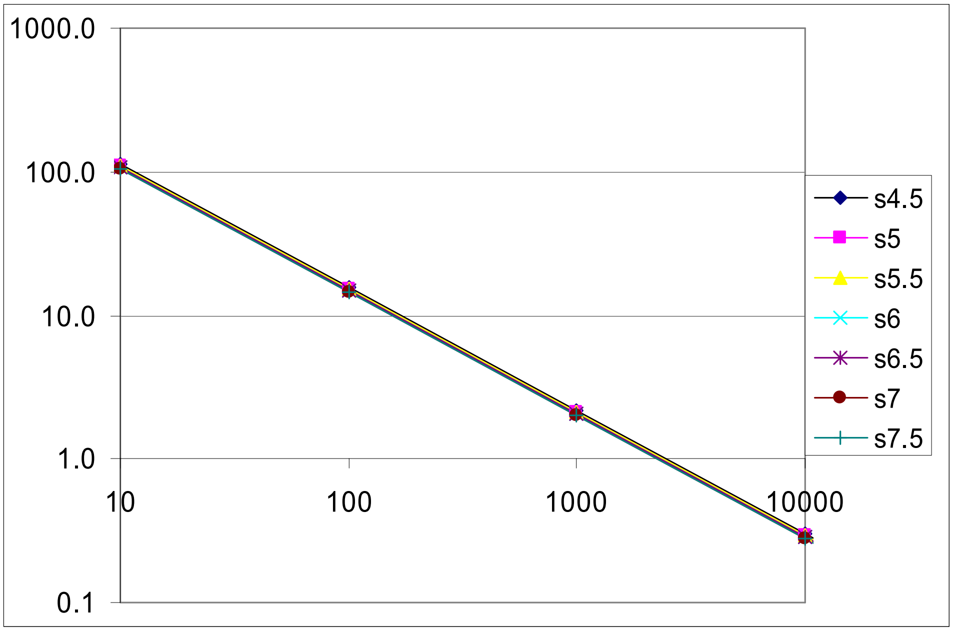


Figure 3: Power law distribution for 1997 U. S. airport connections

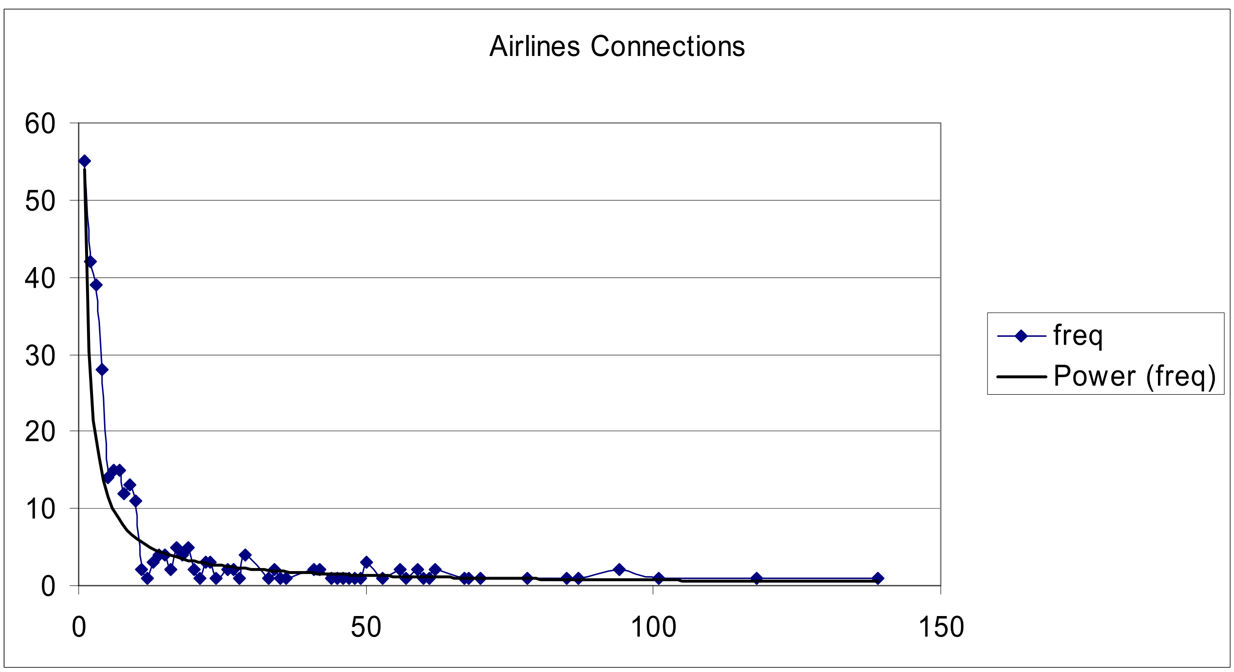


Figure 4: Stretched exponential distribution for 1997 U. S. airport connections

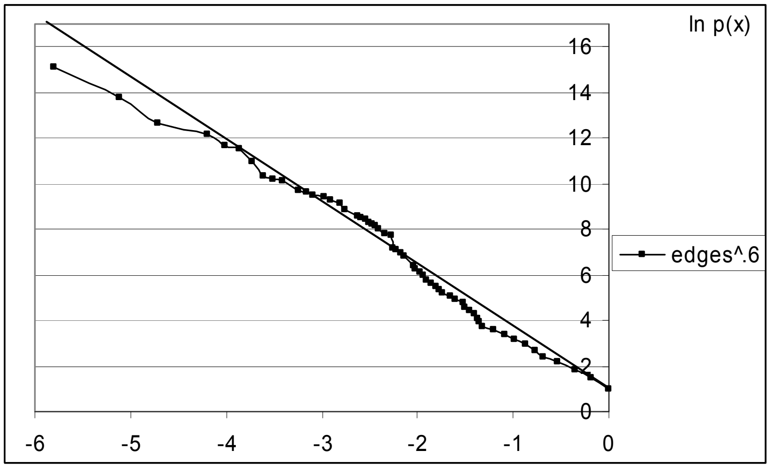
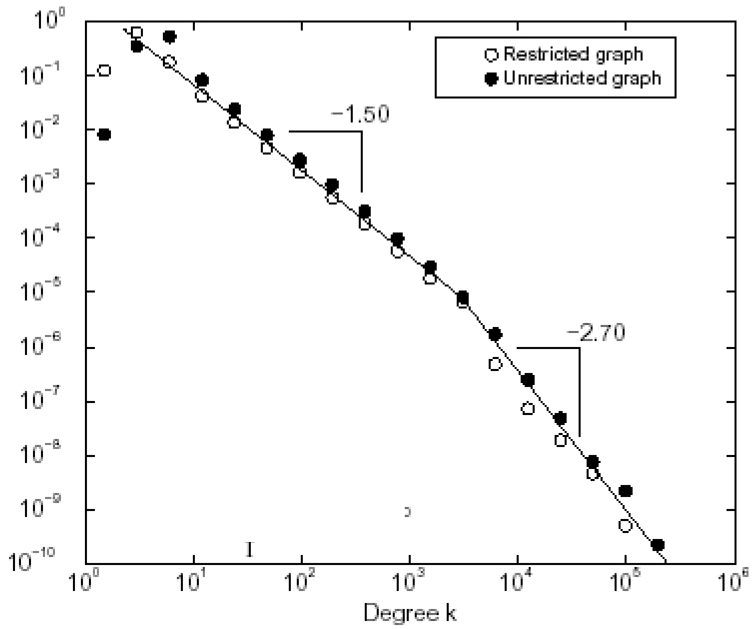


Figure 5: The broken regime for Word Co-occurrence
(from Ferrer-i-Cancho, R. and Solé, R. V. 2001b).



Appendix 1: Top 1/6th of Airports as Hubs

22	3	0.9036	280	84.3373	Greater Buffalo Intl
23	3	0.9036	283	85.2410	Sacramento Metropolitan
24	1	0.3012	284	85.5422	Honolulu Intl
26	2	0.6024	286	86.1446	Chicago Midway
27	2	0.6024	288	86.7470	William P Hobby
28	1	0.3012	289	87.0482	Port Columbus Intl
29	4	1.2048	293	88.2530	Anchorage Intl
33	1	0.3012	294	88.5542	New Orleans Intl/Moisant Fld/
34	2	0.6024	296	89.1566	General Mitchell Intll
35	1	0.3012	297	89.4578	Kansas City Intl
36	1	0.3012	298	89.7590	San Diego Intl-Lindbergh Fld
41	2	0.6024	300	90.3614	Portland Intl
42	2	0.6024	302	90.9639	Washington National
44	1	0.3012	303	91.2651	Memphis Intl
45	1	0.3012	304	91.5663	Cleveland-Hopkins Intl
46	1	0.3012	305	91.8675	John F Kennedy Intl
47	1	0.3012	306	92.1687	Miami Intl
48	1	0.3012	307	92.4699	Washington Dulles Intl
49	1	0.3012	308	92.7711	Mc Carran Intl
50	3	0.9036	311	93.6747	General Edward Lawrence Logan
53	1	0.3012	312	93.9759	Baltimore-Washington Intl
56	2	0.6024	314	94.5783	Nashville Intll
57	1	0.3012	315	94.8795	Seattle-Tacoma Intl
59	2	0.6024	317	95.4819	Salt Lake City Intl
60	1	0.3012	318	95.7831	Phoenix Sky Harbor Intl
61	1	0.3012	319	96.0843	Cincinnati/Northern Kentucky I
62	2	0.6024	321	96.6867	Philadelphia Intl
67	1	0.3012	322	96.9880	Newark Intl
68	1	0.3012	323	97.2892	San Francisco Intl
70	1	0.3012	324	97.5904	Detroit Metropolitan Wayne Cou
78	1	0.3012	325	97.8916	Minneapolis-St Paul Intl/Wold-
85	1	0.3012	326	98.1928	Stapleton Intl
87	1	0.3012	327	98.4940	Charlotte/Douglas Intl
94	2	0.6024	329	99.0964	Pittsburgh Intll
101	1	0.3012	330	99.3976	The William B Hartsfield Atlan
118	1	0.3012	331	99.6988	Dallas/Fort Worth Intl
139	1	0.3012	332	100.0000	Chicago O'hare Intl

22	3	0.9036	280	84.3373	Greater Buffalo Intl
23	3	0.9036	283	85.2410	Sacramento Metropolitan
24	1	0.3012	284	85.5422	Honolulu Intl
26	2	0.6024	286	86.1446	Chicago Midway
27	2	0.6024	288	86.7470	William P Hobby
28	1	0.3012	289	87.0482	Port Columbus Intl
29	4	1.2048	293	88.2530	Anchorage Intl
33	1	0.3012	294	88.5542	New Orleans Intl/Moisant Fld/
34	2	0.6024	296	89.1566	General Mitchell Intll
35	1	0.3012	297	89.4578	Kansas City Intl
36	1	0.3012	298	89.7590	San Diego Intl-Lindbergh Fld

41	2	0.6024	300	90.3614	Portland Intl
42	2	0.6024	302	90.9639	Washington National
44	1	0.3012	303	91.2651	Memphis Intl
45	1	0.3012	304	91.5663	Cleveland-Hopkins Intl
46	1	0.3012	305	91.8675	John F Kennedy Intl
47	1	0.3012	306	92.1687	Miami Intl
48	1	0.3012	307	92.4699	Washington Dulles Intl
49	1	0.3012	308	92.7711	Mc Carran Intl
50	3	0.9036	311	93.6747	General Edward Lawrence Logan
53	1	0.3012	312	93.9759	Baltimore-Washington Intl
56	2	0.6024	314	94.5783	Nashville Intll
57	1	0.3012	315	94.8795	Seattle-Tacoma Intl
59	2	0.6024	317	95.4819	Salt Lake City Intl
60	1	0.3012	318	95.7831	Phoenix Sky Harbor Intl
61	1	0.3012	319	96.0843	Cincinnati/Northern Kentucky I
62	2	0.6024	321	96.6867	Philadelphia Intl
67	1	0.3012	322	96.9880	Newark Intl
68	1	0.3012	323	97.2892	San Francisco Intl
70	1	0.3012	324	97.5904	Detroit Metropolitan Wayne Cou
78	1	0.3012	325	97.8916	Minneapolis-St Paul Intl/Wold-
85	1	0.3012	326	98.1928	Stapleton Intl
87	1	0.3012	327	98.4940	Charlotte/Douglas Intl
94	2	0.6024	329	99.0964	Pittsburgh Intll
101	1	0.3012	330	99.3976	The William B Hartsfield Atlan
118	1	0.3012	331	99.6988	Dallas/Fort Worth Intl
139	1	0.3012	332	100.0000	Chicago O'hare Intl

Appendix 2: The gamma function

The gamma function can be defined in Euler's integral form (a definite integral, i.e., with upper and lower limits) for positive real numbers z where:

$$\Gamma(z) \equiv \int_0^\infty t^{z-1} e^{-t}\, dt \tag{1}$$

$$= 2\int_0^\infty e^{-t^2} t^{2z-1}\, dt, \tag{2}$$

or

$$\Gamma(z) \equiv \int_0^1 \left[\ln\left(\frac{1}{t}\right)\right]^{z-1} dt. \tag{3}$$