

DISA

Dipartimento di Informatica
e Studi Aziendali

2008/11

Structure of Business Firm Networks and Scale-Free Models

*Maksim Kitsak, Massimo Riccaboni,
Shlomo Havlin, Fabio Pammolli
and H. Eugene Stanley*



UNIVERSITÀ DEGLI STUDI
DI TRENTO

DISA Working Paper

DISA

Dipartimento di Informatica
e Studi Aziendali

2008/11

**Structure of Business Firm Networks and
Scale-Free Models**

*Maksim Kitsak, Massimo Riccaboni, Shlomo Havlin,
Fabio Pammolli and H. Eugene Stanley*



UNIVERSITÀ DEGLI STUDI
DI TRENTO

DISA Working Paper

DISA Working Papers

The series of DISA Working Papers is published by the Department of Computer and Management Sciences (Dipartimento di Informatica e Studi Aziendali DISA) of the University of Trento, Italy.

Editor

Ricardo Alberto MARQUES PEREIRA ricalb.marper@unitn.it

Managing editor

Roberto GABRIELE roberto.gabriele@unitn.it

Associate editors

Flavio BAZZANA	flavio.bazzana@unitn.it	Finance
Michele BERTONI	michele.bertoni@unitn.it	Financial and management accounting
Pier Franco CAMUSSONE	pierfranco.camussone@unitn.it	Management information systems
Luigi COLAZZO	luigi.colazzo@unitn.it	Computer Science
Michele FEDRIZZI	michele.fedrizzi@unitn.it	Mathematics
Andrea FRANCESCONI	andrea.francesconi@unitn.it	Public Management
Loris GAIO	loris.gai@unitn.it	Business Economics
Umberto MARTINI	umberto.martini@unitn.it	Tourism management and marketing
Pier Luigi NOVI INVERARDI	pierluigi.noviinverardi@unitn.it	Statistics
Marco ZAMARIAN	marco.zamarian@unitn.it	Organization theory

Technical officer

Paolo FURLANI paolo.furlani@unitn.it

Guidelines for authors

Papers may be written in English or Italian but authors should provide title, abstract, and keywords in both languages. Manuscripts should be submitted (in pdf format) by the corresponding author to the appropriate Associate Editor, who will ask a member of DISA for a short written review within two weeks. The revised version of the manuscript, together with the author's response to the reviewer, should again be sent to the Associate Editor for his consideration. Finally, the Associate Editor sends all the material (original and final version, review and response, plus his own recommendation) to the Editor, who authorizes the publication and assigns it a serial number.

The Managing Editor and the Technical Officer ensure that all published papers are uploaded in the international RepEc publication database. On the other hand, it is up to the corresponding author to make direct contact with the Departmental Secretary regarding the offprint order and the research fund which it should refer to.

Ricardo Alberto MARQUES PEREIRA

Dipartimento di Informatica e Studi Aziendali

Università degli Studi di Trento

Via Inama 5, TN 38100 Trento ITALIA

Tel +39-0461-882147 Fax +39-0461-882124

E-mail: ricalb.marper@unitn.it

Structure of Business Firm Networks and Scale-Free Models

Maksim Kitsak,¹ Massimo Riccaboni,² Shlomo
Havlin,^{1,3} Fabio Pammolli,^{1,4} and H. Eugene Stanley¹

¹*Center for Polymer Studies, Boston University,
Boston, Massachusetts 02215, USA*

²*Department of Computer and Management Studies, University of Trento, Trento, Italy*

³*Minerva Center and Department of Physics,
Bar-Ilan University, Ramat Gan, Israel*

⁴*IMT Institute for Advanced Studies, Lucca, Italy*

(Dated: December 24, 2008)

Abstract

We study the structure of business firm networks in the Life Sciences (LS) and the Information and Communication Technology (ICT) sectors. We analyze business firm networks and scale-free models with degree distribution $P(q) \propto (q + c)^{-\lambda}$ using the method of k -shell decomposition. We find that the LS network consists of three components: a “nucleus,” which is a small well connected subgraph, “tendrils,” which are small subgraphs consisting of small degree nodes connected exclusively to the nucleus, and a “bulk body” which consists of the majority of nodes. At the same time we do not observe the above structure in the ICT network. Our results suggest that the sizes of the nucleus and the tendrils decrease as λ increases and disappear for $\lambda \geq 3$. We compare the k -shell structure of random scale-free model networks with the real world business firm networks. The observed behavior of the k -shell structure in the two industries is consistent with a recently proposed growth model that assumes the coexistence of both preferential and random regimes in the evolution of industry networks.

PACS numbers: 89.75.Hc

I. INTRODUCTION

Many real-world complex systems are described as graphs or networks [1–7]. Networks appear in various areas of science such as physics, biology, computer sciences, economics and sociology [3–7]. Real networks, despite their diversity, share many common properties. Some real networks have been found to be “small-world” [3, 8–10]: despite their large size, the typical distance between nodes is very small, of order of $\log(N)$ or less [3, 11, 12], where N is the number of nodes. Also, some real networks are scale-free (SF) with a power-law tail in their degree distribution

$$P(q) \propto (q + c)^{-\lambda}, \quad (1)$$

where q is the number of links per node, λ and c are distribution parameters [13–16].

Many statistical physics and mathematical techniques have been successfully applied to study network structure including percolation [12, 17–21], scaling [6, 7, 11], partitioning [22], box covering [23–26], k -core percolation [27, 28] and k -shell decomposition [29]. The latter has been recently used to study the topology of the Internet at the autonomous system (AS) level [30]. It has been found that the AS Internet network can be decomposed into three distinct components:

- (i) The nucleus consists of highly connected nodes (hubs) and constitutes a tiny fraction of the network ($\approx 0.5\%$).
- (ii) The tendrils ($\approx 25\%$) are the subgraphs that connect to the network exclusively via the nucleus.
- (iii) The remaining nodes ($\approx 75\%$) of the network constitute the bulk body. Unlike the tendrils, the nodes in the bulk remain connected even if the nucleus is removed.

The appearance of the three components led to the association of the AS Internet structure with that of a jelly-fish model [30]. For an alternative jelly-fish structure see also Ref. [31]. The nucleus plays a crucial role in data transfer throughout the Internet. If the nucleus is damaged or blocked, nodes in the tendrils cease to communicate with other nodes both in the bulk body and in other tendrils. However, the nodes in the bulk body can still communicate with each other, but less efficiently since the average path length between them almost doubles [30]. The network of workplaces in Sweden [32, 33] was also shown to

possess a jelly-fish structure. Mean field analysis of the k -shell structure [34, 35] focuses on random networks with a given degree distribution $P(q)$.

In the present work we adopt the same methodology to address the long-standing question of how long does a typical leader in an industry maintain its position. This question has attracted continuing attention in the economics and management literatures over the past generation [36]. Two rival views exist. The first asserts that leadership tends to persist for a 'long' time while the second one emphasizes the transience of leadership positions due to radical innovation and "leapfrogging competition". The Life Sciences (LS) sector is generally cited as an archetypal example of stability of industry leaders [37], while on the contrary the Information and Communication Technology (ICT) sector is widely considered as a good example of high market turnover and instability [39]. We study the k -shell structure of the LS and ICT industry in order to test these views and better understand the reasons behind the higher stability of the leading firms in the LS sciences as compared to the ICT sector. The networks of LS and ICT firms are networks of nodes representing firms in the worldwide LS and ICT industries with links representing the collaborative agreements among them [26, 37–39]. We also conduct a numerical analysis of model SF networks in order to better understand their k -shell structure. We study how the size and the connectivity of the three components of random SF model networks depend on the SF parameters. We compare the k -shell structure of SF models with that of the LS and ICT industry networks in order to get a better insight into the growth principles of the industry networks.

The rest of the manuscript is organized as follows: in Section II we define the k -shell decomposition and apply it to both real and model SF networks. We then analyze the leadership in the industry networks from the k -shell perspective. In Section III we conduct a systematic analysis of the properties of the three components of model SF networks. In section IV we compare the results for SF models with those observed in the two industrial networks. We conclude our manuscript in Section V with a discussion and summary.

II. K -SHELL, K -CORE, AND K -CRUST

We start the process of the k -shell decomposition [30, 34] on a network by removing all nodes with degree $q = 1$. After the first iteration of pruning, there may appear new nodes with degrees $q = 1$. We keep on pruning these nodes until only nodes with degree $q \geq 2$ are

left. The removed nodes along with the links connecting them form the $k = 1$ shell. Next, we iterate the pruning process for nodes of degree $q = 2$, thereby creating the $k = 2$ shell. We further continue the k -shell decomposition for higher values of q until all nodes of the network are removed. As a result each node in the network is assigned a k -shell index k . The largest shell index is called k_{\max} , which is also the total number of shells in the network, provided all shells below k_{\max} exist. The k -crust is defined as the union of all k -shells with indices smaller than k . Similarly, the k -core is defined as the union of all nodes with indices greater or equal k . (See Fig. 1 for demonstration.) As we explain later, the set of nodes in the k_{\max} shell is called nucleus provided there is a large number of nodes (tendrils) connected to the network exclusively via the k_{\max} shell.

We analyze the k -shell structure of LS and ICT industrial sectors in time periods between 1990 and 2002 and between 1990 and 2000 respectively. Data have been collected from several sources including: Repec, Windhover, Bioscan, Pharmadeals, company press releases and news. The LS network expanded linearly since the mid-1970s while the ICT network took off in the 1990s and grew exponentially for a decade [see Fig. 2(a) and the inset]. The total number of firms in the LS is $N = 6,776$ and in the ICT is $N = 7,759$. These sizes refer to the largest connected component of each of the networks at the last year of observation. Both industrial networks feature SF degree distributions with $\lambda \approx 2.5$, $c \approx 4$. (LS) and $\lambda \approx 3.4$, $c \approx 6$. (ICT) which we estimate with the maximum likelihood method [see Fig. 2(b)]. We use the Kolmogorov-Smirnov test in order to examine the goodness of fit of the degree distributions. The obtained p -values for 1000 trials are 0.24 and 0.33 for the LS and the ICT industry networks respectively.

We next apply the k -shell decomposition procedure to the LS and the ICT networks. For each k -crust we calculate N_0 , the total number of nodes in the k -crust, N_1 , the size of the largest connected component in the k -crust and N_2 , the size of the second largest component of the k -crust. As seen in Fig. 3, the LS network consists of $k_{\max} = 19$ shells while the ICT network (which consists of comparable number of firms) has only $k_{\max} = 6$ shells. The size of the largest cluster in the k -crust starts growing rapidly after $k = 4$ and $k = 2$ in the LS and the ICT networks respectively. At these values of k , the size of the second largest cluster in the k -crust reaches a maximum. The above behavior of the k -crust components is consistent with the existence of a second order phase transition in the k -crust structure [30]. The type of the phase transition taking place at the k -shell decomposition is similar

to that of targeted percolation in SF networks [21, 40].

Unlike in the ICT network, the size of the largest connected component N_1 in LS network undergoes a large jump $N_1(k = 19) - N_1(k = 18) = 745$, while the total size N_0 of the crust experiences only a small change $N_0(k = 19) - N_0(k = 18) = 43$ (See Fig. 3(a)). This can be explained as follows [30]. Approximately 10% of the LS network are firms (which we call “tendrils”) that prefer to sign collaborative agreements exclusively with the 43 firms that form the shell $k = k_{max} = 19$ (which we call the nucleus). Typically, each of these firms (tendrils) signs a small number of agreements with firms in the nucleus and, therefore, has small degree. Thus, the tendrils are removed in the decomposition of the first few shells. However, being connected exclusively to the nucleus in the $k_{max} = 19$ shell, the tendrils do not contribute to the largest connected component of the k -crust until the last k_{max} shell is decomposed. It is the inclusion of the tendrils firms into $N_1(k)$ at the decomposition of the k_{max} shell that results in the observed jump in $N_1(k)$. The appearance of the three components—the nucleus, the tendrils and the bulk body—allows one to associate the structure of the network of LS firms with that of the jelly-fish similar to the Internet at the Autonomous System level [30]. Interestingly, we do not observe the a similar jump in $N_1(k)$ in the ICT network (See Fig. 3(b)).

In both LS and ICT industry sectors the k_{max} shells include market leaders such as Pfizer, GSK, Novartis, J&J, Sanofi-Aventis, Bayer in the LS network and Microsoft, IBM, AT&T, Yahoo, Cisco, AOL, Time Warner and Google in the ICT network. We find that the LS network tendrils are mostly composed of the new start up firms and their university partners. On the one hand, start up firms preferentially attach to market leaders. On the other hand, market leaders compete to sign exclusive deals with new and promising start up firms. We also notice the remarkable stability of the LS nucleus: once a particular firm enters the nucleus it is very likely to remain there for many years.(See Fig. 3c) On the contrary, in the ICT sector there is more emphasis on how to integrate different technologies and markets [39]. Thus, there are more deals between firms of the same size and age but in different technological and market areas. As a consequence, the k -shell of the ICT network is more unstable and heterogeneous which may explain why we do not detect the emergence of a nucleus such as in the LS sector.

As discussed above, one can in general calculate the size of the nucleus S_n as the total number of nodes in the k_{max} shell

$$S_n = N_0(k_{\max}) - N_0(k_{\max} - 1). \quad (2)$$

The increase in $N_1(k)$ at $k = k_{\max}$ is comprised by the inclusion of tendrils S_t and the nucleus S_n . Thus, the size of tendrils can be calculated as

$$S_t = N_1(k_{\max}) - N_1(k_{\max} - 1) - S_n. \quad (3)$$

As seen in Fig. 2(b), both the LS and the ICT industrial networks exhibit a SF degree distribution. Hence, in order to better understand the substructures of real networks—the nucleus, the tendrils, and the bulk components—we analyze the k -shell structure of the random SF models which were generated using the configurational approach [41]. We calculate the k -shell structure of random SF networks with $c \geq 0$, and degree distribution exponent $\lambda \in [2, 3]$. For our simulations (Fig. 4) we choose networks sizes of $N = 8,000$ (which is comparable to the size of the LS and ICT industry networks) and $N = 10^6$ in order to test the influence of finite size effects.

Fig. 4 shows that both the number of shells k_{\max} and the jump in the largest connected component $N_1(k)$ decrease as the exponent λ increases. As the jump becomes less pronounced it becomes harder to detect it motivating us to introduce a quantitative criterion for the emergence of the three distinct components. We define the rate of change of the largest connected component size as

$$R(k) \equiv N_1(k) - N_1(k - 1). \quad (4)$$

We compare the increase of the largest connected component $R(k)$ at $k = k_{\max}$ with that at $k = k_{\max} - 1$. The jump in $N_1(k)$ results in

$$R(k_{\max}) > R(k_{\max} - 1). \quad (5)$$

We use Eq. (5) as a criterion for the existence of a nucleus and tendrils.

By examining $R(k)$ plot we observe in the case of $N = 8,000$ (Fig. 4c), that SF models with $\lambda > 2.5$ already do not have a nucleus and tendrils. However, for $N = 10^6$ we do observe nucleus and tendrils in SF models for $\lambda \leq 2.7$. (Fig. 4d). These observations suggest that all SF networks with $2 < \lambda < 3$ have a nucleus and tendrils provided N is sufficiently large. The above observations for SF models agree with the fact that we observe jelly-fish topology for LS network ($\lambda = 2.5$) and do not observe it for ICT network $\lambda = 3.4$.

However, SF model networks with $N = 8000$ with $\lambda = 2.5$ and $\lambda = 3.4$ have $k_{max} = 5$ and $k_{max} = 2$ respectively, while the measured k_{max} in LS and ICT networks (which have the same λ values) are $k_{max} = 19$ and $k_{max} = 6$. The observed difference in k_{max} between industry networks and SF models with similar parameters suggests to further explore the k -shell structure of SF networks and consider SF model networks with $c > 0$ values in $P(q)$ (as found in the LS and the ICT networks).

III. K -SHELL PROPERTIES OF SCALE-FREE NETWORKS

SF networks may or may not have a jelly-fish structure, depending on the degree distribution $P(q)$ and size N . In order to better understand the k -shell structure of SF networks, we calculate the sizes of the k -shells as a function of N and λ . For each pair of values N and λ we generate $10^3 - 10^4$ realizations of SF models and calculate the average number of shells k_{max} constituting the network as well as their average size S_k (see Fig. 5). For small k values, S_k is proportional to N . As the size of the network increases new shells start to appear. When the size increases further, the new shell growth stabilizes and becomes also proportional to N (See Fig. 5a,c). This result indicates that the size of each shell constitutes a certain finite fraction of the network and this fraction decreases with increasing k . The analytical analysis of k -core structure [34] leads to

$$S_k \propto k^{-\delta}, \quad (6)$$

where

$$\delta = \frac{2}{3 - \lambda}. \quad (7)$$

Hence $\delta \approx 2.2$ for $\lambda = 2.1$ and $\delta \approx 4.0$ for $\lambda = 2.5$, which agrees with our simulations. [Figs. 5(b) and 5(d)]. The appearance of $c > 0$ in the SF degree distribution significantly increases k_{max} [see Figs. 5(e) and 5(f)]. However, the asymptotic dependence of k -shell sizes seems to remain the same as k approaches k_{max} : $S_k \sim k^{-\delta}$ [Figs. 5(e) and 5(f)].

One can estimate the total number of shells in a random SF network of size N as follows. Since every shell constitutes a fixed fraction of N it follows that $S_k \propto Nk^{-\delta}$. The last shell k_{max} needs to possess at least one node $S_{max} \equiv S_{k=k_{max}} \sim 1$, which leads to

$$k_{max} \leq N^{1/\delta}, \quad (8)$$

Indeed, the total number of shells k_{\max} seems to increase as a power law with the network size N [Fig. 6(a)]. The smaller is λ the faster is the growth of k_{\max} . As seen from Fig. 6(c), the estimated exponents $1/\delta$ agree with those predicted by Eq. (8). Note that Eq. (8) together with

$$\delta = 2/(3 - \lambda) \quad (9)$$

is consistent with the fact that networks with $\lambda > 3$ do not have a k -shell structure.

The dependence of S_{\max} on N can be regarded as a crossover from the small N regime with $N < N_c(\lambda)$, where there is no nucleus, to the power-law regime for $N > N_c(\lambda)$ where $S_{\max} \sim N^{\tau(\lambda)}$ (Fig. 6b). We relate the observed crossover with the emergence of the nucleus and tendrils in SF networks for $N > N_c(\lambda)$. The critical size of SF networks, $N_c(\lambda)$, corresponding to the emergence of the nucleus and tendrils in SF networks seems to increase as λ increases. In the $N < N_c$ regime the size of the last shell S_{\max} increases with λ , which can be explained by the fact that SF networks with higher λ have fewer shells. On the other hand, in the power-law regime, the size of the last shell, S_{\max} , (which now becomes the nucleus S_n) is smaller for larger values of λ .

IV. EVOLUTION AND STRUCTURE OF THE LS AND ICT INDUSTRY NETWORKS

We further analyze the k -shell structure of the LS and ICT industry networks. The number of shells k_{\max} in the LS industry grows as a power-law function of its size,

$$k_{\max} \sim N^\theta, \quad (10)$$

and reaches $k_{\max} = 19$ in 2002. Our estimates yield $\theta \approx 0.6$ (Fig. 7a). We find that the number of shells in ICT sector, k_{\max} , also grows and reaches $k_{\max} = 6$ in 2000. Note that S_n for the LS network exhibits fluctuations for $N \in [2000, 5000]$ and stabilizes for $N > 5000$ (See Fig. 7b). As seen in Fig. 7(c,d), the sizes of shells S_k decrease as a function of their index k in both networks. As we notice in Section II, the observed shell sizes S_k as well as the number of shells k_{\max} , measured for the LS and ICT networks, deviate from those obtained from SF models with $c = 0$. We expect for random SF models with $c = 0$, $\lambda = 2.5$ and $\lambda = 3.4$ and similar sizes as LS and ICT to find $k_{\max} = 5$ and $k_{\max} = 2$ respectively. Also, a SF network with $\lambda = 2.5$ is expected to have $S_k \sim k^{-4}$ and $k_{\max} \sim N^{0.25}$ which deviates

from the observed shell sizes. The measured Kolmogorov-Smirnov statistic is $D \approx 0.137$. [see Fig. 7(c)]. The observed differences can be explained by taking into account the offset c in SF degree distribution of both networks. The adjustment of the degree distribution of random SF models with $c = 4.0$ and $c = 6.0$ allows one to obtain similar patterns S_k which are in fair agreement with the industry networks, as seen in Figs. 7(c) and 7(d) (The Kolmogorov-Smirnov statistic in this case, $D \approx 0.074$).

As seen above, the offset $c > 0$ in the SF degree distribution plays a crucial role in the formation of the k -shell structure of industry networks. A possible reason for the emergence of $c > 0$ in growing networks is the combination of preferential attachment with random attachment in network evolution [42]. The coexistence of preferential attachment regime with random collaborative agreements was suggested to take place in industry networks [37]. The random component is caused by the fact that sometimes firms choose exclusive relationships and novelty, and do not prefer to make deals with hub firms. Even though the coexistence of both preferential and random regimes seems to be crucial in the formation of the industry networks, it does not fully reproduce the k -shell structure of LS and ICT industries. We believe that a better understanding of the evolution of the LS and ICT networks may be achieved by further improvements of the modeling.

V. DISCUSSION AND SUMMARY

We use the k -shell decomposition to analyze the structure of the LS and ICT industry networks. We find that the firms in the LS industry can be naturally divided into three components: the nucleus, the tendrils and the bulk body. The nucleus of the LS industry consists mostly of the market leader firms while the tendrils are typically comprised of small start-up firms that preferentially make deals with leading firms which are in the nucleus. We show that the nucleus of the LS industry exhibits remarkable stability in time. In contrast, the ICT industry does not have a nucleus. We also analyzed the dependence of the k -shell structure of SF model networks on N , λ and c . We observed the formation of the nucleus and the tendrils in SF networks only for $\lambda < 3$. The number of shells k_{\max} and the size of the nucleus S_n are larger for SF networks with $c > 0$ compared to those with $c = 0$. Our results can partly explain the k -shell structure of LS and ICT industry networks. The coexistence of preferential and random attachment leads to the appearance of the offset $c > 0$ in the

SF degree distribution $P(q) \sim (q + c)^{-\lambda}$ [42]. Thus, the appearance of $c > 0$ in the degree distribution of LS and ICT networks might be explained by the interplay of random and preferential agreements among firms in the industries [37].

Acknowledgments

We thank ONR, European project DAPHNET, the Israel Complexity Center, Merck Foundation (EPRIS project) and Israel Science Foundation for financial support. We thank L. Braunstein, S. Carmi and L. K. Gallos for valuable discussions.

-
- [1] P. Erdős and A. Rényi, *Publ. Math. Inst. Hung. Acad. Sci.* **6**, 290 (1959). [I](#)
 - [2] P. Erdős and A. Rényi, *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17 (1960).
 - [3] B. Bollobas, *Random Graphs* (Cambridge University Press, 2001). [I](#)
 - [4] S. Wasserman and K. Faust, *Social Network Analysis*, (Cambridge University Press, 1999).
 - [5] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks* (Oxford University Press, 2003).
 - [6] M. Newman, A. L. Barabási and D. J. Watts, *The Structure and Dynamics of Networks* (Princeton University Press, 2006). [I](#)
 - [7] S. Bornholdt and H. G. Schuster *Handbook of Graphs and Networks* (WILEY-VCH, New York, 2001). [I](#), [I](#)
 - [8] S. Milgram, *Psychol. Today* **2**, 60 (1967). [I](#)
 - [9] D. J. Watts and S. H. Strogatz, *Nature* **393**, 440 (1998).
 - [10] R. Albert, H. Jeong, and A.-L. Barabási, *Nature* **401**, 130 (1999). [I](#)
 - [11] R. Cohen and S. Havlin, *Phys. Rev. Lett.* **90**, 058701 (2003). [I](#), [I](#)
 - [12] R. Cohen and S. Havlin *Complex Networks: Structure, Stability and Function*(Cambridge University Press, Cambridge, in press 2008). [I](#), [I](#)
 - [13] H.A. Simon, *Biometrika* **42**, 425 (1955). [I](#)
 - [14] D. de S. Price, *J. Amer. Soc. Inform. Sci.* **27**, 292, (1976).
 - [15] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
 - [16] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002). [I](#)

- [17] S. Kirkpatrick, *Rev. Mod. Phys.* **45**, 574 (1973). [I](#)
- [18] D. Stauffer and A. Aharony, *Introduction to Percolation Theory, 2nd Edition* (Taylor and Francis, New York, 2004).
- [19] A. Bunde and S. Havlin, *Fractals and Disordered Systems, 2nd Edition* (Springer, Berlin, 1996).
- [20] R. Cohen, K. Erez, D. B. Avraham, and S. Havlin, *Phys. Rev. Lett.* **85**, 4626 (2000).
- [21] D. S. Callaway, M. E. J. Newman, S. H. Strogatz and D. J. Watts, *Phys. Rev. Lett.* **85**, 5468 (2000). [I](#), [II](#)
- [22] G. Paul, R. Cohen, S. Sreenivasan, S. Havlin and H. E. Stanley, *Phys. Rev. Lett.* **99**, 115701 (2007); Y. Chen, E. Lopez, S. Havlin, and H. E. Stanley, *Phys. Rev. Lett.* **96**, 068702 (2006). [I](#)
- [23] C. Song, S. Havlin, and H. Makse, *Nature* **433**, 392 (2005). [I](#)
- [24] C. Song, S. Havlin, and H. Makse, *Nature Physics* **2**, 275 (2006).
- [25] K. I. Goh, G. Salvi, B. Kahng, and D. Kim, *Phys. Rev. Lett.* **96**, 018701 (2006).
- [26] M. Kitsak, S. Havlin, G. Paul, M. Riccaboni, F. Pammolli, and H. E. Stanley, *Phys. Rev. E* **75**, 056115 (2007). [I](#)
- [27] B. Bollobas, *Graph Theory and Combinatorics: Proceedings of the Cambridge Combinatorial Conference in honor of P. Erdős*, 35 (Academic, New York, 1984) [I](#)
- [28] S. B. Seidman, *Social Networks*, **5**, 269 (1983). [I](#)
- [29] J. I. Alvarez-Hamelin, L. Dall'Asta, A. Barrat and A. Vespignani, *Networks and Heterogeneous Media* **3**, 371 (2008). [I](#)
- [30] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir, *Proc. Natl. Acad. Sci. USA* **104**, 11150 (2007). [I](#), [II](#), [3](#)
- [31] L. Tauro, C. Palmer, G. Siganos, and M. Faloutsos, *Global Internet* (Nov. 2001). [I](#)
- [32] WWW.SCB.SE [I](#)
- [33] Y. Chen, G. Paul, R. Cohen, S. Havlin, S. P. Borgatti, F. Liljeros and H. E. Stanley, *Phys. Rev. E* **75**, 046107 (2007). [I](#)
- [34] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Phys. Rev. Lett.* **96**, 040601 (2006). [I](#), [II](#), [III](#)
- [35] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, *Phys. Rev. E* **73**, 056101 (2006). [I](#)
- [36] J. Sutton, *American Economic Review*, **97**, 222 (2007). [I](#)

- [37] L. Orsenigo, F. Pammolli, and M. Riccaboni, *Res. Policy* **30**, 485 (2001). [I](#), [IV](#), [V](#)
- [38] F. Pammolli and M. Riccaboni, *Small Business Economics* **19(3)**, 205 (2002).
- [39] M. Riccaboni and F. Pammolli, *Research Policy* **31**, 1405 (2002). [I](#), [II](#)
- [40] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, *Phys. Rev. Lett.*, **86**, 3682 (2001). [II](#)
- [41] M. Molloy and B. Reed, *Random Struct. Algorithms* **6**, 161 (1995). [II](#)
- [42] R. Albert and A. L. Barabási, *Phys. Rev. Lett.* **85**, 5234 (2000). [IV](#), [V](#)

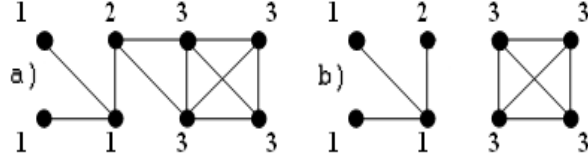


FIG. 1: Illustration of the k -shell decomposition method. (a) Original network. Nodes are marked by corresponding k -shell indices. Note that the k -shell index does not coincide with the node degree. (b) The 3-crust (left) and the 3-core (right) of the original network.

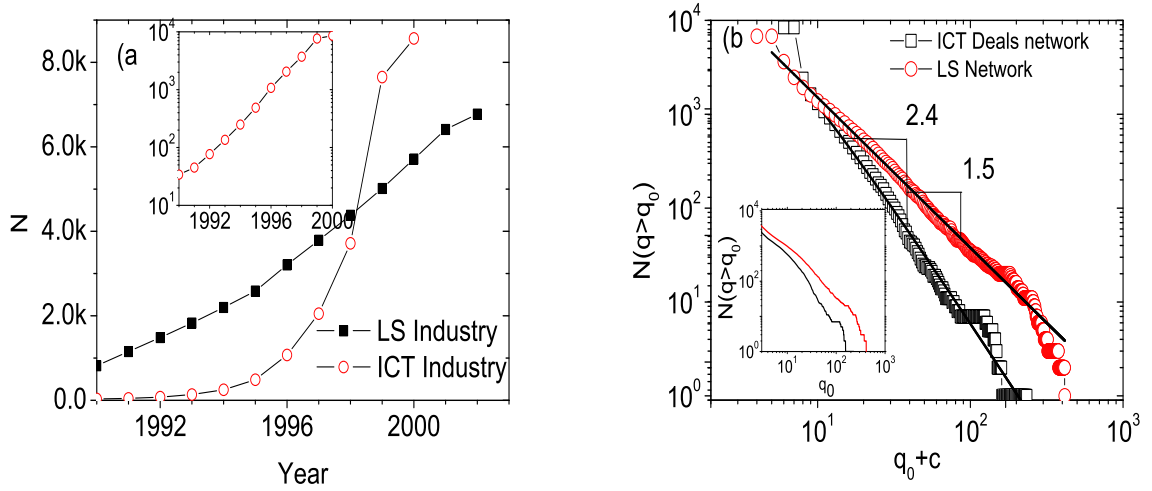


FIG. 2: (a) The growth of the largest connected component of LS and ICT industries. The LS industry expands almost linearly while the ICT industry exhibits exponential growth (see inset, a semi-log plot of the ICT network size as a function of time). (b) Cumulative degree distribution $N(q > q_0)$ of LS and ICT networks. The inset displays cumulative degree distribution $N(q > q_0)$ of LS and ICT network without offset c .

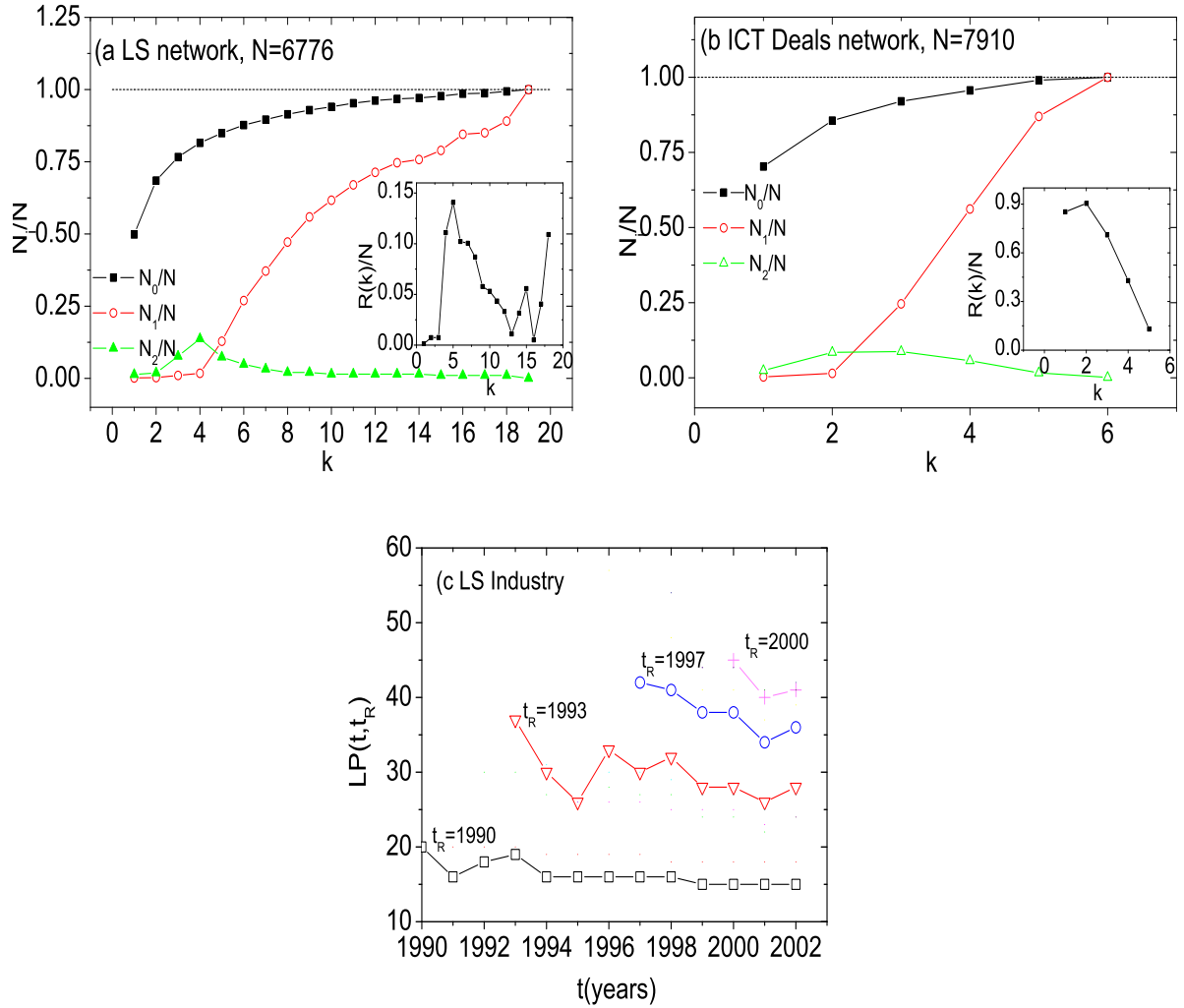


FIG. 3: (a) The k -shell structure of the LS network and (b) of the ICT network. Plots of the size of the k -crust, N_0 ; the size of the largest connected component of the k -crust, N_1 ; and the size of the second largest connected component of the k -crust, N_2 , as a function of the k -crust index k . The sizes are normalized with the total number of nodes in the network, N . N_2 is multiplied by 100. The transition of the k -crust at shell $k = 18$ of the LS network reveals a jelly-fish topology [30]. Since the change in $F_1 \equiv N_1/N$ at the $k = k_{\max}$ is smaller in the ICT network than in the previous shell $k = k_{\max} - 1$, the ICT does not have a jelly-fish structure. The insets of (a,b) show the rate R of the largest connected component change as a function of shell index k . (c) The leadership persistence $LP(t, t_R)$ as a function of years in the LS industry. We define $LP(t, t_R)$ as a number of firms that were in the nucleus both at time t_R and t . Note that most of the LS firms remain in the nucleus after they first enter it.

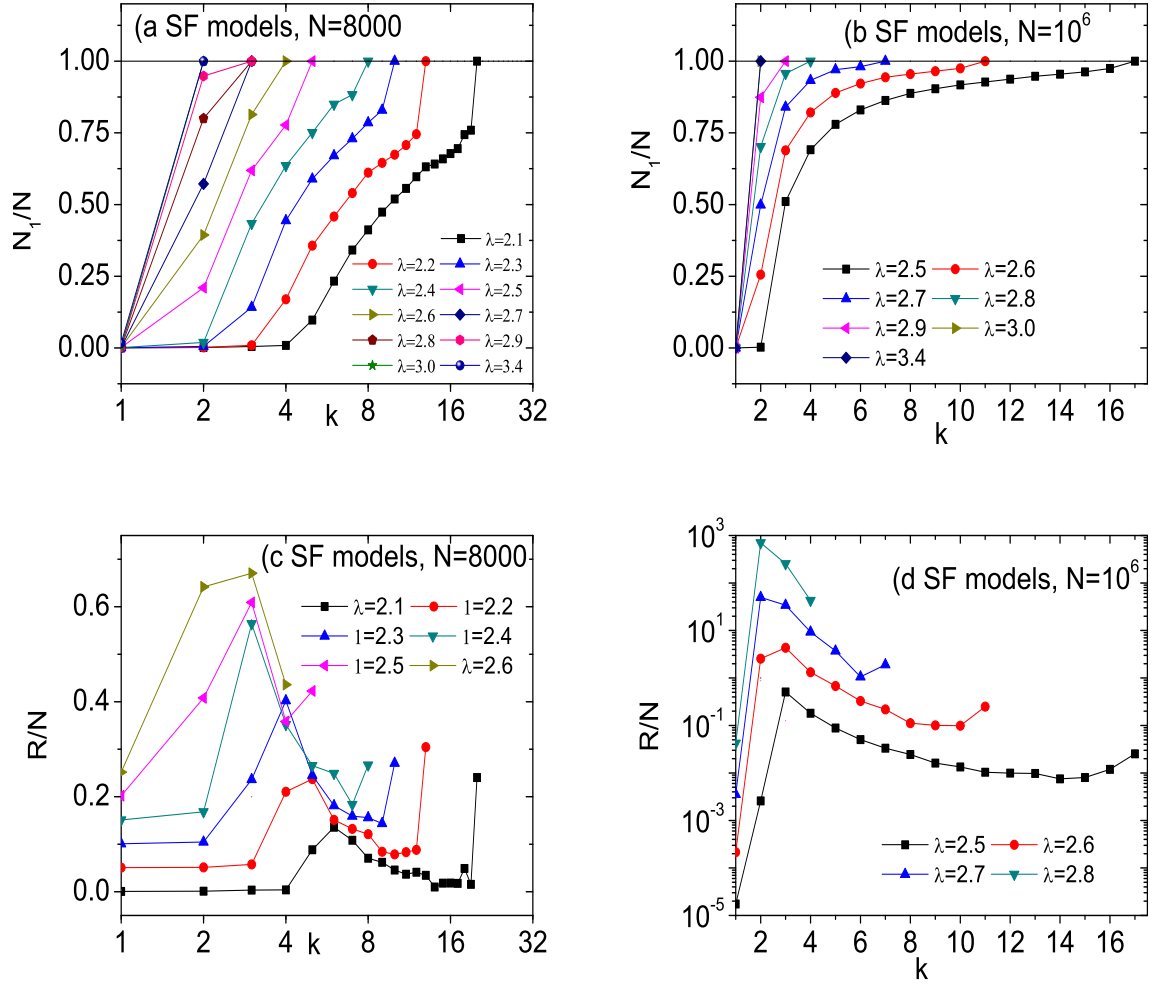


FIG. 4: (a,b) The k -shell structure of random SF models with (a) $N = 8000$ and (b) $N = 10^6$ nodes are shown for comparison. Note that $\lambda = 3.4$ curve overlaps with that of $\lambda = 3.0$. (c,d) The rate of the largest connected component change, R , as a function of shell index k for (c) $N = 8000$ and (d) $N = 10^6$. Note that in order to avoid the overlap of curves in (c,d) we subsequently shift the plots with respect to each other by the additive factor 0.05 in (c) and the multiplicative factor of 10 in (d). The sizes of the nucleus and the tendrils decrease as λ increases. The nucleus and the tendrils disappear in (a,c) for $N = 8000$ at $\lambda > 2.5$ and in (b,d) for $N = 10^6$ at $\lambda > 2.7$.

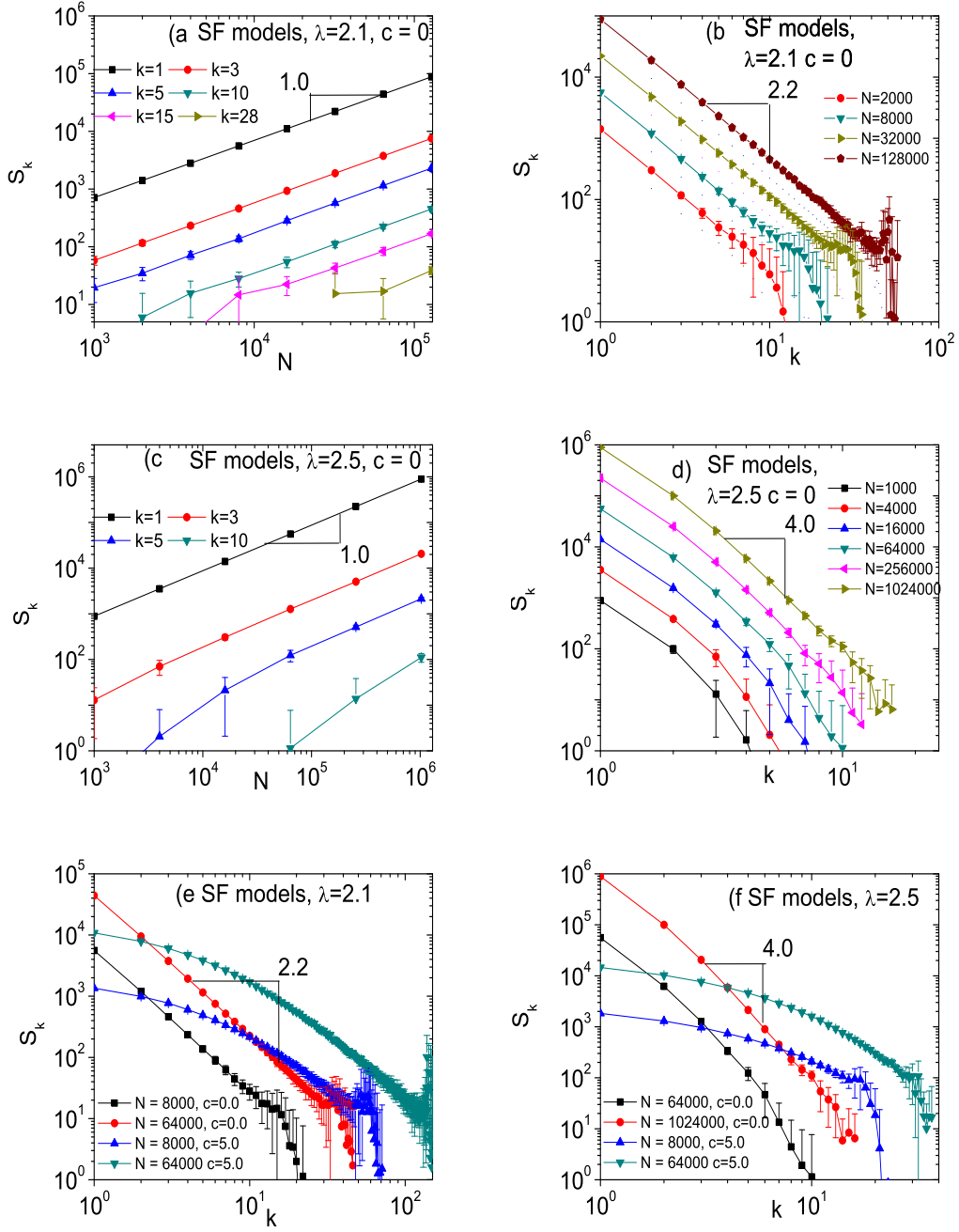


FIG. 5: Sizes of k -shells, S_k , for SF random models with $c = 0$ (a,b) $\lambda = 2.1$ and (c,d) $\lambda = 2.5$ as a function of (a,c) N and the k -shell index (b,d) k . Note that sizes of k -shells increases proportionally to N . (e,f) Sizes of k -shells, S_k , for SF networks with $c = 5.0$ and (e) $\lambda = 2.1$, (f) $\lambda = 2.5$. SF models with $c = 5.0$ have significantly larger number of shells, k_{\max} , compared to SF models with the same λ and $c = 0$. S_k in SF models with $c = 5.0$ decreases significantly slower as a function of k compared to SF models with the same λ and $c = 0$ in the small k region. In the large k region both types of SF models with $c = 0.0$ and $c = 5.0$ seem to have similar behavior $S_k \sim k^{-\delta}$, where $\delta = 2/(3 - \lambda)$.

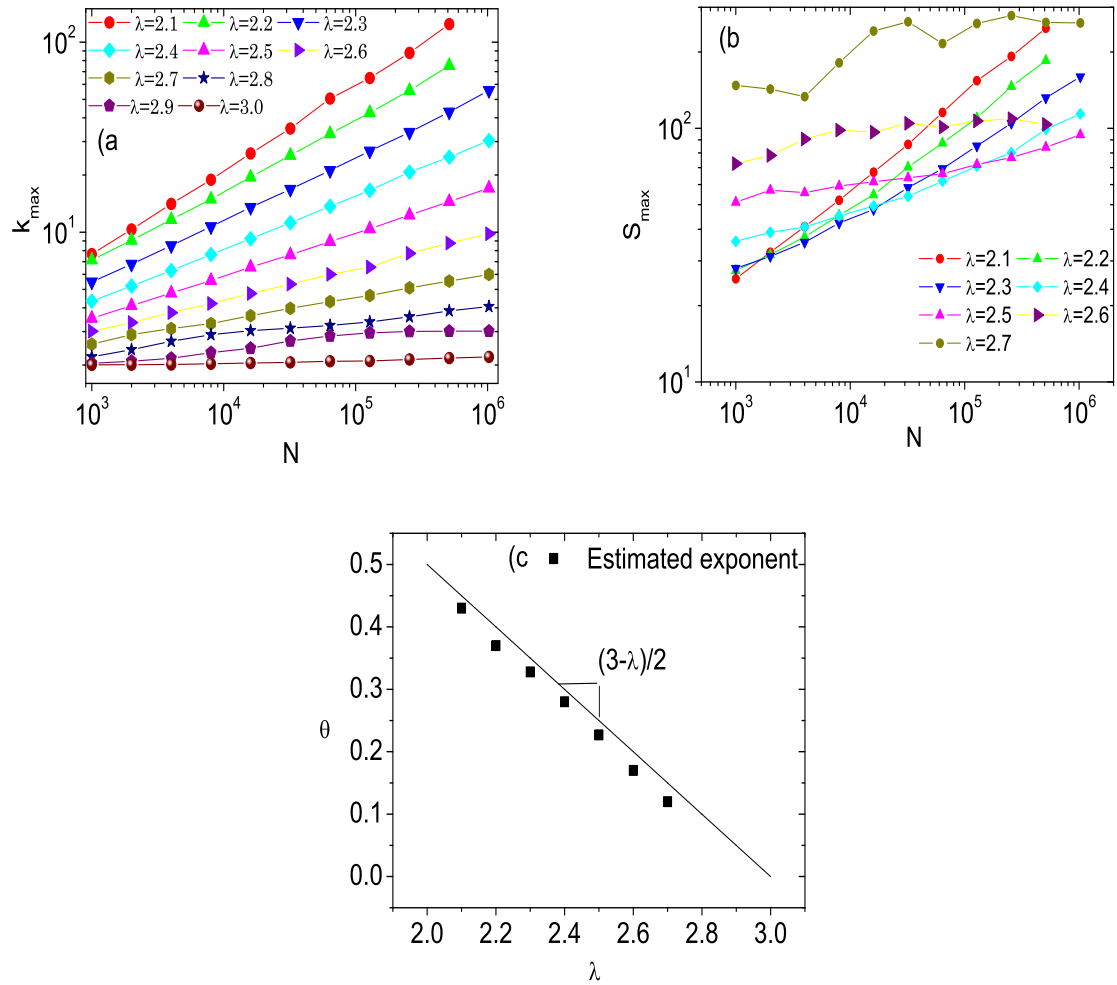


FIG. 6: (a) The total number of shells, k_{\max} , in SF models as a function of N . Note that $k_{\max} \propto N^\theta$. (b) The size of the last shell S_{\max} in the SF model as a function of N . Each curve crosses over into a power law regime for $N \geq N_c(\lambda)$, where $N_c(\lambda)$ increases with λ . (c) The calculated exponent θ as a function of λ (symbols). Our calculated values of δ agree with the mean field theory result $\delta \approx 2/(3 - \lambda)$ (solid line).

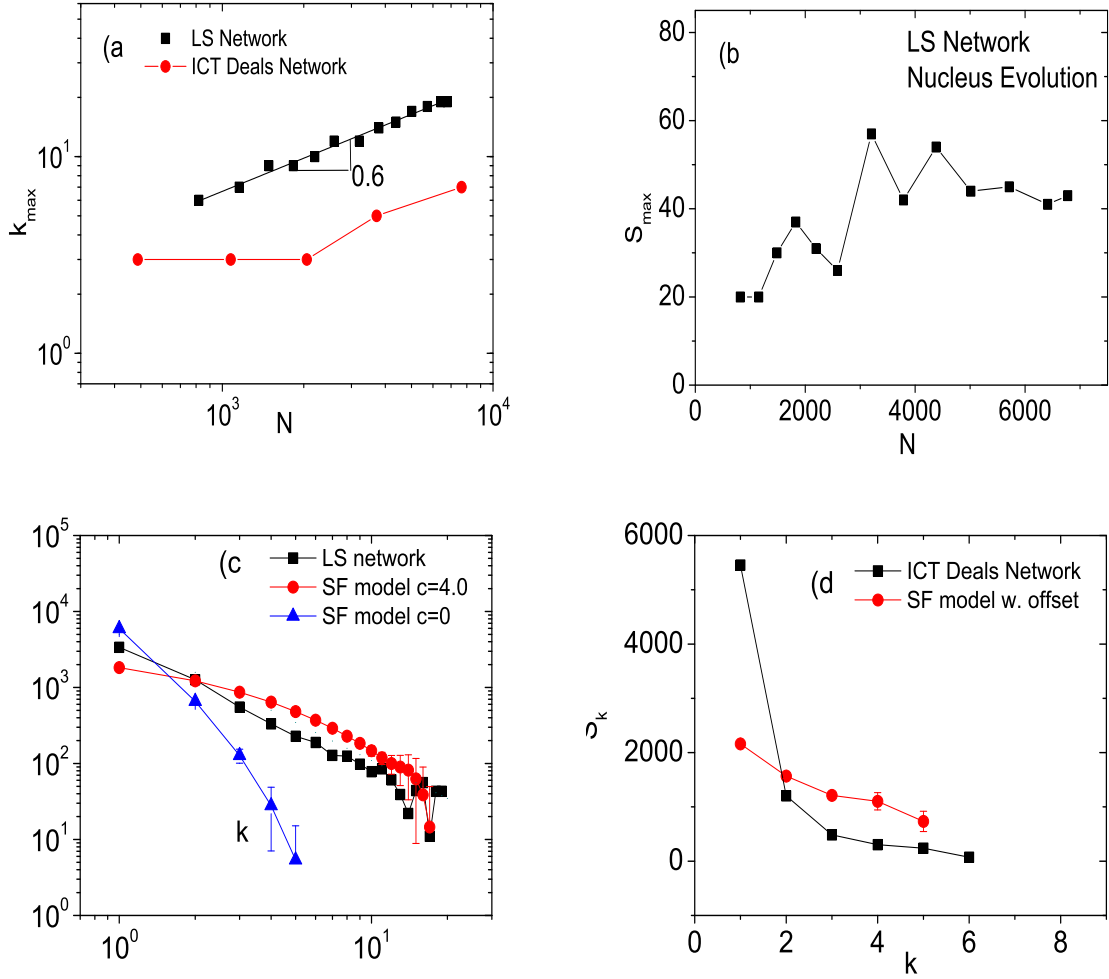


FIG. 7: (a) The largest shell index k_{\max} of the LS and the ICT networks as a function of N calculated for different years. The number of shells k_{\max} of the LS network increases approximately as a power-law function of the network size $k_{\max} \sim N^\theta$, where $\theta \simeq 0.6$. (b) Size of the nucleus of the LS network, S_n , as a function of N . S_n exhibits fluctuations as LS network grows. Unlike in the analyzed SF models, S_{\max} becomes stable for $N > 5000$. (c) S_k as a function of the k -shell index k for the LS network (squares). Shell sizes S_k as a function of k decrease significantly slower than $S_k \sim k^{-4}$, which is expected for a random SF model with the same λ and $c = 0$ (triangles). However, the offset introduction of $c = 4.0$ in the SF degree distribution, $P(q)$, mimics the k -shell structure of the LS network (circles). (d) S_k as a function of k -shell index k for the ICT network (squares). SF model with $\lambda = 3.4$ and $c = 0$ does not possess a k -shell structure. However, the introduction of $c = 6.0$ in the SF degree distribution yields similar k -shell structure (circles), but we attribute it only to finite size effect.

