Modern Applied Science
www.ccsenet.org/journal.html
Vol. 1, No. 4
November 2007

The Value of a Network

Nick Wills-Johnson (corresponding author)
Centre for Research in Applied Economics
Curtin University of Technology
GPO Box U1987
Perth, Western Australia 6845
Tel: 61-8-9266 3759 E-mail: n.wills-johnson@curtin.edu.au
Peter Hornby
CSIRO Exploration & Mining
26 Dick Perry Avenue
Kensington, Western Australia 6151
Tel: 61-86436 8640 E-mail: Peter.Hornby@csiro.au

Abstract

This paper presents a simple mechanism for quantifying the value which a network creates based upon the way in which it is connected. The mechanism is grounded in a process of information sharing by nodes and is, in a sense, an extension of Bonacich's (1972) centrality measure.

Keywords: Networks, Value, Super-additivity

1. Introduction

When parties join together in a network, they are able to create a whole greater than the sum of their parts, a superadditivity of synergies. This paper examines a means of calculating superadditivity based upon the connection pattern of a network and the quasi-static results of an information transfer game. Section Two outlines the literature on network attributes in social networks and game theory. Section Three outlines the measure of value, some of its characteristics and some extensions to the model. Section Four concludes.

2. Network Structural Analysis

Wasserman and Faust (1994) suggest five levels of network analysis:

The individual level.

The dyadic or tiered level.

The triadic level.

The subgroup level.

The global level.

The individual level examines the relationships between nodes relative to one another. A common individual-level analytic methods is that of centrality. Centrality measures the relative importance of a node in a network, compared to all other nodes in that network. Freeman (1979) discusses a number of centrality measures including:

Degree Centrality: The number of edges incident upon a node, with more edges indicating greater importance, and the node being more 'embedded' in a network.

Closeness Centrality: The minimum geodesic distance from each node to each other node in the network, which is useful when the concern is the speed of transfer across a network. The centre of a star is the central node by this measure.

Betweeness Centrality: The extent to which a node lies between other nodes, and is hence a conduit for information, or to which a node can control the flow of information.

Importance might also be judged not only on how many nodes a node is joined to, but how important those nodes are. Bonacich (1972) calculates this by considering a centrality measure based upon the largest eigenvalue of the adjacency matrix. This measure is related to our own, as is discussed below.

Stephenson & Zelen (1989) suggest that information between nodes might follow many paths at the same time. They propose a measure of centrality based upon information theory in statistics, by maximising the 'information' contained in a weighted sum of all paths between each node (where the weighting depends on the variance in the path distances between nodes).

Burt (1992) suggests that what is important in networks is the strategic location of 'holes' within a network. Linkages are costly to form, and hence nodes should only link to other nodes when they provide access to new information and, moreover, when the links allow them to become arbitrageurs between two densely connected local networks. Burt suggests that, in strongly connected networks, information flows rapidly between all members, soon becoming common knowledge, and hence providing little opportunity for strategic behaviour. This contrasts with the view of Coleman (1988), who suggests that highly interconnected networks are beneficial, because they provide strong linkages and stability to their members. Benassi & Gargiulo (2000), in a study of the Italian computer industry, suggest that each type of network has a role in certain circumstances, with strong linkages important where stability is valuable, and sparse networks important where flexibility is necessary.

Dyadic-level analysis focuses on ties between two nodes and the paths between them. Milgrom's (1967) small worlds hypothesis is one example of this. Monge & Contractor (2003) suggest two key concepts in relation to dyadic analysis; reachability and redundancy. Reachability refers to the shortest path between two points, whilst redundancy refers to the number of paths between two points. One extension of our model (see below) falls loosely within the ambit of reachability, considering the way in which information generated in one part of the network reaches the remainder of the network.

Triadic-level analysis examines nodes three at a time. Where ties are directional, Monge and Contractor (2003) note the importance of transitivity; A linked to B, B linked to C and A linked to C implies transitivity. Where links are non-directional, transitivity is the same as cyclicality. Cycles in graphs and their effects on the value of nodes in a network are examined in a game theoretic framework by Gomez, Gonzalez-Aranguena, Manuel, Owen, del Pozo & Tejada (2003) and their concept of a connected hull. Triads are also the first point of stability in a growing network in a dynamic extension of our measure (see below).

A more general form than triadic-level analysis is analysis of sub-groups, where collections of nodes strongly connected to each other but weakly connected to the rest of the network for the focus of analysis. Granovetter's (1973) analysis of strong and weak ties social networks and their importance in finding a job are an early example of sub-group analysis. Clique or sub-group structures are widely observed in real-world networks and this has lead to a large literature on dynamic evolutionary algorithms which replicate such networks. Goh, Eom, Kanhg & Kim (2006), and Moriera, Paula, Costa-Filho & Andrade (2006) are two recent examples of this approach, from the physics literature where much of it has developed. Cliques and sub-groups are also an area of significant interest in game theory. Conley & Smith (2005) survey the literature in the public goods field, where it has been widely applied to the study of Tiebout equilibria (Note 1). Bloch (2005) examines application in industrial economics, where it has been used to study cartels, as well as more benign forms of industrial networks, such as research and development collaborations. Adler & Smilowitz (2006) examine an application in airline mergers.

The final level of analysis is the global level, which considers the network as a whole. Our measure operates at this level. Monge & Contractor (2003) suggest two global measures; network density and network centralisation. The first measures the number of ties in a network compared to the total possible number of ties, whilst the second measures the dispersion in centrality scores. In the game theory literature, the global value of a network is considered as a product of the analysis of the dynamic growth of the network. Jackson (2005) presents a recent review of this literature. Jackson & Wolinsky (1996) suggest that network growth is a function of each pair of nodes in the network examining the benefits and costs of forming a link and forming the link if each experiences and increase in utility from doing so. By summing the extra utility created by each link formed in a round of a network growth game, one might ascertain the overall value of the network to its participants.

The notion that the value of a network is related to the value it creates for each individual node seems intuitive. Indeed, there are parallels between Jackson & Wolinski's (1996) model and our own, at the conceptual level. However, it is also very intensive in information; in order to calculate network value, one needs to know the utility functions of all nodes and the gains and losses which would be made by each party when a link is formed. This paper suggests a simpler approach, based on the pattern of connectivity and the quasi-static results of an information transfer game.

3. The Measure

We motivate our measure by considering, like Stephenson & Zelen (1989), that networks represent patterns of information flow between the individuals. Each node joins the network with a stock of private information, which may be complementary to the private information stocks of others in the network creating synergies through information-sharing. This process of creating synergies allows those in the network to create a value which is greater than the sum of the values each is able to create independently, using their private stocks of information. Knowledge of where such synergies lie underpins decisions about whom each node will choose to connect to. If one considers information as underpinning utility, then our approach is roughly analogous to that of Jackson & Wolinsky (1996). We suggest that nodes aiming to join or expand a

network examine their private information stocks and what they know about the private information stocks of others and make decisions about with whom to connect directly. Once these decisions are made, the network is formed, its members communicate and it produces its output. We suggest that the network is able to produce this output precisely because of the way in which it was connected, and the synergies thus created. The synergies represent the amount of network superadditivity which is created through information sharing.

The measure is perhaps best understood by considering it as the end product of a dynamic information transfer process between the nodes of a network which transfer private information stocks between each other through a series of rounds of interaction. The dynamic representation of the 'information transfer game' has, at time *k*, the vector of information x^k, which can be described thus:

$$x^k = x^0 + Ax^{k-1} \tag{1}$$

Where x^0 describes the initial information vector, and *A*, the adjacency matrix, describes the way information is updated each 'round' of the process. Over a period of infinite updating the final (or quasi-static) information state is described thus:

$$x^\infty \cong (I - A)^{-1} x^0 \tag{2}$$

If one considers the information stocks held by each node of the network through the process, the eigenvectors can be taken to represent some measure of the amount of each type of information held by each node in the new quasi-static equilibrium after the private information stocks held at the outset have been augmented by the information sharing process. The eigenvectors are described thus:

$$e_i = (1 - \alpha_i)^{-1} e_i \tag{3}$$

where α_i represents the eigenvalues of the adjacency matrix *A* and e_i appears on both the left-hand side and right-side of the equation because the eigenvectors are the same for *A* and $(I\text{-}A)^{-1}$. If λ_i represents the eigenvalues of the matrix $(I\text{-}A)^{-1}$ then this may be expressed:

$$e_i = \lambda_i e_i \tag{4}$$

Each eigenvalue λ_i is summarising how far the eigenvector e_i has been 'stretched' along the ith dimension of the underlying information space through the augmentation process of information-sharing. If all eigenvalue-eigenvector pairs are considered, we have a measure of the amount of 'stretch' in each dimension in the information-space which has resulted from the communication activities along the paths laid out by the network pattern. Each of the $\lambda_i e_i$'s are orthogonal. If they are multiplied together, they will provide an indication of the volume of the underlying space. If the eigenvectors are normalised such that their length is unity, the eigenvalues provide an indication of the length of the underlying information space in the given dimension indexed by that particular eigenvalue. With e_i normalised to one, the result becomes the product of the λ_is, or the determinant of the matrix $(I\text{-}A)^{-1}$.

This is true if and only if the eigenvalues of the adjacency matrix are less than one, otherwise the dynamic state oscillates, rather than converging. In an adjacency matrix of zeroes and ones, the largest eigenvalue will often be larger than one. The matrix must thus be normalised to prevent oscillation. We do so by choosing a scalar factor ε with $0 < \varepsilon < 1$, which, when multiplied by the adjacency matrix A, produces a new matrix A' which does not have any eigenvalues greater than one and which hence produces dynamic convergence. The use of ε implies that not all information held by a node is passed on in one round of communication. For ε we choose the reciprocal of the Frobenius Norm of the adjacency matrix. We could have chosen any matrix norm, but the Frobenius Norm in a zero-one matrix is simply the square root of the sum of the non-zero entries, making it easy to calculate.

One might also consider our measure as an extension of Bonacich's (1972) measure of centrality. Bonacich's measure can be developed as follows: if the adjacency matrix is multiplied by itself, the result, rather than describing which nodes are connected to which, describes how many two step connections exist between the ij^{th} node. If the adjacency matrix is multiplied by itself again, it describes how many three step connections exist and so on. The series of such connections is $I+A^2+A^3+\ldots\ldots+A^n$ which can be summed as follows:

$$\sum_{i=0} A^i = (I - A)^{-1} \tag{5}$$

This relationship is true where the eigenvalues of the adjacency matrix are all less than one. If they are not, oscillation, rather than convergence occurs, and the adjacency matrix must be normalised.

Bonacich suggests that his measure, which endeavours to obtain the closest measure of the actual 'bond' between two nodes by minimising the sum of the squared differences between the estimated and actual (but unknown) matrices representing bonds, is akin to principal components analysis in statistics, in that the use of a (suitably standardised) eigenvector as the centrality measure preserves as much as is possible, the maximum variance in the original. Mathematically (following Bonacich, 1987), if *R* is the matrix of relationships, and *e* the measure of 'centrality', the centrality of unit *i* is given by:

where λ is constant, chosen so the equations have a non-zero solution. This can be expressed as:

$$\lambda e = Re \tag{7}$$

where λ represents the eigenvalue and *e* its associated eigenvector. Bonacich considers the largest eigenvalue/eigenvector pair as these preserve the most variation from the original data in *R* but (Note 2), as he notes,

any eigenvector/eigenvalue pair can be used as a centrality measure, although other pairs will not preserve as much variation as the largest pair. However, as noted by Hadi & Ling (1998), despite containing the largest amount of variation present in the original data, the first principal component (or indeed, the first *p* of them) may not in fact be 'closest' to the underlying data. This means it may be appropriate to use difference eigenvalues, or even combinations of eigenvalue/eigenvector pairs. Our measure effectively uses all principal components. This would not be useful from a statistical perspective, but the focus here is on the value of a network not statistical estimation.

3.1Characteristics of the Measure

For an empty network of size *n*, the score is always one, which accords with the intuitive notion that a network with no connections would not be able to produce any more than is possible for its members acting alone, as there is no scope for synergy. As nodes are added to the network, the scores rise, although not in a manner strictly monotonic with link addition. A commonly studied network shape is the star, which is often used as a test for the validity of a measure of network centrality. For our measure, however, stars score relatively lowly, and the score increases as nodes are added to the star. This is shown in Table One.

Note in Table One that all stars score two. In the general case of an *n* node network with an *(n-s)* node star, all stars score two. This is a characteristic of stars and is not a general result for network shapes in expanding networks. For example, the score of an *(n-s)* node box or circle (where each node is connected only to its immediate neighbours) in an *n* node network decreases in size as the number of nodes in the box increases. Some shapes appear less scaleable than others as networks grow.

Table One here

The results of Table One seem intuitive; when considering the network as a whole, it may not be particularly efficient to have all nodes transferring information through a central node. Few real-world networks are stars (see Barabasi, 2002) and, as Jackson & Wolinski (1996) suggest, if the apportionment of network value is somehow related to one's position in a network, then peripheral nodes have a strong incentive to link directly with each other rather than communicate through the central node. Our measure reflects this. However, Table One raises an issue; in every case the highest score is attained by the fully-connected network. This raises the question of whether a fully-connected network will necessarily be what one might expect to occur in each case. As Barabassi (2002) notes, empirically, most networks exhibit a power-curve relationship, whereby some nodes are densely connected and many are sparsely connected. Something prevents a network from evolving to the fully-connected state and that something, missing from Table One, is the cost of link formation. For our measure to be useful in the analysis of network growth, it needs to be incorporated into a model which includes notions of the cost of link formation.

The main aim of this paper is to introduce and discuss our measure and not to discuss network evolution (Note 3). However, it is worthwhile considering the utility of the measure in a simple model of network evolution. A very simple notion of cost is that of the veto, whereby network participants can allow or disallow a link to form depending upon the impact of that link upon themselves. This establishes a tension between the desires of the group to maximise what it can produce and the desires of individuals within the group to maximise their share of production. The tension seems an intuitively useful notion; if a group relied upon some relative measure of prestige to apportion gains (like the centrality of each of its members) and the formation of a link resulted in a net gain for the group, but one which diverted gains away from an actor who currently had high status in the group, then it does not seem unreasonable to suspect that that actor would endeavour to prevent the new link from forming. One could model this formally by considering veto power to be a function of the centrality of each node in the network, and considering an exercise of veto power between two extremes:

Strong veto: a given node will allow no linkages to be made anywhere in the network which reduces its centrality.

Weak veto: a given node will allow no connections to itself which would reduce its centrality.

Network formation thus proceeds along a path whereby connections form which increase the value of the network as a whole but which are not subject to veto due to the effects on the centrality of members of the network. Using Bonacich centrality as the centrality score underpinning the exercise of veto, it becomes apparent that the notion of strong veto may be too strong; experimentation with some small networks suggests that there are very few cases where the addition of a link does not violate strong veto rules. On the other hand, it would appear that weak veto rules may be too weak; experimentation with small networks suggests that the weak veto almost always allows a given network to move to a fully connected state. The exception to this is when a sub-network becomes full connected, at which point the weak veto stops all further

network growth.

In reality veto within a given network is likely to be more complex; some nodes might have a strong veto, some a weak veto and some none at all. Alternatively, the amount of veto power a node has might be dependent upon its place in the network, with more central nodes being able to exercise greater veto than less central nodes. There could be an overarching hierarchy of decision-making, such as that suggested by Palamara & Shapley (2000) which determines veto power. The precise nature of the veto wielded by different nodes in a given network may depend upon context, but as a basic notion, the idea of a tension between group and individual desires seems useful.

3.2Extensions

There are a number of relatively simple extensions which could be made to the measure to take into account a greater richness of knowledge about a given network. For example, if something were known about the relative resistances to information flow between each of the pairs of nodes (for example, if two nodes are joined, but joined by email, they might be less able to communicate than if they were located physically adjacent to one another), then it may be possible to replace the ones and zeroes with numbers between one and zero. Similarly, one could account for asymmetric flows of information between two nodes.

A more fundamental extension, however, is to change the learning model. The above description of information flow, is essentially a 'no-learning' model (Note 4); each node takes into the model the information it possesses and, whilst it transfers information from other nodes as part of the process of being in the network, it never adds to its own stock of knowledge during the course of the game. This is realistic in a quasi-static world, but less so in one where one wishes to trace the path of knowledge acquisition, rather than simply plot the end point of resultant value. In particular, whilst it shows the equilibrium at which a network will eventually arrive, it does not show whether the journey there involves oscillation, nor how long it might take. Both of these are possible, and would add richness to the analysis. In general, the process is as follows: rather than the adaptations process above of:

$$x^k = x^0 + Ax^{k-1} \tag{8}$$

consider the other extreme of the 'perfect learning' case:

$$x^k = x^{k-1} + Ax^{k-1} \tag{9}$$

which can also be written:

$$x^k = (I + A)x^{k-1} \tag{10}$$

and gives rise to a $\lambda_i=(1+\alpha_i)^{-1}$. If the adjacency matrix is composed of zeroes and ε, where $0< \varepsilon<1$, rather than zeroes and ones, the perfect learning and no learning cases converge as ε approaches zero. ε could represent the proportion of one's knowledge that is passed on in a single iteration. Another alternative is to express:

$$x^k = f(x^0) + Ax^{k-1} \tag{11}$$

which captures the fact that information is passed on imperfectly in a more realistic way than the use of the Froebenius Norm of the adjacency matrix.

In this dynamic setting, the determinant is not longer a particularly useful measure, because it is too crude to capture the pattern of adaptation which occurs as iterations progress. However, examination of the individual eigenvalues and their associated eigenvectors provides a wealth of information about evolving networks.

4. Conclusions

The measure developed in this paper endeavours to quantify, in a relatively simple fashion, the value of a network, connecting it to the pattern of the network and basing the measure in a theoretical framework of information transfer.

The simplicity of the measure suggests it may be attractive to network analysts, particularly in comparing networks and examining the evolution of networks over time, as part of a broader model. Further work establishing the empirical validity of the measure by comparing the score of a series of networks with their actual output, would be useful if the analysis is to be cardinal rather than ordinal, but the brief analysis of the characteristics of the measure discussed here suggests it represents a potentially useful addition to the suite of analysis tools used in the examination of network attributes.

References

Adler, N. & Smilowitz, K. (2006). Hub and spoke network alliances and mergers: Price-location competition in the airline industry. *Transport Research Part B*, 41(4), 394-409.

Barabasi, A.L. (2002). *Linked: The new science of networks.* Cambridge Massachusetts: Perseus Publishing.

Benassi, M. & Gargiulo, M. (2000). Trapped in your own net? Network cohesion, structural capital and the adaptation of social capital. *Organization Science*, 1(2), 183-96.

Bloch, F, 2005, "Group and Network Formation in Industrial Organization: A survey" in G. Demange & M. Wooders (Eds.) *Group formation in economics: Networks, clubs and coalitions*, (pp.335-53) Cambridge UK: Cambridge University Press.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2, 113-120.

Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92(5), 1170-82.

Burt, R.S. (1992). *Structural holes: The social structure of competition*. Cambridge Massachusetts: Harvard University Press.

Coleman, J.S. (1988). Social capital and the Creation of Human Capital. *The American Journal of Sociology*, vol 94 (supplement), ppS95-S120.

Conley, J. & Smith, S. (2005). Coalitions and clubs: Tiebout equilibrium in large economies. in G. Demange & M. Wooders (Eds.) *Group formation in economics: Networks, clubs and coalitions*, (pp. 246-65) Cambridge UK: Cambridge University Press.

Faust, K. & Wasserman, S. (1994). *Social networks analysis: Methods and applications.* Cambridge, UK: Cambridge University Press.

Freeman, L.C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks*, 1, 215-39.

Goh, K.I., Eom, Y.H., Kanhg, B. & Kim, D. (2006). Structure and evolution of online social relationships: Heterogeneity in unrestricted discussions. *Physical Review E*, 73(6): 066123-1 to 8.

Gomez, D., Gonzalez-Aranguena, E., Manuel, C., Owen, G., del Pozo, M. & Tejada, J. (2003). Centrality and power in social networks: A Game Theoretic Approach. *Mathematical Social Sciences*, 46, 27-54.

Granovetter, M.S. (1973). The strength of weak ties. *American Journal of Sociology*, 78, 1360-80.

Hadi, A.S., & Ling R.F. (1998). Some cautionary notes on the Use of Principal Components Regression. *The American Statistician* 52(1), 15-19.

Jackson, M.O. & Wolinski, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory*, 71, 44-74.

Jackson, M.O. (2005). A survey of network formation models: Stability and efficiency. in G. Demange & M. Wooders (Eds.) *Group formation in economics: Networks, clubs and coalitions*, (pp. 11-57) Cambridge UK: Cambridge University Press.

Milgrom, S. (1967). The small world problem. *Psychology Today*, 2, 60-7.

Monge, P.R. & Contractor, N.S. (2003). *Theories of communication networks*. New York: Oxford University Press.

Moriera, A.A., Paula, D.R., Costa-Filho, R.N. & Andrade, J.S. (2006). Competitive cluster growth in complex networks. *Physical Review E*, 73(6), 065101: 1-4.

Palamara, J.R. & Shapley, .LS. (2000) Simple games and authority structures. *UCLA Working Paper No 796* March 2000.

Stephenson, S. & Zelen, K. (1989). Rethinking centrality: methods and examples. *Social Networks*, 11, 1-37.

Notes

Note 1. The local provision of public goods and the formation of a society into jurisdictions based upon preferences for collections of such locally provided (and funded) public goods.

Note 2. Bonacich, by basing his centrality scores on the largest eigenvector/eigenvalue pair, is effectively taking the amount of 'stretch' in the dimension where the most stretch has occurred, and then apportioning that across nodes according to the elements of the eigenvector to measure centrality.

Note 3. Jackson (2005) and Monge & Contractor (2003) contain two surveys of models of network evolution.

Note 4. Bonacich's (1972) measure is similarly 'non-learning', as it has the same basis as ours.

Table 1. Network Scores from Stars to Full Connection

Network Size	Scores										
	Star	+1arc	+2arcs	+3arcs	+4arcs	+5arcs	+6arcs	+7arcs	+8arcs	+9arcs	+10arcs
3 node	2	2.74788									
4 node	2	2.34062	2.67731	3.48776							
5 node	2	2.18936	2.39427	2.58431	2.92571	3.35392	4.22552				
6 node	2	2.11541	2.25201	2.39252	2.53009	2.69853	2.91907	3.18196	3.54051	4.03791	4.96244