1. **INTRODUCTION**
   1. **Review papers**
      1. **A description of the state-of-the-art use of networks in finance**

A network can simply be described as a set of nodes and connections between them known as the edges. Earlier memories of the application of network theory have been about its application to areas such as economics, social networks, and transport networks etc. But its venture into finance was inspired by various crisis that have bewitched the financial systems globally, network theories are now researched and applied on financial systems. The concept of network theory in its relation to finance is such that the nodes in a network refer to financial institutions and edges refer to the links, financial transactions, between the financial institutions.

There are several types of connections in the financial system but all are reflected in the Asset and Liability segments of a financial institutions balance sheet (Allen and Babus, 2008). A sample balance sheet is provided in the Appendix figure 1.1.

In the past, Network theory was mostly used to understand issues pertaining to financial stability and contagion. Recent research and applications of Network theory in finance has now extended to financial issues such as the analyses of investment banking, the role of social networks in investment decisions and corporate governance, and to understand how interbank markets can freeze.

* + 1. **importance of network analysis for the study of systemic risk and contagion in financial networks.**

According to Caldarelli et al (2013), the degree of contagion risk of a financial system may be strongly linked to the topology of the web of relationships linking the institutions. It is therefore important to understand how the topology either impedes or foster the spread of such epidemic (Ganesh et al,). The analysis of such Network can help to answer questions with regards to network effects and its formations (Allen and Babus, 2008). Network analysis can also be useful to identify systemically important institutions or cliques which may be channels to propagate risk to the entire financial system. It provides supervisory bodies with the means to test the resilience of a financial network to potential systemic risk and it acts as an empirical tool for testing the effectiveness of macro-prudential policies (Kanno, 2015).

* 1. **A description of the research carried out on Japan Banking Network.**

The paper by Kanno (2015) titled “*The network structure and systemic risk in the Japanese interbank market”* assessed the network structure of bilateral exposures in the Japanese interbank market. The Japan interbank market consists mainly of call and bankers acceptance markets. Japan’s financial system had suffered a few crisis in the last 20 decade, for example the Heisei great recession of 1997-1998 and the global financial crisis, which lead to some large financial institutions declaring bankruptcy. Japan’s Banking regulatory body together with the world bankers authorities (Basel Committee) had since been awakened to the need for a proper monitoring and measuring system for systemic risk in financial systems.

Kanno analysed the systemic risk in Japanese financial network through the interbank network based on claims that the interbank network represented a significant component in the analysis of systemic risk in Japan. The analyses was based on several network measures such as the centrality measures, degree distributions, and modified susceptible – infected- recovery (SIR) models e.t.c.

The study was performed on interbank data for 2009 and 2013. The bank data analysed was 127 and 121 banks for year 2009 and 2013 respectively. These banks comprised of major banks, regional banks and second-tier banks. The interbank network was presented as a directed graph where the source nodes represents the set of banks with liabilities in the interbank market and the target nodes represents the set of banks with claims in the interbank market.

Betweenness centrality for the two years identified three mega-banks in japan whose soundness is central to the Japanese financial system stability. This result coincides with the selection of the banks by the Japan banking supervisory body as Global systemically important banks in accordance to Basel 3. Closeness centrality produced similar result of the top banks to those of betweenness centrality apart from the second-tier banks. In contrast, eccentricity centrality results were different from the previous two centralities. The centrality results are provided in Table 1 of the Appendix.

In terms of the structure of the network, the interbank network was observed to have characteristics of scale-free networks. The in-degree distribution had both small-world and scale-free characteristics while the out-degree distribution had scale-free characteristics as seen in figure 1.2 in the Appendix. This indicates the dominance of the major banks in the interbank network. These banks are the hubs in the network as seen in figure 1.3 in the Appendix. A list of those banks were collected as domestic and Global systematically important banks. The modified SIR model identified a 10 percent increase in the number of contributors to systemic risk of the Japanese interbank market which affirms the increase in number of hubs between 2009 and 2013.

* + 1. **Some additional analysis that can be done in the study**

Further analysis of the network can be done first by calculating additional centrality measures such as the eigenvector and subgraph centralities for the Japanese interbank network. Unlike the previous measures that described the relationship between each node and its nearest neighbours, this spectral measures can help to extend the analysis to a more global neighbour in a network. For example, Subgraph centrality can be used to detect how important each nodes are in all subgraphs in the network.

Apart from analysing the effect of individual nodes in the Japan interbank network, Kanno did not mention if some nodes formed cliques in the network. Taking account of the modularity of the network will detect communities in the interbank network. Also, it will be useful to investigate if banks relate with each other based on some similarities by calculating the pearson coefficient of the network. This may help to explain the kind of relationship and topology of the network.

1. **METHOD**

**2.1. Erdős-Rényi Network and Barabási - Albert Network Creation**

In order to further analyse the Japanese interbank network studied by Kanno, this report will attempt to simulate networks with properties to similar to the Japanese interbank network studied. Two networks G1 and G2 will be created. G1 will be an Erdős-Rényi network (Random network) and G2 a Barabási-Albert (Scale-free) network.

**Random Network (G1)**

The Erdős-Rényi model G(n, p) created by Paul Erdős and Alfréd Rényi is a model used for generating random networks with n nodes based on a probability p that an edge may be set between each pair of nodes in the network. The density of edges of the network is the probability that a pair of nodes are connected in the network. G1 was created with 500 nodes and a network density of 13% similar to the interbank network observed in the work of Kanno.

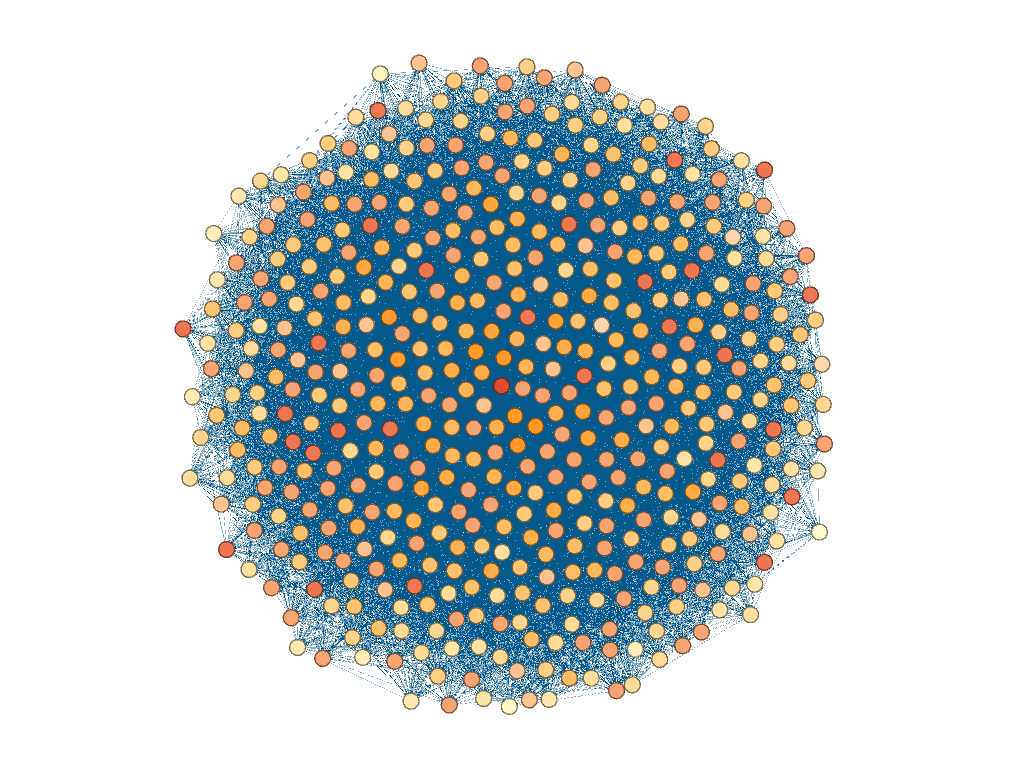


Figure 2.1 G1 network plotted with Gephi software

Various random networks were generated and the number of in order to identify the average number of edges for appropriate for a random network with that size of nodes. As shown in (figure 2.2), the number of iterations of random networks generated was increased and the number of connections generated came close to a normal distribution with a mean of 16215 connections.

Table 2.1 Average number of edges for a number of iterations of random networks

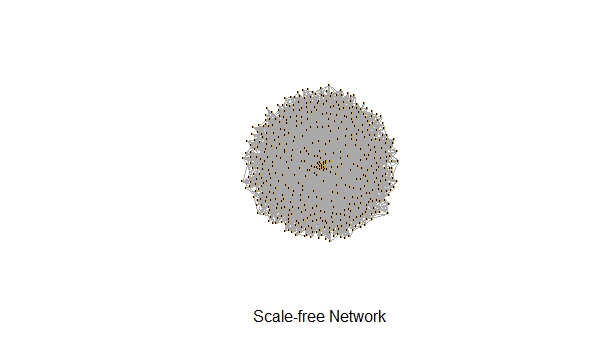
|  |  |
| --- | --- |
| **Number of iteration** | **Average connections** |
| 1 | 16107 (varies) |
| 10 | 16215.55 |
| 100 | 16215.59 |
| 1000 | 16215.42 |
| 10000 | 16215.74 |

|  |  |
| --- | --- |
| Rplot01.jpeg | Rplot2.jpeg |
| Rplot3.jpeg | Rplot04.jpeg |

Figure 2.2 multiple iterations of the random network

**Barabási-Albert Model (G2)**

Barabasi-Albert model is a type of random network model that is based on preferential attachment. This implies that new nodes are more likely to connect to nodes with a high degree. In generating this network, a node with high degree of edges have more tendency to add another edge than nodes with low degree of edges [4]. G2 was created using the Barabasi-Albert model G (n, m). G2 network was created with 500 nodes and m = 18 based on properties of the interbank network observed in the work of Kanno



.

Figure 2.3 Barabási-Albert (G2) network

**3.0 S-I-R MODEL**

**3.1 SIR**

SIR model is one of the models of network theory that tries to represent and explain the propagation of an epidemic in a network. SIR models three types of individuals in a network, the susceptible, infected and the recovered. The susceptible S are the number of individual that are likely to get an infection. The infected I are the carriers of such disease while the recovered R are the number of individuals that have recovered and are immune to such infection in a population N (Kermack and McKendrick, 1927; Anderson and May, 1991).

Where S + I + R = N

|  |
| --- |
|  |

= Rate infected individual gives rise to new infections

= Rate of recovery once infected

This defines the epidemic threshold.

If this threshold is less than 1 the epidemic dies out but if the threshold is above 1 the epidemic remains.

disease dies out

disease epidemic remains

In the simulation of propagation of epidemics in the G1 and G2 networks, the total population size N will be the number of nodes 500. While the number of susceptible, infected and the recovered will vary.

**3.2. S-I-R EPIDEMIC SIMULATION ON G1 NETWORK**

The number of infected was highest when the recovery rate was less than the rate of infection. From figure 3.1 it can be seen that the epidemic dies out faster when the recovery rate was higher than the infection rate. Similarly with the G2 network except that the epidemic spreads faster in G1 than in G2. In some cases the epidemic dies out in G2 after just a few nodes.

Table 3.1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Beta** | **Gamma** | **Ro** | **Av. Max. Number of Infected** | **Peak time** |
| 0.25 | 0.25 | 1 | 455 | 0.6825506 |
| 0.25 | 0.5 | 0.5 | 416 | 0.6316207 |
| 0.25 | 1 | 0.25 | 362 | 0.5981288 |
| 0.5 | 0.25 | 2 | 477 | 0.3608231 |
| 0.5 | 0.5 | 1 | 455 | 0.3425026 |
| 0.5 | 1 | 0.5 | 415 | 0.3149752 |

|  |  |
| --- | --- |
| **R2.5.jpeg** | **R1.25.jpeg** |
| **R1.jpeg** | **R0.25.jpeg** |

Figure 3.1. SIR simulations for G1 network

**3.3. S-I-R EPIDEMICS SIMULATION ON G2 NETWORK**

**Table 3.2**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Beta** | **Gamma** | **Ro** | **Av. Max. Number of Infected** | **Peak time** |
| 0.25 | 0.25 | 1 | 409 | 1.15688 |
| 0.25 | 0.5 | 0.5 | 349 | 1.047771 |
| 0.25 | 1 | 0.25 | 261 | 0.8705393 |
| 0.5 | 0.25 | 2 | 445 | 0.6395224 |
| 0.5 | 0.5 | 1 | 412 | 0.588713 |
| 0.5 | 1 | 0.5 | 344 | 0.5074977 |

|  |  |
| --- | --- |
|  |  |
|  |  |

Figure 3.2. SIR simulations for G2 network

**4.0. PROPERTIES OF THE G1 AND G2 NETWORKS**

**4.1. Degree distribution**

The degree distribution of a network is an ordered list of the degrees of every node in that network, usually visually represented in a chart. It is used to determine how skewed the degree of nodes are across a network. It is represented by the probability of connections nodes in a network has.

**4.1.1. Erdos-Renyi (G1) network**

Observing the G1 network as a directed graph similar to the work paper. The random network was observed to follow a poisson distribution for both the in- degree and out-degree of G1. The in - degree and out - degrees of nodes in the G1 network are skewed to the right but the in-degree of nodes were more skewed than the out-degree.

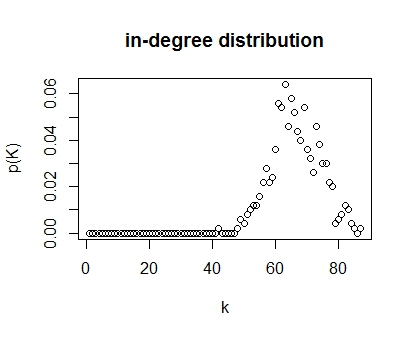
****

Figure 4.1 in-degree distribution of G1 network

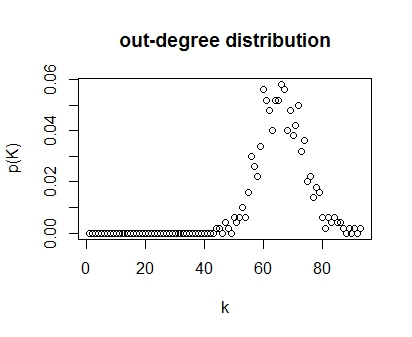
****

Figure 4.2. out-degree distribution of G1 network

**4.1.2. Barabasi-Albert (G2) Network**

|  |  |
| --- | --- |
|  |  |

For G2 network, the in-degree had the form of a power law distribution with an exponent of 2.17714 while the out degree followed a power law distribution with an exponent of 1.627322. Indicating that only a few nodes have huge numbers of edges, a large number of nodes have a decent amount of edges, but significantly fewer, and most nodes are connected by very few edges.

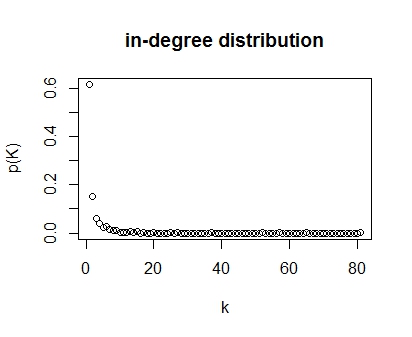


Figure 4.3 in- degree distribution of G2 network

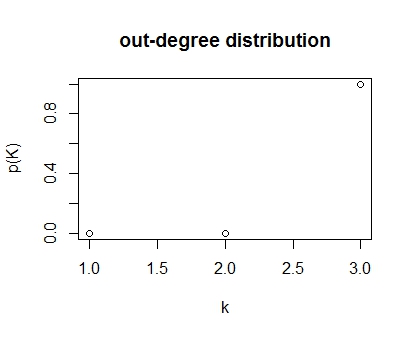


Figure 4.4 out-degree distribution of G2 network

**4.2. Average path length**

A path in a network between any two vertices is the number of edges required to connect them without repetition of nodes and edges. Shortest path also known as geodesic between any two vertices is the least number of edges required to connect them.

Let u and v be any two vertices in a network G such that ,

and paths between u and v are

the shortest path equals to the least .

**4.2.2.Result**

Therefore an Average path length is the mean of all the shortest paths within a network. Average path length gives an indication as to how fast information or an epidemic can be propagated within a network. For G1 network, the Average path length is 1.870669 while G2 has an average path length of 1.862254. Both G1 and G2 have average path length of 2 edges indicating there are at least two edges between every pair of vertices in both networks.

**4.3.Watts - Strogatz clustering coefficient**

Clustering coefficient measures the level of interaction between the neighbours of a node [3]. It is calculated as the actual number of edges between the neighbours divided by the possible number of edges.

For the entire network, the degree of clustering (average wattz – strogatz clustering coefficient) is calculated as the average of all the local clustering coefficients in the network.

**4.3.2. Results**

The average watts-strogatz clustering coefficient for G1 is 0.2418317 while G2 transitivity of 0.1577381. The results show that G1 and G2 both have small amounts of clusters in their respective networks but there are fewer clusters in G2 than in G1. However in both networks the chances that any two neighbours of a certain node are connected is quite small.

**4.4.Centralities**

Centrality is a measure of importance of nodes or edges in a network. Nodes or edges within the network are ranked based on the level of importance (Waseerman and faust, 1994).The node or edge with the highest centrality is very essential in the network so much that they can either facilitate or impede the spread of an information or epidemic in the network also removing them may disconnect the network.

**4.4.1.Degree centrality**

The degree centrality is simply the degree of edges each node has in the network. It measures the ability of a node to communicate directly with a node. Mathematically, the degree of a node i in a simple network G can be defined as a adjacency matrix A, as

**4.4.2.Closeness centrality** measures how close a node is from all other nodes in a network. It is a function of proximity a node (Kanno, 2015). It can be written mathematically as

where

is the sum of all shortest path distances *d(u,v)* between the node u and any other node in the network (Estrada, ).

**4.4.3. Betweenness centrality** is a measure of the degree to which a node lies on paths between pairs of nodes (freeman, 1977). Nodes with high betweeness centrality are potential targets to external attacks on a network (Puzis et al, 2008).

Where ρ(i, j) is the number of shortest paths from node i to node j, and ρ(i, k, j) is the number of these shortest paths that pass through node k in the network (Estrada, 2009 ).

**4.4.4. Eigenvector centrality** measures the quality of the connections a node has, that is, how important the neighbour nodes of a node is within the network.

If and then the elements of x and y give the right and left eigenvector centralities, respectively.

**4.4.5.Subgraph centrality**

Subgraph centrality measures the role each node plays in all subgraphs of a network. It measures participation of the nodes by assigning more weights to participation in smaller subgraphs than larger subgraphs in a network. Estrada (2009) argues that subgraph centrality discriminates better than the others. It can expressed mathematically as

Where is the number of closed walks of length k starting and ending on vertex i in the network and are equal to the i th diagonal entry of the k th power of the adjacency matrix, A of a network (Estrada, 2009 ; Estrada and Rodríguez-Velázquez, 2005).

**4.4.6. Centrality Results**

**Centrality result for G1 network**

For the G1 network in Table 4.1 ten nodes were selected based on their centrality ranking on several centrality measures. First was ranking based on direct contact and closeness to neighbour nodes. Ten odes were ranked in the same order by their both degree and closeness centralites ranking nodes 46, 173 and 267 as the first, second and third most important nodes in the G1 network. However ranking the nodes in terms of the shortest paths they lie on- the betweenness centrality, their ranking changed as node 46 still remained the most important but 173 came 5th and 267 was 7th. In terms of how important nodes connected to them are in the network, the eigenvector centrality ranked 173 as first and 46 as second node 267 was nowhere on the list a new node 137 made the list as the third ranked. To further observe the ranking of nodes based on participation in all subgraphs within the G1 network. The ranking completely was different with nodes 290,226, and 274 taking the first, second and third positions and the node 46 was ranked 10th on the list.

TABLE 4.1. Centralities for top Ten nodes in the G1 network

|  |  |  |  |
| --- | --- | --- | --- |
| **Nodes** | **Betweenness Centrality** | **Nodes** | **Eigenvector Centrality** |
| 46 | 710.855382841707 | 173 | 1 |
| 461 | 634.116663035892 | 46 | 0.969853883482225 |
| 435 | 623.759027483727 | 137 | 0.916110791448059 |
| 64 | 607.235322693292 | 330 | 0.903295376336459 |
| 173 | 602.442016305915 | 444 | 0.903227258751071 |
| 108 | 600.023326847409 | 461 | 0.901171283554898 |
| 267 | 598.162074986408 | 237 | 0.898532742048116 |
| 303 | 597.179794820725 | 293 | 0.894730562920119 |
| 137 | 596.522830597861 | 314 | 0.888336148056804 |
| 296 | 595.9988683249 | 110 | 0.879594810908488 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Nodes** | **Degree centrality** | **Nodes** | **Closeness Centrality** | **Nodes** | **Subgraph Centrality** |
| 46 | 161 | 46 | 0.00118203309692671 | 290 | 4.6149621710524e+25+0i |
| 173 | 157 | 173 | 0.00117096018735363 | 226 | 4.05245148773207e+25+0i |
| 267 | 154 | 267 | 0.00116959064327485 | 274 | 3.85899371488408e+25+0i |
| 303 | 154 | 303 | 0.00116959064327485 | 113 | 3.81209779378045e+25+0i |
| 461 | 152 | 461 | 0.00116959064327485 | 108 | 3.78390350081142e+25+0i |
| 137 | 152 | 137 | 0.00116822429906542 | 230 | 3.73770745209893e+25+0i |
| 407 | 152 | 407 | 0.00116822429906542 | 191 | 3.71840887876898e+25+0i |
| 64 | 151 | 64 | 0.00116686114352392 | 279 | 3.69344458981609e+25+0i |
| 108 | 151 | 108 | 0.00116686114352392 | 427 | 3.69102764466866e+25+0i |
| 290 | 151 | 290 | 0.00116414435389988 | 46 | 3.6545893136748e+25+0i |

**Centrality result for G2**

For the G2 network in Table 4.2, the list of top ten ranked nodes for direct contact with immediate neighbour was the same for the proximity ranking (closeness centrality) and in the same order. Nodes 2, 20,16 and 7 remained the first, second, third and fourth ranked nodes in that order across all measures of centrality except betweenness centrality where the ranking changed. Node 2 still remained first and most central but node 7 was ranked as the second most central node on the shortest path between all pairs of vertices within the network. Nodes 16 and 20 came third and fourth.

TABLE 4.2. Centralities for top Ten nodes in the G2 network

|  |  |  |  |
| --- | --- | --- | --- |
| **Nodes** | **Degree Centrality** | **Nodes** | **Closeness Centrality** |
| 2 | 144 | 2 | 0.00117096018735363 |
| 20 | 139 | 20 | 0.00116414435389988 |
| 16 | 138 | 16 | 0.00116279069767442 |
| 7 | 136 | 7 | 0.00116009280742459 |
| 23 | 134 | 23 | 0.00115740740740741 |
| 1 | 131 | 1 | 0.00115340253748558 |
| 10 | 130 | 10 | 0.00115207373271889 |
| 6 | 125 | 6 | 0.0011454753722795 |
| 8 | 122 | 8 | 0.00114025085518814 |
| 15 | 118 | 15 | 0.00113636363636364 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Nodes** | **Betweenness Centrality** | **Nodes** | **Eigenvector Centrality** |
| 2 | 3012.89490363648 | 2 | 1 |
| 7 | 2814.33919746579 | 20 | 0.966495886574699 |
| 16 | 2798.2368423417 | 16 | 0.94877191580041 |
| 20 | 2677.92544222972 | 7 | 0.942071667057397 |
| 23 | 2625.60955797193 | 1 | 0.941616221043923 |
| 1 | 2467.95868632008 | 10 | 0.927125510223354 |
| 10 | 2270.75158067759 | 23 | 0.92404624329002 |
| 8 | 2248.84511491384 | 6 | 0.913457170984825 |
| 6 | 2009.26204792141 | 11 | 0.889381405486946 |
| 33 | 2003.08664635614 | 3 | 0.86625885555901 |

|  |  |
| --- | --- |
| **nodes** | **Subgraph Centrality** |
| 2 | 1.78642436176751e+21 |
| 20 | 1.66872453999066e+21 |
| 16 | 1.60808230965235e+21 |
| 7 | 1.58544988086325e+21 |
| 1 | 1.58391727499887e+21 |
| 10 | 1.53554198223585e+21 |
| 23 | 1.52535891325108e+21 |
| 6 | 1.49059963891501e+21 |
| 11 | 1.41306039183919e+21 |
| 3 | 1.34054070997388e+21 |

**4.5. COMPARISON WITH THE JAPANESE INTERBANK NETWORK**

Assuming the Japan interbank network follows a random network- G1’s pattern of formation and a central node in the network (almost every node is important because of the high connections between nodes unlike G2) is infected with some potential credit default or bankruptcy, it can pose a huge treat to the entire network because on average each node in the network will be connected to as much as 33 other nodes causing a quick spread of an epidemic. As seen from the SIR model, an infection spreads within a random network G1 faster than in the Barabasi - Albert network model and it quickly reaches its peak in the network.

However, if the Japan interbank network has a topology in the form of a BaraBasi-Albert network G2, there will be a few nodes in the network called hubs that are central to the network. If these hubs get infected then the entire system gets infected but if the larger population of nodes with a few connections get faced with an epidemic of credit default risk or bankruptcy chances are that it does not spread all through the network. In other cases the risk dies out after contact with a few least connected nodes before spreading across the entire network.

**5.0. CONCLUSION**

The G1 network taken as a directed graph followed a poisson distribution both for the in-degree and the out-degree while the G2 network followed a power- law distribution for both the in and out degrees distributions. The random network aided in propagating epidemics faster than the G2 network which had a few nodes with very high degrees. It is easier to identify and monitor central nodes with a potential to cause systemic risk within a network in a G2 network than in G1 network because in the G2 network all the measures of centrality identified the same set of nodes as very central in the network regardless of their ranking but for G1 network, each centrality measure identified a different set of central nodes in the network only a few measures produced sets of central nodes containing similar node entities.

**REFERENCES**

Allen, Franklin, and Ana Babus. "Networks in finance." (2008). <http://anababus.net/research/survey.pdf>

Anderson RM, May RM (1991) Infectious diseases of humans: Dynamics and control. Oxford: Oxford University Press. 757 p.

Chinazzi, Matteo, and Giorgio Fagiolo. "Systemic risk, contagion, and financial networks: A survey." Contagion, and Financial Networks: A Survey (April 2, 2013) (2013). <http://www.econstor.eu/bitstream/10419/89353/1/740929119.pdf>

Ganesh A, Massoulie L, Towsley D. The effect of network topology on the spread of epidemics. INFOCOM 2005. 24th Annual Joint Conf. IEEE Computer Communications Societies. 2005;2: 1455–466. Availabel at : <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1498374>

Ernesto Estrada, Juan A. Rodríguez-Velázquez: Subgraph centrality in Complex Networks. Physical Review E71, 056103 (2005).

Ernesto Estrada About the Discriminant Power of the Subgraph Centrality and Other Centrality Measures (Working paper). Available at: <https://arxiv.org/ftp/arxiv/papers/1305/1305.6836.pdf>

Kanno, Masayasu. "The network structure and systemic risk in the Japanese interbank market." Japan and the World Economy 36 (2015): 102-112. <http://www.sciencedirect.com/science/article/pii/S0922142515000365>

Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. Proc R Soc Lond A 115: 700–721.

**APPENDIX**

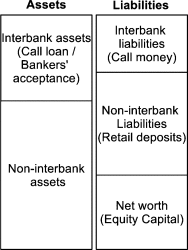
****

Figure 1.1 a sample balance sheet for financial network analysis

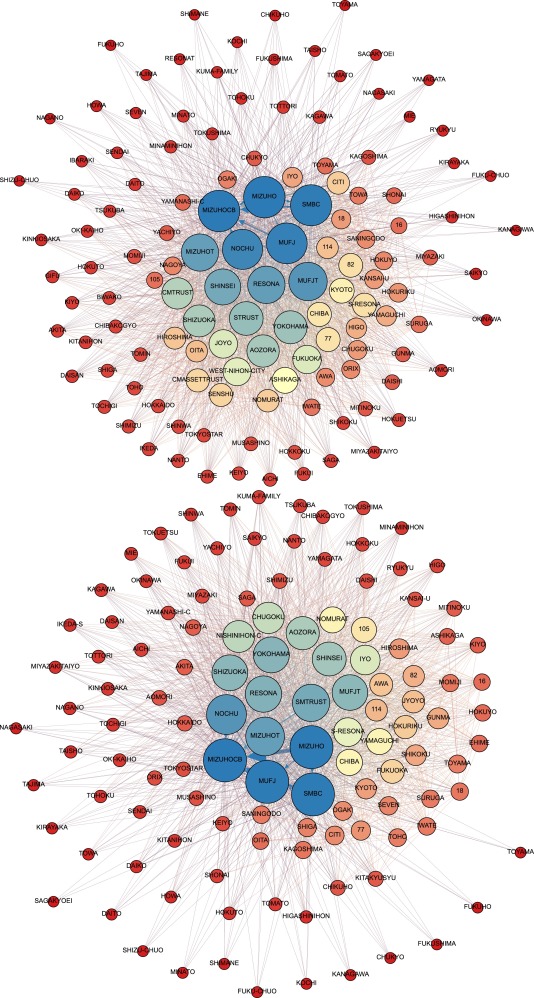


Figure 1.3 Upper network represents the Japanese interbank network for the year 2009 and lower network represents the Japanese interbank network for the 2013

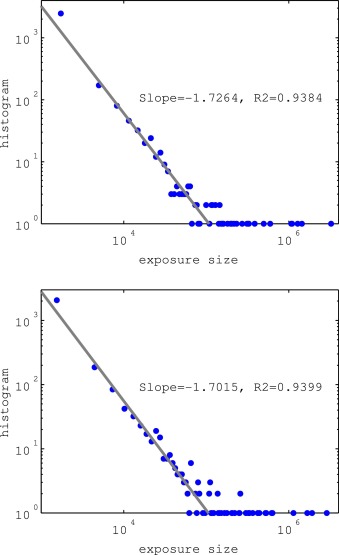


Figure 1.2 The fitted distribution for Japanese banking network for the years 2009(upper graph) and 2013(lower graph)

**Table 1**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Top 20 Japanese banks ranked in terms of betweenness centrality. | | | | | | | | | |
| March 2009 |  |  |  |  |  | March 2013 | |  |  |
| Id | Label | Betweenness (%) | Closeness | Eccentricity | Id | Label | Betweenness (%) | Closeness | Eccentricity |
| 1 | Mizuho | 100.0 | 1.00 | 1 | 1 | Mizuho | 100.0 | 1.00 | 1 |
| 16 | MizuhoCB | 100.0 | 1.00 | 1 | 16 | MizuhoCB | 89.3 | 1.02 | 2 |
| 9 | SMBC | 82.9 | 1.03 | 2 | 5 | MUFJ | 81.3 | 1.04 | 2 |
| 5 | MUFJ | 74.8 | 1.05 | 2 | 9 | SMBC | 81.3 | 1.04 | 2 |
| 30 | Nochu | 67.9 | 1.05 | 2 | 30 | Nochu | 47.3 | 1.11 | 2 |
| 288 | MUFJT | 48.1 | 1.11 | 2 | 294 | SMTrust | 30.7 | 1.14 | 2 |
| 10 | Resona | 42.0 | 1.11 | 2 | 289 | MizuhoT | 28.6 | 1.32 | 2 |
| 289 | MizuhoT | 31.1 | 1.37 | 2 | 10 | Resona | 22.1 | 1.21 | 2 |
| 397 | Shinsei | 28.5 | 1.24 | 2 | 138 | Yokohama | 19.2 | 1.25 | 2 |
| 294 | STrust | 21.0 | 1.21 | 2 | 288 | MUFJT | 19.2 | 1.23 | 2 |
| 138 | Yokohama | 20.2 | 1.18 | 2 | 149 | Shizuoka | 14.4 | 1.40 | 2 |
| 398 | Aozora | 19.9 | 1.15 | 2 | 397 | Shinsei | 12.2 | 1.32 | 2 |
| 149 | Shizuoka | 13.0 | 1.32 | 2 | 398 | Aozora | 9.2 | 1.39 | 2 |
| 291 | CMTrust | 9.7 | 1.42 | 2 | 17 | Saitama Resona | 7.2 | 1.23 | 2 |
| 177 | Fukuoka | 8.3 | 1.24 | 2 | 174 | Iyo | 4.8 | 1.42 | 2 |
| 130 | Joyo | 3.8 | 1.52 | 2 | 134 | Chiba | 2.7 | 1.32 | 2 |
| 158 | Kyoto | 3.7 | 1.18 | 2 | 170 | Yamaguchi | 2.5 | 1.32 | 2 |
| 190 | Nishi-Nippon City | 2.9 | 1.55 | 2 | 190 | Nishi-Nippon City | 1.8 | 1.68 | 2 |
| 401 | Citibank | 2.3 | 1.03 | 2 | 177 | Fukuoka | 1.1 | 1.30 | 2 |
| 134 | Chiba | 2.1 | 1.29 | 2 | 168 | Chugoku | 1.0 | 1.70 | 2 |

R code used for the assignment

|  |
| --- |
| ######### CODE FROM IGRAPH http://igraph.org/r/doc/00Index.html####  #### Edited by BROWN NATHANIEL BODUNRIN #######  ##### CONTRIBUTOR SCOT BEE #######  install.packages("devtools")  library(devtools)  install.packages("igraph")  install.packages("tckl")  install.packages("rgl")  library(rgl)  install.packages("rgexf")  library(rgexf)  # create a sample igraph object  g <- erdos.renyi.game(500, 0.13, directed = FALSE,loops = FALSE)  plot(pref)  # construct the nodes and edges data for gexf conversion  nodes <- data.frame(cbind(V(g), as.character(V(g))))  edges <- t(Vectorize(get.edge, vectorize.args='id')(g, 1:ecount(g)))  # do the conversion  gephiER<-write.gexf(nodes, edges)  print(gephiER, file="mygraph1.gexf")  nodes <- data.frame(cbind(V(pref), as.character(V(pref))))  edges <- t(Vectorize(get.edge, vectorize.args='id')(pref, 1:ecount(pref)))  # do the conversion  gephiER<-write.gexf(nodes, edges)  print(gephiER, file="pref.gexf")  ## G1 random networks used for the assignment  ##using density of 13%  ##set.seed is used to reproduce the same network  set.seed(14)  G1=erdos.renyi.game(500, 0.13 , type = "gnp", directed = T,loops = FALSE)  gsize(G1)  ## preferential attachment with m= 18  ## working paper used 127 nodes with network density of 13%  ## this implies 2080 edges in the network and approximately m=18 to achieve this  ## number of connections  set.seed(13)  G2= barabasi.game(500, m= 18, directed = F)  gsize(G2)  ## ploting the G1 and G2  plot(G2, vertex.label= NA, edge.arrow.size=0.02,vertex.size = 0.5, xlab = "Scale-free Network")  ### the code below was used to plot an histogram of total number of edges in for nultiple iterartions of the random network  replicate(10000, erdos.renyi.game(500, 0.13 , type = "gnp", directed = FALSE,loops = FALSE), simplify = FALSE) %>%  vapply(gsize,0) %>%  hist(main="10000 iterations of random networks", xlab="number of connections")  ### to find the mean of multiple iterations of a generated random network  connections<-NULL  n.times <- 10000  for (i in 1:n.times) {  r <- erdos.renyi.game(500,0.13, type = "gnp", directed = FALSE,loops = FALSE)  connections[i]<-gsize(r)  }  mean(connections)  hist(connections, main="10000 iterations of random networks", xlab="number of connections")  #### to verify the density of the network  graph.density(G1)  ### Average path length  mean\_distance(G1, unconnected = TRUE)  ##average clustering coefficient Transitivity of the Networks  transitivity(G2, type = "global")  ##### The degree distribution of the network  ## for undirected network, remove the mode argument  outDegree = degree(G2, v = V(G2),mode= "out",loops = FALSE, normalized = FALSE)  inDegree = degree(G2, v = V(G2),mode= "in",loops = FALSE, normalized = FALSE)  degtable= data.frame(InDegree = c(inDegree), OutDegree=c(outDegree))  subdeg=head(degtable[order(degtable$InDegree, decreasing=TRUE), ], 10)  write.table(subdeg, file = "degreetable.csv", sep = ",", col.names = NA,  qmethod = "double")  fit1 <- fit\_power\_law(inDegree+1)  fit2 <- fit\_power\_law(inDegree+1, implementation="R.mle")  fit1$alpha  stats4::coef(fit2)  dist=degree.distribution(G1, v = V(G1), mode= "in",loops = FALSE)  plot(dist, main="in-degree distribution", xlab="k" ,ylab= "p(K)")  ### Centralities  DC=centr\_degree(G1)  EC=eigen\_centrality(G1, directed = TRUE)  CC=closeness(G1, vids = V(G1), mode = c("total"),  weights = NULL, normalized = FALSE)  BC=betweenness(G1, v = V(G1), directed = TRUE, weights = NULL,  nobigint = TRUE, normalized = FALSE)  SC=subgraph\_centrality(G1, diag = F)  nodes=1:500  data= data.frame(nodes= c(nodes), DegCen = c(DC$res), CloseCen=c(CC), BetwCen=c(BC),EigenCen=c(EC$vector),SubCen=c(SC))  head(data)  ##we select the top ten centralities for all measures except for subgraph where we  ##choose the tail ten  subdata=head(data[order(data$DegCen, decreasing=TRUE), ],10)  ## where c(1,5) means the first(nodes) and fifth(eigenvector centrality) columns  ##in data  ms=subdata[,c(1,2)]  ## to write the centralities into an excel file  write.table(ms, file = "deg.csv", sep = ",", col.names = NA,  qmethod = "double")  ###Cliques and modularity  oc <- cluster\_optimal(rand)  modularity(oc)  modularity(g, membership(oc))  modularity(rand, membership, weights = NULL)  ##### Generate multiple iterations of the SIR simulations  ## undirected graphs were used for the sir simulation  ##due to igraph's warnings that "Edge directions are ignored in SIR model"  result<-NULL  results<-NULL  n.times <- 1000  Matrix<-matrix(ncol=2 , nrow=n.times)  colnames(Matrix) <- c("No of Infected","Time")  sm <- sir(G1, beta=0.25, gamma=1, no.sim=1000)  for (i in 1:n.times) {  result[i]<-(max(sm[[i]]$NI))  results[i]<-sm[[i]]$times[head(which(sm[[i]]$NI==max(sm[[i]]$NI)), n=1)]  Matrix[i,1]<-as.numeric(result[i])  Matrix[i,2]<-as.numeric(results[i])  }  tsummary=apply(Matrix, 2, mean)  tsummary  ### plotting the graph of a the SIR simulations  plot(sm, comp = "NI", median = FALSE, main="G2 Ro = 1",  quantiles = NULL, color = "Red", median\_color = NULL,  quantile\_color = NULL, lwd.median = 2, lwd.quantile = 2,  lty.quantile = 3, xlim = NULL, ylim = NULL, xlab = "Time",  ylab = "Population")  lines(sm[[1]]$times, sm[[1]]$NR, col = "green", lwd = 2)  lines(sm[[1]]$times, sm[[1]]$NS, col = "blue", lwd = 2)  legend('topright', c("Number of susceptible","Number of infected","Number of recovered"),lty=1, col=c('blue','red','green'), bty='n', cex=.75) |