‘Collective innovation’ in a model of network formation with preferential meeting

Nicolas CARAYOL[1] and Pascale ROUX[2]

[1] BETA (UMR CNRS 7522), Université Louis Pasteur, 61, avenue de la Forêt Noire, F-67085 Strasbourg Cedex, tel:+33(0)388352204, fax:+33(0)390242071 carayol@cournot.u-strasbg.fr

[2] LEREPS, Université de Toulouse 1, Manufacture des Tabacs, 21 Allée de Brienne, F-31000 Toulouse, tel:+33(0)561128707, fax:+33(0)561128708 roux@univ-tlse1.fr

Summary. In this paper, we present a model of ‘collective innovation’ built upon the network formation formalism introduced by Jackson and Wolinski (1996) and Jackson and Watts (2002). Agents localized on a circle benefit from knowledge flows from some others with whom they are directly or indirectly connected. They also support costs for direct connections which are linearly increasing with geographical distance separating them. The dynamic process of network formation departs from the specialized literature in that it exhibits preferential meeting for close agents. The results concern the set of stochastically stable networks selected in the long run. Their architectures are compared to the ones obtained in the simple ‘connections model’. Our main result is to show under what circumstances pairwise stable “small worlds” networks are stochastically selected.

1 Introduction

There is an increasing consensus in the economic literature to say that network structures are significantly influencing the outcomes of many social or economics activities. Recent formal economic contributions highlighting how -both individual and collective- behaviors and performances are grounded in networks which are often in turn shaped by agents. Predictions concern various contexts such as information diffusion on job opportunities [9, 10], firms’ organizational design [8, 19, 30], R&D collaborations [17, 18], market organization [37], etc. The very originality of the economic approach to networks resides in the focus on network formation. A theoretical framework has been proposed by [22] based on a two-sided network formation game. Their approach is also usually called “mixed approach” since it is halfway between the cooperative [31] and the non-cooperative ones [3]. That simply means that two agents have to agree simultaneously to become directly connected while only one defection breaks an existing link. The contribution of [22] also constitutes an important point of departure to analyze and model endogenously emerging structures. Such a line of research has been further developed by Jackson and Watts [21] (initiated in [35]) who introduce the notion of stochastically stable networks.

Built upon this line of investigation, this paper aims to study 'collective innovation' in a dynamic model of network formation. But what is meant by 'collective innovation'? There is an important body of empirical literature focusing on the importance of network relations in determining firms innovation rates. Far from being the outcome of isolated agents efforts, innovation is usually described as a collective process [2, 29, 34]. Turning to the theoretical contributions to the issue, several previous works examine how network structures matter for innovation dynamics through information, knowledge or technology diffusion [13, 32, 39]. However, they are not concerned with network formation which remains a crucial issue for knowledge dynamics and innovation. This question is of interest: If the network structure has obviously much to say about innovative performance, then one may naturally wonder about the circumstances that allow various network structures to emerge.

In our network based model of 'collective innovation', agents benefit from knowledge flows from agents with whom they are directly or indirectly connected. The higher the distance in the relational network, the weaker the spillover. We also introduce an external metric which allows us to consider the costs of direct links [24] as a function of -geographic- distance between agents. Altogether, our model can be interpreted as an extension of the well known *connections model* introduced by [22]. Since our main concern is with the dynamic formation of networks, we make use of the stochastic (Markov) process introduced by [21] based on notions and results initially proposed by [25, 40]. Nevertheless our model departs from theirs in that we enrich the meeting process: We introduce a *preferential meeting* rule which influences links formation, assuming that agents meet easily other agents in their neighborhood. This way we simply reject the uniform meeting probability and weight the probability that two unconnected agents meet with both the inverse of their distance (on both metrics: The relational and the geographic ones).

We expect that the underlying Markov chain will select pairwise equilibria that have some in common with the empirical literature on networks. This way we come to another body of literature which emerged recently in Physics dealing with the structure of large networks as evidenced by web sites links, relational networks, coauthoring scientific paper [4, 5, 28, 36] [3]. Watts and Strogatz [36] define *Small Worlds* as being highly clustered and having some long distant connections. Such structures are called as such because the av-

[3] As a matter of fact, our preferential meeting rule has some in common with the so called "preferential attachment" process which has recently been highlighted as crucial for generating networks characterized with skew vertices distribution. Several models have been introduced [6, 7, 23, 27, 38]. For a complete review one can refer to [1]. However these models have a very weak description of agents' behaviors.

erage distance between agents is usually small (known as the "six degree of separation", Milgram [26]). While the economic literature on network formation has not dedicated much attention to network characterization, focusing mainly on the compatibility between networks efficiency and stability [20], the selected equilibria we obtain here cannot fall anymore systematically under the usual categories (empty net, complete net, cycle, star). Therefore, we will compute several indexes that capture interesting features of the graphs. Our main result is that the collective innovation model tends to select in the long run pairwise stable networks which share "small worlds" properties.

The paper is organized as follows. Section 2 presents the static features of our 'collective innovation' model in the network formation formalism. Section 3 introduces the dynamic process. Section 4 presents the results. Section 5 concludes.

2 Network formation and the 'collective innovation' model

Each agent is assumed to increase its knowledge through internal capacities and/or by communicating directly through costly relationships with other agents. Direct connections between agents, which are called *pairwise links* since the willingness of both the two agents is necessary to establish and maintain a link, form the relational network which is represented as a non-directed graph. In this model, agents can also benefit from indirect (and costless) connections, through the relational network of their partners, but in a decreasing manner *i.e.* the benefits deteriorate with the *relational distance*. We then consider that the rate at which agents innovate is deduced from their knowledge accumulation rate which is in turn obtained through their relational network.

We begin with some basic notions in network formation. In this respect, the point of departure is the network formation formalism introduced by [22]. We then turn to the description of the innovation process.

2.1 Basic notions in network formation

Properties and typical structures of graphs

Consider a finite set of n agents, $N = \{1, 2, ..., n\}$ with $n \geq 3$, and let i and j be two members of this set. Agents are represented by the nodes of a non-directed graph which edges represent the links between them. The graph then constitutes the relational network between the agents. A link between two distinct agents i and $j \in N$ is denoted ij. A graph g is a list of non ordered pairs of connected and distinct agents. Formally, $\{ij\} \in g$ means that ij exists in g. We define the complete graph $g^N := \{ij \mid i, j \in N\}$ as the set of all subsets of N of size 2, where all players are connected with all the others.

Let $g \subseteq g^N$ be an arbitrary collection of links on N. We define $G = \left\{g \subseteq g^N\right\}$ as the finite set of all possible graphs between the n agents.

Let $g' = g + ij = g \cup \{ij\}$ and $g'' = g - ij = g \backslash \{ij\}$ be respectively the graph obtained by adding ij and the one obtained by deleting ij to the existing graph g. The graphs g and g' are said to be *adjacent* as well as the graphs g and g''. For any g, we define $N(g) = \{i \mid \exists j : ij \in g\}$ and $\#N(g)$ the set and the number of agents who have at least a link in the network g. We also define $N_i(g)$ and $\eta_i(g)$, the set and the number of the links agent i has, that is: $N_i(g) = \{ij \mid \exists j : ij \in g\}$ and $\eta_i(g) = \#N_i(g)$. The latter is also called the *degree* of node i. The total number of links in the graph g is: $\eta(g) = \#g = \frac{1}{2}\sum_{i \in N} \eta_i(g)$

A *path* in a non empty graph $g \in G$ connecting i to j, is a sequence of edges between distinct agents such that $\{i_1 i_2, i_2 i_3, ..., i_{k-1} i_k\} \subset g$ where $i_1 = i$, $i_k = j$. The length of a path is the number of edges it contains. Let $i \longleftrightarrow j$ be the set of the path connecting i and j. The set of *shortest paths* between i and j noted $i \widetilde{\longleftrightarrow} j$ is such that if $\forall k \in i \widetilde{\longleftrightarrow} j; k \in i \longleftrightarrow j$ and $\#k = \min_{h \in i \longleftrightarrow j} \#h$. We define $d_g(i,j) = d(i,j)$ as the number of links of the shortest path(s) between i and j, also called *geodesic distance*. When there is no path between i and j then their geodesic distance is conventionally infinite: $d(i,j) = \infty$.

An external metric is also introduced, representing for example, the geographic position of agents [24]. Such external metrics defines a distance operator denoted $d'(i,j)$. In our model, we consider that agents are located on a circle (or a ring). Without loss of generality, agents are assumed to be ordered according to their index, such that i is the immediate geographic neighbor of agent $i+1$ and agent $i-1$ but agent 1 and agent n who are neighbors. The geographic distance may simply be obtained by $d'(i,j) = \min\{|i-j|\,; n - |i-j|\}$.

Finally, a graph $g \subseteq g^N$ is said to be *connected* if there exists a path between any two vertices of g. Notice that a *cycle* on g is a path for which $\{i_1 i_2, ..., i_{k-1} i_k\} \subset g$ is such that $i_1 = i_k$. A graph is said to be acyclic if it contains no cycle.

Hence several typical graphs can be described. Let $i, j \in N$. First of all, the *empty graph*, denoted $g^{\emptyset}$, is such that it does not contain any links. A *ring* is a connected graph composed of exactly one path. Formally, we call a network $g \in G$ a ring (also a chain) if g is connected and if :

- for all $i < j : ij \in g$, there does not exist h such that $i < h < j$ and
- for all $i > j : ij \in g$, there does not exist h such that $j < h < i$.

Such a graph is denoted g°. It is a regular network of order $k = 1$, in which all agents are connected and only connected with their two closest geographic neighbors. The double ring denoted $g^{2\circ}$ is a regular network of order $k = 2$ such that all agents are connected and only connected with their four closest neighbors. Finally, a non empty graph $g \in G$ is a (complete) *star*, denoted $g^{\star}$, if there exists $i \in N$ such that if $jk \in g^{\star}$, then either $j = i$ or $k = i$. Agent

i is called the center of the star. Notice that there are n possible stars, since every node can be the center.

Network formation, stability and efficiency

Over time, pairs of agents meet and decide to form, maintain or break links. The formation of a link requires the consent of both the two agents but not its deletion which can emanate from one of them unilaterally. Moreover, agents are myopic which means that they take decisions on the basis of their impacts only on their current payoffs i.e. according to the state of the current network. Let $\pi_i(g_t)$ be the individual payoffs that agent i receives from the graph g_t. Jackson and Wolinski [22] introduce the notion of *pairwise stability* which can be distinguished from the one of Nash equilibrium since the process of network formation is both cooperative and non cooperative. The formal definition of this notion is the following.

Definition 1. *(based on [22]) A network $g \subseteq g^N$ is pairwise stable if: (i) for all $ij \in g$, $\pi_i(g) \geq \pi_i(g - ij)$ and $\pi_j(g) \geq \pi_j(g - ij)$, and (ii) for all $ij \notin g$, if $\pi_i(g + ij) > \pi_i(g)$ then $\pi_j(g + ij) < \pi_j(g)$.*

The efficiency of a network is computed by the *total value* of the corresponding graph g, which is a function $\pi : \{g \mid g \subseteq g^N\} \to R$, with $\pi(\emptyset) = 0$. At a given period t, it is given by:

$$\pi(g_t) = \sum_{i \in N} \pi_i(g_t) \tag{1}$$

Definition 2. *(based on [22]) A network $g \subseteq g^N$ is efficient if it maximizes the value function $\pi(g)$ on the set of all possible graphs $\{g \mid g \subseteq g^N\}$ i.e. $\pi(g) \geq \pi(g')$ for all $g' \subseteq g^N$.*

2.2 Knowledge flows and innovation

Let us now turn toward describing how knowledge is diffused through the network connections. Let us assume that knowledge is accumulated both through internal (fixed) capacities of the agent and through the direct and indirect connections that allow him to have access to others' (new) knowledge. Thus the total knowledge accumulated at period t may be obtained as follows:

$$\Delta k_i^t = \Delta k_i(g_t) = \omega_i + \sum_{j \in N \setminus i} \delta^{d(i,j)} \omega_j \tag{2}$$

where g_t is the state of the current network (which is invariant on $[t, t+1[$), ω_i and ω_j are respectively the knowledge created by agents $i, j \in N$ during one unitary period of time and which are assumed to be exogenous and constant over time and agents. Thus the second component of the expression (2) is traducing the flow of knowledge absorbed by i, which emanates simultaneously from other agents j (assuming no time lag for simplicity), through direct and indirect interconnections between i and agents j. Parameter δ represents the transferability factor that is the share of new knowledge produced which is effectively directly or indirectly transmitted through each edge. Thus, we consider that the communication is not perfect: the positive externality deteriorates with the relational distance of the connection. Hence, we assume that $\delta \in \left]0,1\right[$. For instance, if i and j are indirectly connected through a third agent, each will get δ^2 of the flow of knowledge each creates.

We assume that the expected number of innovations generated by i during any unitary period t is given by:

$$\theta_i^t = \lambda \Delta k_i^t \tag{3}$$

where λ is a non null positive parameter [4].

Let us now define the (expected) payoff function which is deduced from the shape of the graph:

$$\pi_i^t = \theta_i^t V - c_i^t \tag{4}$$

where θ_i^t is the expected number of innovations seen above (3), V is the net profit generated by an innovation and c_i^t is the costs incurred by i, computed as follows:

$$c_i^t = c_i\left(g_t\right) = C + \sum_{j \in N_i(g_t)} cd'\left(i,j\right) \tag{5}$$

It is thus potentially affected by a fixed cost and by the costs spent for being connected to his direct neighbors[5].

The net profit generated by any agent i at period t, may be thus understood as a function of the graph and the position i occupies in it. That value may thus be written as $\pi_i^t = \pi_i\left(g_t\right) = \pi\left(\Delta k_i\left(g_t\right), c_i\left(g_t\right)\right)$. Compiling expressions (2), (4) and (5) one gets:

$$\pi_i\left(g_t\right) = \lambda V \left(\omega_i + \sum_{j \in N \backslash i} \delta^{d(i,j)} \omega_j\right) - C - \sum_{j \in N_i(g_t)} cd'\left(i,j\right) \tag{6}$$

[4] For a more detailed description of the innovation process based on knowledge accumulation, one can refer to [12].

[5] Relying on Debreu's hypothesis [15] according to which closely located players incur less cost to establish communication, Johnson and Gilles [24] have first extended the connections model of [22] introducing a spatial cost topology in their network formation approach. Links costs are increasing with geographic distance between agents. The traditional assumption is that it's less costly to establish and maintain relationships when agents are geographically close.

Remark 1. Our formulation of the payoff function is voluntarily very close to the connections model first introduced in [22]. If we arbitrarily fix $\lambda = 1/V$, $C = 0$, *and let* $d'(i,j) = 1, \forall i,j$, *thus we have the same formulation as theirs. Notice that if we have* $\omega_i = \omega_j = \lambda V = C = 1$, *then one gets the simple connections model which is well known in the network formation literature. One can also observe that when reintroducing geographic distance in link costs, then one obtains the same payoffs specification as the one of [24], who first introduced some external metric (theirs is the line instead as the circle in our model).*

3 Dynamic network formation

This section is dedicated to the presentation of our perturbed stochastic process of network formation. We begin with the first step of the dynamic settings, namely the meeting process. We will consider that the probability for a given pair of unconnected agents to be selected is not fixed but varies across agents according to their relative position on the current relational graph. Then, we turn towards the last features of the dynamic process and present its generic properties.

3.1 The preferential meeting process

In most of the works investigating the evolution of network (for example in [21, 35]), it is assumed that any pair of agents have the same probability to meet at each period: It thus constitutes an implicit assumption of an uniform meeting probability: $\forall i,j \in N, p_{ij}^t = p^t$. This assumption is threefold: *i)* every pair of unconnected agents have the same probability to meet; *ii)* every pair of (directly) connected agents have the same probability to reconsider their relation, and *iii)* connected agents reconsider their relations with the same frequency as unconnected agents meet.

Here we reject the *i)* part of the assumption while trying to preserve the *ii)* and *iii)* ones for symmetry reasons. Therefore, if we write P^t the probability that the chosen two agents are unconnected at period t, we assume that:

$$P^t = \sum_{ij \notin g_t} p_{ij}^t = \frac{\# g^N - \eta(g_t)}{\# g^N} \tag{7}$$

Together with considering that $\forall ij \in g_t,\ p_{ij}^t = p^t$, this implies that the probability that two connected agents reconsider their relationship at each period is such that $p^t = \frac{1}{\# g^N}$.

Moreover, we do not consider that unconnected people may meet with constant and time independent probabilities. Indeed, this assumption can be justified in the case of anonymous market interactions when the number of agents considered is very large. Here, we introduce a preferential meeting process, considering that the probabilities for a pair of unconnected agents to be selected is not independent across agents and vary according to their relative position on the current relational graph. Hence, we consider that the less is the relational distance between two unconnected agents, the greater will be the probability of their selection. Moreover, we consider that this probability increases with their geographic proximity, which is invariant. This ensures that the probability of any two unconnected agents is never null (which is a necessary condition to preserve the ergodicity property of the stochastic process presented below).

Formally, we introduce a preferential meeting process for unconnected agents which is captured by the simple following formula:

$$p_{ij}^{t} = d\left(i,j\right)^{-\gamma} + d'\left(i,j\right)^{-\beta}, \forall ij \notin g_t \tag{8}$$

where γ and β are two positive parameters capturing the relative importance of relational indirect connections and geographic proximity in the probability that two unconnected agents meet each other. This expression is also subject to normalization such that (7) is respected.

3.2 The limit behavior of the perturbed stochastic process

The dynamic process can be described as follows. At each time period t, two agents i and $j \in N$ are selected by the preferential meeting process described above. Then, if the selected two agents are directly connected, they can jointly decide to maintain their relation or unilaterally decide to sever the link between them. If they are not connected, they can jointly decide to form a link or renounce unilaterally. Formally, those two situations are the following:

(i) if $ij \in g_t$, the link is maintained if $\pi_i(g_t) \geq \pi_i(g_t - ij)$ and $\pi_j(g_t) \geq \pi_j(g_t - ij)$. Otherwise, the link is deleted.

(ii) if $ij \notin g_t$, a new link is created if $\pi_i(g_t + ij) \geq \pi_i(g_t)$ and $\pi_j(g_t + ij) \geq \pi_j(g_t)$,with a strict inequality for one of them.

The stochastic process introduced here can be defined as a Markov chain which finite states correspond to the "current" network at the end of a given period. In other words, the state of the system at time t (with $t = 0, 1, 2, ...$) is given by the graph structure $g_t \in G$. The evolution of the system $\{g_t, t \geq 0\}$ can be described as a discrete-time stochastic process with state space G.

Following [21], we then introduce small random perturbations ε ($\varepsilon \in (0, a]$) which invert agents' right decisions in creating, maintaining or deleting links. These perturbations may be understood as mistakes or as mutations. The characterization of the asymptotic behavior of this process is due to [40]. For small but non null values of ε ($\varepsilon \in (0, a]$), it can be shown that the discrete-time Markov chain being irreductible and aperiodic, has a unique corresponding stationary distribution. Such perturbed stochastic processes are said to be ergodic. Intuitively ergodicity implies that it is possible to transit directly or indirectly between any chosen pair of states in a potentially very long period of time (which also means that any state of the system can be directly or indirectly reached from any given one)[6]. Moreover, when ε goes to zero, the stationary distribution converges to a unique *limiting stationary distribution*. The states that are in the support of this limiting stationary distribution are called *stochastically stable* and are either pairwise stable (cf. Definition 1) either part of a close cycle of states[7]. Notice that the ergodicity property is quite interesting since it allows us to run numerical simulations in order to examine the long run behavior of the system [33]: We can then compute the unique limiting stationary distribution of the process.

4 Networks selection: Results

This section introduces our results on the shapes of the stochastically stable networks obtained in a simple version of our collective innovation model. For that purpose, we first present the numerical settings used and the graph indexes we compute for characterizing various networks architectures.

4.1 Numerical setting and indexes

In this section, we present the results obtained for a simplified version of the collective innovation model presented in Section 2. Our general profit given in (6) may be simplified as: $\omega_i = \omega_j = \lambda V = C = 1$. Moreover, for simplification purposes, we will consider that $c = \frac{2}{n}$ for even values of n, and $c = \frac{2}{n-1}$, otherwise. Recall also that the dynamic process used is based on the preferential meeting principle introduced in Section 3. For simplification purposes again, we use a simple rule assuming that $\gamma = \beta = 1$.

We numerically simulate the unique limiting stationary distribution of the perturbed dynamic process of [21] (for which the error term is decreasing down to zero) by the following simple rule:

$$\varepsilon^t = \begin{cases} 0.02 & if\ t < 50 \\ 1/t & otherwise \end{cases} \tag{9}$$

Thus we ensure that errors affect the dynamics while they are decreasing down to zero when time increases: $\lim_{t\to\infty} \varepsilon^t = 0$.

Let us now introduce several indexes. The average number of neighbors gives us a measure of the *network density*: $\overline{\eta}(g)$. Next, we will compute the

[6] It allows the long run state of the system to become independent of its initial conditions. Indeed, processes that are non-ergodic are said to be “path dependent” [14] since their limiting behavior is dependent on the initial state of the system.

[7] Such process is called a regular perturbation of the initial stochastic process (without trembles). Definitions, properties and some proofs are examined in [11]. Initial contributions are the ones of [16, 25, 40], and [21].

maximal and minimal sizes of neighborhoods: $\max_{i\in N}\eta_i(g)$ and $\min_{j\in N}\eta_j(g)$. We also compute the two indexes introduced by [36]. The first one is simply computing the average distance of (directly or indirectly) connected agents. It is given by:

$$d(g) = \frac{1}{n(n-1)} \sum_{i\in N} \sum_{j\in N\backslash i: i\longleftrightarrow j\neq\emptyset} d(i,j) \tag{10}$$

The second is the *average cliquishness* which indicates to what extent the neighborhoods of connected people overlap. It is given by:

$$c(g) = \frac{1}{n(n-1)} \sum_{i\in N} \sum_{jl:j,l\in N_i(g)} \frac{\Delta(l,j)}{\eta_i(g)} \tag{11}$$

with $\Delta(l,j)$ defines such that $\Delta(l,j) \equiv 1$ if $j \in N_l(g)$ and 0 otherwise. Lastly, we examine to what extent the geographic distances and the relational connections overlap. Let us propose an *average (geographic) distance of direct connections* index which will indicate us to what extent the endogenous network and the exogenous metrics discorrelate:

$$D(g) = \sum_{j:ij\in g} \frac{d'(i,j)}{\eta(g)} \tag{12}$$

In the following subsection, we study the limit distribution of states and show that *small world*-like networks may be selected through the process of network formation.

4.2 Features of the limit distribution of networks

The first goal is to study the limit distribution of the process in one simple numerical situation. For that purpose, we ran $1,000$ simulations of $10,000$ periods[8] with the empty graph as initial condition and with $n = 20$ and $\delta = 0,7$. Nodes degree distribution peaks at 6 neighbors, being slightly asymmetric. No agent has less than four neighbors: This is because establishing direct links with geographically close agents generates low costs. In the meantime, there is no agent having more than eight neighbors because no one is intending to support the high costs of many direct links. The network self-organizes itself in a shape which has some in common with regular networks. One may observe in the descriptive statistics obtained on such distribution (presented in Table 1) that the cliquishness coefficient $c(g)$ is quite high: Nearly as high as the one of the double ring $g^{2\circ}$. More, these clustered networks are correlated to the geographic metric: The average geographic distance between connected pairs of agents is quite small ($D(g) \simeq 2,5$). However, the network departs from such regular structure in that the average path length is singularly lower than for the single ring ($1,84 < d(g^{\circ}) \simeq 5,26$).

[8] Time series analyses conduced over more than 100,000 periods showed that the process nearly always converged to a given pairwise stable state after 10,000 periods.

Table 1. Some descriptive statistics on the graph indexes computed for the limit graph distribution approximated through 1,000 experiments.

	mean	median	max	min	var
Av. #neighbors: $\overline{\eta}(g)$	5.67	5.70	6.20	5.3	0.03
Min #neighbors: $\min_{i\in N} \eta_i(g)$	4.04	4.00	5	4	0.04
Max #neighbors: $\max_{i\in N} \eta_i(g)$	7.06	7	8	6	0.17
Av. path length: $d(g)$	1.84	1.84	1.94	1.69	0.00
Av. cliquishness: $c(g)$	0.042	0.042	0.053	0.03	0.00
Av. geo dist. of links: $D(g)$	2.47	2.46	2.77	2.26	0.01
Activity	116.15	115	171.78	78	218.1

In order to provide a better understanding of these results, we represent in Figure 1 below two networks structures selected in the long run. The first structure is obtained with our model whereas the second is obtained with the simple connections model of [22], that is both without preferential meeting and without link costs increasing with geographic distance. It should be noticed that both networks are pairwise stable for their respective payoffs function. The left graph clearly exhibits small world features: High clusterization while some distant connections remain. We also computed the limit stochastic distribution for the simple connections model by using a similar experimental protocol as the one used for the simple distributed innovations model. We find that networks are less dense ($\overline{\eta}(g) = 4.4$), have similar mean values for the average path length (1.84) while the cliquishness coefficient is not significantly different from zero.

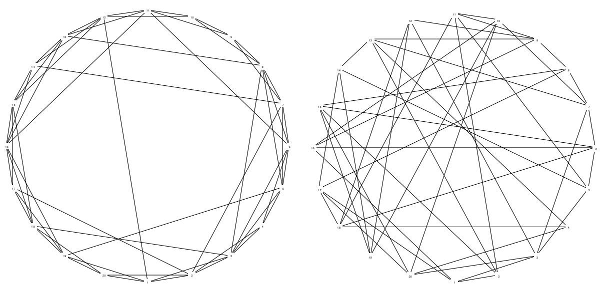

Fig. 1. Two limit typical stable networks selected by the stochastic process in the simple collective innovation model (left graph) and in the simple connections model of [22] (with 20 agents, after $10,000$ periods, obtained for $\delta = 0.7$)

4.3 How does the limit distribution vary with the decay parameter?

We now turn to examine more systematically how the limit distribution varies with the decay parameter δ. To do so we perform a set of (10) experiments for each small increment of δ over its value space ($]0,1[$). In Figure 2 below, we represent the average path length and cliquishness of the limit networks as functions of delta. We also represent the values of the two indexes for the corresponding random graphs. By *corresponding random graphs*, we mean random graphs that have the same number of vertices than the networks obtained with our model for any given value of δ. The indexes values of such random graphs were numerically computed with a thousand experiments on each given number of edges.

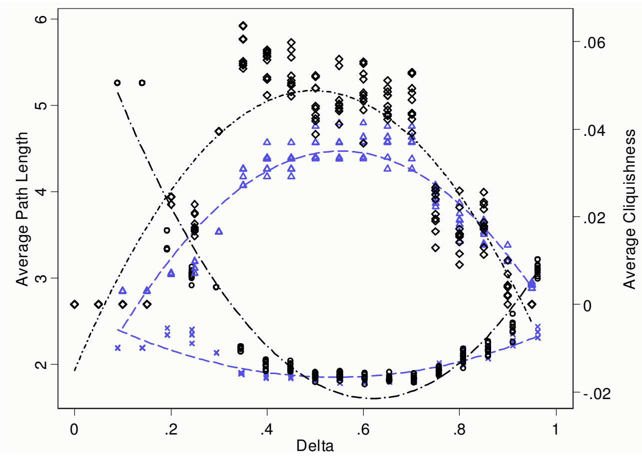

Fig. 2. Average path length (circles) and average cliquishness (diamonds) of the limit networks when $\delta \in]0,1[$ varies (20 agents, 10,000 periods). We also computed the values of these indexes for random networks having a fixed number of links (1,000 experiments for each one). For each limit network (selected for a given value of δ), we represent the associated values of these indexes (crosses and triangles) for the corresponding random graph. The curves account for fitted values of the indexes based on a quadratic specification.

We observe that the average path length of the selected networks decreases from $\delta = 0.1$ to 0.4, then remains nearly constant and increases for high values of δ (≥ 0.75). Average cliquishness suddenly increases from zero at $\delta = 0.2$ and reaches his maximum again very rapidly for $\delta = 0.35$. For this value of δ, we nearly have the weakest average path length while average cliquishness is already at its maximum. Moreover, the average path length of the selected networks is nearly as small as the one obtained for the corresponding random graph, whereas the average cliquishness is significantly larger. Such a situation, where both high cliquishness and low path length are obtained, presents strong similarities with the small world network structure as defined formally by [22]. Finally, $c(g)$ decreases slowly and stabilizes until $\delta = 0.75$, from where it goes down to 0 when δ becomes close to 1.

We also studied the precise shape of the networks selected. We found that the empty graph $g^{\emptyset}$ is selected when $\delta \leq c = 0.1$. When $c < \delta \leq 2c$ the geographic ring g° emerges: in this case, all agents are connected to their two closest neighbors. When δ is 0.3, agents are nearly always connected to their four closest geographic neighbors. This situation corresponds to the double geographic ring $g^{2\circ}$. From $0.4 \leq \delta \leq 0.7$, we found a very 'stable' situation (plateau) characterized by flat maximum neighborhood sizes which decrease from there. For a large range of parameter δ values ($0.35 \leq \delta \leq 0.9$), we thus obtain networks which exhibit small world features. The most characteristic networks are obtained when $\delta \simeq 0.35$.

5 Conclusion

In this paper, we examined a dynamic stochastic process of network formation. In our network based model of 'collective innovation', agents benefit from knowledge flows by communicating with agents with whom they are directly or indirectly connected. We also introduced heterogenous cost of linking and a preferential meeting rule governing the dynamic process of links formation, which consisted in weighting the meeting probability between any two agents by the inverse of their relational and geographic distances. We studied the characteristics of the long term selected graphs computing standard statistical indexes (average path length, clustering coefficient, etc.). Their architectures were compared to the ones obtained in the simple connections model. Our main result is to show that this model generates stochastically stable networks which share features with "small worlds" unlike the initial simple connections model.

References

1. Albert R, Barabási AL (2002) Statistical mechanics of complex networks. Reviews of Modern Physics 74:47–97
2. Allen RC (1983) Collective invention. Journal of Economic Behavior and Organization 4:1–24
3. Bala V, Goyal S (2000) A non-cooperative model of network formation. Econometrica 68:1181-1229
4. Barabási AL, Albert R (1999) Emergence of scaling in random networks. Science 286:509–512
5. Barabási AL, Albert R (2000) Topology of evolving networks: Local events and universality. Physical Review Letters 85:5234–5237

6. Barabási AL, Jeong H, Néda Z, Ravasz E, Schubert A, Viscek T (2001) Evolution of the social network of scientific collaborations. Physica A 311:590–614
7. Bianconi G, Barabási AL (2001) Competition and multiscaling in evolving networks. Europhysics Letters 54:436–442
8. Bolton P, Dewatripont M (1994) The firm as a communication network. Quarterly Journal of Economics 99(4):809–839
9. Calvó-Armengol A (2003) Job contact networks. Mimeo University of Barcelona and CERAS-ENPC
10. Calvó-Armengol A, Jackson MO (2001) Social networks and the resulting dynamics and patterns of employment and wages. Mimeo California Institute of Technology, Pasadena
11. Carayol N, Roux P (2003) Behavioral Foundations and Equilibrium Notions for Social Network Formation Processes. Mimeo BETA Strasbourg
12. Carayol N, Roux P (2004) Self-organizing innovation networks: When do Small Worlds emerge? To appear in European Journal of Economic and Social Systems
13. Cowan R, Jonard N (2001) Knowledge creation, knowledge diffusion and network structure. In: Kirman A, Zimmermann JB (eds) Economies with Heterogenous Interacting Agents. Springer 503:327-343
14. David PA (1985) Clio and the economics of Qwerty. American Economic Review, Papers and Proceedings 75:332–337
15. Debreu G (1969) Neighboring economic agents. La Décision 171:85–90
16. Freidlin M, Wentzell A (1984) Random perturbations of dynamical systems. Springer Verlag, New York
17. Goyal S, Joshi S (2000) Networks of collaboration in oligopoly. Discussion Paper TI 2000-092/1, Tinbergen Institute, Amsterdam-Rotterdam
18. Goyal S, Moraga JL (2001) R and D networks. Mimeo, Tinbergen Institute, Rotterdam
19. Guimerà R, Arenas A, Díaz-Guilera A, Vega-Redondo F (2001) Information processing and optimal organizational structures. Working paper
20. Jackson MO (2003) A survey of models of network formation: Stability and efficiency. Mimeo California Institute of Technology, Pasadena
21. Jackson MO, Watts A (2002) The evolution of social and economic networks. Journal of Economic Theory 106(2):265–295
22. Jackson MO, Wolinsky A (1996) A strategic model of social and economic networks. Journal of Economic Theory 71:44–74
23. Jeong H, Néda Z, Barabási AL (2003) Measuring preferential attachment in evolving networks. Europhysics Letters 61(4):567–572
24. Johnson C, Gilles RP (2000) Spatial social networks. Review of Economic Design 5:273–299
25. Kandori M, Mailath G, Rob R (1993) Learning, mutation and long run equilibria in games. Econometrica 61:29–56
26. Milgram, S (1967) The small world problem. Psychology Today 2:60–67
27. Newman MEJ (2001) Clustering and preferential attachment in growing networks. Cond-mat/0104209
28. Newman MEJ, Watts DJ, Strogatz SH (2001) Random graphs models of social networks. Preprint, to appear in Proceedings of the National Academy of Science

29. Powell WW, Koput KW, Smith-Doerr L (1996) Interorganizational collaboration and the locus of innovation: Networks and learning in biotechnology. Administrative Science Quarterly 41:116–145
30. Radner R (1993) The organization of decentralized information processing. Econometrica 61(5):109–1146
31. Slikker M, van den Nouweland A (2000) Network formation models with costs for establishing links. Review of Economic Design 5:333–362
32. Valente T (1996) Social network thresholds in the diffusion of innovations. Social Networks 18:69–89
33. Vega-Redondo F (2002) Building up social capital in a changing world. Mimeo, University of Alicante
34. Von Hippel E (1989) Cooperation between rivals: Informal knowhow trading. In: Carlsson B (ed) Industrial Dynamics, Technological, Organizational and Structural Changes in Industries and Firms. Kluwer, Boston 157–176
35. Watts, A (2001) A dynamic model of network formation. Games and Economic Behavior 34:331–341
36. Watts DJ, Strogatz SH (1998) Collective dynamics of ‘small worlds’ networks. Nature 393:440–442
37. Weisbuch G, Kirman A, Herreiner D (2000) Market organization. Economica 110:411–436
38. Yook S, Jeong H, Barabási AL, Tu Y (2001) Weighted evolving networks. Physical Review Letters 86(25):5835–5838
39. Young HP (2002) The diffusion of innovations in social networks. Working Paper, SantaFe
40. Young HP (1993) The evolution of conventions. Econometrica 61(1):57–84