World Journal of Management
Volume 2. Number 3. September 2010. Pp. 81 – 100

The role of institutions and firm characteristics in the networks of firm publications. An analysis of the Italian biotech sector through the Social Network Analysis

Rosamaria d'Amore*, Roberto Iorio*^ and Agnieszka Stawinoga**

In this paper we analyse, through the instrument of the social network analysis, the network of co-authorships in the publications of the firms belonging to the Italian life-science biotech sector. The aim of the paper is to analyse the role of the different institutions (universities, research centres, hospitals and firms) inside such network. We try to conduct this analysis in depth, taking into consideration also the nature and characteristics of the biotech firms. We use a classification of the biotech firms according to the criteria suggested by OCED and we analyse the characteristics of the networks generated by the publications of each typology of firm, focusing the attention on the role played by the different institutions inside each of them. Therefore, crossing two dimensions of the analysis (nature of the institutions, characteristics of the firms) we try to shed new light on the architecture of the sectoral system of innovation, that may be considered a relevant part of the whole Italian national system of innovation. The analysis shows that the Italian innovation system is based on a balance among different institutions, each of them having a prevalent function. The central role is covered by the universities, particularly of the great universities in the Northern Italy, the more industrialised part of the country; their prominent role, among the different kinds of institutions, is common to all the typologies of biotech firms, with particularly strong ties with R&D biotech firms and a limited exception in the network of publications of the so-called targeted biotech firms that, being wholesaler, are more interested in collaboration with institution oriented to the "application" of biotechnologies. Research centres have an important role in bridging different institutions, as it is possible to infer from their frequent presence in large and heterogeneous networks of co-autorship. Collaboration with hospitals is less systematic, but they have a prominent role in more applied research, as it is demonstrated by the central role they have in the network of publications of the targeted firms. This analysis also shows a possible point of weakness of the Italian system of innovation, represented by the low degree of collaboration among different firms, at least on the point of view of co-autorship of scientific publication in the biotech sector.

Key-words: Innovation in biotech sector; sectoral system of innovation; firm publications; networks of co-authorship; OECD classification of biotech sector.

1. Introduction

Biotech is a strongly science-based sector, where the production of new knowledge and new products is absolutely usual. As nowadays it commonly happens, such production of new knowledge happens as the result of collaborations, among firms or between firms and the institutions devoted to the "production" of science (universities, research centres, etc.). Another characteristic that is nowadays common to many science-based sectors is that at least part of the new knowledge produced by the firm is frequently disclosed through the instrument of the "open science" (publications on scientific journals, conferences, etc.). As a joint result of this two points, firms often publish co-authored papers.

*University of Salerno;** University of Naples-Federico II
^corr. auth.: riorio@unisa.it

In this paper we analyse the network of co-authorships in the publications of the firms belonging to the Italian life-science biotech sector. We take into consideration the institutions the authors of the papers belong to, classifying them into four categories: universities, research centres, hospitals and firms. The aim of the paper is to analyse the role of the different institutions inside the network of publications, through the instrument of the Social Network Analysis (SNA).

We try to develop this analysis in depth, taking into consideration also the characteristics of the firms. We use to this purpose a classification of the Italian biotech firms done by D'Amore and Vittoria (2008, 2009), according to the criteria suggested by the OECD, that are based on the kind of activity mainly conducted by a biotech firm. Biotech firms are therefore classified in six categories (plus a residual one) and we are able to analyse the seven networks of co-authorships of publications for each typology of firm. We are therefore able to compare the characteristics of such networks, and the role played by the different institutions inside each of them.

In this way, crossing two dimensions of the analysis (nature of the institutions, characteristics of the firms) we try to shed new light on the architecture of the sectoral system of innovation (Malerba, 2002), that, given the high level of research conducted in the biotech sector and its high level of innovativeness, may be considered as a relevant part of the whole Italian national system of innovation.

This kind of analysis may have important policy implications. Facing the economic crisis, many countries have the need to rationalise the expenditure in education and research. Because of the crucial importance of this kind of expenditure for economic and civil growth, there is the need to have a deep knowledge of the research and innovation system, in order to take correct decisions. This kind of study, focused on a country, like Italy, that shares with many advanced countries the condition of a high level of technology, but not at a leadership level, may be for many countries an useful example of such attempts to understand in depth important parts of the national innovation system.

The paper is articulated in five section. After this introduction, the second section reports a review of the literature about the relevant issues for this paper: the first sub-section is about the generation and exchange of knowledge in the biotech sector; as such exchange often takes the form of co-publication, in the second sub-section we focus the attention on the use of the SNA to study the phenomenon of the co-authorship; the third sub-section illustrates the OECD classification of the biotech sector. The third section describes the sources of the data and introduces some methodological principles of the SNA. The fourth section reports the results of the empirical analysis: after some descriptive statistics, the second sub-sections illustrates the results of the SNA; a brief synthesis of the results follows. Some final considerations conclude the paper.

2. Literature review

2.1. Research collaborations and publications in the biotech sector

The biotech sector is characterized by a complex knowledge base, where the sources of expertise are widely dispersed. Network relations are frequently used to access this knowledge. As Powell et al. (1996) argue, the locus of innovation will be found in networks rather than in individual firms. Biotech rely mostly on inter-organizational collaborations. There are many organizations where it is possible to find the knowledge, the expertise useful for the firm: it is possible to find it in the universities, in the research centres, in the hospitals. Therefore the innovative networks usually include all the actors of the specific sectoral system of innovation, that, given the scientific and commercial importance of the biotechnologies, is an important part of the “national systems of innovation”. According to the triple helix vision, there are intensive scientific collaborations between universities, industrial organization and government agencies (Etzkowitz and Leydesdorff, 2000; Etzkowitz et al., 2000) and particularly universities may increasingly function as a locus of national knowledge intensive network.

Biotech sector is not only multi-disciplinary, but it is multi–institutional as well. In fact, in addition to research universities, both start-up and established firms, government agencies, non profit research institutes and leading hospitals play a key role in conducting and funding research (Powell et al., 1996). Notwithstanding this articulated institutional framework, universities keep a key role: a large fraction of biotechnology firms originated from universities or at least depend on linkages with universities for their competitive success (Audretsch and Stephan, 1996; Powell et al., 1996; Zucker et al., 1998, 2002; Stuart and Sorenson, 2003; Owen-Smith and Powell, 2004; Stuart et al., 2007).

The new knowledge generated by these collaborations not only takes the form of industrial innovations, but is often disclosed trough the scientific publications: research collaborations often generate co-authored publications. Over two-thirds of even formal alliance partners in this field also appear as partners in scientific publications (Gittelman, 2006) and there is a close link between successful patents and scientific publications in this field (Gittelman and Kogut, 2003; Murray and Stern, 2007).

Therefore, if the aim is to study the dynamics of the knowledge exchanges and of the innovative networks inside a technological field, considering that data on publications are usually of high quality and easy to access, it is possible to study the publications of the firms.

2.2. The use of the Social Network Analysis to study the co-author relationship

The SNA is a tool useful to analyse, in many situations, how individuals or organizations are related. It is a multidisciplinary methodology, developed mainly by sociologists and researchers in social psychology in the 1960s and 1970s. The SNA is based on the assumption of the importance of relationships among interacting units or nodes. Trough the shape of a SNA we can determine a network’s usefulness to its individuals, understand the linkages among social entities and the implications of these linkages.

One way of studying such networks in academic research communities is to conduct co-citation analysis, where the links are established through the way authors refer to one others’ research and publications (Horn et al., 2004; Lin, 1995). Another good way to study similar networks was observed by Newman (2001a,b; 2004) who studied co-authorship networks and research collaborations within academic research communities to understand collaboration network patterns and characteristics.

In this paper we adopted the co-authorship analysis rather than the co-citation analysis, because the co-authorship more directly reflects the nature and structure of formal relationships among members of a research community (Newman, 2001b). Co-authorship (i.e., collaboration in researching and publishing journal article) is an important primary descriptor of scientific publications. Fox and Faver (1984) indicated that the average number of authors per paper has risen to nearly two in business literature and that 67% of papers have more than one author, paralleling a similar increase in the rate of co-authorship throughout academia. Studying co-authorships and their patterns can be an useful mean for understanding collaborative relationships in a research community.

The study of scientific collaboration helps to establish group and work networks that can be analyzed and evaluated through bibliometric techniques and represented in what some authors call co-authorship networks or bibliometric maps. These analyses, applied to the study of co-authorship and collaborative relationships between institutions for scientific publications, allow the existing relations between the social agents responsible for the publications to be identified and represented graphically, setting out the number of members in the network, the intensity of the relationships existing between them and who the most relevant members are with respect to a wide range of measures or indicators.

The peculiarity of our study is that it is conducted at an institutional level: an example of an empirical study, trough the SNA, of co-autorships networks at an institutional level may be found in Chinchilla-Rodriguez et al. (2008).

2.3. Identifying firm typologies: the OECD classification of the biotech firms

One of the aims of this paper is to verify if different kinds of firm generate different networks of knowledge exchange. The biotech sector may be considered as intrinsically heterogeneous, therefore we can observe the effects of different firm characteristics, even inside a unique technological domain.

One of the main characteristics of the biotech sector is its multidisciplinarity. Biotechnologies cover a broad range of knowledge fields; in fact there are many different definitions of the sector. The most frequently used definition is given by OECD: "Biotechnology consists in the use of scientific and engineering principles (based on microbiology, genetic, biochemistry, chemical and biochemical engineering) to transform materials using biological agents (such as micro organism, enzyme, animal or vegetable cells) with the purpose to obtain good and service" (OECD, 1989).

The OECD Statistical Framework for technology (OECD 2005 a, b) also defined biotech activities, identifying six classes. The main distinction is between *production* and *service* activities; than, among production activities, it distinguishes between active, innovative and dedicated biotech firms, in order to identify activities more or less focused on biotech. More in detail: a b*iotechnologically active firm* (BAF) is defined as a firm engaged in key biotechnology activities, like the application of at least one biotech technique to produce goods or services and/or the performance of biotechnology R&D; a *dedicated biotech firm* (DBF) is a BAF whose predominant activity involves the application of biotech techniques to produce goods or services and/or the performance of biotech R&D; an *innovative biotech firm* (IBF) is defined as a BAF that applies biotech techniques for the purpose of implementing new products or processes. Among service activities, the distinction is between R&D, market and other service oriented firms. More in detail: a *biotechnology R&D firm* is a firm completely focused on R&D, with no product sales; it is classified by Italian national statistical offices into the R&D service industry category; a *targeted firm* is a firm classified as wholesaler, for instance local operations of large foreign pharmaceutical firms, whose local affiliate performs biotechnology research, but acts mainly by a wholesale distributor; an *other types of services firm* uses biotech techniques for the purpose of providing a services (for example waste management and environmental remediation).

3. Data and methodology

3.1. Sources of data

In order to build a database of scientific publications in the biotech sector we made an intersection of two databases: *i*) RP Biotech data base; *ii*) ISI Web of Science. They are briefly described in the following.

RP Biotech data base. It is a collection of potentially all the Italian firms belonging to the biotech sector according to the OECD definition (see description and detail in the previous paragraph), active at December 2005. This database collects 865 firms; 501 of them are for profit firms, 364 are no profit firms: in this study we considered only the 306 life-science for profit firms. The firms have been classified according to the OECD criteria described before; those firms that it was not possible to classify in one of the OECD typology have been classified in the residual category of the “out” firms.
ISI databases, especially the Science Citation Index®, and the web-based version Web of Science® (in the following pages WoS), provide the best source of information to identify the basic research activity across all countries and fields of science. It is a detailed bibliometric database of journal articles and citations of worldwide research literature, that contains 14 000 international peer-reviewed scientific and technical journals.

We obtained information about publications of the selected firms, across the period 2003-2005. The record of each publication in ISI-Web of knowledge reports, among other kinds of information, the name of the authors and the name of the institutions the authors belong to. We extracted all the publications where the name of at least one of the selected firms (Italian life-science for-profit biotech firms) appeared among the institutions of affiliation. Then, in order to develop our analysis at the institutional level, we divided the institution in five categories (universities, research centres, hospitals, Italian life-science for-profit biotech firms[1], other firms) and established what category each institution belonged to.

3.2. Methodology for the Social Network Analysis

The primary step to represent our data as a network was creating an affiliation network, where the set of actors were composed by: the set of publications written by authors affiliated to the Italian biotech firms; the institutions (universities, research centres, hospitals, firms) all the authors of those publications are affiliated to. Since we were interested in analysing collaboration between Italian biotech firms and other institutions, we converted a binary incidence matrix A (the rows represent publications and the columns actors) into an adjacency matrix $Z=A^TA$. The element $Z_{(i,j)}$ gives information about the number of publications written by both actor *i* and actor *j*. The diagonal element $Z_{(i,i)}$ informs about the number of publications of actor *i*. For further analysis we decided to dichotomize matrix Z and set diagonal elements to 0. Since we did not consider a direction of ties, we obtain an undirected co-authorship network of 900 nodes and 4729 ties.

We used a categorical variable "Type of institution" to classify nodes in the five institutional categories described at the end of the previous section. To describe the scientific collaboration network among Italian biotech firms and other institutions we focused on the analysis of the network structure and on the position of actors.

In order to better understand the network structure we decided to firstly measure the density which gives the general information about the level of connectedness of the network. Next, to see if our graph is connected or not and to individualize subgroups of nodes in which the network may be partitioned, we were interested in analysing the components and the cliques of the network.

One of the essential question in network analysis is to determine the importance of a particular node or edge in a network. Several metrics have been proposed to quantify centrality of nodes and/or connectivity, the most commonly used metrics are the degree centrality, the closeness centrality and the betweenness centrality (Freeman 1979). The degree centrality considers nodes with the higher number of adjacent edges (higher degree). In the collaboration network, if we consider a binary matrix, degree is equal to the number of collaborators that an author has.

In the case of a valued matrix the degree is equal to the number of collaborations. As the valued degree measures the number of interactions, it seems to give a more interesting information than the "non-valued" degree, and this is the reason why we give more attention to the valued degree in our analysis; anyway to compare values of centralization indices they must be calculated on binary matrix.

Closeness centrality is a global metric based on the average length of the paths linking a node to others and reveals the capacity of a node to be reached. Since our whole network is disconnected, we could not obtain the value for this index for the whole graph, so we decided to obtain it for the main component of the network.
Betweenness centrality is a metric based on geodesic distances counts; it represents the nodes ability to influence or control communication in the network. The betweenness centrality focuses on the capacity of a node to be an intermediary between any two other nodes.

Centralization is a measure of the variability or dispersion of the actor centralities indices, since it compares each actor index to the maximum attained value. To study the graph centralization we compared the degree, closeness and betweenness centralization indices.

To understand if the different groups of institutions collaborate within and between each other we examined the homophily of ties in the network. We used the E-I index, which is based on comparing the numbers of ties within groups and between groups. Values of this index can range from -1, when all ties are within members of the group, to 1, when all ties are external to the group. The E-I index can be applied at three levels: the whole network, each group, and each node.

To study and represent networks we used following software for network analysis: Ucinet 6.221 (Everett, Borgatti, Freeman, 2002) and NetDraw 2.089 (Borgatti, 2002).

4. Empirical results

4.1 Some descriptive statistics

115 of the considered firms made at least one publication during the period 2003-2005. The total number of publications is 1053. The total number of the affiliation institutions of the authors is 900; besides the 115 Italian biotech firms, we identified 218 universities, 289 hospitals, 134 researcher centres and 114 other firms. The institutional co-operation in publication is very frequent: in 918 on the total number of 1053 publication (87.18%) the authors belong to more than one institution; in the others 135 publications the only institution of affiliation is one of the biotech firms. The average number of institutions per paper is 3.43. There are only two firms which did not write any papers in collaboration. As said before, the firms are classified in the categories defined by OECD. Table 1 gives some important information number of the firms belonging to each OECD typology, their attitude to publish and to make publications in collaboration.

Table 1: Propensity to publish and to cooperate by category of Italian biotech firm

	Number of firms	Nr. of publications	Average nr. of publications	Average nr. of co-authoring institutions in each paper
Innovative	42	555	13.214	3.363
Active	8	57	7.125	3.115
Dedicated	18	143	7.944	3.629
R&D	23	249	10.39	3.612
Targeted	10	33	3.3	3.303
Other services	7	23	3.286	3.13
Out	7	22	3.143	2.864
TOTAL	115	1053	9.156	3.43

We can observe that the average number of publications (third column) is quite different across the OECD classes, but the average number of institutions involved in each publication (fourth column) is quite similar.

The innovative firms have on average the highest number of publications (13.12), followed by R&D firms (10.39). Innovative firms are usually big firms, and this may partly explain the high quantity of research done, therefore the high number of publications; on the other side, R&D are on average of a smaller size, but they have research as their specific goal.

Firm size may have, on the contrary, an ambiguous effect on collaborations: big firms, particularly involved in research (as innovative firms are) on one side can do much and high-level research, involving many partners; on the other side they may have a research centre inside, where there are generally many researchers devoted to this kind of activity, therefore it is possible to think they do not need a lot of resources outside. In fact innovative firms have an intermediate average number of co-authors per publication.[2]

Table 2 shows the number of each kind of institutions the biotech firms (again divided according the OECD typology) have collaborated with.[3]

Table 2 : Co-authoring institutions by category of Italian biotech firm

	Universities	Hospitals	Research centres	Biotech firms	Other firms	Total nr. of coauthoring institutions
Innovative	152	153	78	43	78	504
Active	33	53	6	8	12	112
Dedicated	57	61	45	18	31	212
R&D	84	57	36	24	31	232
Targeted	16	49	5	10	2	82
Other services	6	7	2	7	2	24
Out	17	17	5	7	2	24

The nature of the research done by each category of biotech firm may influence the kind of institutions they collaborate more: a more basic nature of the research done by R&D firms may clearly justify the frequency of co-authorships with university, while the more applied activity done by targeted firms explains the high number of collaborations with hospitals.

An analysis with SNA may give more information about such patterns of collaborations.

4.2. Results of the Social Network Analysis

In this section we present some results of our analysis, firstly for the whole network and then for the different networks obtained dividing our firms according to their OECD typology.[4] To describe the scientific collaboration network among Italian biotech firms and other institutions we focused on the analysis of the network structure and of the position of actors.

To understand better the structure of our network we calculated the density and we individualized subgroups of nodes. The density results 0.0117 and there appears to be 11 different components, a large one of 875 vertices, and only some components in between, which range from 2 to 5 vertices, and only 2 isolates.

There are 1288 cliques with minimum size equals to 3, the biggest clique includes 20 nodes and it is created by the biotech firm Bracco Imaging and its collaborators.

Figure 1 shows the undirected network of co-authorships (900 nodes and 4729 ties) with 11 different components in evidence. There are 5 different figures for different types of institutions: Circle-Italian biotech company; Square-University; Down Triangle-Research Centre; Up Triangle-Hospital, Box-other company.

Figure 1: Undirected network of co-autorships

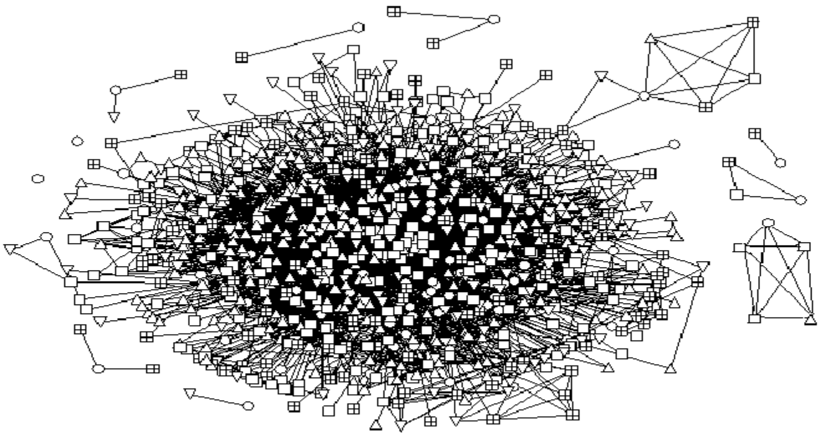


Figure 2 shows the same undirected network (900 nodes and 4729 ties) with evidence of different types of institutions and links among. The figures for different types of institutions are the same as in Figure 1.

Figure 2: Undirected network of co-autorships. Evidence of different types of institutions and links among

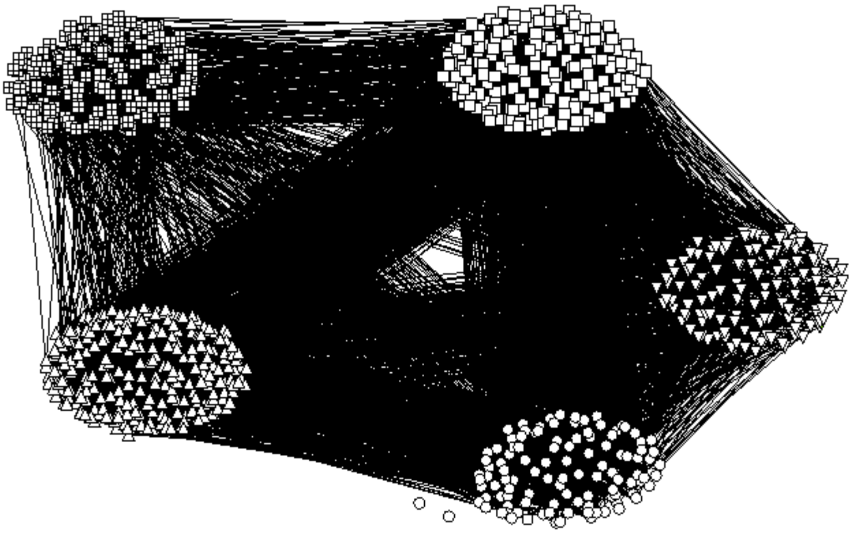


Table 3 shows some statistics regarding the whole network. As we classified our nodes in some institutional categories, we have been able to obtain the mean values of such statistics for each of these institutional categories.

Table 3: Degree, closeness and betweenness centrality and E-I index by kind of institution

	Degree Centrality (binary network)		Degree Centrality (valued network)		Closeness Centrality*		Norm Betweenness centrality		E-I index
	Mean (Std Dev)	Max	Mean (Std Dev)	Max	Mean (Std Dev)	Max	Mean nBet	Max nBet	
Whole network	15.127 (27.325)	132	10.509 (14.896)	313	32.228 (4.489)	49.323	0.228 (0.858)	10.511	0.407
Italian biotech firms	22.41 (33.33)	108	13.50 (17.68)	223	33.829 (4.391)	44.752	0.566 (1.167)	8.291	0.995
Universities	23.16 (40.75)	132	14.89 (21.72)	313	33.077 (5.369)	49.323	0.422 (1.353)	10.511	0.271
Research Centres	10.93 (16.82)	73	7.75 (10.24)	106	31.553 (4.197)	44.32	0.077 (0.32)	2.928	0.711
Hospitals	12.61 (19.22)	113	10.14 (10.91)	264	32.489 (3.705)	45.687	0.074 (0.414)	6.229	0.087
Other firms	6.11 (5.86)	22	4.78 (4.18)	36	29.758 (3.758)	37.462	0.018 (0.111)	1.291	0.634

*this index was calculated for the main component of the network

The whole network has a degree centralization of 13.54% and a between centralization of 10.29%. The closeness centralization calculated on the main component of the network is 34.25%.

For the whole network, the mean degree centrality overall is 10.509 (each subject co-authored papers on average with 10.509 subjects), with a maximum of 132, belonging to the Public University of Milan (Università Statale di Milano). The mean valued degree centrality overall is 15.127 (each subject has done on average 15.127 co-authorships), with a maximum value of 313, always belongs to the Public University of Milan. Considering the institutional categories, the highest value of mean degree belongs to the universities (23.16), even higher than the average value for biotech firms (22.41), that was expected to be rather high, as in every publication there is at least one of the biotech firms, because our database is based on our firms publications. Therefore we can say that the universities cover a central, key-role in networks of publications. At a certain distance we find the hospitals (12.61), followed by research centres (10.93) and other firms (6.11). It is interesting to observe that the total number of the hospitals (289) is higher than the total number of universities (218), while the order in terms of mean degree is the opposite: a lot of hospitals collaborate with firms, but the collaboration with every single hospital appears as more occasional than the collaboration with every single university: in fact three universities have the highest values of degree in binary network (the public University of Milan is followed by the University of Turin and University of Rome "La Sapienza"); the first hospital (Hospital San Raffaele, Milan) comes at the fourth place. In the valued network the hospital San Raffaele takes second place after University of Milan and it is followed by University of Rome “La Sapienza” and University of Turin.

These values and considerations underline, therefore, the prominence of the universities in the biotech system of innovation, respect to other locus where research is done (research centres and hospitals), while co-authorships involving firms belonging to different sectors are quite unusual.

We also calculated, for the whole network, the values regarding betweenness. Considering the average values for each kind of institution, we observe the highest value for the Italian biotech firms (normalized value: 0.566), followed by universities (0.422), research centres (0.077), hospitals (0.074) and other firms (0.018). We can therefore observe two inversions (universities/biotech firms; hospitals/research centres) respect to the order for the mean degree: while co-authorships with research centres are less frequent than with hospitals, research centres are more “able” than hospitals to “bridge” different partners.

With regard to closeness centrality, University of Milan is the closest node to other institutions; among the biotech firms Bracco Imaging is the closest. Considering the average values for each kind of institution, we observe that the mean values of closeness centrality for different institutions do not range very much. The lowest value belongs to the group of other firms.

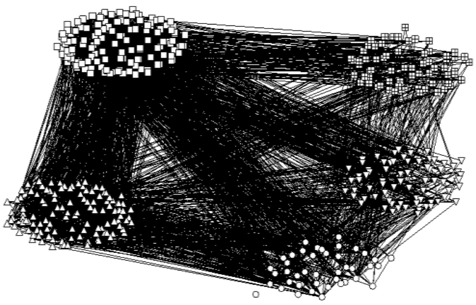
On another point of view, this property of the research centres to bridge different institutions emerges from another kind of statistics, the E-I index: the average value for research centres is 0.711, for hospitals is 0.087: we can interpret such result in the sense that, if a paper is co-authored by a biotech firm with a research centre, there are frequently one or more authors of different kind too (low level of homophily); the opposite happens if there is a co-authorship with an hospital (the presence of partners of another institutional type is much less frequent). The value for this index is next to the maximum (0.995) for the Italian biotech firms: this derives from the very low degree of inter-firm collaborations; this also holds for other firms, as the high value of this index (0.634) shows; universities have, on average, a value of 0.271.

Now we turn to the analysis of the networks deriving from the classification of the biotech firms according to the OECD typology. As a consequence of the different number of firms belonging to the different categories, we have networks of different size. The largest network is generated by the publications of the 42 innovative firms, including 504 nodes (institutions) and 2429 ties; it is followed by the networks of the R&D firms (23 firms, 223 nodes, 796 ties), of the dedicated firms (18 firms, 212 nodes and 852 ties), of the active firms (8 firms, 112 nodes, 381 ties), of the targeted firms (10 firms, 82 nodes, 219 ties), of the “out” firms (7 firms, 50 nodes, 126 ties) and of the other services firms (7 firms, 24 nodes, 35 ties).

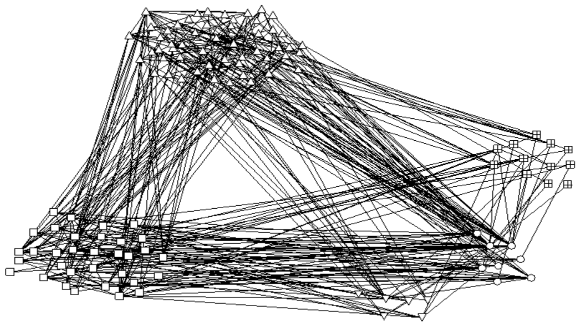
The Figures 3(a-g) shows the graphs of co-authorship networks for the innovative (3a), active (3b), dedicated (3c), R&D (3d) and targeted (3e), other service (3f) and “out” (3g). The figures for different types of institutions are the same as in Figure 1.

Figure 3: Networks for different groups of OECD classification with evidence of different types of institutions and links among.

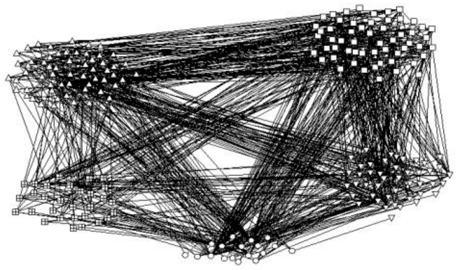
3a



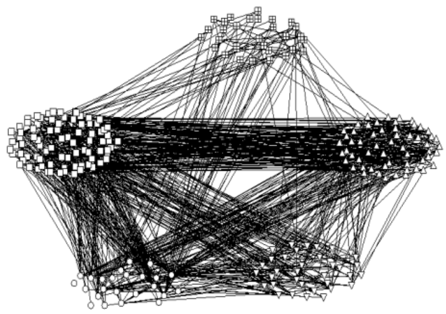
3b



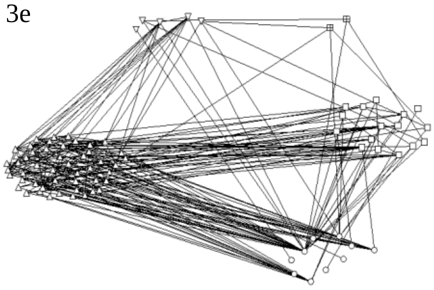
3c



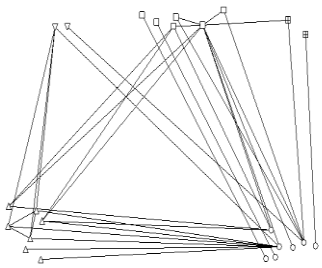
3d



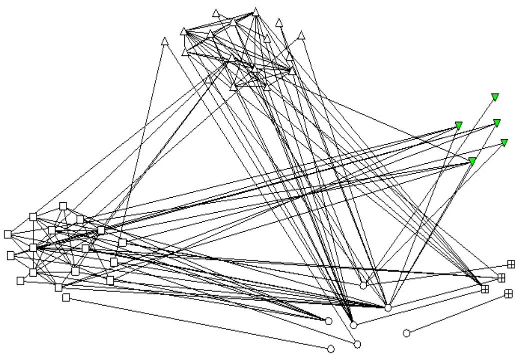
3e



3f



3g



The values of density, degree centralization and betweenness centralization of such networks are negatively correlated with their dimension. This is an expected result if the characteristics of the various network are overall similar. The regularity in the negative correlation is in fact broken only by the network of the "out" firms; this is a small network, but the value of betweenness centralization is low: in this network there are in fact six little components.

Let now consider the institutional analysis: we may observe that the order of "importance" among the different institutions, expressed by the mean valued degree and by the mean betweenness, is the same across the different sub-groups, with a few limited exceptions.

As regards the mean valued degree, Italian biotech firms have the highest value, followed by universities, hospitals, research centres and other firms. The only exceptions concern the following cases: universities have an higher value then Italian biotech firms in the network of other services firms; other firms come before research centres in the network of the "out" firms; in the network of the targeted firms, the order is: Italian biotech firms, hospitals, research centres, universities and other firms; this last case is interesting, as this exception to the general trend appears consistent with the nature and goal of this kind of firms: targeted firms are suppliers, therefore they are less interested in "basic" research (that typically happens with universities) than in collaboration with institution oriented to the "application" of biotechnologies, as hospitals are.

It has to be underlined that in these networks it is possible to observe the same peculiarity of the whole networks: while the number of hospitals is similar, and often higher than the number of universities, the mean degree is always higher for universities, indicating a more frequent interaction with each university.

As regards betweenness, in the networks of innovative and active firms the order among the institutions is the same than in the whole network: Italian biotech firms come first, then we have universities, research centres, hospitals and other firms. In the other sub-networks, the more relevant exception is represented by the hospitals, that are at the second place.

Observing the E-I index in the different sub-groups, in every network research centres show the highest level of heterophily: only in the network of the dedicated firms there is another type of institution (other firms) with an higher value of the index, while the lowest values belong to hospitals and universities[5].

4.3. Synthesis of the results

From our analysis it emerges clearly the central role occupied by the universities in the architecture of the Italian life-science biotech innovation system. In particular, two great universities situated in the North-West of Italy (University of Milan and Turin) have a prominent position, followed by an university from the Centre of Italy (University of Rome "La Sapienza"). There is likely a relationship between this observation and the high number of biotech firm situated in the region of Milan (Lumbardy), Turin (Piedmont) and Rome (Lazio).

Hospitals are important actors of the publication networks as the collaboration with them is frequent. The collaboration with research centres is less systematic, but their role appears important too, because of their capacity to operate as a bridge among different institutions. A negative point of the Italian system of innovation is represented by the low degree of collaboration among different firms.

The analysis developed taking into consideration the characteristics of the firms revealed a substantially homogeneous behaviour of the biotech sector across its internal articulation. Anyway the exceptions to the regularities appears consistent with the nature and the goals of the firms, like the particularly important role covered by the hospitals for the targeted firms and by the universities for the R&D firms.

5. Conclusions

In recent years Italy, in order to face the economic crisis but also the problem of a large public dept, has known a process of rationalization, or even drastic cuts, of the public expenditure in the fields of education and research. The crisis has been generalised and the problems of budget constraints, even if not severe as in Italy, concern many countries; therefore many countries have to take similar difficult decisions. There is therefore a general need to understand in depth the points of strength and weakness of the research and innovation system, in order to operate reasonable choices. In this paper we tried to give an example of such kind of analysis, focusing an important aspect of the Italian system of research and innovation: through the instrument of the SNA we investigated the networks of co-authorships in the publications of the firms belonging to a highly innovative sector, the life-science biotech. We analysed the role and importance of the different kinds of institutions that constitute the Italian system of innovation (universities, research centres, hospitals and firms). This analysis reveals unambiguously the central role covered by the universities, particularly by some specific academies (like University of Milan, Turin and Rome). Anyway the other institutions have an important role too: the research centres often participate in large and heterogeneous networks, having the role to bridge different institutions; a lot of hospitals make research in collaboration with firms, publishing the results. A point of weakness seems to be the infrequent collaborations among firms. These conclusions seem to hold even if we take into consideration the main characteristics of the firm, that is their belonging to one or another typology (as defined by OECD); anyway this deeper analysis shows some peculiarities, like the particularly frequent ties between universities and the biotech firms focused on R&D or the central role covered by the hospitals in the network of the *targeted* firms: considering the nature of such kind of firms, therefore the kind of research mainly conducted by them, we can conclude that universities are by large the main point of reference for basic research, while it is noteworthy the importance of the hospitals for the applied research.

If a policy suggestion may be derived from such analysis, we may conclude that each kind of institution has its peculiar and fundamental role in the system of innovation, therefore it seems that should be avoided to valorise some institutions penalising others. Other suggestions concern the links among firms that should be incentivized, not only on a productive or commercial point of view, but also in the production of new knowledge destined to the "open science". Finally, it emerges the need to consider, if a policy is designed, the peculiarity of the design of the innovation system in each sub-sector, also inside a same technological field.

Acknowledgements

This paper derives by a strict and continuous collaboration among the authors. Anyway sections 2.1 and 3.1 may be mainly attributed to R. Iorio, sections 2.2 and 2.3 to R. d’Amore, section 3.2 to A. Stawinoga. Sections 4, introduction and conclusions have been jointly written.
We are particularly grateful to prof. Pasquale Persico, who gave us many precious ideas about this study, to prof.ssa Maria Rosaria d’Esposito, as she greatly stimulated us to realize this paper, and to prof. Giuseppe Giordano because of his clarifying suggestions.

Notes

[1] For the sake of brevity, in the following this group will be simply called “Italian biotech firms”.

[2] D’Amore and Iorio (2010), through an econometric analysis on the same publication data but on a larger period of time (2001-05), clarify that both the firm size and the OECD typology have an effect on firm propensity to publish and collaborate in publication. Controlling for firm size, R&D firms publish more than innovative firms, while targeted firms have the higher propensity to collaborate.

[3] This number obviously do not equal the total number of collaborations: it is possible to collaborate many times with the same institution. For instance, innovative firms have collaborated in publications with 504 institutions, but they have a total number of 1799 collaborations.

[4] For example, one network is obtained extracting from the whole set of publications those which have an innovative biotech firm among the institutions of affiliation of the authors; another network is obtained considering the publications of the active firms, and so on.

[5] The complete statistics for the seven networks are reported in the tables A1-A7 in the Appendix.

References

Audretsch, D.B., Stephan, P.E.: Company-Scientist locational links: the case of biotechnology. *American Economic Review* **86** (3), 641–652 (1996)
Borgatti S.P. : Netdraw Network Visualization. Analytic Technologies: Harvard, MA (2002)
Borgatti, S.P., Everett, M.G. and Freeman, L.CUcinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies (2002)
Chinchilla-Rodríguez, Z.; Moya-Anegón, F.; Vargas-Quesada, B.; Corera-Álvarez, E.; Hassan-Montero, Y. : Inter-institutional scientific collaboration: an approach from social network analysis. PRIME International Conference (2008).
d’Amore R., Iorio R. : Is OECD classification of the biotech firms useful to explain different behaviours in research collaboration and publication? An analysis of the Italian case.” Quaderno di Ricerca IRAT/CNR n.47, Enzo Albano Editore, Napoli (2010).
d’Amore R., Vittoria P. : Il nuovo settore delle imprese biotecnologiche. Fonti informative, indicatori statistici e ambiti di policy, *L’industria*, **2**, 329-346 (2008)
d’Amore R., Vittoria P. : Assessing statistical standards for emerging industries. Applying OECD statistical codes to Italian biotech population lists”, *World Review of Science, Technology and Sustainable Development* **6** (2/3), 233-243 (2009)
Etzkowitz H., Leydesdorff L. : The dynamics of innovation: from National Systems and “Mode 2” to a Triple Helix of university–industry–government relations, *Research Policy* **29** (2), 109–123 (2000)

Etzkowitz, H.A., Webster, C., Gebhardt, B., Terra, R.C.: The future of the university and the university of the future: evolution of ivory tower to entrepreneurial paradigm. *Research Policy* **29**, 313–330 (2000)

Fox, M.F., Faver, C.A. : Independence and cooperation in research: the motivation and costs of collaboration. *The Journal of Higher Education* **55** (3), 347–359 (1984)

Freeman, L.C. : Centrality in Social Networks: Conceptual Clarification. *Social Networks* **1,** 215-239 (1979)

Gittelman, M.: National institutions, public-private knowledge flows, and innovation performance: A comparative study of the biotechnology industry in the United States and France", *Research Policy* **35** (7),1052-1068 (2006)

Gittelman, M., Kogut, B.: Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. Management Science **49** (4), 366–382 (2003)

Horn, D.B., Finholt, T.A., Birnholtz, J.P., Motwani, D., Jayaraman, S.: Six degrees of Jonathan Grudin: a social network analysis of the evolution and impact of CSCW research. In: Paper Presented at the Computer Supported Cooperative Work (CSCW) (2004)

Lin, C.H.: The cross-citation analysis of selected marketing journals. *Journal of Management* **12** (3), 465–489 (1995)

Malerba, F.: "Sectoral system of Innovation and Production". *Research Policy*, **31**, 247-264 (2002)

Murray, F., Stern, S. :Do formal intellectual property rights hinder the free flow of scientific knowledge: an empirical test of the anti-commons hypothesis. *Journal of Economic Behavior and Organization* **63** (4), 648-687 (2007)

Newman, M.E.J.: Scientific collaboration networks. I. Network construction and Fundamental results. *Physics Review* E **64** (1), 0161311–0161318 (2001)a

Newman, M.E.J.: The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences* **98** (2), 404–409 (2001) b.

Newman, M.E.J.: Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences* **101** (S1), 5200–5205 (2004)

OECD: OECD, Biotechnology: Economic and Wider Impact. (1989)

OECD, A Framework for Biotechnology Statistics. Organisation for Economic Co Operation and Development, Paris (2005) a

OECD. Statistical Definition of Biotechnology (2005) b

Owen-Smith, J., Powell, W.: Knowledge networks as channels and conduits: the effects of spillovers in the Boston biotechnology community. *Organization Science* **51**, 5–21 (2004)

Powell W., Koput K., Smith-Doerr L.: Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology, *Administrative Science Quarterly* **41**, 116-145 (1996)

Stuart, T.E., & Sorenson, O.: The geography of opportunity: Spatial heterogeneity in founding rates and the performance of biotechnology firms. *Research Policy* **32**, 229-253 (2003)

Stuart TE, Ozdemir SZ, Ding WW.: Vertical alliance networks: the case of university-biotechnology pharmaceutical alliance chains. *Research Policy* **36**, 477–491(2007)

Wasserman S., Faust K.: Social Network Analysis: Methods and Applications, Cambridge University Press (1994)

Zucker, L., Darby, M., Brewer, M. Intellectual human capital and the birth of US biotechnology enterprises, *American Economic Review* **88,** 290-306 (1998)

Appendix

Table A1: Degree, betweenness centrality and E-I index for the network of publications of innovative firms

	Degree Centrality (binary network)		Degree Centrality (valued network)		Norm Betweenness centrality		E-I index
	Mean (Std Dev)	Max	Mean (Std Dev)	Max	Mean nBet	Max nBet	
Whole network	9.639 (12.298)	108	14.135 (23.213)	223	0.395 (1.527)	18.776	0.395
Italian biotech firms	17.37 (22.70)	108	34.88 (50.3)	223	1.997 (3.583)	18.776	0.989
Universities	12.39 (15.36)	100	18.51 (27.04)	212	0.251 (1.07)	14.428	0.21
Research Centres	5.76 (5.97)	33	8.44 (10.46)	65	0.088 (0.37)	2.747	0.786
Hospitals	9.29 (7.31)	39	10.97 (9.5)	56	0.064 (0.193)	1.219	0.13
Other firms	4.58 (3.94)	18	6.08 (5.39)	25	0.016 (0.063)	0.397	0.686

Table A2: Degree, betweenness centrality and E-I index for the network of publications of active firms

	Degree Centrality (binary network)		Degree Centrality (valued network)		Norm Betweenness centrality		E-I index
	Mean (Std Dev)	Max	Mean (Std Dev)	Max	Mean nBet	Max nBet	
Whole network	6.804 (5.986)	43	8.696 (8.41)	63	1.669 (4.642)	33.574	0.339
Italian biotech firms	15.88 (14.40)	43	22.5 (20.67)	63	11.04 (12.25)	33.574	1
Universities	7.12 (4.72)	19	9.09 (5.66)	21	2.13 (3.54)	15.67	0.234
Research Centres	7.17 (4.12)	15	8.33 (4.27)	15	1.038 (2.54)	6.22	0.907
Research Centres	7.17 (4.12)	15	8.33 (4.27)	15	1.038 (2.54)	6.22	0.907
Hospitals	5.83 (4.17)	21	7.38 (6.08)	30	0.409 (1.83)	0.268	-0.01
Other firms	4.0 (2.17)	7	4.42 (2.31)	8	0.04 (0.096)	0.268	0.833

Table A3: Degree, betweenness centrality and E-I index for the network of publications of dedicated firms

	Degree Centrality (binary network)		Degree Centrality (valued network)		Norm Betweenness centrality		E-I index
	Mean (Std Dev)	Max	Mean (Std Dev)	Max	Mean nBet	Max nBet	
Whole network	8.038 (8.596)	53	11.802 (17.462)	158	1.027 (3.13)	25.586	0.406
Italian biotech firms	13.56 (16.64)	52	23.22 (30.79)	120	4.916 (7.414)	25.586	1
Universities	9.42 (8.45)	35	14.51 (15.0)	58	1.11 (2.045)	10.047	0.322
Research Centres	7.13 (7.11)	24	9.87 (10.95)	36	0.51 (1.866)	12.05	0.564
Hospitals	7.18 (7.28)	53	10.13 (20.55)	158	0.683 (2.69)	18.058	0.224
Other firms	5.29 (4.41)	22	6.29 (6.53)	36	0.041 (0.159)	0.0804	0.61

Table A4: Degree, betweenness centrality and E-I index for the network of publications of other services firms

	Degree Centrality (binary network)		Degree Centrality (valued network)		Norm Betweenness centrality		E-I index
	Mean (Std Dev)	Max	Mean (Std Dev)	Max	Mean nBet	Max nBet	
Whole network	2.917 (2.308)	10	4.5 (5.252)	26	3.08 (8.862)	34.321	0.675
Italian biotech firms	3.00 (2.77)	8	5.29 (4.03)	11	5.242 (11.062)	29.644	1
Universities	3.33 (3.44)	10	6.5 (9.71)	26	5.742 (14.00)	34.321	0.4
Research Centres	2.50 (2.12)	4	2.5 (2.12)	4	0	0	1
Hospitals	3.00 (1.41)	4	3.43 (2.07)	7	0.395 (1.046)	2.767	0.429
Other firms	1.50 (0.71)	2	1.5 (0.71)	2	0	0	1

Table A5 Degree, betweenness centrality and E-I index for the network of publications of the "out" firms

	Degree Centrality (binary network)		Degree Centrality (valued network)		Norm Betweenness centrality		E-I index
	Mean (Std Dev)	Max	Mean (Std Dev)	Max	Mean nBet	Max nBet	
Whole network	5.040 (3.412)	17	6.8 (5.589)	24	0.364 (1.296)	7.313	0.111
Italian biotech firms	6.29 (5.82)	17	8.71 (8.44)	24	1.735 (3.009)	7.313	1
Universities	4.94 (2.61)	12	7.94 (5.95)	20	0.31 (0.918)	3.571	-0.143
Research Centres	3.20 (1.64)	5	3.2 (1.64)	5	0	0	1
Hospitals	5.59 (3.43)	10	6.76 (4.8)	14	0.046 (0.074)	0.17	-0.347
Other firms	3.25 (2.87)	7	3.25 (2.87)	7	0	0	1

Table A6 Degree, betweenness centrality and E-I index for the network of publications of R&D firms

	Degree Centrality (binary network)		Degree Centrality (valued network)		Norm Betweenness centrality		E-I index
	Mean (Std Dev)	Max	Mean (Std Dev)	Max	Mean nBet	Max nBet	
Whole network	6.862 (6.723)	48	9.44 (10.869)	74	0.974 (2.94)	23.062	0.435
Italian biotech firms	11.79 (12.40)	48	19.25 (21.14)	74	4.177 (6.306)	23.062	0.986
Universities	6.57 (6.29)	33	9.89 (10.8)	57	1.162 (2.809)	13.662	0.326
Research Centres	5.50 (5.10)	26	6.64 (6.28)	27	0.231 (0.733)	3.787	0.737
Hospitals	8.05 (5.10)	27	9.09 (6.49)	42	0.259 (1.15)	8.193	0.12
Other firms	3.23 (2.23)	12	4.52 (3.33)	13	0.162 (0.9)	5.011	0.32

Table A7 Degree, betweenness centrality and E-I index for the network of publications of targeted firms

	Degree Centrality (binary network)		Degree Centrality (valued network)		Norm Betweenness centrality		E-I index
	Mean (Std Dev)	Max	Mean (Std Dev)	Max	Mean nBet	Max nBet	
Whole network	5.341 (4.636)	36	6.39 (6.154)	46	1.941 (7.701)	61.98	0.279
Italian biotech firms	7.70 (11.06)	36	17.7 (10.56)	46	8.936 (19.245)	61.98	1
Universities	4.75 (3.09)	11	4.5 (2.63)	11	2.312 (5.318)	18.627	0.632
Research Centres	5.60 (4.16)	10	4.8 (3.77)	11	0	0	0.857
Hospitals	5.10 (2.81)	11	4.98 (3.14)	11	0.67 (3.252)	21.481	-0.136
Other firms	3.50 (0.71)	4	3.5 (2.12)	5	0	0	1