

D'Ignazio, Alessio; Giovannetti, Emanuele

Conference Paper

Clustering, connectivity and hierarchies in the internet global supply chain networks

23rd European Regional Conference of the International Telecommunications Society (ITS),
Vienna, Austria, 1st-4th July, 2012

Provided in Cooperation with:

International Telecommunications Society (ITS)

Suggested Citation: D'Ignazio, Alessio; Giovannetti, Emanuele (2012) : Clustering, connectivity and hierarchies in the internet global supply chain networks, 23rd European Regional Conference of the International Telecommunications Society (ITS), Vienna, Austria, 1st-4th July, 2012, International Telecommunications Society (ITS), Calgary

This Version is available at:

<https://hdl.handle.net/10419/60372>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Clustering, Connectivity and Hierarchies in the Internet Global Supply Chain Networks

Preliminary Conference Version

Alessio D'Ignazio, Banca D'Italia

Emanuele Giovannetti¹, IIMP- Anglia Ruskin University, Cambridge and
DSE -University of Verona

Abstract

Internet Service Providers compete for customers while exchanging traffic flows to provide a complete, end to end, service to final users. This requires reliable interconnections among competitors that form multiple *Global Supply Chain Networks* (GSCNs) for the delivery and exchange of Internet traffic.

Interconnection decisions form the architecture of the Internet supply chain as they design the rules of the game played by the operators, in terms of reciprocal access pricing and quality and modalities of traffic exchanged. From a provider's point of view, the strategic assessment of its direct interconnection environment is crucial in defining the competitive and complementary elements of its extended GSCN. This paper focuses on the relationship between a provider's connectivity and the degree of mutual connectivity among the operators this provider is connected to. The strategic relevance of this relationship is clearly explained as follows: the better connected a provider is, the easier it is to deliver its traffic with high *Quality of Service* and low costs, while the less interconnected among themselves a provider's neighbours are, the easier it is, for the provider, to exert its bargaining power over them. This bargaining power, of a well-connected provider over its poorly connected network-neighbours, shows when contracting over quality standards, access pricing and interconnection terms. This paper estimates two separate econometric models showing that the connectivity features of the GSCN display significant differences in network hierarchy and complexity depending on whether they are observed from a European, North American or Rest of the World observation point.

Key words: Internet; Complexity; Global Supply Chain Networks; Clustering; Connectivity

1 Introduction

Worldwide Internet Service Providers compete for customers while exchanging traffic flows to provide a complete, end to end, service to final users. This requires reliable interconnections among competitors that form multiple *Global Supply Chain Networks* (GSCNs) for the delivery and exchange of Internet traffic. These GSCNs are often characterised by conflicts of interest as their independent components compete, at different network levels, and face complex problems of revenue allocation mechanisms across the value chain.

¹Corresponding Author. We would like to thank Amogh Dhamdhere (CAIDA) and Constantine Dovrolis, (Georgia Tech) for pointing us to the new data used in the econometric estimation of the clustering Model. The responsibility for any mistake remains of course our own.

The set of existing interconnections among ISPs, forming the primary links of the supply chain, represents the outcome of their commercial policies as they are the physical expression of hundreds of specific managerial decisions of whether to interconnect with another ISP, that can be, at the same time, a competitor or a complementor.

These decisions on interconnection represent the key strategic steps defining the architecture of the global Internet supply chain and in doing so they design the rules of the game played by Internet operators, in terms of reciprocal access pricing and/or quantity and modalities of traffic exchanged. The key implication is that interconnection choices, while essential from a technological point of view, are also primarily of a strategic nature, as their outcomes define the scope and space where competition for the process of global delivery of electronic information through the Internet takes place.

An understanding of the hierarchies and bargaining power within the GSCNs resulting from the bilateral interconnection choices is of immediate relevance both for competition policy and for assessing the profitability of the commercial strategies of the ISPs.

From a policy perspective, the analysis of the market structure is a precondition for the assessment of potential anticompetitive effects of proposed mergers among operators, for the assessment of the risk of abuse of dominant position practices, such as discrimination in interconnection policies, and for the detection of collusive behaviour, always a concrete risk when of cooperation among competitors is required, as it happens along a GSCN.

The impact of potential anticompetitive effects affects at least three levels of economic activities relevant for public policy:

- The price that final consumers pay for Internet services, and hence the speed and reach of Broadband penetration,
- The evolution of the industry structure and the associated global supply chain, towards a more concentrated morphology, and the implications that this process may bring in stifling innovation.
- The macroeconomic implications in terms of reduced economic growth due to a potential slowdown in the expansion process of the Broadband Infrastructure.

From a provider's point of view, the strategic assessment of its neighbouring environment, described by the set of interconnection policies, is crucial in defining the competitive and complementary elements of its supply chain network, and is vital for:

- A proper assessment of its cost structure, which is affected by the mix of free and the paid for interconnection agreements;
- The Quality of Service (QoS) implications due to the prevalence of direct or indirect interconnections agreements,
- The traffic and demand effects that different connectivity architectures may induce.

From both a researcher's and a business point of view, one of the main problems in understanding the strategic setting of the GSCNs for Internet connectivity lies in the lack of available market data to assess its structure. Interconnection data are often protected by confidentiality agreements as the market for interconnection does not involve product or services sold to final users. This paper addresses these problems by using an indirect inference approach based on data mining on interconnection policies derived from Internet topology metrics usually employed to assess Internet network technical reliability.

More in details, this paper focuses on the strategic relevance of the relationship between the connectivity of a provider and the level of mutual connectivity among the operators this provider is connected to. The relevance of this trade-off is clearly explained as the better connected a provider is, the easier it is to deliver its traffic achieving both high QoS and low costs, while the less interconnected among themselves a provider's neighbours are, the easier it is for this provider to exert its bargaining power over them and to appropriate most of the value otherwise generated along the entire supply chain network.

This bargaining power, of a well-connected provider over its poorly connected network-neighbours, is likely to manifest when contracting over quality standards, access pricing and terms of interconnection.

The main objectives of this paper are: (i) to review and analyse the evolution of the GSCN delivering Internet connectivity; (ii) to describe the appropriate complexity metrics that affect the functionality and features of this value chain (iii) to understand the role that Internet Exchange Points (IXPs) play into redefining the rules of interconnections among providers and (iv) to assess whether the hierarchical metrics and relationships characterising the GSCN show significant differences for IXPs located in Europe, North America or in the rest of the World.

The structure of the remainder of the paper is organised as follows: Section two discusses Connectivity in value chain networks. Section three describes the relevant connectivity metrics used in the analysis. Section four introduces the Empirical Analysis of the connectivity for the EuroIX Members. Section five introduces and compares two different econometric specifications capturing the hierarchical features of the Internet GSCNs. Section six concludes the paper indicating possible extensions of the present research.

2 Connectivity and Complexity in value chain networks

2.1 Connectivity as implementation of supply chain management

The Business models, considered in this paper, focus on firms' roles within a network of interrelations and hierarchies. The value of the traffic exchanged on the Internet is of a network's nature, as it is generated by the value added at the

different *hops* by many different network nodes. For this reason, interconnected ISPs provide a typical example of *Value Chain Networks* (VCNs): organizations the performance of which revolves around the modalities of partners selection (Talluri, Baker and Sarkis, (1999)).

More generally, *networked organizations*, are considered as something in-between anonymous *markets* and *internal organizations*, whose success is due to the combined shortcomings: of the *markets* in organising complementary resources and capabilities, and of *individual-organizations* to allow for the degree of flexibility required in highly dynamic global supply chains (Lo Nigro and Abbate (2011)). Hallikas et al. (2004) rationalise the benefits obtained from *networked organizations* in terms of both: reduced financial and technological risks and improved competitive advantage through specialization, while Link and Marxt, (2004) define as the *cooperation risks* necessarily associated to *networked organizations* the dependency of every *network company* on other network companies' resources.

This paper focuses on the usage of *complexity metrics* for the study of GSCNs capabilities derived from network topology models. A similar approach was proposed by Xiao, et al. (2012) for the analysis of the eco-industry supply chain. They describe each step of a complex supply chain as a node of a network and analysed their connectivity properties based on a modification of the original *Random Connections Model* of Barabasi Albert (1999). M'Chirgui and Penard (2011) apply similar complexity metrics for the study of Internet connectivity, emphasizing the analogies with both the airline and the credit card industries and focusing on the different Internet hierarchical levels, linking the emerging topological features to the supply chain management of alliances across operators.

The economic interpretation, and a taxonomy, for different Internet network structures was discussed in D'Ignazio and Giovannetti (2007). This paper goes further by considering the connectivity metrics at an IXP level of disaggregation. Clusters of connected ISPs at IXPs are seen as alliances: a form of governance to efficiently manage both traffic exchanges but also network externalities benefits and the reduction of opportunistic risks of free-riding. These alliances then extend globally across IXPs at different geographic locations. Network alliances are essential to create values, for example by improving the reliability of service levels agreements (SLAs) they can provide to their globally distributed customers, acquiring, as a consequence, a competitive advantage towards the Internet demand for quality-sensitive applications.

2.2 *Connectivity as building block of the global supply chain networks*

The Internet is made up of several thousand Internet Service Providers (ISP) who are linked up together. Each ISP runs their own network in an autonomous fashion, managing the data according to their own policies, hence these networks are called autonomous systems (ASes). These ASes are linked up at one or more

points and each of them is linked to one or several other ASes. Two main types of connections are frequently used:

- Direct link: where the two ISPs establish a direct connection between two routers at the edge of their network. Such an interconnection is sometimes also called *private*, and
- Internet Exchange link: where the ISPs connect to a centralized Internet Exchange Point (IXP), which is able to relay data transmission between any of the connected ISPs. This type of interconnection is also called *public*.

Establishing a direct link has the main disadvantage, that it is costly and time consuming. To allow for greater flexibility, IXPs were founded, such that the ISP only needs to connect to this IXP, and then the IXP network is able to establish the links dynamically between ASes through its network configuration. These low technical overheads to establishing a peering link at an IXP enable the cost-effective clustering among geographically closely related ASes.

The internal connectivity of a network and the way its existing connections are organised and structured can be captured in alternative ways and with different metrics. Below, the two main metrics on network connectivity used in this paper are introduced. These are:

1. The degree, or connectivity, k_i , of a node i , representing the number of connections linking this node with other nodes of the same network under study. In this paper k_i describes the number of direct connections agreements AS i , has with other ASes within a given IXP, and

2. The clustering coefficient c_i , introduced by Watts and Strogatz (1998), given by the ratio between the existing links e_i among all the k_i nodes connected to node i , over the maximum potential number of such interconnections:

$$c_i = \frac{e_i}{2(k_i(k_i - 1))}$$

In our paper the clustering coefficient provides the ratio between the existing interconnections among all ASes that are directly connected to AS i , over their maximum potential number.

Both these variables, k_i and c_i , are key indicators of the strategic environment of a particular AS, as they jointly represent the *local supply chain networks* (LSCNs), the connectivity environment of a provider, with its associated strategic complementarities and competitive threats.

While connectivity, k_i , can be interpreted as a measure of an AS's strength, since it provides the set of alternative next *hops* in its traffic routing, the clustering coefficient, c_i , captures the *interconnection relationships* among providers located at adjacent stages of the *LSCN*, and determines the influence these providers can exert on AS i , whether as its upstream suppliers, downstream buyers, or parallel peers along AS i 's LSCN.

These network relationships show a specific complexity, for the Internet as different ASes play, at the same time, alternating roles along their network's global supply chain; being simultaneously distributors and originators of each other's traffic (Giovannetti 2002).

2.3 Connectivity and competitive advantage within the global supply chain networks

The relationship between the connectivity k_i and the clustering coefficient c_i , of a provider i , can be interpreted within a "Six forces framework" analysis (Brandenburger and Nalebuff (1996) as the natural extension of the five forces analysis (Porter 2008) to capture the many dimensions of strategic interaction taking place along the GSCN defining the process of transmitting Internet traffic. More in detail, the relationship between the two parameters, k_i and c_i , expresses the *bargaining power* of:

1. the *input suppliers*, as the neighbours of provider i , can be *upstream* along its GSCN;
2. the *buyers*, as these providers can be purchasers of provider's i connectivity along its GSCN;
3. the *competitors*, as the neighbours of provider i , can also sell connectivity to providers' i customers (business stealing along the provider i 's GSCN);
4. the *demand substitutors*, as they can provide alternative forms of connectivity, than that supplied by provider i , one, they may, for example, offer alternative indirect routes along its GSCN instead than competing for the same route offered by provider i ;
5. the *upstream complementors*, as the neighbours of provider i may ease providers' i costs conditions, by aggregating demand towards third party input suppliers, along its GSCN, so that provider i , will have lower costs in presence of demand economies of scale for its inputs.
6. The *downstream complementors* as the neighbours of provider i may produce network externalities increasing provider's i final demand along its GSCN.

The empirical analysis of the relationships between *clustering* and *connectivity* also captures the role that the emergence of new traffic exchange modalities, such as the appearance and growth of IXPs, may exert and their strategic impact on the *global supply chain network* of relationships on the Internet backbone.

2.4 Connectivity and Performance along the global supply chain networks

The unit of analysis chosen is at the Autonomous System, AS, level. The difference between a graph of AS and one having routers as unit of analysis is very significant, because the AS provide a level of abstraction that conceals many underlying network properties. For the technical aspects, the connectivity at the

router level is important, whereas at the AS level the connectivity is more important for the economic aspects of the Internet.

2.4.1 Latency

The latency depends to a large degree on the number of traversed nodes for a particular flow along the GSCN in absence of congestion. Thus, the shortest path length for a particular route in a graph of routers can be used to indicate the transmission latency of a data packet. If the GSCN has a higher overall connectivity, this will reduce the average of shortest path lengths and therefore lead to on average shorter transmission latencies.

2.4.2 Fault Tolerance

The original purpose for the development of the Internet was to create a fault tolerant communication system that is able to detour around lost connections and damaged nodes. Despite the fact that the Internet diverted from this original aim in the direction of a commercial conglomeration of networks, the Internet has proved to be more disaster resilient than other communication infrastructures. Most notably, during the attacks on the World Trade Center that took out several core Internet routers, the dynamic routing protocols circumvented the network gaps as much as possible within about half an hour. This kind of fault tolerance requires higher degrees of connectivity and a clustering coefficient significantly above zero along the GSCN. For this kind of failures, we ought to pay attention to router connectivity, but if one also takes into account failures that affect whole companies, AS level connectivity becomes important as well.

2.4.3 Quality of Service

In order to establish some contract that ensures certain Quality of Service (QoS) properties for data transmission, all the nodes along the GSCN path need to agree on the terms and protocols for this QoS. The main difficulty with introducing QoS to the Internet lies in the interfaces between all the different ISPs. It would be reasonably easy to enable QoS within one network domain, but it becomes increasingly difficult, the more different network providers have to participate in this QoS path, the longer the GSCN is. If we have a high degree of connectivity at the AS level, it is likely that

the transmission path spans only few different ASes, and therefore it might be easier to offer a QoS for the whole GSCN transmission path.

2.4.4 Connectivity and survival

The degree, or connectivity, k_i of an AS, i , has been studied in Pastor Satorras and Vespignani (2004) in relationship to both birth and death rates for ASes, the processes defining the market entry and exit dynamics. In particular, these authors find that the less connected a network is, the higher the probability of

ceasing to exist, while there are very few observations of networks having high connectivity from birth. This indicates that ASes connectivity is the outcome of a long process in which the individual AS needs to establish itself as valuable partner of other potential peers.

3 IXPs connectivity and global supply chain network structure

Consider the set of ISPs members of a given IXP, $S = \{1,2,3,\dots,N\}$. From the IXP Peering Matrix one can verify the set of reciprocal peering relationships between these ISPs. Whenever a cell of the Peering Matrix displays a “1”, the two ISPs corresponding to column and row defining the cell coordinates, are peering. The set of “1s” in the peering Matrix corresponds to the set of links, or the set of pairs of ISPs (i,j) connected by a peering relationship. The maximum number of peering relationships among the N members of an IXP is given by: $\binom{N}{2}$. The

Peering Matrix of an IXP, can be defined as an $N \times N$ matrix $\underset{N \times N}{[\mathbf{A}]}$, where the single element A_{ij} is binary:

$$A_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E \\ 0 & \text{if } (i, j) \notin E \end{cases}$$

3.1 Peering connectivity index and average connectivity

An immediate way to explore the issue of connectivity within a certain IXP is to look at the *peering connectivity index* $\bar{a} = \frac{1}{n(n-1)} \sum_i \sum_j A_{ij}$. This coefficient given

by the ratio of all active peering decisions, the sum of all the “1” in the Peering Matrix over the set of all possible peering decisions, the sum of all the Matrix cells, apart from the main diagonal, provides the probability that any two ISPs randomly chosen are linked by a peering relationship.

The Peering Matrix, defining a set of vertices and binary links among them also defines an undirected graph. The number of peering agreements, vertex, of an ISP is of course given by the sum of 1s in the peering matrix, corresponding to the row characterizing the ISP: $k_i = \sum_j A_{ji}$. The statistical distribution of k_i

across ISPs is an important first step in the analysis of the Peering Matrix as a Graph, and can be used to infer the probability that a randomly chosen member of an IXP will have k peering agreements at that IXP. From the complete set of connectivity values, the connectivity of each node, $\{k_1, \dots, k_i, \dots, k_N\}$ it is possible to obtain two other indexes focussing on the average connectivity behaviour at IXP level:

1. The *average connectivity* of the IXP: $\langle k_i \rangle_{IXP} = \frac{1}{N} \sum_{i \in N} k_i$ showing the average connectivity of the providers, and

2. The *peering connectivity index*: $\bar{a} = \frac{1}{N(N-1)} \sum_i \sum_j A_{ij} = \frac{1}{N(N-1)} \sum_i k_i$

showing the average connectivity of the Exchange as a whole, IXP.

The statistical distribution of k_i across ISPs provides important information about the possibility of concentration in interconnection agreements, however, since many different connection architectures are compatible with the same level of average connectivity, further measures of connectivity are needed to better represent the “GSCN architecture”. One of the most used measures is given by the clustering coefficient.

3.2 Clustering

The clustering coefficient of a node (AS) i was defined in section 2 as the ratio between the peering agreements, e_i among the nearest neighbors of node i out

of the potential number of such agreements: $c_i = \frac{e_i}{2(k_i(k_i-1))}$.

The clustering coefficient, thus, measures the probability that two peers of an AS i , are also connected between them. Given the peering Matrix the actual number of peering agreements among the peers of an AS i , can be computed as

$e_i = \frac{1}{2} \sum_{j,l} A_{ij} A_{jl} A_{li}$. In other words, e_i is the number of peering agreements

characterising the sub-matrix whose row (and column) elements are the peers of the node i . Indeed the clustering coefficient, c_i , of an individual AS describes how well locally interconnected the set of its neighbours is.

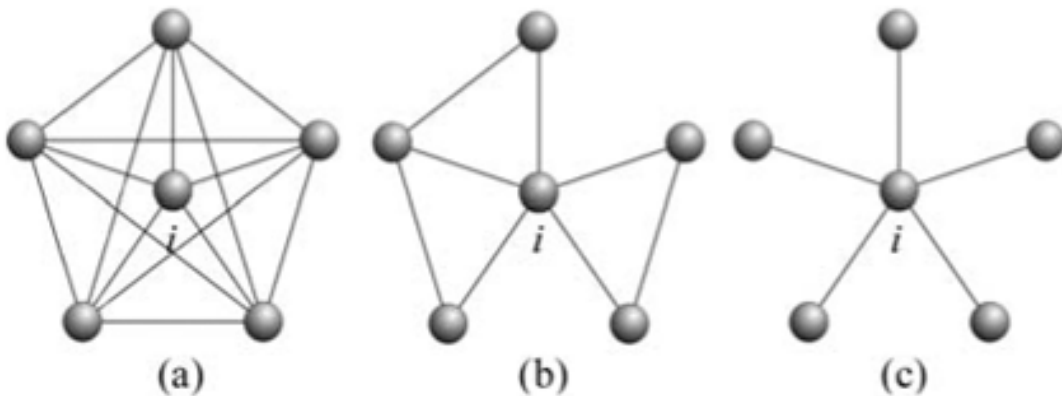


Figure 1. Networks: (a) High Clustered, (b) Low Clustered, (c) Non Clustered

Also for this metrics it is of extreme interest to characterize the clustering structure of the entire IXP, to which the individual ASes belong. This is done by

calculating the *IXP average clustering coefficient* easily obtained from those of its

individual members: $\langle c_i \rangle = \sum_{i \in N} \frac{c_i}{N}$.

Vazquez et al. (2002) studied average connectivity for the entire Internet for three years 1997, 1998, and 1999 showing that the average number of connections, per node, has been fairly stable, in these years, at around 3.6 connections per AS. These authors also calculate the values of the *average clustering coefficient*, for the entire Internet, finding it significantly higher than the one usually associated to a random graph, where links between ASes are randomly generated with a given probability p . This higher level of neighbourhood clustering was explained by considering the influence of geographical distance between ASes, and its role in determining linkage costs and therefore affecting the decision of whether to establish a new link between any two ASes. The authors interpreted this high value of the average clustering coefficient as an indicator of an Internet architecture whereby nodes, in a given geographic region, have many connections between themselves. This clustering feature is indeed typical of networks arranged following a regular lattice topology, in which every AS is only directly linked to its nearest lattice neighbours.

However the analysis of another indicator, the average shortest path between any two nodes (is the smallest number of connections that link the nodes i and j) showed that existing distances between any two ASes are usually much shorter, than those required to travel between any two nodes on a network, with a regular lattice topology. Moreover short average paths are a typical indicator of *small world* property since any two nodes of the set can be joined by a small number of hops (Watts and Strogatz (1998)).

The co-presence of both high local clustering and short distances across network nodes (ASes) provides a clear clue about the real structure of the Internet architecture as a GSCN where a crucial hub role is played by the backbones. These indeed provide connectivity for ASes across long distances, thereby visibly reducing the number of hops required to connect any two random ASes in the entire Internet.

3.3 Distribution of Connectivity

A relevant feature which has emerged from early Internet studies by Faloutsos et al (1999), based on the Oregon Route Views data², is that the distribution of connectivity at the AS level scales following a power law behaviour, i.e. it is best linearly approximated in a double logarithmic system of coordinates. The interpretation of this scale-free behaviour is straightforward: with a power law scaling it is not very unlikely to observe large, otherwise rare, events, in this case ASes with high connectivity, large number of peering agreements, or hubs. This

type of scaling behaviour describes GSCN features of a majority of poorly connected ASes, with a non-negligible presence of very well connected ones, the hubs and backbones, which play a crucial role and have major bargaining power in the revenue allocation process along the values chain of providing Internet connectivity. Translating this scaling property for the data set analysed in this paper, it implies that if the distribution of the number of peering agreements per AS should follow a power law, implying that the *average connectivity* of the IXP,

$$\langle k_i \rangle_{IXP} = \frac{1}{N} \sum_{i \in N} k_i, \quad \text{is no longer a very representative statistics for the}$$

distribution, since this would have “fat tails” and scale free properties. This feature is a mark for self-similarity and under these conditions it has been proposed to use a characterization of network heterogeneity with the ratio between second and first moment of the peering distribution, i.e. by using a

$$\text{parameter: } \langle\langle k \rangle\rangle = \frac{\langle k^2 \rangle}{\langle k \rangle}.$$

A further issue of interest is the analysis concerning the connectivity of sub-networks composed by the best connected, highest k_i , ASes. If these are more interconnected among themselves than the least connected ones, clustering among the top tier will appear.

Following Vazquez et al 2002 the focus of our analysis will investigate the relationship between a node connectivity, k_i , and its neighbours' connectivity by considering the behaviour of the *average clustering coefficient* for a given vertex value k , the clustering coefficient averaged across all ASes members of an IXP, having the same peering connectivity value.

These authors found that the *average clustering coefficient* scales, with a negative exponent, as a *power law* function of the peering connectivity, k , hence indicating that the more connected ASes have a lower clustering coefficient, and therefore a less clustered local neighborhood. This is a strong mark of hierarchical differentiation in the GSCN, since the many, ASes connected to the few *big ones* (with a high k), will be not well connected amongst themselves: therefore high connectivity together with low clustering implies a more hierarchically differentiated GSCN, and would be also of interest and concern for competition authorities.

4 Empirical Analysis of the connectivity for the EuroIX Members

This section introduces the statistical properties of both connectivity and clustering for each single IXP of the Euro-IX. First the issue of connectivity within the IXPs is analysed, then the statistical relationships between the clustering coefficient and the connectivity, are studied using the data contained in the Peering matrices of the 33 Members of Euro-IX.

4.1 Peering connectivity behaviour

The first step in analysing the features of the Internet GSCN requires to consider the degree of internal connectivity of each IXP; this is done by calculating each IXP's internal efficiency at providing its members ISPs with connectivity through the *peering connectivity index* introduced in Section 2:

$$\bar{a} = \frac{1}{N(N-1)} \sum_i \sum_j A_{ij} = \frac{1}{N(N-1)} \sum_i k_i.$$

However, the process of delivering Internet traffic along the entire GSCN, does not necessarily conclude within the borders of an IXP. If the final destination of an Internet packet is not reached by peering within an Internet Exchange the next step in the analysis of the GSCN requires to focus on how the different IXPs are interconnected among themselves. From intra-IXP connectivity metrics a complete analysis of the GSCN needs to include the inter-IXP connections as well. The most direct way of capturing inter-IXPs interconnections is by calculating the percentage of ISPs, members of an IXP, that also peer at a different IXPs. Table 1, below gives us the Euro-IX member ranking accordingly to their *peering connectivity index*, and reports membership levels, and percentages of their members being also peers at other, different, IXPs.

Table 1. IXP Ranking According to Internal Connectivity (Euro-IX members)

	Peering connectivity index	members	% of members that are also peering at other IXPs
catnix	0.882	18	0.5
lix	0.818	11	0.27
ndix	0.800	6	0.5
namex	0.771	15	0.67
inex	0.750	9	0.56
manap	0.700	29	0.69
lonap	0.695	43	0.81
espanix	0.693	28	0.68
vix	0.677	80	0.39
nix_cz	0.634	54	0.31
de-cix	0.618	139	0.62
cixp	0.609	24	0.67
topix	0.603	13	0.43
madix	0.600	6	0.83
nix	0.596	49	0.15
linx	0.589	161	0.78
tix	0.577	51	0.55
lipex	0.570	53	0.89
gigapix	0.563	20	0.25
bnix	0.553	47	0.62
parix	0.539	38	0.58
mix	0.528	59	0.47

netnod	0.503	46	0.5
amsix	0.496	214	0.55
ficix	0.486	21	0.33
aix	0.476	15	0.07
bix	0.371	48	0.08
gn-ix	0.362	21	0.5
xchangepoint	0.315	124	0.59
ronix	0.195	22	0
bcix	0.162	23	0.2
msk-ix	0.150	139	0.02

The data on the Euro-IX members show that on average, IXPs have a high *peering connectivity index*. Indeed, fourteen are characterized by a peering index greater than 0.6. On the other hand, only three of them show an index lower than 0.3. The overall average peering index for the Euro-IX is thus quite high, with a value of 0.56.

The frequency distribution of the *peering connectivity indexes* across different IXPs is characterized by a bell-shaped pattern with high symmetry (mode is approximately equal to mean and median value). The graph below indicates that the internal level of connectivity is quite similar among all the Euro-IX members (21 IXPs have values among 0.5 and 0.7).

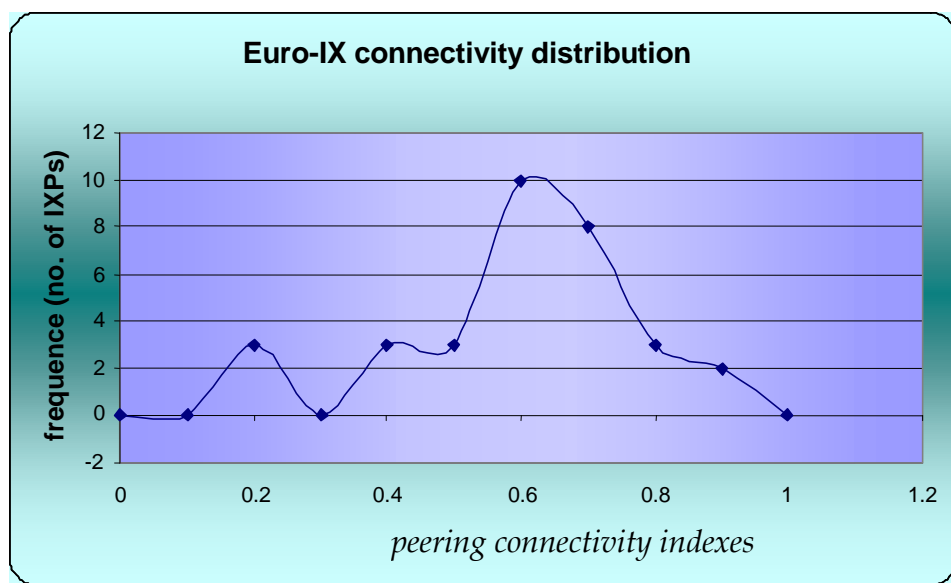


Figure 2. Euro-IX peering index distribution

It is interesting to investigate the possibility of the existence of a relationship between the internal *peering connectivity indexes* and the dimension of the single IXP, expressed by the number of its members. The scatter diagram, reported below in Figure 3 is very dispersed. Indeed, the linear correlation coefficient between these two magnitudes equals -0.28, suggesting possibly difficulties in

coordination arising exponentially with the number of members. Anyway this value is too low to safely assume the existence of a linear relation.

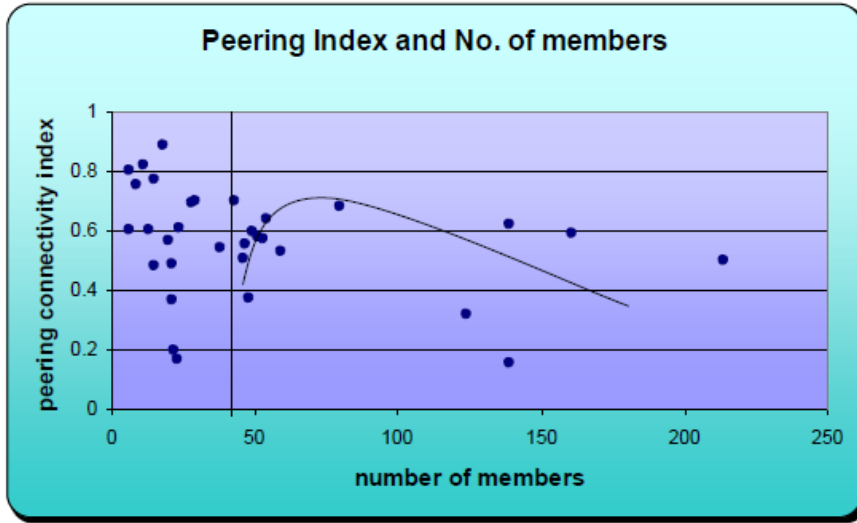


Figure 3. Euro-IX peering connectivity – membership number

Figure 3 seems to suggest the opportunity to consider this relationship in two separated subsamples of IXPs, according to their membership numbers. For the sample of IXPs with low number of members, peering connectivity indexes are very dispersed, going from 0.18 to 0.9; in this case the peering index seems not influenced at all by the size of the IXP.

On the other hand, the second sample of IXPs with larger memberships shows a more interesting feature: the peering index appear to follow a quadratic function of the number of ISPs members, first increasing and then decreasing after reaching a peak. In general, the peering index (or internal connectivity) seems not to be scale free. Several statistics are reported in the Appendix table A1.

5 An econometric model on the functional relationship between clustering and connectivity

The main aim of this paper is to use complexity metrics to understand the relationships linking the specific steps of the value chain associated to the process of transmitting internet traffic, across a multiplicity of providers that are linked in multiple strategic relationships of competition and cooperation. To this aim this section focuses on how peering strength, represented by k_i , affects local clustering, exploring how these relationships translate into different network hierarchies among the ASes for specific World Areas.

In particular, for each IXP the *clustering coefficient* is obtained from those of its

AS members: $\langle c_i \rangle = \sum_{i \in N} \frac{c_i}{N}$ (see table A2 for some descriptive statistics).

In order to test econometrically the relationship between peering connectivity and clustering at the IXP level we estimated two alternative models by OLS: Model 1, the results of which are reported in Table 2, based on a linear functional

specification; and Model 2, the results of which are reported in Table 3, based on a logarithmic one. In order to assess how the different structure and competition degree that characterizes the Internet upstream market across the world affects the connectivity features of the GSCN, we run our estimates separately for IXPs based in Europe, North America and rest of the World.

$$\text{Model 1 : } \langle c_i \rangle_{IXP} = \beta_0 + \beta_1 \langle k_i \rangle_{IXP} + \beta_2 N_{IXP} + \varepsilon$$

Table 2 Estimation of linear effects of IXP Connectivity on IXP Clustering: Europe, North America and Rest of the World

$\langle c_i \rangle_{IXP}$ (IXP_clustering)	Europe	North America	Rest of the World
$\langle k_i \rangle_{IXP}$ (Average connectivity index)	-0.00605*** (0.00148)	-0.00263 (0.00157)	-0.00396*** (0.00109)
N_{IXP} (IXP Members)	0.00338*** (0.000775)	0.00438** (0.00175)	0.00613*** (0.00184)
Constant	0.447*** (0.0797)	0.250** (0.106)	0.329*** (0.0958)
Observations	72	42	81
R-squared	0.409	0.296	0.425

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

$$\text{Model 2 : } \log \langle c_i \rangle_{IXP} = \beta_0 + \beta_1 \log \langle k_i \rangle_{IXP} + \beta_2 N_{IXP} + \varepsilon$$

Table 3 Estimation of Logarithmic effects of IXP Connectivity on IXP Clustering: Europe, North America and Rest of the World

$\log \langle c_i \rangle_{IXP}$ (log(IXP clustering))	Europe	North America	Other
$\log \langle k_i \rangle_{IXP}$ (log(Average connectivity index))	-0.148*** (0.0389)	-0.0582 (0.0651)	-0.115*** (0.0409)
N_{IXP} (IXP Members)	0.00200*** (0.000438)	0.00294** (0.00126)	0.00348** (0.00149)
Constant	0.679*** (0.133)	0.329 (0.244)	0.525*** (0.177)
Observations	72	42	81
R-squared	0.410	0.319	0.433

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

From the results we can see that the average connectivity is negatively correlated with clustering coefficient both in Europe and in the Rest of the World, and both in Model 1 and Model 2. On average, more connected IXPs are characterized by a lower level of clustering, hinting at the presence of some bargaining power across the members of IXPs. However, the effect is not significant for the North American parameter. When comparing the two different model specifications we notice that the R-squared are higher for Model 2 than for Model 1, pointing towards the selection of the logarithmic equation as a preferable representation of the underlying relationship between clustering and connectivity.

This finding signals that power law scaling, between these two variables, is present at IXP level in Europe and in the Rest of the World indicating a more hierarchical values chain, in these two continents and that there are differences in this behaviour between the IXP based in Europe and in the rest of the world compared to the North Americans ones.

6 Conclusions

Connectivity constitutes the building blocks of the different stages in the value chain of the transmission of Internet traffic and is relevant to many quality-related Internet parameters. Therefore, by focussing on connectivity at IXPs, we can assess the specific contribution that these are providing to the quality of the Internet as a whole. Moreover connectivity and clustering are casting new light on further issues of economic interest: as entry and exit rates into the ISP market, while the statistical relationship between clustering and connectivity provides tremendous insights into the Internet upstream market architecture. While this work contains a preliminary analysis of these data, it was intended to introduce the relevance of the underlying complexity metrics for the subsequent market analysis, and the assessment of power relationships shaping the strategic exchanges along the value chain of the Internet traffic transmission sector.

A preliminary analysis of the statistical relationship between connectivity and clustering was performed to understand the internal hierarchy of each IXP, and to monitor for the emergence of self-similar power law type of relationships which have often been indicated as markers for a specifically hierarchical Internet architecture. A predominantly inverse relationship between *local clustering and peering strength* was found confirming that, even within the members of a given IXP, the most important ones are not those whose direct neighbours are better connected among themselves. The comparison between power law and linear decay was clear for Europe and the rest of the World but not for North America, showing that other features of the Internet value chain may be relevant in explain this geographic differences.

Finally, further research should focus on meta-connectivity, considering linkages among IXPs as an additional covariate. This could be done through the study of cross-membership of their ISPs, helping to better understand GSCN integration under the special point of view of Internet traffic flows.

References

- Barabasi A .L., Albert,R., 1999. Emergence of scaling in random networks. *Science* 286(5439), 509–512
- Brandenburger,A.M., and Nalebuff, B.J., Co-opetition, Doubleday, New York (1996).
- D’Ignazio, A., and Giovannetti, E. (2006), "Antitrust Analysis for the Internet Upstream Market: A Border Gateway Protocol Approach." *Journal of Competition Law & Economics* , Vol 2-1, pp 43-69.
- D’Ignazio, A., and Giovannetti, E. 2007.From Exogenous to Endogenous Economic Networks: Internet Applications", *Journal of Economic Surveys*, (2006) Vol. 20, No. 5, pp. 757-796.
- D’Ignazio, A., and Giovannetti, E. 2007, "Spatial Dispersion of Peering Clusters in the European Internet", *Spatial Economic Analysis*, Vol. 2:3, pp. 219 - 236
- D’Ignazio A., and Giovannetti E., 2009.Asymmetry and Discrimination in Internet Peering. Evidence from the LINX *International Journal of Industrial Organization*, Vol.27, pp. 441- 448.
- Faloutsos,M., Faloutsos, P., and Faloutsos, C.,"On power law relationships of the internet topology", ACM SIGCOMM, 1999
- Euro-IX website <http://www.euro-ix.net/>
- Giovannetti, E., 2002. Interconnection, Differentiation and Bottlenecks in the Internet. *Information Economics and Policy*. vol 14 n. 3 pp 385-404.
- Hallikas,J.,Karvonen,I.,Pulkkinen,U.,Virolainen,V.-M.,Tuominen,M., (2004). "Risk management processes in supplier networks." *International Journal of Production Economics* 90(1),47–58.
- Link, P., Marxt, C., 2004. Integration of risk and chance management in the co-operation process. *International Journal of Production Economics* 90, 71–78.
- Lo Nigro, G.,and Abbate, L., 2011.Risk assessment and profit sharing in business networks, *International Journal of Production Economics*, Volume 131, Issue 1, May, Pages 234-241
- M’Chirgui, Z.,and Penard,T., 2011. How to provide Quality of Service guarantees in the Internet? Lessons learnt from the airline and card-based payment sectors, *European Management Journal* , 29, 306– 318
- Oregon Route Views, <http://archive.routeviews.org>

- Pastor-Satorras, R., and Vespignani, A., 2004. *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge University Press, Cambridge.
- Pero, M., Rossi T, Noe' C, and A. Sianesi, 2010, An exploratory study of the relation between supply chain topological features and supply chain performance. *International Journal of Production Economics*, 123 266–278.
- Porter, M. E..2008. The Five Competitive Forces that Shape Strategy, *Harvard Business Review* p.86-104.
- Talluri, S., Baker, R.C., Sarkis, J., 1999 A framework for designing efficient value chain networks, *International Journal of Production Economics*, Volume 62, Issues 1–2, Pages 133-144,.
- Vazquez, A., Pastor-Satorras, R., and Vespignani, A., 2002. Large-scale topological and dynamical properties of the Internet. *Phys. Rev. E* 65, 066130.
- Watts, D. J., and Strogatz, S. H., 1998. Collective dynamics of small-world networks *Nature*, 393:440-442.
- Xiao, Z., G., Zhou and B. Wange, 2012. Using modified Barabasi and Albert model to study the complex logistic network in eco-industrial systems. *International Journal of Production Economics*