Social Network Theory, Broadband and the World Wide Web

Daniel Sgroi *
Faculty of Economics and Churchill College, University of Cambridge

Keywords: sopcial network theory, stability, broadband, world wide web

Contact Details:

Faculty of Economics
Austin Robinson Building
Sidgwick Avenue
Cambridge
CB4 9DD
United Kingdom

Email: daniel.sgroi@econ.cam.ac.uk
Tel: +44 (0)1223 335244
Fax: +44 (0)1223 335299

* Funded by the EU FP6 Specific Support Action "Competition, Contents and Broadband for the Internet in Europe" IST-2004-2012 (D8).

Social Network Theory, Broadband and the World Wide Web

Abstract

This paper aims to predict some possible futures for the World Wide Web based on several key network parameters: size, complexity, cost and increasing connection speed thorough the uptake of broadband technology. This is done through the production of a taxonomy specifically evaluating the stability properties of the fully-connected star and complete networks, based on the Jackson and Wolinsky (1996) connections model modified to incorporate complexity concerns. We find that when connection speeds are low neither the star nor complete networks are stable, and when connection speeds are high the star network is usually stable, while the complete network is never stable. For intermediate speed levels much depends upon the other parameters. Under plausible assumptions about the future, we find that the Web may be increasingly dominated by a single intermediate site, perhaps best described as a search engine.

1 Introduction

Recently in economics, a literature has emerged looking at how networks develop in society. A good survey can be found in Jackson (2003) focusing on social networks, and Bloch (2003) which provides a useful list of applications to industrial organization. The potential range of application is considerable, from social networks, information networks in firms, through to physical networks, such as the Internet or broadband networks, though the Internet has so far received surprisingly sparse attention. The possible exception is in peering and transit in the Internet, where there is some work, for example Badasyan and Chakrabarti (2003), but still not as much as the size and importance of the Internet might warrant. A key feature of the current literature on social network formation is that networks are not necessarily planned centrally, or even prone to much control, they often simply emerge and evolve, which is characteristic of the Internet and the World Wide Web. A second related literature looks at what makes a network optimal, which represents a possible end-state to this process of evolution. Perhaps the key criterion linking these two literatures is the notion of network stability. Should an individual user be able to create a new link in a network and should this improve his or her utility then it seems reasonable to assume that it will be done, at least in the medium to long-run. Eventually however we might reach a state where no one would benefit from a change to the network structure and this seems a good candidate for the long-term resting point of the network formation process.

A dichotomy exists between the literature on social network theory which tends to deal with links between individuals or firms which are often cheap to initiate, and so the *de facto* cost may come in terms of congestion or complexity, and physical networks where links are often very expensive to build and maintain and hence have a high direct cost. Take for instance the decision to make a work colleague aware of your area of expertise, versus the decision to build a new road, gas pipeline, or railway line. The big cost in the first case might be the concern that once identified you might face a greater work-load, in terms of the second case there is a considerable cost even to building the link. The Internet provides an interesting hybrid. It is a genuinely physical network, but one where connecting to a given existing network is relatively cheap in terms of direct cost, with much of the analysis taking place in terms of associated costs like complexity or network externality effects on others which might seem more relevant concerns for the social network literature.

This paper seeks to use some of the existing tools of social network theory to analyze the world wide web and broadband access. We develop a taxonomy of predictions for the future of the web in which the impact of high speed broadband access can be judged. To give an overview, we first modify the Jackson and Wolinksy (1996) connections model, an archetypal social network model, to incorporate differing connection speeds and complexity costs. Second, we apply this new model to the world wide web and formulate a taxonomy of predictions for the future of the web. Finally, a residual, but important, aim of the paper is to show how social network theory can assist in assessing the future impact of broadband access, and is by no means an exhaustive attempt to complete the task, rather it presents a number of starting points for research. Section II develops a connections model for the Web, defining notation, and examining the archetypal networks to be used throughout, and ends with a particularly important list of limitations and potential extensions to the underlying model used in this paper. Section III examines the key concept of stability, and sections IV and V apply this concept to the star network and complete network in turn. Section VI turns to efficiency and section VII summarizes the main findings of the paper. Section VIII concludes.

2 A Connections Model of the Web

A user accessing the World Wide Web can derive utility by accessing the web-sites of others, or more specifically the information contained on those sites. The Internet allows connection to distant resources, which raises the value of the Internet as more and more users join and allow further interconnection. We will consider web-sites to offer two basic services: either they provide information; or they provide access to another web-site which in turn may provide information, or further access. We can then imagine the web as a series of connected nodes, each of which is of value for the information contained in each node, and the access to further nodes which is granted. In this way the web is well modelled as a collection of nodes connected by edges on a graph: the classic *modus operandi* of the social network literature.

Crucially in the social network literature the cost of forming links is relatively low, and hence there is a clear distinction between road and rail networks for which the cost of establishing a link is so significant that notions of stability are very different. In particular it may not be possible for users to connect to each other; the network may be under the direct control of others who are not themselves nodes. The world wide web with its series of interconnected web-sites is much more like a social network in practice, perhaps typified by a network of familial or social connections, or a firm with a series of connections across different workplaces. On the web users can and do control the links on their own web-sites, and can and will alter those links if they wish to re-optimize, making notions of individual stability important at the level of individual nodes.

We begin with some simple definitions of what we mean by nodes, edges, links, etc. applied to the Internet, and also some clear definitions of what we mean by the Internet, world wide web, broadband, etc. Then we move to two examples of how social network theory can be applied to Internet usage, and how the move from slow connection speeds to faster broadband links alter these decisions.

Here we directly apply a modified version of the model of Jackson and Wolinsky (1996) to the World Wide Web. We can simply characterize the Web as a complex graph in which nodes represent web sites, and edges represent hyper-links between sites. One major issue for a Web user is how quickly the relevant information can be obtained. Here we also focus on the number of clicks required to reach a page containing the information needed, and the level of complexity a user must negotiate when searching for information on the web. In each case we proxy these considerations using the standard tools of social network theory plus a few modifications.

Consider a user who wishes to find a certain piece of information and so decides to search the web. For simplification we assume that each user begins his search on his own home-page, which may itself contain some information which might make it the end point of someone else's quest for knowledge. The user is assumed to know exactly where the information rests and indeed the shortest route to travel. The user then has several considerations: firstly the number of pages which have to be traversed; secondly, the level of complexity of each page which might slow the search, and finally the cost of maintaining any links which enable other web-pages to be accessed. To summarize the key parameters:

1. We measure distance between nodes as the minimum number of clicks between a user and the information he seeks;

2. Complexity is measured in terms of the number of superfluous hyper-links on each site, which can be time-consuming to navigate;

3. Direct cost is a constant measure of the cost associated with keeping and maintaining direct links from the user's home-page to other sites;

4. Finally, connection speed between nodes is of considerable importance and acts together with the distance between nodes.[1]

Each of these measures will be introduced and explained. As we shall see we can use this model to address questions about the network, such as how broadband technology might impact on the network, through a unilateral increase in the speed parameter. Note that throughout we assume the perfect reliability of links, so our notion of superfluous links does not factor in any benefits from redundancy.

Figure 1 gives an example of a simple representation through which a single user can gain a required piece of information. We might imagine any site with many links, especially the centre of the star like structure in Figure 1 to represent a search engine, which potentially links to all sites on the network. This is complex to navigate but ensures a link to virtually anywhere a user wishes to go.

[insert FIGURE 1 here]

[1] In the internet if we wish to travel from one node to another it is quite possible that the information we wish to access will actually travels via an alternative route. However, we are more interested in the actual experience of the user. To give an example, consider a user wishing to navigate from his home-page, node A, to node C, via node B. First, the user finds the link to node B on his home-page and clicks this link. Then the user waits for node B, the new webpage, to appear, and must navigate this site searching for the link to node C. Once this is found the user clicks the link and waits for node C to appear. So the user has traversed from A to C via B, and even if the connection speeds when clicking the links to B and C were such that connection was instantaneous, the user still had to search through the websites to find the links to B and C, and it is this searching time which provides one of the novel aspect of this paper. It is important to note that this paper is mainly concerned with these metaphorical travels by users through the web, and not necessarily the flow of information from node to node.

2.1 The Basic Model

Let us start with the network model given in Jackson and Wolinsky (1996). Let $N = \{1, 2, ..., N\}$ be the finite set of nodes on the World Wide Web (or simply the Web) with the cardinality of this set of nodes (or network size) as N^S. Examples of such nodes include web-sites containing information of use, the home-pages of users themselves, ISP home-sites to which users subscribing to those services initially connect, and search engines, which are represented as nodes with a particularly large number of edges or *central stars* (representing their aim to allow connection to a wide variety of sites, for example on 10th May 2004, Google claimed to be able to search 4,285,199,774 web pages). We start be defining the graph and edges.

Definition 1 *The complete graph, denoted g^N is the set of all subsets of N of size 2. The set of all possible graphs on N is then $\{g \mid g \subset g^N\}$.*

Definition 2 *We denote the subset of N containing nodes i and j as the (hyper)link or edge ij.*

So if $ij \in g$ then the nodes or sites i and j are directly connected while if $ij \notin g$ then the nodes or sites are not directly linked, though they may be connected indirectly (via other sites). We denote the addition or subtraction of a node or site from a network via the notation $g + ij = g \cup \{ij\}$ or $g - ij = g \backslash \{ij\}$ respectively. Let $N(g) = \{i \mid \exists j \text{ s.t. } ij \in g\}$. The next important concept is the path connecting nodes.

Definition 3 *A path in g connecting i_1 and i_n is a set of nodes $\{i_1, i_2, ..., i_n\} \subset N(g)$ s.t. $\{i_1i_2, i_2i_3, ..., i_{n-1}i_n\} \subset g$. The graph $g' \subset g$ is a component of g, if for all $i \in N(g')$, $i \neq j$, there exists a path in g' connecting i and j, and for any $i \in N(g')$ and $j \in N(g)$, $ij \in g$ implies that $ij \in g'$.*

Jackson and Wolinsky focus on the total productivity of the graph or network, and how this is allocated among nodes, using a *value function* and an *allocation function.*

Definition 4 *The value is represented by $\upsilon : \{g \mid g \subset g^N\} \to \mathbb{R}$, given as the aggregate of all individual utilities, $\upsilon(g) = \Sigma_i u_i(g)$, where $u_i : \{g \mid g \subset g^N\} \to \mathbb{R}$. The set of all such value functions is V.*

This naturally leads to a definition of efficiency.

Definition 5 *A network or graph is efficient if $\upsilon(g) \geq \upsilon(g')$ for all $g^{'} \subset g^N$, so it returns the maximal possible total value.*

We next define the allocation rule.

Definition 6 *The allocation rule $Y : \{g \mid g \subset g^N\} \times V \to \mathbb{R}^N$ describes how the value associated with each network is distributed to the individual users. $Y_i(g, \upsilon)$ is the payoff to user i from graph g under the value function υ.*

Finally, we end with definitions of stability and defeat.

Definition 7 *The graph g is pairwise stable with respect to υ and Y if: (i) for all $ij \in g$, $Y_i(g, \upsilon) \geq Y_i(g - ij, \upsilon)$ and $Y_j(g, \upsilon) \geq Y_j(g - ij, \upsilon)$; and (ii) for all $ij \notin g$, if $Y_i(g, \upsilon) < Y_i(g + ij, \upsilon)$ then $Y_j(g, \upsilon) > Y_j(g + ij, \upsilon)$.*

Definition 8 *We shall say that g is pairwise defeated by g' if $g' = g - ij$ and (i) is violated for ij, or if $g' = g + ij$ and (ii) is violated for ij.*

Condition (ii) embodies the assumption that, if i strictly prefers to form the link ij and j is just indifferent about it, then it will be formed. The notion of pairwise stability is not dependent on any particular formation process. That is, we have not formally modeled the procedure through which a graph is formed. Pairwise stability is a relatively weak notion among those which account for link formation and as such it admits a relatively larger set of stable allocations than might a more restrictive definition or an explicit formation procedure (see Jackson and Wolinsky, 1996 for more discussion of this). As such we will need the similar but alternative requirement of individual stability.

Definition 9 *The graph g is individually stable with respect to v and Y if: (i) for all $ij \in g$, $Y_i(g, v) \geq Y_i(g - ij, v)$; and (ii) for all $ij \notin g$, $Y_i(g, v) \geq Y_i(g + ij, v)$.*

This is simply the requirement that any individual node i in the network does not strictly benefit from breaking a link or establishing a link with any other node $j \neq i$. Therefore we can also adjust the definition of defeat:

Definition 10 *We shall say that g is individually defeated by g' if $g' = g - ij$ and (i) is violated for ij, or if $g' = g + ij$ and (ii) is violated for ij.*

Both individual stability and individual defeat are the more relevant concepts for a network in which each individual i can make or break a link ij without the express permission of individual j.

2.2 Costs and Benefits

Now we move further from the Jackson and Wolinsky model in order to tailor the model to our needs. In particular we need to modify the utility function in the Jackson and Wolinsky model to take account of the complexity of web-browsing. Furthermore, we are considering mainly directed links, so the network will look slightly different from each user's perspective, which partly explains why we use individual stability, and why issues of efficiency are less relevant, as we shall see in later sections.[2]

Consider the shortest distance t_{ij} between i and j to be the main measure of distance, and $\delta \in (0, 1)$ to be a connection speed variable, so as δ rises the connection speed rises, rendering distance less important, but as δ falls sites several links away become hard to reach and provide relatively low utility. This is represented by the function $\delta^{t_{ij}}$. In particular while we restrict the speed variable to lie in the unit interval under normal circumstance, we will model a broadband high speed connection by simply allowing $\delta \to 1$ and comparing the impact with lower levels of δ achievable under a dialup connection.[3]

[2] Allowing users to return along their path of travel is fine given the existence of a "back" key on most browsers. This explains why we do not consider the return path back to the last site as a superfluous link, and therefore exclude going back as part of the complexity cost, since we can differentiate hitting the "back" button, rather than following a new link. It is also clearly not superfluous in any sense. For a search engine, you might of course encounter the site you came from, but we can reasonably assume you would recognize and instantly discount this site.

[3] Note of course that $\delta = 0$ literally means zero connection speed, and $\delta = 1$ literally means instantaneous connection speed. However, by considering the limits $\delta \to 0$ and $\delta \to 1$ we can proxy the impact of slowing connection speeds right down, or moving towards instantaneous connections, and that is the approach taken in this paper. Of course, we also consider intermediate speeds throughout, but the extremes are often useful in revealing the ultimate impact of increasing or decreasing speeds.

The notion of the cost of setting up links is a more complex procedure than simply setting a fixed cost per link of c_{ij} or c, as the main cost comes in terms of the ensuing complexity of sites with large numbers of links, which add text and time to negotiate the site. Even fast connection times will not alter the need to spend time and effort negotiating a complex site. On the other hand the direct cost to setting up a link is small, though positive. Therefore for simplicity we imagine a fixed cost of $c_{ij} = c$ for all links, but add to this a complexity cost $(1 - \beta^{h_{ij}})$ where $\beta \in (0, 1)$ and h_{ij} is a direct measure of the number of irrelevant links on intermediate sites (including i) on the path between i and j. For example, a simple line from i to j with no other links contains no complexity, so $h_{ij} = 0$, and $\beta^{h_{ij}} = 1$, therefore producing an overall complexity cost of 0. However as we add a link to node k at node i this renders the node i more complex to negotiate, so the complexity cost rises to $1 - \beta$, and so on as we add links from intermediate nodes on the geodesic between i and j.[4] Figure 2 gives an example of the minimally complex route between i and k in network A, and a much more complex route between i and k in network B, even though the distance, $t_{ij} = 2$, remains the same across both networks. Note that we do not include "doubling-back" in the measure of complexity since it seems reasonable that a user can recognize their own site, or where they have come from, so will not need to worry about doubling-back in error. Hence for network A we have no complexity cost as $h_{ij} = 0$ and so $\left(1 - \beta^0\right) = 0$, but for network B we have positive complexity as $h_{ij} = 4$ and so $\left(1 - \beta^4\right) > 0$.

[insert FIGURE 2 here]

Denote the total number of links from i as N_i^L and $N^L = \sum_{i \in g} N_i^L$ as the total number of links in the network (which we can regard as a possible proxy measure of the overall complexity of the network), and the total number of nodes or sites in g as N^S (which we defined earlier as the cardinality or total size of the network).[5] We can therefore identify superfluous links as the number of links along the shortest route from i to j minus the minimal distance t_{ij}, so

$$h_{ij} = \left(N_i^L - 1\right) + \left[\left(N_{k_1}^L - 2\right) + ... + \left(N_{k_{t_{ij}-1}}^L - 2\right)\right] = N_i^L + 1 - 2t_{ij} + \sum_{z=1}^{t_{ij}-1} N_{k_z}^L \qquad (1)$$

Where $k_1, ..., k_{t_{ij}-1}$ are the $t_{ij} - 1$ intermediate nodes which lie on the shortest path between i and

[4] Note that since we assume the perfect reliability of links thus there is no gain from the existence of irrelevant links. It is of course feasible that an alternative route, which would require superfluous links, would be worthwhile, but for simplicity we ignore this.

[5] The formulation of complexity costs given below is here concave in N, and so quite suits the inclusion of search engines. Once the number of linked sites becomes large, the increase in costs are going to be very small, and hence the user would prefer to use the engine rather than trawl through a whole list of sites. This is confirmed by the analysis below which shows that the network architecture based most strongly on a well-connected search engine, the fully-connected star network, is especially stable for large N.

j.[6] The total utility to i from network g can therefore be given as:

$$u_i(g) = \sum_{j \neq i} \delta^{t_{ij}} - N_i^L c - \sum_{j:ij \in g} \left(1 - \beta^{h_{ij}}\right) \quad (2)$$

2.3 Specific Network Architectures

We next move on to define some standard network architectures which will form the bulk of our analysis.

Definition 11 *In a (fully-connected) star all sites connect to the centre star (cs) and no other node, so $N_i^L = 1$ for all $i \neq cs$, and $N_{cs}^L = N^S - 1$. The shortest distances between i and j is*

$$t_{ij} = \begin{cases} 2 & \text{for } i \neq j, cs; j \neq cs \\ 1 & \text{for } i = cs; j \neq cs \\ 1 & \text{for } i \neq cs; j = cs \end{cases} \quad (3)$$

and by equation 1 the number of superfluous links is

$$h_{ij} = \begin{cases} N^S - 3 & \text{for } i \neq j, cs; j \neq cs \\ N^S - 2 & \text{for } i = cs; j \neq cs \\ 0 & \text{for } i \neq cs; j = cs \end{cases} \quad (4)$$

So for a star, movement to cs is very short and not complex, movement via (or from) cs is also short but complex. Next we examine a symmetric network.

Definition 12 *In a symmetric network every node (site) has the same number edges (links) of, $N_i^L \equiv \overline{N}$ for all i, and therefore by equation 1 the number of superfluous links between the shortest distance between i and j are simply given by*

$$h_{ij} = N_i^L + 1 - 2t_{ij} + \sum_{z=1}^{t_{ij}-1} N_{k_z}^L = \left(\overline{N} - 2\right) t_{ij} + 1 \quad (5)$$

One example of a symmetric network is a circle, where every node has two edges, another special type of symmetric network called the complete network is one in which every node is connected to every other node, so in particular $N_i^L = N^S - 1$.

Definition 13 *In a complete network every node (site) links to every other node (site). Therefore for all $i \neq j$ the shortest distance between i and j is $t_{ij} = 1$ and by equation 1 the number of superfluous links is*

$$h_{ij} = N_i^L - 1 = N^S - 2$$

[6] Note that where required we assume a lexicographic ordering whereby firstly a shortest distance is determined and secondly complexity is calculated, no user is assumed to opt for a longer route no matter how much lower the complexity cost. If there are multiple shortest distances then the least complex is used. Note that we can justify this assumption on two grounds. Firstly, we might argue that users may have some idea of the distance, but not necessarily the complexity of the route. Secondly and far more significantly, we can note that under the two key architectures to be used this requirement is not important. For the star network there is effectively no alternative route, and when a new route is constructed through an additional link ij it provides both a shorter and less complex route from i to j. For a complete network removing when we consider multiple routes from i to j when the direct link ij is removed, the shortest route is always no more complex than any other.

In a complete network all sites are very close, but direct costs are high. Next we will examine the stability properties of certain key network architectures, and how these change as the main parameter values change.

2.4 Limitations and Extensions

Finally, it is important to stress the limitations of this paper, which are address in a series of inter-linked sections.

2.4.1 Network Architectures

This paper is very focused on predicting the shape of the web, and in particular limiting itself to the two most well-known possible resting places for the evolution of any network: the complete network and the fully-connected star. It is of course feasible that other long-term solutions can and will exist. It is also clear that changes in technology may render a network unstable from an initial point of stability. Should technological change occur at a rate greater than the speed of movement towards a stable architecture, then stability may never occur, with changes constantly altering the end-point of the process, so it is never reached.

2.4.2 Pairwise and Individual Stability

Simply removing complexity as an issue does not reduce this paper to a retelling of Jackson and Wolinsky (1996) in a different setting, as the stability conditions are different. This paper assumes that anyone can link to anyone else's site, or browse the information contained on that site. This may seem a reasonable approximation, but many sites can and do restrict access through passwords, or the blocking of access from certain IP addresses. If this is increasingly the case, then it makes sense to move to pairwise stability as the operating paradigm.

To compare the stability concepts, under pairwise stability two nodes must both agree to break or make a link. Therefore, when considering the stability of an arbitrary network architecture, it becomes clear how these concepts differ. In a complete network the key issue for stability is breaking links, which is easier under individual stability, and so stability is less likely under individual stability. Under a fully-connected star network, the key condition for all non-central nodes, concerns adding a link, which is also easier to foresee under individual stability, so under that assumption stability is also less likely to be obtained in a star network. For the central node the key issue is cutting links, and so as with a complete network, this is once-again simpler with individual stability, and so stability is harder. Therefore our general findings for stability may be tougher than those imposed by pairwise stability.

2.4.3 The Centre of the Star

In many ways the central node in a fully-connected star network is very different in this paper, than simply a node that happens to be at the centre of the network. We can see this by considering the argument that especially as $N \to \infty$ why should the central node wish to link to every single website? Costs might become prohibitive, and traffic through the site might also reach prohibative levels. Our answer for the purposes of this paper is that the central node is implicitly considered to be different from all other nodes: it has a rationale for existing that is not based on sharing information with other nodes, rather it is itself a firm attempting to win business as the central node and it has a profit function that is increasing in the number of sites it links. This is most clear when glancing at Google's main page in which it proudly declares the count of how many pages are it "searches". The various sites offered by candidate central nodes such as Google, Yahoo and others, provide ample justification for the quest to have a higher searching capacity than rival sites.

In future work, we might wish to look more closely at the rationale behind the actions of the central node in a star network, or "search engines", and treat them very differently. We might wish our model to include an explanation for the appearance of such a site in a network, either through some evolution of a "standard node", or through the direct insertion of a search engine into the network. Historically there have been many such sites, though often only a small number have been dominant. In many respects the simplicity of this paper cannot justify the long-run existence of more than one such site, so we implicitly assume that only one will become dominant. In reality, at least in the short to medium run, multiple such sites do continue to exist. We might even consider meta-search engines, which choose between existing search engines to be the long-run centre, with a group of search engines clustered around these, something plausible, but not considered here.

2.4.4 Efficiency

The main issue for this paper concerns stability, as we are mainly focused on making long-run predictions. However efficiency is traditionally at least as important a concern. One major feasible point (made at least in Jackson and Wolinsky) concerns the contradictions between these two concepts. In particular we might find that the most efficient network is not stable, or the only stable network is not efficient. Future work will and should examine the implications for the shape of the world wide web. Since the web evolves, rather than being planned, stability may well be the more important predictive concept, but if a final prediction involves an inefficient structure there may be important policy implications.

2.4.5 Small Worlds

Recent related work (Goyal, 2004), provides an alternative framework from Jackson and Wolinsky (1996). Goyal starts with the notion of co-authors working together, with some highly regarded authors who link to others providing something like the central nodes of star networks. This provides some interesting results, which recall the small world argument, that no two people (or authors, or nodes) are likely to be too distant in equilibrium, because of the existence of highly regarded authors. Even lowly economists might co-author with someone who has co-authored with someone who has co-authored with a highly regarded economist. This might equally well be a good basis for a model of web-links, and in particular incorporates a form of congestion cost (based on the limit of what the central authors can hope to achieve) that might relate well to web-browsing.

2.4.6 Speed and Content

Finally, this paper makes an assumption that broadband is well-modelled in terms of a large speed increase. Add to this the assumption that broadband will become increasingly popular and we are left with the assumption that speeds will rise over time. Taken to its logical extreme, the paper moves towards predicting what the web will look like when connection times are instantaneous. However, throughout this we are assuming that content does not change, or at least if it does, not sufficiently quickly to counterbalance increases in speed. It may well be that this is incorrect. Greater use of video streaming, online gaming and other high bandwidth content, can undoubtedly limit the rises in connection speed produced by greater use of broadband. To attempt to second guess the advances of such content is beyond the scope of this paper, but it will undoubtedly provide an upper limit to effective speed increases.

3 Stability in an Example Network

We allow any node to link to any other, or similarly any node may break a link to another. We examine undirected networks and then examine their stability properties under the assumption that any node can break or make additional links. The biggest impact here is to make the relevant notion of stability, individual stability, and not pairwise stability. We might wonder why stability should be a very important consideration for the world wide web, which is necessarily changing and growing as time progresses. However, especially in terms of the impact of faster connection speeds, stability might provide a reference point from which to begin predicting the long-term future shape of the web.

In this section we give show the method of determining stability for an example network before going on in the following sections to examine the stability properties of the star network and the complete network.

Figure 3 shows a network g_1, and we define $g_2 = g_1 + ij$ (i.e. g_2 is the network g_1 in addition to the dotted link ij).

[insert FIGURE 3 here]

We wish to determine the individual stability requirement for g_1 and one way to do this would be to consider the addition of a single link at ij. The gain in terms of proximity to j for i would be given by the difference $\delta - \delta^2$, while there is an extra cost both direct, c, and indirect through added complexity. Denote the gain in terms of distance from the change from g_1 to g_2 as $\Delta D(g_1, g_2)$. In this case the gain in distance is simply

$$\Delta D(g_1, g_2) = \delta^{t_{ij}(g_2)} - \delta^{t_{ij}(g_1)} = \delta - \delta^2$$

Denote the change in direct cost and complexity as $\Delta C(g_1, g_2)$ which is

$$\begin{aligned} \Delta C(g_1, g_2) &= c + \left(1 - \beta^{h_{ij}(g_2)}\right) + \left(1 - \beta^{h_{ik}(g_2)}\right) + \left(1 - \beta^{h_{il}(g_2)}\right) \\ &\quad + \left(1 - \beta^{h_{im}(g_2)}\right) - \left(1 - \beta^{h_{ij}(g_1)}\right) + \left(1 - \beta^{h_{ik}(g_1)}\right) + \left(1 - \beta^{h_{il}(g_1)}\right) + \left(1 - \beta^{h_{im}(g_1)}\right) \\ &= c + 3\left(\beta^2 - \beta^3\right) \end{aligned}$$

The net stability condition requires that the change in cost outweighs the benefit in reduced distance, so we have the condition

$$g_1 \text{ is stable } \Leftrightarrow \Delta C(g_1, g_2) > \Delta D(g_1, g_2) \Leftrightarrow c + 3\left(\beta^2 - \beta^3\right) > \delta - \delta^2$$

The first term c is just the direct cost, which added to the term $3(\beta^2 - \beta^3)$ gives the additional complexity for i of the new network. Put simply, when i wishes to travel to any node other than j, a slight increase in complexity is faced as a new superfluous link has been added. On the other hand, j is closer by $(\delta - \delta^2)$, and the route to j has remained as complex as before and so is not considered. Now we can note that as $\beta \to 1$ and $\delta \to 0$ so the connection speed rises and concern for complexity falls, then g_1 is stable for any $c > 0$. Similarly for all other extreme combinations such as $\{\beta \to 0, \delta \to 0\}$, $\{\beta \to 1, \delta \to 1\}$ and $\{\beta \to 0, \delta \to 1\}$. For intermediate values of β and δ the network can become unstable for low values of c. Given that c (the direct cost of establishing a hyper-link on a web-site) is likely to be quite low, we can only be certain of a stable network with extreme values of β and δ. Note that by switching from dialup to broadband, we would expect $\delta \to 1$ and so move some way towards stability.

To complete the analysis of stability we would need to consider removing a link, and also examine the possible actions of nodes other than i, but rather than go through this process for a very specific network, we will instead switch our attention to two important classes of networks: the star and the complete network.

4 Stability in the Star Network

Figure 4 below shows a network g_3, and we define $g_4 = g_1 + ij$. Here and unless otherwise said, we consider the star network to have only one component and to encompass everyone.

[insert FIGURE 4 here]

We wish to consider whether the network is individually stable. Firstly, we consider the impact of adding a new link, and then follow with deleting a link. Should both of these actions result in a lower utility for the node considering the addition or subtraction of a link, then the network is individually stable.

4.1 Adding a Link

Starting with the addition of a new link ij results in a shorter distance between the nodes i and j, so

$$\Delta D\left(g_3, g_4\right) = \delta^{t_{ij}(g_4)} - \delta^{t_{ij}(g_3)} = \delta - \delta^2$$

However, the addition of ij also adds a direct cost and a greater complexity

$$\begin{aligned}\Delta C\left(g_3, g_4\right) &= c + \left(1 - \beta^{h_{ij}(g_4)}\right) + \left(1 - \beta^{h_{ik}(g_4)}\right) + \left(1 - \beta^{h_{il}(g_4)}\right) \\ &\quad + \left(1 - \beta^{h_{im}(g_4)}\right) - \left(1 - \beta^{h_{ij}(g_3)}\right) + \left(1 - \beta^{h_{ik}(g_3)}\right) + \left(1 - \beta^{h_{il}(g_3)}\right) + \left(1 - \beta^{h_{im}(g_3)}\right) \\ &= c + 1 - 2\beta + 3\beta^2 - 2\beta^3\end{aligned}$$

The total impact of the change from g_3 to g_4 is therefore

$$\Delta D\left(g_3, g_4\right) - \Delta C\left(g_3, g_4\right) = \left(\delta - \delta^2\right) - c - \left(1 - 2\beta + 3\beta^2 - 2\beta^3\right)$$

The first term represents the shorter distance, the second the direct cost c, the third term the extra complexity cost of the new network. Before we continue, note the following useful lemma.

Lemma 1 *$0.5p^{n-1} + 0.5p^{n+1} > p^n$ for all $n > 0$ and $p \in (0,1)$.*

Proof. Since $dp^n/dn = p^n \ln p < 0$ and $d^2p^n/dn^2 = p(\ln p)^2 > 0$ therefore $p^n - p^{n+1} < p^{n-1} - p^n$ which immediately implies the result. ■

Now, since $\beta \in (0,1)$, we know that $2\beta^2 > 2\beta^3$, and from lemma 1 we also know that $1 + \beta^2 > 2\beta$, and therefore it must be that $(1 - 2\beta + 3\beta^2 - 2\beta^3) > 0$. Hence g_4 is definitely more complex than g_3, as we would expect with the addition of a new link at i. Now we have our condition for stability

$$g_3 \text{ is stable } \Rightarrow c + 1 - 2\beta + 3\beta^2 - 2\beta^3 > (\delta - \delta^2) \tag{6}$$

We can generalize expression 6 to give us stability conditions for any star network. Any new link ij where neither i nor j are the centres of the star, necessitate a direct cost rise of c and a rise in complexity costs of $\left(1 - \beta^{N^S-1}\right) - \left(1 - \beta^{N^S-2}\right)$ for roots between i and $k \neq j, cs$. There are $\left(N^S - 3\right)$ such nodes. Complexity costs fall between i and j by $\left(1 - \beta^{N^S-2}\right) - (1-\beta)$ and rise between i and the cs by $(1-\beta)$, leaving a final stability condition for a general star network of

$$\begin{aligned}
\text{stable} \;\; &\Rightarrow \;\; c + \left(N^S - 3\right)\left[\left(1 - \beta^{N^S-1}\right) - \left(1 - \beta^{N^S-2}\right)\right] + (1-\beta) \\
&> \;\; \left(\delta - \delta^2\right) + \left[\left(1 - \beta^{N^S-2}\right) - (1-\beta)\right] \\
&\Rightarrow \;\; c + \left(N^S - 3\right)\left(1 - \beta^{N^S-1}\right) + 2(1-\beta) > \left(\delta - \delta^2\right) + \left(N^S - 2\right)\left(1 - \beta^{N^S-2}\right) \\
&\Rightarrow \;\; c + 1 - 2\beta + \left(N^S - 2\right)\beta^{N^S-2} - \left(N^S - 3\right)\beta^{N^S-1} > \left(\delta - \delta^2\right) \qquad (7)
\end{aligned}$$

While this is only a necessary condition, and not sufficient for stability, we will consider some extreme values of the relevant parameters in brief now, and look in depth again when the necessary and sufficient condition is produced below. Note that this collapses to expression 6 when we set $N^S = 5$.

As complexity concerns grow more and more important, so $\beta \to 0$ then condition 7 becomes a requirement for $1 + c > \delta - \delta^2$, which must be true as $1 > \delta - \delta^2$, for $\delta \in (0,1)$. As complexity becomes less and less important, so $\beta \to 1$, we have the required condition $c > \delta - \delta^2$, which is recognizably the same condition as in Jackson and Wolinsky (1996) which makes sense as their model is similar, but with no complexity concerns. Note that when $\beta \to 1$ and $\delta \to 1$, i.e. a virtually instantaneous connection, stability is trivially satisfied for any c. As network size shrinks down to a small size, for example $N^S = 4$, we have the stability condition $1 + c + 2\beta^2 > \left(\delta - \delta^2\right) + 2\beta + \beta^3$. Since we know that $\beta^2 > \beta^3$, and from lemma 1 we know that $1 + \beta^2 > 2\beta$, we know that this is a weaker condition than $1 + c > \delta - \delta$ and is certainly satisfied for $\delta \to 1$.

Usually we will consider $N^S = 4$ to be the smallest non-trivial network, since $N^S = 3$ shifts from star to complete with the addition of a link, and $N^S = 2$ shifts to a the empty network with the subtraction of a link. However, just to check the example for $N^S = 3$, we have the condition of $1 + c > \delta - \delta^2 + \beta$ which is also trivially satisfied for $\delta \to 1$. For $N^S \to \infty$ we have the condition $1 + c > \left(\delta - \delta^2\right) + 2\beta$. Finally, if $N^S \to \infty$ and $\delta \to 1$ then the condition becomes $1 + c > 2\beta$.

4.2 Subtracting a Link

The impact of subtracting a link is immediate for $i \neq cs$, since this will remove i from the network, so we need only consider whether the network itself provides a positive utility. In effect then we need $u_i(g) > 0$ as our second criterion for stability. In general this gives us $u_i(g) = \sum_{j \neq i} \delta^{t_{ij}} - N_i^L c - \sum_{j:ij \in g} (1 - \beta^{h_{ij}})$. We know that in the star network there is only a single link at $i \neq cs$, so $N_i^L = 1$. Furthermore, the distance to any node $j \neq cs$, is simply 2, while the distance to $j = cs$ is simply 1, so we have

$$\sum_{j \neq i} \delta^{t_{ij}} = \left(N^S - 2\right) \delta^2 + \delta$$

Finally, since there is only one link at $i \neq cs$ but $N^S - 1$ links at cs then $h_{ij} = N^S - 3$ for $i \neq cs$, and since it costs nothing in terms of complexity to reach cs, we know the total complexity costs are

$$\sum_{j:ij \in g} \left(1 - \beta^{h_{ij}}\right) = \left(N^S - 2\right) \left(1 - \beta^{N^S - 3}\right)$$

Totalling these we have a second stability condition

$$\text{stable} \Rightarrow u_i(g) > 0 \Rightarrow \left(N^S - 2\right) \delta^2 + \delta - c - \left(N^S - 2\right) \left(1 - \beta^{N^S - 3}\right) > 0 \tag{8}$$

Now once again let us consider some extreme cases, noting of course that this is a necessary but not sufficient condition for stability. Firstly, as $N^S \to \infty$, so network size grows large, we have $(1 - \beta^{N^S - 3}) \to 1$, so $u_i(g) < 0$. Put simply, as the network size grows the complexity of the central star becomes so great that journeys via the central star are simply too costly. However, since all journeys must go via the central star (aside from visiting the centre itself) the entire network becomes too costly for each node, and each node will have an incentive to drop out. This is only avoided in the case where $\beta = 1$, i.e. where there is no concern for complexity at all, or when $\delta \to 1$, so connection tends towards instantaneous, in both cases $u_i > 0$ and certainly if $\beta = \delta \to 1$ all of our earlier stability conditions hold, so the star is stable.

4.3 Overall Stability

In order to be stable a star network needs to meet both condition 7 and condition 8, and hence the overall stability condition incorporating both addition and subtraction of a link is

$$\begin{aligned} \text{stable} \quad \Leftrightarrow \quad & 0 < \min\{\left(N^S - 2\right) \delta^2 + \delta - c - \left(N^S - 2\right) (1 - \beta^{N^S - 3}), 1 + c \\ & -2\beta + (N^S - 2)\beta^{N^S - 2} - (N^S - 3)\beta^{N^S - 1} - \left(\delta - \delta^2\right)\} \end{aligned} \tag{9}$$

Now we can examine what this final necessary and sufficient condition looks like for various extreme parameter values.

Appendix 1 works through all the various possible combinations, and show how the stability condition responds to various parameter values. Here we summarize the main results. With no concern for complexity (so $\beta = 1$), stability is not possible for low speed networks (where $\delta \to 0$), guaranteed for extremely high-speed networks (where $\delta \to 1$), and possible for intermediate speed levels. With extreme complexity concerns ($\beta = 0$), stability is also not possible for networks where connection speeds tend towards zero (where $\delta \to 0$), and also guaranteed for high speed networks (where $\delta \to 1$). With extreme complexity concerns and intermediate speeds stability is unlikely, though possible for smaller network sizes, and not possible as $N^S \to \infty$. If the concern for complexity is intermediate, then we have a similar pattern with no stability with low speeds, definite stability at high speeds, and possible stability at intermediate speeds, though only if N^S is finite. It seems that the most important parameter is probably speed, with high speed guaranteeing stability regardless of complexity concerns and low speeds ruling out complexity. However, with intermediate speeds then the level of complexity costs, direct costs and network size become important.

5 Stability in the Complete Network

Figure 5 shows a network g_5, and we define $g_6 = g_1 + ij$. Once again we wish to consider whether the network is individually stable.

[insert FIGURE 5 here]

5.1 Adding a Link

For the complete network if we consider the impact for i of adding a link this must result in the duplication of an existing link, which adds an additional cost of c, and additional complexity, but with no gain, and is therefore clearly not a beneficial change for any parameter values except for $c = 0$, and $\beta = 1$, which would render the analysis trivial.

5.2 Subtracting a Link

More importantly, we consider the impact on the utility to i of cutting the link ij. The gain here is in terms of a reduced direct cost, c, a change in overall complexity, and a fall in the shortest distance between i and j. Note that all other nodes have a direct link, so the complexity cost across the entire network would be relatively easy to calculate, since it only impacts on journeys starting from i and the journey from j to i. Focusing on i the shortest distance to j falls from δ to δ^2, and direct cost falls by c.

The complexity of the link ij changes from $1 - \beta^{N^S-2}$ to $1 - \beta^{2(N^S-3)}$. Initially, the route involves a direct movement from i to j. At i there are $N^S - 1$ links (the network is complete, but there is no link from i to i) and one of these links is essential, so there are $N^S - 2$ superfluous links. When we cut the ij path we force movement from i to j to go via an intermediate site k. At i there are now $N^S - 3$ superfluous links since one further link has been removed. At k there are also $N^S - 3$ superfluous links, since of the $N^S - 1$ links at k, one is essential, and one would involve doubling-back. So we have $2N^S - 6$ superfluous links for the route ij in $g - ij$, as opposed to $N^S - 2$ in g.

Finally, there is an impact on all other routes from i, which fall in complexity from $1 - \beta^{N^S-2}$ to $1 - \beta^{N^S-3}$ for all but the ij link (so for $N^S - 2$ other routes). Overall for a general complete network it makes sense to cut the ij link if:

$$c + \left(N^S - 2\right)\left(\beta^{N^S-3} - \beta^{N^S-2}\right) > \left(\delta - \delta^2\right) + \left(\beta^{N^S-2} - \beta^{2N^S-6}\right)$$

Which gives us a stability condition

$$\text{stable} \Rightarrow \left(\delta-\delta^2\right)-c+\left(\beta^{N^S-2}-\beta^{2N^S-6}\right)-\left(N^S-2\right)\left(\beta^{N^S-3}-\beta^{N^S-2}\right)>0 \quad (10)$$

5.3 Overall Stability

Since adding a link is not sensible, the conditions relating to removing a link are both necessary and sufficient to establish individual stability and hence we can replace condition 10 with the necessary and sufficient condition:

$$\text{stable} \Leftrightarrow \left(\delta-\delta^2\right)-c+\left(\beta^{N^S-2}-\beta^{2N^S-6}\right)-\left(N^S-2\right)\left(\beta^{N^S-3}-\beta^{N^S-2}\right)>0 \quad (11)$$

Of course, $\delta-\delta^2>0$ and $c>0$. Clearly, since $\beta\in(0,1)$ it must be the case that $\beta^{N^S-3}>\beta^{N^S-2}$, for all $N^S>2$. Finally, $\beta^{N^S-2}=\beta^{2N^S-6}$ for $N^S=4$, and $\beta^{N^S-2}>\beta^{2N^S-6}$ for all $N^S>4$. By considering restrictions on β, δ and N^S we can find the set of parameter values for which expression 11 is met, and the complete network is stable.

A full examination of alternative parameter values is found in Appendix 2, and here we restrict ourselves to a summary. With no concern for complexity so $\beta=1$, network size is not relevant for stability, and extreme values of δ guarantee instability. Intermediate levels of connection speed moving towards $\delta=1/2$, can produce stability, but only if c is low. With extreme concern for complexity, so $\beta=0$, the complete network may be stable, again only for intermediate connection speeds. For intermediate levels of concern for complexity stability is possible once again only for intermediate levels of δ. It appears that regardless of level of concern for complexity, direct costs or network size, if connection speeds are very fast or very slow then stability is not possible.

6 Efficiency

While we are predominantly interested in the long-run shape of the web, and so especially concerned about *stability*, in this section we consider the most *efficient* network for a given set of network parameters $\{\delta,\beta,c\}$, using definition 5. The inclusion of complexity costs, and the relationship between complexity and the architecture of the network makes it harder to go into as much detail about efficiency as stability. In fact, we cannot produce general conditions without considering the specific network architecture. However, there are still various interesting observations which can be made. We begin by looking at networks with low costs, and then move on to examine networks with high costs both with and without complexity considerations.

6.1 Low Cost Links

Under the conditions for efficiency of the network, it is clear that if the overall cost of building a new link are lower than the gain to be made, we would expect more links to be added. In the limit a complete network is the best that could be achieved if this were the case, and there is clearly no point adding duplicate links. In particular, if their exists no link ij in g, there are no complexity concerns, $\beta=1$, and $c<\delta-\delta^2$ then a new link between i and j will be added, and this will continue until the complete network is formed.

With complexity costs things get more complex since a direct link ij may well reduce the complexity of a journey to j from i and *vice versa*, but will raise the complexity of a journey from i and j to all other nodes. Consider the network g_7 when $ij \notin g_7$, and the alternative $g_8 = g_7 + ij$. Our condition which takes into account complexity costs is therefore:

$$\text{construct } ij \Leftrightarrow c + \sum_{k:ik \in g_7, k \neq j} \left[\left(1 - \beta^{h_{ik}(g_8)}\right) - \left(1 - \beta^{h_{ik}(g_7)}\right) \right] < \left(\delta - \delta^2\right) + \left(1 - \beta^{h_{ij}(g_7)}\right) - \left(1 - \beta^{h_{ij}(g_8)}\right)$$

Note that $h_{ik}(g_8) = h_{ik}(g_7) + 1$, in the summation. We can simplify to:

$$\text{construct } ij \Leftrightarrow c + \sum_{k:ik \in g_7, k \neq j} \left(\beta^{h_{ik}(g_7)} - \beta^{h_{ik}(g_7)+1}\right) < \left(\delta - \delta^2\right) + \left(\beta^{h_{ij}(g_8)} - \beta^{h_{ij}(g_7)}\right)$$

Now we need to consider this condition for the shift from g_7 to g_8, but also the updated condition for any further change from g_8 to g_9, etc., yielding the general condition:

$$\text{construct } ij \Leftrightarrow c + \sum_{k:ik \in g_n, k \neq j} \left(\beta^{h_{ik}(g_n)} - \beta^{h_{ik}(g_n)+1}\right) < \left(\delta - \delta^2\right) + \left(\beta^{h_{ij}(g_{n+1})} - \beta^{h_{ij}(g_n)}\right)$$

Should the general condition continue to be satisfied as n rises until the complete graph is formed, then the uniquely efficient network must be the complete graph, as links will continue to be added until there are no more non-duplicate links remaining.

6.2 High Cost Links

One the other hand if the costs of adding new links are prohibitively high then we may find that a complete graph is not the most efficient network, indeed it may be that $\upsilon(g) < 0$ for all $N^S \geq 2$, and hence there is little sense anyone linking to anyone else. We could imagine that this might be the case if c is sufficiently high, or if complexity costs are sufficiently high.

6.2.1 Without Complexity

Let us once again start by considering $\beta = 1$, i.e. no complexity concerns, which actually gives us the situation as described in Jackson and Wolinsky (1996), and so the unique efficient network is a star encompassing everyone if $\delta - \delta^2 < c < \delta + 0.5(N^S - 2)\delta^2$ and there will be no links if $c > \delta + 0.5(N^S - 2)\delta^2$. The proof of this is a straightforward slight modification of Jackson and Wolinsky's proof of their Proposition 1 (pages 49-50). First consider a component of g containing m individuals. Let $k \geq m - 1$ be the number of links in this component. The value of these direct links is $k(2\delta - 2c)$. This leaves at most $m(m-1)/2 - k$ indirect links. The value of each indirect link is at most $2\delta^2$. Therefore, the overall value of the component is at most

$$\upsilon_c \equiv k(2\delta - 2c) + (m(m-1) - 2k)\delta^2$$

If this component is a star then its value would be

$$\upsilon_c^s \equiv (m-1)(2\delta - 2c) + (m-1)(m-2)\delta^2$$

Since $k \geq m-1$ and $c > \delta - \delta^2$ the difference between these two expressions yields

$$(k-(m-1))(2\delta - 2c - 2\delta^2) \leq 0 \tag{12}$$

We can replace the weak inequality in expression 12 with a strict inequality if $k > m-1$. Note that the value of this component can equal the value of the star only when $k = m-1$. Any graph with $k = m-1$, which is not a star, must have an indirect connection which has a path longer than 2, getting value less than $2\delta^2$. Therefore, the value of the indirect links will be below $(m-1)(m-2)\delta^2$, which is what we get with star. We have shown that if $c > \delta - \delta^2$, then any component of an efficient graph must be a star. Note that any component of an efficient graph must have non-negative value. In that case, a direct calculation of v_c^s shows that a single star of $m+n$ individuals is greater in value than separate stars of m and n individuals. Thus if the efficient graph is non-empty, it must consist of a single star. Again, it follows from the value of v_c^s that if a star of n individuals has non-negative value, then a star of $n+1$ individuals has higher value.

Finally, we should note that a star encompassing everyone has positive value only when $\delta + ((N^S - 2)/2)\delta^2 > c$, so if this condition fails then we can expect no links at all.

6.2.2 With Complexity

Now if $\beta \in (0,1)$, so complexity matters we need to modify this result. Once again, we need to consider not only the impact on complexity along the ij route if a new link ij is established, but also the impact for the entire set of nodes linked to both i and j. Unfortunately since the complexity cost is not the same across nodes we cannot easily show the superiority of the star encompassing everyone. For example, in any component of g containing m individuals, with $k \geq m-1$ as the number of links in this component, the value of these direct links, varying by node is $2k(\delta - c - \sum_{j:ij \in g}(1-\beta^{h_{ij}}))$. We know that the overall value of the component if a star is

$$\widehat{v}_c^s \equiv 2(m-1)(\delta - c - (1-\beta^{m-2})) - (m-1)(m-2)(\delta^2 - (1-\beta^{m-2}))$$

However, for a general component, the value is

$$\widehat{v}_c \equiv k(2\delta - 2c) + (m(m-1) - 2k)\delta^2 - \sum_{i=1}^{k} \sum_{j:ij \in g} (1-\beta^{h_{ij}})$$

This can only be simplified under specific assumptions about network architecture because of the nature of the complexity cost. We can note that it is no longer clear that a star encompassing everyone is efficient for intermediate cost levels since a star network involves relatively high complexity costs of

$$2(m-1)(1-\beta^{m-2}) + (m-1)(m-2)(1-\beta^{m-2})$$

Other network architectures such as the complete network might well involve lower complexity costs. To motivate these high complexity costs consider the route from $i \neq cs$ to $j \neq cs$. This passes through the cs and at that point requires a highly complex navigation, whereas a route which bypassed the star would necessarily avoid reaching an intermediate point containing $m-2$ irrelevant links. This mirrors our findings in the stability section when we found that for $\beta < 1$, the star network is not stable, since users always have an incentive to attempt to directly link to other nodes to avoid the complexity of going through the centre.

7 A Taxonomy of Stability Results

Here we summarize the stability and efficiency results of the last two sections and find a taxonomy of predicted network architectures, including the implications for the move towards broadband access for all Internet users. We also compare our findings with those of the unmodified Jackson and Wolinsky model.

Of main concern in this paper is the long-run stability of certain archetypal networks: the star network and the complete network, which have received considerable attention in the literature as possible long-run stable architectures.

In each case these propositions have already been proved in the main text of sections V and VI. For a Star Network we have:

Proposition 1 *For $\beta < 1$, a star network encompassing everyone is stable if and only if* $\left(N^S - 2\right)\delta^2 + \delta - c - \left(N^S - 2\right)\left(1 - \beta^{N^S-3}\right) \geq 0$ *and* $1 + c + N^S\beta^{N^S-2} + 3\beta^{N^S-1} \geq \delta - \delta^2 + 2\beta + 2\beta^{N^S-2} + N^S\beta^{N^S-1}$.

We can add some of the sharpest results for certain extreme parameter values of δ and network size for the star network, all proved in section V.

Proposition 2 *For $\beta = 0$, a star network is never stable if $\delta \to 0$, and always stable if $\delta \to 1$, regardless of network size. If $\beta = 0$ and $\delta \in (0,1)$ then the star network is not stable if $N^S \to \infty$, and may be stable for smaller network sizes.*

Proposition 3 *For $\beta = 1$, a star network is never stable if $\delta \to 0$, and always stable if $\delta \to 1$, regardless of network size. If $\beta = 1$ and $\delta \in (0,1)$ then the star network may be stable depending upon parameter values and will certainly be stable if $c \in \left(\delta - \delta^2, \delta + 2\delta^2\right)$.*

Proposition 4 *For $\beta \in (0,1)$, a star network is never stable if $\delta \to 0$. When $\delta \to 1$, the star is always stable for the extremes of network size, and for intermediate networks will certainly be stable if $1+c > 2\beta$, and may be unstable otherwise. If $\beta \in (0,1)$ and $\delta \in (0,1)$ a star network will not be stable if $N^S \to \infty$, and may be stable for lower N^S depending upon the values of δ, β and c.*

For a complete network the main result as proved in section VI, part C is:

Proposition 5 *For $\beta < 1$, a complete network is stable if and only if:*

$$\left(\delta - \delta^2\right) - c - \left(N^S - 2\right)\left(\beta^{N^S-3} - \beta^{N^S-2}\right) > 0.$$

We can add some results for certain extreme values of δ and network size for the complete network, all proved in section VI.

Proposition 6 *For $\beta = 0$ and $\beta = 1$ a complete network is not stable if $\delta \to 0$ or $\delta \to 1$, and for $\delta \in (0,1)$ is stable if and only if $c < \delta - \delta^2$.*

Proposition 7 *For $\beta \in (0, 1)$, a complete network is never stable if $\delta \to 0$ or $\delta \to 1$, and may be stable for intermediate values of δ if β and c are sufficiently low, in particular if $c > 1/4$ the complete network cannot be stable even for intermediate δ and low values of β.*

[insert TABLE 1 here]

Finally, extracting these results we have a taxonomy of stability results as described in Table 1. Among the most interesting features present in Table 1 are the fact that in many cases only one of the two standard structures looks likely to survive into the long-run for a given set of parameter values. For example, if users have an over-riding concern to keep complexity costs down, so $\beta = 0$, then at slow connection speeds neither network looks like a good candidate for long-term stability, and for intermediate connection speeds the star and complete networks may work for given other parameter values, with the star only a viable candidate if $\delta \to 1$. As the concern for complexity declines, the complete network is only feasible for $\beta = 1$, with an intermediate value of δ, and only if $c < \delta - \delta^2$, with higher values of c favoring the star network, until c becomes so great that neither looks to be stable. The best use for the table is if a vague idea of the likely parameter values can be estimated. For example, if we believe that the number of nodes in the system is likely to become extremely large (noting that there are already many billions of nodes, we might feel that $N^S \to \infty$, provides a good approximation). We might also believe that with the widespread adoption of broadband technology connection speeds will only get higher and we might expect $\delta > 0$, but with content becoming more complex we might not think $\delta \to 1$, is reasonable and so we opt for $\delta \in (0, 1)$, and finally we might expect some concern for complexity, but perhaps not over-riding concern, so set $\beta \in (0, 1)$. Glancing at the relevant element in the table this tells us that the complete network looks more likely to be stable than the star network, and so we can expect more and more complex network structures with more nodes inter-linked. On the other hand if we think broadband technology will enable almost instantaneous connection speeds, and complexity concerns will fall as users become more and more used to the complexities of web navigation, then we might prefer $\delta \to 1$ and $\beta = 1$, and so favor the star network over the complete network for all but trivially small networks.

8 Conclusions

We have seen that social network theory often using fairly restrictive assumptions can provide some insights into the stability and efficiency properties of networks, and the web is one such network that can be analyzed. The seminal paper by Jackson and Wolinsky (1996) is a useful starting point for a discussion of the long-term shape of the world-wide-web, though we need to add some complexity costs, and make predictions about the future size of the web, and the future speed of Internet connections. This paper makes such predictions and modifies their model to incorporate complexity costs and thereby generates a taxonomy of predictions.

To give some examples of how this taxonomy might be used, let us restate some of the results in more emotive language. Let us say a network where $N^S \to \infty$ is 'large' and small otherwise, a network where $\beta < 1$ involves concern for complexity, but not if $\beta = 1$, and a network where $\delta \to 1$ as having a fast connection (say, broadband), and otherwise slow (say, dialup). Now, reading from the table, we can say that as networks grow in size and connection speeds increase (say everyone has access to broadband, and broadband speeds continue to rise) social network theory under the assumptions detailed in this paper would predict that with concern for complexity a star is a more likely end result. With very low connection speeds neither the complete network nor the star seem plausible stable structures.

Making some general assumptions about future trends, as speeds seem likely to increase, especially with the widespread adoption of broadband, and with network sizes also increasing the only real uncertainty is how users deal with the increased complexity of the web, with the star network being far less sensitive to complexity costs than the complete network it may be our best guess for the future, which explains why complexity costs are an important and necessary addition to the Jackson and Wolinsky model if we wish to use social network theory to analyze the world wide web. Alternatively, if we believe that concern about complexity will stay high, as other costs fall, and speeds do not universally increase, perhaps the demand for broadband access falling off, then the complete network might well be able to survive into the long-run.

What this means for the web in practice requires that we return to examine what we mean by stars and complete networks. A star is given here as a model of the web based around many users connecting to each other's web-sites via a central clearing house, well represented by a single ISP or an all-encompassing search engine, with no need for any other hyper-linking between pages, or even use of favorites or book-marking. A complete network is very different, a much more decentralized structure with many large sites containing hyper-links to all pages of interest, we might imagine many users travelling via hyper-links on their own pages or choices from their favorites menus directly to where they want to go, with minimal use of search engines, or centralized sites. The implications of whether either of these archetypes become dominant are considerable given the role which major online presences such as ISPs, search engines and all other forms of centralized structures have to play. The more centralized and star-like the structure, the more power rests in the hands of fewer sites, while the more complete a network the less controlling influence such sites can wield over the majority of users. Perhaps the single most powerful prediction is that *as speeds increase with the greater adoption of broadband technology, the star network will be the more likely architecture to survive into the long-run, and so we might expect to see an increasingly centralized structure.* Whether this is seen as a problem from a social planner or government perspective, is left as a question for future research, but may well depend upon a subjective set of priorities. It should also be stressed that the model used in this paper has several limitations which need to be understood before accepting the final results and considering policy implications, and the final part of section III provides a list of such difficulties, together with some possible extensions.

References

Badasyan, N. and Chakrabarti, S. (2003). 'Private Peering among Internet Backbone Providers', Mimeo: Virginia Tech.

Bloch, F. (2003). 'Group and Network Formation in Industrial Organisation: A Survey', Mimeo: GREQAM.

Goyal, S. and van der Leij, M. (2004). 'Economics: An emerging small world', Mimeo: Univeristy of Essex.

Jackson, M. O. (2003). 'A Survey of Models of Network Formation: Stability and Efficiency', forthcoming in Group Formation in Economics: Networks, Clubs, and Coalitions, edited by G. Demange and

M. Wooders, Cambridge University Press: Cambridge.

Jackson, M. O. and Wolinsky, A. (1996). 'A Strategic Model of Social and Economic Networks', *Journal of Economic Theory* 71, pp. 44-74.

Appendix 1: The Star Network

Here we present the full impact of different assumed parameter values on the stability of the star network.

Part A: No Concern for Complexity

We will start with the model with no complexity concerns, so $\beta = 1$. Now condition 9 becomes

$$\text{stable }|_{\beta=1} \Leftrightarrow 0 < \min\left\{N^S\delta^2 - 2\delta^2 + \delta - c, c - \left(\delta - \delta^2\right)\right\} \tag{13}$$

Now we will consider condition 13 for some variations of the other parameter values.

(i) $\delta \to 0$. When connection speeds are very slow, condition 13 becomes

$$\text{stable }|_{\beta=1,\delta\to 0,N^S=4} \Leftrightarrow 0 < \min\{-c, c\} \Leftrightarrow c < 0$$

Which cannot hold, regardless of network size, so the network is definitely not stable. This is immediate given the nature of the star network: with speeds tending towards zero, a user will wish to build a direct link to j since it is simply to slow to travel via the central node.

(ii) $\delta \in (0, 1)$. Next we consider an intermediate level of speed, firstly, combined with a small network, which produces the stability condition

$$\text{stable }|_{\beta=1,\delta\in(0,1),N^S=4} \Leftrightarrow 0 < \min\left\{2\delta^2 + \delta - c, c - \delta + \delta^2\right\}$$

Note that we therefore need $c < \delta + 2\delta^2$ and $c > \delta - \delta^2$, so we will have stability only if $c \in \left(\delta - \delta^2, \delta + 2\delta^2\right)$. This is feasible, for example when $\delta = 1/2$ we have the maximal range of $c \in (1/4, 1)$. Though for other values of δ the range is more prohibitive. If we consider both intermediate speed levels and an intermediate network size we of course simply have condition 13, but ruling out extremes

$$\text{stable }|_{\beta=1,\delta\in(0,1),N^S\in(4,\infty)} \Leftrightarrow 0 < \min\left\{N^S\delta^2 - 2\delta^2 + \delta - c, c - \left(\delta - \delta^2\right)\right\} \tag{14}$$

Note that since network size is likely to be considerable, we might find the results for $N^S \to \infty$, to be most compelling, which gives us

$$\text{stable }|_{\beta=1,\delta\in(0,1),N^S\to\infty} \Leftrightarrow 0 < c - \left(\delta - \delta^2\right)$$

Therefore we need $c > \delta - \delta^2$ for stability. However, even for lower network sizes the second half of condition 14 is likely to be the binding constraint, and so we are likely to want $c > \delta - \delta^2$ for stability. In effect then, for intermediate network sizes the star network will be stable, if direct costs are high relative to the savings to be gained from building a direct route from i to j.

(iii) $\delta \to 1$. With instantaneous connection speeds, and the smallest non-trivial network, condition 13 becomes

$$\text{stable }|_{\beta=1,\delta\to 1,N^S=4} \Leftrightarrow 0 < \min\{1 - c, c\}$$

Again, with immediate connection, and an intermediate network condition 13 becomes

$$\text{stable}\mid_{\beta=1,\delta\to1,N^S\in(4,\infty)}\Leftrightarrow 0<\min\left\{N^S-1-c,c\right\}$$

Finally, with a combination of immediate connection times and an extremely large network condition 13 becomes

$$\text{stable}\mid_{\beta=1,\delta\to1,N^S\to\infty}\Leftrightarrow 0<c$$

All three conditions are trivially satisfied since $c\in(0,1)$, so the network must be stable when $\delta\to1$, if there is no concern for complexity. Put simply, with speeds tending towards instantaneous connection, there is very little to be gained from building a direct route from i to j as going via the central node is so quick, but there is a direct cost c to be incurred. Hence the star network is stable when speeds are extremely high.

Part B: Some Concern for Complexity

When there are some complexity concerns, we wish to consider $\beta\in(0,1)$, and more especially values of β which are not too close to either extreme. Our basic stability condition is not refined any further here, so we start with condition 9, and will refine it further as we add additional constraints.

(i) $\delta\to0,N^S=4$. When connection speeds tend towards zero and the network is extremely small condition 9 becomes

$$\text{stable}\mid_{\beta\in(0,1),\delta\to0,N^S=4}\Leftrightarrow 0<\min\{\beta-c-2),1+c-2\beta+2\beta^2-\beta^3\}$$

Clearly since $\beta<1$ and $c<1$, then $\beta-c-2<0$, and so the network is not stable.

(ii) $\delta\to0,N^S\in(4,\infty)$. Once again connection speeds tend towards zero, but now we consider intermediate network size. Condition 9 becomes

$$\begin{aligned}\text{stable}_{\beta\in(0,1),\delta\to0,N^S\in(4,\infty)}\quad\Leftrightarrow\quad &0<\min\{-c-\left(N^S-2\right)(1-\beta^{N^S-3}),\\ &1+c-2\beta+\left(N^S-2\right)\beta^{N^S-2}-\left(N^S-3\right)\beta^{N^S-1}\}\end{aligned}$$

Now consider the first half of the condition. We need $-c-\left(N^S-2\right)(1-\beta^{N^S-3})>0$. However for $N^S>4$ this will be negative, and so the network is not stable.

(iii) $\delta\to0,N^S\to\infty$. Now we consider $N^S\to\infty$, with $\delta\to0$ and therefore we have the stability condition

$$\text{stable}\mid_{\beta\in(0,1),\delta\to0,N^S\to\infty}\Leftrightarrow 0<\min\lim_{N^S\to\infty}\{-\left(N^S-2\right)(1-\delta^2)+\delta-c,1+c-2\beta-\delta+\delta^2\}$$

Which clearly achieves negative infinity in the first half of the condition and so results in a network which is not stable.

(iv) $\delta\in(0,1),N^S=4$. Next we consider an intermediate level of speed, combined with a small network, which produces the stability condition

$$\text{stable}\mid_{\beta\in(0,1),\delta\in(0,1),N^S=4}\Leftrightarrow 0<\min\{2\delta^2+\delta+2\beta-c-2,1+c-2\beta+2\beta^2-\beta^3-\delta+\delta^2\}$$

A stable network looks unlikely here, but it is possible. For example if $c>1/4\geq\left(\delta-\delta^2\right)$, then the second half of the condition must be satisfied, and with a high enough value of δ and β the first condition is met, so we have a stable network.

(v) $\delta \in (0,1), N^S \in (4,\infty)$. If we consider both intermediate speed levels and an intermediate network size we of course simply have condition 9. Stability is possible or not, depending upon the parameter values. For example, as just seen in (iv) with a low enough value of N^S, and various other conditions in the other parameters, stability is possible.

(vi) $\delta \in (0,1), N^S \to \infty$. With intermediate speed and a very large network size condition 9 becomes

$$\text{stable} \,|_{\beta\in(0,1),\delta\in(0,1),N^S\to\infty} \Leftrightarrow 0 < \min \lim_{N^S\to\infty} \{-\left(N^S-2\right)\left(1-\delta^2\right)+\delta-c, 1+c-2\beta-\delta+\delta^2\}$$

Now since the first half of the condition tends towards negative infinity, the network is not stable.

(vii) $\delta \to 1, N^S = 4$. With virtually instantaneous connection speeds, and the smallest non-trivial network, condition 9 becomes

$$\text{stable} \,|_{\beta\in(0,1),\delta\to1,N^S=4} \Leftrightarrow 0 < \min\{1-c+2\beta, 1+c-2\beta+2\beta^2-\beta^3\}$$

Since $c > 0$, it must be that $1-c+2\beta > 0$, so the condition becomes

$$\text{stable} \,|_{\beta\in(0,1),\delta\to1,N^S=4} \Leftrightarrow 0 < 1+c-2\beta+2\beta^2-\beta^3$$

Since $\beta < 1$, we know that $\beta^2 > \beta^3$ and furthermore, by Lemma 1, we know that $1+\beta^2 > 2\beta$, and therefore it must be that $1+2\beta^2 > 2\beta+\beta^3 \Rightarrow 1+c-2\beta+2\beta^2-\beta^3 > 0$. Therefore the network is stable.

(viii) $\delta \to 1, N^S \in (4,\infty)$. With virtually instantaneous connection speeds, and an intermediate network condition 9 becomes

$$\begin{aligned}\text{stable} \,|_{\beta\in(0,1),\delta\to1,N^S\in(4,\infty)} &\Leftrightarrow 0 < \min\{1-c+\left(N^S-2\right)\beta^{N^S-3},\\ &1+c-2\beta+\left(N^S-2\right)\beta^{N^S-2}-\left(N^S-3\right)\beta^{N^S-1}\}\end{aligned}$$

The first term, $1-c+\left(N^S-2\right)\beta^{N^S-3}$, must be positive, so we need only consider the second

$$\text{stable} \,|_{\beta\in(0,1),\delta\to1,N^S\in(4,\infty)} \Leftrightarrow 0 < 1+c-2\beta+\left(N^S-2\right)\beta^{N^S-2}-\left(N^S-3\right)\beta^{N^S-1}$$

For low values of N^S, the result in (vii) shows us that we will see a stable network, but for larger values of N^S, we cannot be certain. Since $\left(N^S-2\right)\beta^{N^S-2} > \left(N^S-3\right)\beta^{N^S-1}$, we do know that $1+c > 2\beta$ provides a sufficient condition for stability regardless of the value of N^S. So if $\beta < 1/2$ we will certainly have a stable network. This reveals that a low value of β or a high value of c can ensure stability.

(ix) $\delta \to 1, N^S \to \infty$. Finally, with the combination of virtually instantaneous connection speeds and an extremely large network condition 9 becomes

$$\text{stable} \,|_{\beta\in(0,1),\delta\to1,N^S\to\infty} \Leftrightarrow 0 < \min\{1-c, 1+c-2\beta\}$$

Since $1-c > 0$, we have

$$\text{stable} \,|_{\beta\in(0,1),\delta\to1,N^S\to\infty} \Leftrightarrow 0 < 1+c-2\beta$$

So as in (viii) $1+c>2\beta$ ensures stability, though here it is a necessary condition as well as being sufficient.

Part C: Extreme Concern for Complexity

Now we consider extreme concern for complexity, which we will model by setting $\beta=0$, in condition 9 which gives us

$$\text{stable} \Leftrightarrow 0<\min\{\delta-c-\left(N^S-2\right)\left(1-\delta^2\right),(1+c)-\left(\delta-\delta^2\right)\}$$

However, since $\left(\delta-\delta^2\right)<1$, we know that $(1+c)>\left(\delta-\delta^2\right)$, and so the condition becomes, simply

$$\text{stable} \Leftrightarrow 0<\delta-c-\left(N^S-2\right)\left(1-\delta^2\right) \tag{15}$$

Note that this will be a tough condition to satisfy, failing for high values of N^S and c.

(i) $\delta\to 0, N^S=4$. When connection speeds tend towards zero and the network is extremely small condition 15 becomes

$$\text{stable}\mid_{\beta=0,\delta\to 0,N^S=4} \Leftrightarrow 0<-c-2$$

Since $c>0$, this condition fails and the network is not stable.

(ii) $\delta\to 0, N^S\in(4,\infty)$. Once again connection speeds tend towards zero, but now we consider intermediate network size. The condition 15 becomes

$$\text{stable}\mid_{\beta=0,\delta\to 0,N^S\in(4,\infty)} \Leftrightarrow 0<-c-\left(N^S-2\right)$$

With $c>0$ and $N^S>4$, this condition will fail and the network cannot be stable.

(iii) $\delta\to 0, N^S\to\infty$. Now we consider $N^S\to\infty$, with $\delta\to 0$ and therefore we have same stability condition

$$\text{stable}\mid_{\beta=0,\delta\to 0,N^S\to\infty} \Leftrightarrow 0<\lim_{N^S\to\infty}\left\{-c-\left(N^S-2\right)\right\}$$

The condition fails and so the network is not stable..

(iv) $\delta\in(0,1), N^S=4$. Next we consider an intermediate level of speed, combined with a small network, which produces the stability condition

$$\text{stable}\mid_{\beta=0,\delta\in(0,1),N^S=4} \Leftrightarrow 0<0<\delta+2\delta^2-c-2$$

This condition can be met, for sufficiently high values of δ and low enough c, though it seems unlikely. For example, stability is not possible if $\delta<\frac{1}{4}\sqrt{8c+17}-\frac{1}{4}$. Even with $c=0$, we still need a value of δ of almost 0.79 or higher, and with a positive c the condition becomes even more difficult. However, it is worth noting that stability is only effectively ruled out if $c=1$, as for any lower value of c there does exist a value of δ high enough to enable stability.

(v) $\delta\in(0,1), N^S\in(4,\infty)$. If we consider both intermediate speed levels and an intermediate network size we of course simply have condition 15 and so for stability we need $N^S\delta^2+2+\delta>N^S+c+2\delta^2$. We have just seen in (iv) that for the very lowest values of N^S this is a difficult condition to achieve and becomes more difficult as N^S rises.

(vi) $\delta\in(0,1), N^S\to\infty$. With intermediate speed and a very large network size condition 15 becomes

$$\text{stable}\,|_{\beta=0,\delta\in(0,1),N^S\to\infty}\Leftrightarrow 0<\lim_{N^S\to\infty}\left\{\delta-c-\left(N^S-2\right)\left(1-\delta^2\right)\right\}$$

So the condition fails and the network is not stable.

(vii) $\delta \to 1, N^S = 4$. With virtually instantaneous connection speeds, and the smallest non-trivial network, condition 15 becomes

$$\text{stable}\,|_{\beta=0,\delta\to 1,N^S=4}\Leftrightarrow 0<1-c$$

The condition is met since $c < 1$ and so the network is stable.

(viii) $\delta \to 1, N^S \in (4,\infty)$. With virtually instantaneous connection speeds, and an intermediate network size, condition 15 is $1 - c > 0$, as in (vii), and so once again the network is stable.

(ix) $\delta \to 1, N^S \to \infty$. Finally, with virtually instantaneous connection speeds and an extremely large network size the stability condition remains $0 < 1 - c > 0$ as in (vii) and the network remains stable.

Appendix 2: The Complete Network

Here we present the full impact of different assumed parameter values on the stability of the complete network.

Part A: No Concern for Complexity

We will start with the model with no complexity concerns, so $\beta = 1$. Now condition 11 becomes

$$\text{stable}\,|_{\beta=1}\Leftrightarrow \delta-\delta^2>c \tag{16}$$

Note that this is not dependent upon N^S, so we will only consider condition 16 for variations of δ.

(i) $\delta \to 0$. When connection speeds tend towards zero condition 16 becomes

$$\text{stable}\,|_{\beta=1,\delta\to 0}\Leftrightarrow 0>c$$

This is not the case, and hence the network is not stable under extremely slow connection speeds. Simply put, connection speeds are so low that being connected to another link is too costly.

(ii) $\delta \in (0, 1)$. Next we consider an intermediate level of speed, combined with a small network, which produces the stability condition

$$\text{stable}\,|_{\beta=1,\delta\in(0,1)}\Leftrightarrow \delta-\delta^2>c$$

Since $\inf_\delta \left(\delta - \delta^2\right) = 0$, for $\delta \in (0, 1)$, we know that as δ approaches the extreme values of 0 and 1 there will be no level of cost sufficiently small to enable the network to be stable. However, $\max_\delta \left(\delta - \delta^2\right) = 1/4$, when $\delta = 1/2$, so for values of δ approaching 1/2, there are levels of cost sufficiently low to enable the network to be stable.

(iii) $\delta \to 1$. With virtually instantaneous connection speeds, condition 16 becomes

$$\text{stable}\,|_{\beta=1,\delta\to 1}\Leftrightarrow 0>c$$

Since $c > 0$, this condition cannot hold, and the network is not stable. Here the speed is so high that lengthening the path from i to j makes no effective difference, but since cutting the direct link saves the cost c, it is worthwhile.

Part B: Some Concern for Complexity

When there are some complexity concerns, we wish to consider $\beta \in (0, 1)$, and more especially values of β which are not too close to either extreme. Our basic stability condition is not refined any further here, so we start with condition 11, and will refine it further as we add additional constraints.

(i) $\delta \to 0, N^S = 4$. When connection speeds tend towards zero and the network is extremely small (say $N^S = 4$) condition 11 becomes

$$\text{stable} \mid_{\beta\in(0,1),\delta\to 0,N^S=4} \Leftrightarrow -c - 2\left(\beta - \beta^2\right) > 0$$

This will fail as $\beta \in (0, 1)$ and $c > 0$. Considering larger networks, results in a condition

$$\text{stable} \mid_{\beta\in(0,1),\delta\to 0,N^S\in(4,\infty)} \Leftrightarrow -c + \left(\beta^{N^S-2} - \beta^{2N^S-6}\right) - \left(N^S - 2\right)\left(\beta^{N^S-3} - \beta^{N^S-2}\right) > 0$$

Which again, fails. Finally, when considering $N^S \to \infty$, with $\delta \to 0$, failure is even more clear, since stability requires

$$\text{stable} \mid_{\beta\in(0,1),\delta\to 0,N^S\to\infty} \Leftrightarrow -c > 0$$

Therefore with some concern for complexity, and a network where connection speeds tend towards zero, the complete network will not be stable, irrespective of network size. To understand this we have to consider the great advantage of the complete network: every node is close to every other. This becomes irrelevant when connection speeds are so slow that even a single link is prohibitively slow to navigate.

(ii) $\delta \in (0, 1), N^S = 4$. Next we consider an intermediate level of speed, combined with a small network, which produces the stability condition

$$\text{stable} \mid_{\beta\in(0,1),\delta\in(0,1),N^S=4} \Leftrightarrow \left(\delta - \delta^2\right) - c - 2\left(\beta - \beta^2\right) > 0$$

A stable network is possible for low enough values of β and c, if δ is sufficiently close to the middle of the range. The best case scenario for stability is where $\delta = 1/2$, where the stability condition is

$$\text{stable} \mid_{\beta\in(0,1),\delta=1/2,N^S=4} \Leftrightarrow 1/4 > c + 2\left(\beta - \beta^2\right)$$

If we consider both intermediate speed levels, so $\delta \in (0, 1)$, and an intermediate network size, so $N^S \in (4, \infty)$ we of course simply have condition 11. Stability is possible or not, depending upon the parameter values. For example, as just seen for $N^s = 4$, with a low enough value of N^S, β and c stability is possible. As seen in (i), and (iii) if $\delta \to 0$ or $\delta \to 1$ then stability will fail, so much relies on a connection speed which is not too fast or too slow. If we consider intermediate speed levels, so $\delta \in (0, 1)$, but consider an extremely large network, so $N^S \to \infty$, then condition 11 becomes

$$\text{stable} \mid_{\beta\in(0,1),\delta\in(0,1),N^S\to\infty} \Leftrightarrow \left(\delta - \delta^2\right) - c > 0$$

Which is met if c is sufficiently low and δ is not too close to 0 or 1. Hence, with some concern for complexity and intermediate speeds, the complete network may be stable, but much depends upon the parameter values.

(iii) $\delta \to 1, N^S = 4$. Note that since the only expression in δ is $\delta - \delta^2$, the case where $\delta \to 1$, is exactly as in the case where $\delta \to 0$. In particular, when $\delta \to 1$ and we consider the smallest non-trivial network $N^S = 4$ condition 11 becomes

$$\text{stable } |_{\beta \in (0,1), \delta \to 1, N^S = 4} \Leftrightarrow -c - 2\left(\beta - \beta^2\right) > 0$$

As in case (i) stability fails. For intermediate network sizes, we have

$$\text{stable } |_{\beta \in (0,1), \delta \to 1, N^S \in (4,\infty)} \Leftrightarrow -c + \left(\beta^{N^S-2} - \beta^{2N^S-6}\right) - \left(N^S - 2\right)\left(\beta^{N^S-3} - \beta^{N^S-2}\right) > 0$$

So, once again, as in case (i) stability will fail for values of $N^S \in (4, \infty)$. Finally, when $N^S \to \infty$, condition 11 becomes

$$\text{stable } |_{\beta \in (0,1), \delta \to 1, N^S \to \infty} \Leftrightarrow -c > 0$$

So, as in case (i) stability fails. Irrespective of the size of the network, if there is some concern for complexity, then stability will always fail when connection speeds are extremely high. This is perfectly reasonable given that the major advantage of the complete network is the proximity of each node to each other, which becomes less and less important as speeds increase.

Part C: Extreme Concern for Complexity

Now we consider extreme concern for complexity, which we will model by setting $\beta = 0$, in condition 11 to yield a new stability condition

$$\text{stable } |_{\beta = 0} \Leftrightarrow \left(\delta - \delta^2\right) - c > 0 \tag{17}$$

Note that this is exactly the same as condition 16. This condition is most likely to be met where δ is not too close to either 0 or 1, in particular if $\delta = 1/2$, then stability is achieved if $c < 1/4$. For $\delta \to 0$ or $\delta \to 1$, stability is impossible for all $c > 0$, and if $c > 1/4$ then stability is impossible for any $\delta \in (0, 1)$. Note also that N^S is irrelevant when there is extreme concern for complexity. Intuitively we might think this condition is so very difficult simply because for very high costs there is a temptation to exit the network, and for low costs to build new links, therefore we need a very particular combination of low costs and intermediate speed levels, to make it worth being a part of the network but not quite worth building a new link.

			$N^S=4$			$N^S\in(4,\infty)$			$N^S\to\infty$	
		$\delta\to0$	$\delta\in(0,1)$	$\delta\to1$	$\delta\to0$	$\delta\in(0,1)$	$\delta\to1$	$\delta\to0$	$\delta\in(0,1)$	$\delta\to1$
$\beta=0$	Star	No	Maybe	Yes	No	Maybe	Yes	No	No	Yes
		Prop4	Prop3	Prop4	Prop4	Prop3	Prop4	Prop4	Prop4	Prop4
	Complete	No	Maybe	No	No	Maybe	No	No	Maybe	No
		Prop8	$c<\delta-\delta^2 \Leftrightarrow$Yes	Prop8	Prop8	$c<\delta-\delta^2 \Leftrightarrow$Yes	Prop8	Prop8	$c<\delta-\delta^2 \Leftrightarrow$Yes	Prop8
$\beta\in(0,1)$	Star	No	Maybe	Yes	No	Maybe	Maybe	No	No	Yes
		Prop6	Prop3	Prop6	Prop6	Prop3	Prop6	Prop6	Prop6	Prop6
	Complete	No	Maybe	No	No	Maybe	No	No	Maybe	No
		Prop9	Prop7	Prop9	Prop9	Prop7	Prop9	Prop9	Prop7	Prop9
$\beta=1$	Star	No	Maybe	Yes	No	Maybe	Yes	No	Maybe	Yes
		Prop5	$c\in(\delta-\delta^2,\delta+\delta^2) \Rightarrow$Yes	Prop5	Prop5	$c\in(\delta-\delta^2,\delta+\delta^2) \Rightarrow$Yes	Prop5	Prop5	$c\in(\delta-\delta^2,\delta+\delta^2) \Rightarrow$Yes	Prop5
	Complete	No	Maybe	No	No	Maybe	No	No	Maybe	No
		Prop8	$c<\delta-\delta^2 \Leftrightarrow$Yes	Prop8	Prop8	$c<\delta-\delta^2 \Leftrightarrow$Yes	Prop8	Prop8	$c<\delta-\delta^2 \Leftrightarrow$Yes	Prop8

Table 1: A Taxonomy of Stability Results

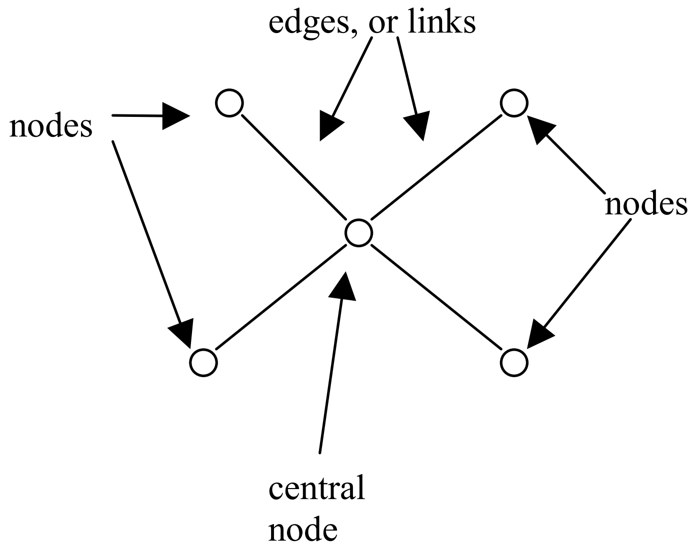

Figure 1: Example Network

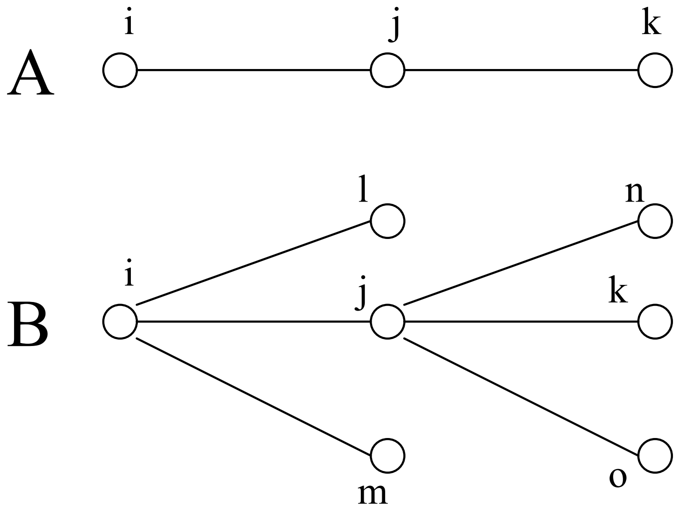

Figure 2: Complexity

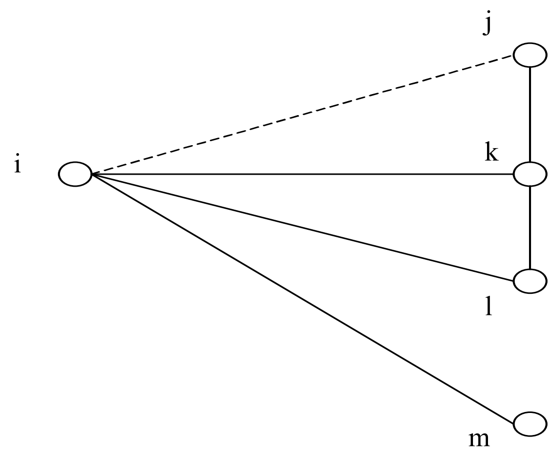

Figure 3: Stability in an Example Network

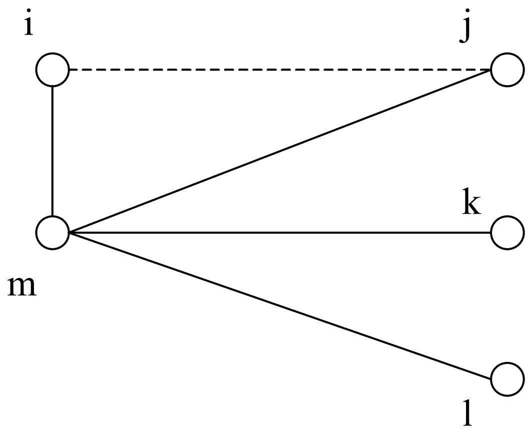

Figure 4: Stability in the Star Network

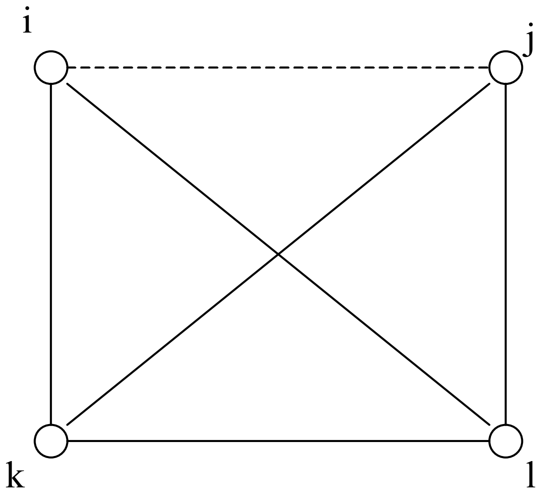

Figure 5: Stability in the Complete Network