Parametric network model inspired by nucleation process

In the previous chapters (Chapter 3 and 4), we discussed about the modeling of the real networks using social theories. Now the question is that: can we adopt network formation process to explain some other processes such as chemical process? In this chapter, we attempt to model nucleation as a network formation process and explore in the direction of modeling of social events of short duration such as workshops and conferences. Recall that complex networks formalism has been used in different and diverse field of academic disciplines including computer science, physics, chemistry, biology, mathematics, and social sciences in order to design and explain different complex systems in real world [Barrat *et al.*, 2008; Harrison, 1999; Jackson *et al.*, 2008; Miyao *et al.*, 2016; Newman, 2010; Prill *et al.*, 2005; Rao *et al.*, 2013] . Specifically, networking approach has been established to be a useful mechanism for understanding a system which consists of a large number of interconnected systems. In this chapter, we make an attempt to deliver a proof of principle for nucleation as a networking process.

In nucleation, individual entities move in the solution or vacuum in Brownian motion. In the process, two or more entities come in physical proximity and thereby produce temporal combined units or clusters. Depending on the prevailing physical conditions and the size of the unit, the unit gains in terms of volume free energy and looses surface free energy. Such clusters, called embryos, may also gather additional units by similar way and loose units by dissolution. If they travel across an energy hill by gathering additional units, the volume free energy gain dominates over surface free energy effect after a certain size. At this stage, embryos give birth to a nucleus which is qualified to grow without dissolution. On the other hand, an embryo may travel in the opposite direction in the energy axes and dissolve completely or partially. In other words, the entities in the solution form network, they have reward and penalty on formation of network which depends on the prevailing conditions and the unit size, and an individual embryo behaves stochastically which is similar to a network formation. Thus, nucleation can be thought of as the formation of either a new thermodynamic phase or a new structure via assembly of participating molecules. The rate of nucleation can be viewed as how long an observer has to wait before significant number of the new phase or structure appears.

The foregoing discussion of 'penalty' and 'reward' in connection to network formation can be related to the overall free energy gain for cluster formation which is given by

$$\Delta F = -\phi s + \psi s^{2/3}$$

where ϕ and ψ are the volume and surface free energy of the cluster (network) composed of s units respectively. Definite expressions of ϕ and ψ are available for a given system. Usually, ϕ is a function of the ratio of the number of units in the system (concentration) to the number of units that shall not participate in the network formation (saturation concentration). The rate of nucleation is an exponential function of ΔF is given by

$$\dot{N} = k_1 \exp\left[\frac{-\Delta F}{kT}\right].$$

In order to interpret the crystallization process as a networking process we consider the atoms (or ions or molecules) as nodes in a given system. We define a link between two nodes when a node has in physical contact with an another node in the system. As we know that the number of physical contacts for a node is limited to a number, which is called co-ordination number (C_n) of the node, we declare that the degree of a node of the 'to be formed network' (after or during crystallization) is restricted by a fixed number for each node. In our simulated results, we consider $C_n = 6$ for any node participating in the crystallization process as observed in many standard crystallization processes, for example, [Hoppe, 1970; Wells, 2012].

In the context of crystallization, we consider that a cubical is occupied by a solution with c unit of concentration. The interaction frequency of the nodes (atoms) of the solute in the solution depends on the concentration of the solute in the cubical. As we know that crystallization is a reversible process, the nodes, which lie in the boundary of the crystals, get dissolved in the solution which depends on the positions of the atom in the crystal. The atoms which are weakly attached with the crystal have higher chances to get dissolved in the solution where a probability $p_d(i)$ for a node i is associated with the crystal. It is needless to mention that it depends on the position of the node inside the crystal.

Once we describe a network formation approach for nucleation based on the aforementioned assumptions, we propose to extend this network formation formalism in to a social network model. Thus we consider the example of formation of groups or communities of researchers when a workshop or conference is organized in an institute. We develop a social network model in spirit of the nucleation process under some suitable assupmtions in the begining of the formation of the social network.

The chapter is organized as follows. In Section 2 we discuss crystal formation with single solute as a network formation process. In Section 3, we propose a social network model inspired by the network approach to crystallization. Finally we conclude our findings.

5.1 NETWORK APPROACH TO CRYSTAL FORMATION WITH SINGLE SOLUTE

In the network formulation for crystal formation we assume that a set V of nodes (atoms) in a solute can take part in the crystallization, that is $|V|$, the number of nodes in V are participating in the crystallization process, whereas, the solution can accommodate maximum $N \, (\gg |V|)$ number of nodes of the solute. Then we define the probability of linking a node to an another node, which is a potential candidate to be a part of the crystal, as

$$p = \frac{|V| - |\mathscr{C}|}{N}$$

where $|\mathscr{C}|$ is the size (the number of nodes) of the crystal formed during the process. Obviously, $|\mathscr{C}| = 0$ in the begining of the network formation. Consequently, the dissolution probability is defined as

$$q(i) = 1 - \frac{k_i}{C_n} \tag{5.1}$$

where k_i is the degree of a node $i \in V$ which represents the number of contacts the node i in the crystal.

In order to introduce the concept of free energy penalty for the nucleation process, we define a constant

$$\triangle E = \sum_{i \in \mathscr{C}} [C_n - k_i] - \sum_{i \in \mathscr{C}} k_i = |\mathscr{C}| C_n - 2 \sum_{i \in \mathscr{C}} k_i \tag{5.2}$$

where $\mathscr{C}$ denotes the crystal and $|\mathscr{C}|$ is the number of nodes (atoms) in the crystal $\mathscr{C}$. Note that, $\sum_{i \in \mathscr{C}} [C_n - k_i]$ corresponds to the surface energy and $\sum_{i \in \mathscr{C}} k_i$ represents the volume energy in the network realization of the nucleation process.

The one-one correspondence of the terms used in standard nucleation and the terms which we just have introduced for its network formation interpretation are given in Table 5.1.

Now we describe the mechanism of a growing network which ultimately explains the nucleation process.

1. Let V be the set of members which are participating in the network formation. Network formation started from a random node which becomes the first member of the crystal $\mathscr{C}$.

2. A random node from $V \setminus \mathscr{C}$ is selected (from the boundary of $\mathscr{C}$), which gets connected with $\mathscr{C}$ with probability p.

3. A random node $i \in \mathscr{C}$ which gets removed from $\mathscr{C}$ with probability $q(i)$.

4. Repeat the steps 2 and 3 until $|V| = |\mathscr{C}|$.

A sample crystal produced by adopting the network formation process is given in Fig. 5.3(b) which is called a crystal-network (CrN). The crystal in Fig. 5.3(b) is generated by considering $|V| = 8 \times 10^2, N = 10 \times 10^2$. The behaviour of the characteristic curves of the classical nucleation process and crystallization through network formation approach are the same which can be observed form the Figs. 5.1 and 5.3(a).

It may be noted that for a fixed N, a different choice of $|V|$ can influence the crystal formation process. Thus, we pose the following question. For a fixed N, how does the concentration of nodes, that is, $|V|$ change the dynamics of crystallization in crystal-network formation? It is well-known in classic nucleation process that the concentration influences the crystallization process, in fact, lower the concentration is, slow the rate of crystallization.

We investigate the effect of p (which depends on $|V|$) in crystal-network by simulating the process for $|V| = 6 \times 10^2$ and $|V| = 8 \times 10^2$ with a fixed $N = 10 \times 10^2$. It is found that lower concentration level increases the time of crystallization under network approach which is standard in classic crystallization process. The conclusion is made by plotting $\triangle E$ after taking average over 5000 crystal-networks, see Fig. 5.4(a).

It is well known from the classical nucleation theory that the molecules or ions continuously get in to crystal and dissolve in to the solution during the crystallization process. During the crystallization process the size of the crystal changes continuously that can be measured in the

Table 5.1 : Correspondence between the symbols used in crystallization and corresponding network formation process.

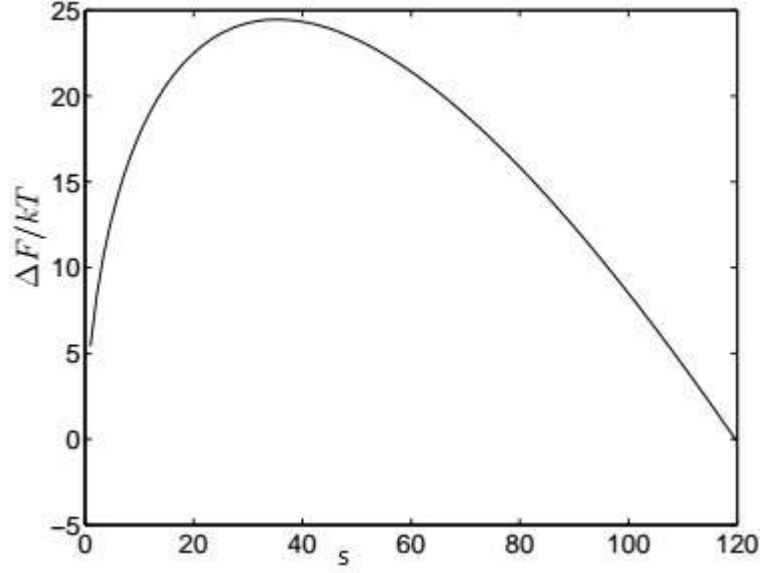| In crystallization	In network formation
the set of atoms which take part in crystallization	V
c	p
$p_d(i)$	$q(i)$
$\triangle F$	$\triangle E$

Figure 5.1: Horizontal axis represents size of the nucleus (s) and vertical axis represents free energy penalty.

context of network formation. The average rate of change ($|\dot{\mathscr{C}}|$) in the size of the network-crystal is given by

$$|\dot{\mathscr{C}}| = p - \bar{q}$$

where $\bar{q} = \frac{1}{|\mathscr{C}|}\sum_{i\in\mathscr{C}}\left(1 - \frac{k_i}{C_n}\right)$ is the average probability of dissolution. Further

$$|\dot{\mathscr{C}}| = \frac{|V| - |\mathscr{C}|}{N} - 1 + \frac{\bar{k}}{C_n}. \tag{5.3}$$

where $\bar{k}$ is the average degree of the nodes inside the crystal networks. From the Eq. (5.2) and (5.3),

$$|\dot{\mathscr{C}}| = \frac{-|\mathscr{C}|}{N} - \frac{\triangle E}{2C_n|\mathscr{C}|} + \left(-0.5 + \frac{|V|}{N}\right).$$

From the Fig. 5.2, it is observed that the process of crystal formation is fastest around the point where $\triangle E = 0$, see red dots in the Fig. 5.2, after that the rate of the change in the size of a crystal decreases slowly. In Fig. 5.2, we plotted the $100 \times |\dot{\mathscr{C}}|$ so that the comparison between the curve of $\triangle E$ and $100 \times |\dot{\mathscr{C}}|$ can be visualized at same scale.

In classical nucleation theory, the rate of nucleation highly depends on concentration. The rate of nucleation is given by a complicated non-linear function of the prevailing saturation concentration as

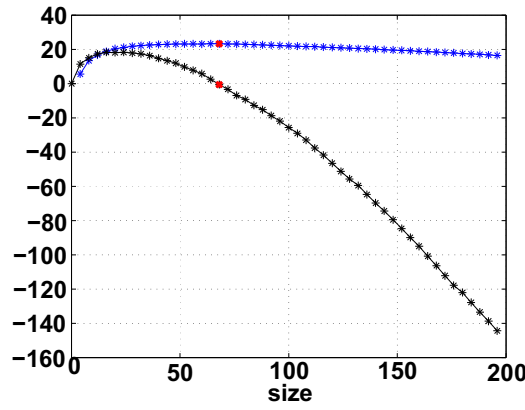$$\dot{N} = k_1 \exp\left[\frac{k_2}{-(ln(c/c_0))^2}\right]. \tag{5.4}$$

Figure 5.2 : Plot in blue represents the average rate of change in the size of the crystal at given size and plot in black is $\triangle E$. $|\dot{\mathscr{C}}|$ is maximum around the point where $\triangle E = 0$. Points in red color.

This function shows a zero rate around the concentration c_0 and suddenly jumps to a very high value after c exceeds a certain concentration. Then the nucleation and associated growth of the cluster reduce the c which brings down the rate again close to zero. This is the general feature of the equation and also observed in the case of homogeneous nucleation.

In the context of nucleation process under network formation approach, we simulated the crystal formation process at different concentration levels of solution. For a given solution, c_0, k_1 and k_2 are constants. We consider $k_1 = k_2 = 1$, $c \in \{0.95, 0.9, 0.85, 0.8, 0.75, 0.7, 0.65, 0.6\}$ and $c_0 = 0.5$. For each value of c, we numerically simulated the average time taken by the nucleation process to generate a crystal nucleus that is related to nucleation rate of the solution. The numerical results are averaged over 1000 simulations. In Fig. 5.4(b), average frequency of nuclei generation at each value of concentration level is plotted against the function of concentration c which is given in right of the Eq. (5.4). The rate of nucleation increases exponentially as the concentration increases. The plot between nucleation rate and quantity $\left(\frac{k_2}{-(ln(c/c_0))^2} \right)$ in Fig. 5.4(b) validate the novelty of the process of nucleation using network formation approach. The behaviour of the rate of nucleation in the context of network formation is highly similar to the classical nucleation process.

## 5.2 A SOCIAL NETWORK MODEL INSPIRED BY NUCLEATION

In the previous section we have studied the crystallization by interpreting it as a network formation process. In this section we propose a social network model which inherit communities in the resultant network inspired by the process. Here, we emphasize that these communities can be thought of as groups of entities with similar interest.

We assume that the atoms are social entities and the social ties between these entities are represented by physical contacts of the atoms after crystallization. Hence we propose the following model.

1. Let V be the set of nodes which are participating in the network formation process among the total number of available nodes $N \gg |V|$ in a system.

2. Select multiple nodes, say, $k (\lll |V|)$, called the initiators for the resultant network that act as the starting points of crystallization simultaneously. In the beginning, a participating node
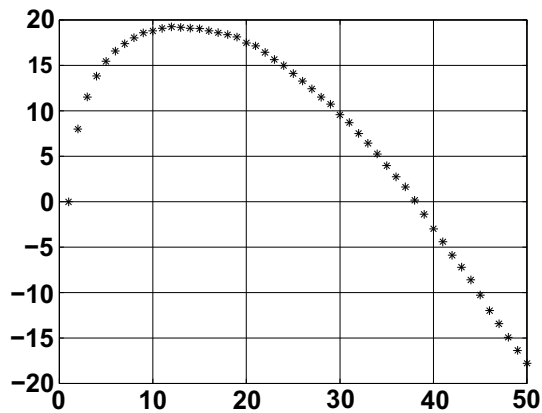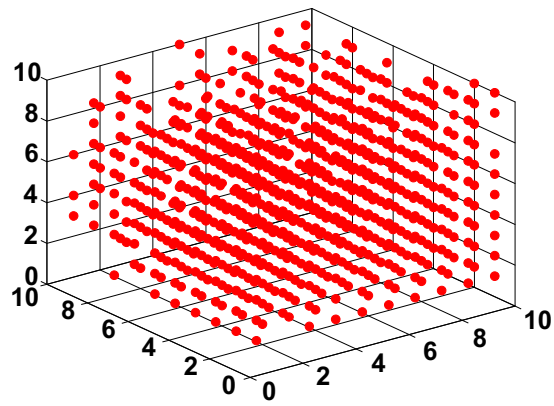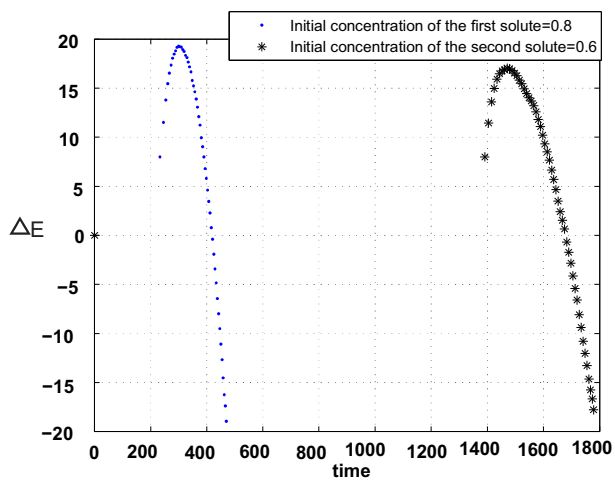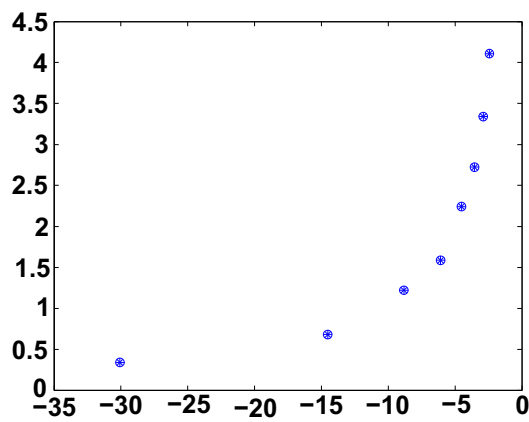
(a)   (b)

Figure 5.3 : In (a) Horizontal axis represents size of the nucleus (s) and vertical axis represents $\triangle E$. In (b) Crystal generated through network formation approach.

(a)   (b)

Figure 5.4 : In (a) Crystallization time while the initial concentration level of solute is different. In (b) Horizontal axis represents the quantity $\left[\frac{k_2}{-(ln(c/c_0))^2}\right]$ and vertical axis represents nucleation rate $\dot{N}$ in the context of network formation process.

i (a node which wants to participate in the network formation process) gets linked with the initiators which are physically close to *i* with probability

$$p = \frac{|V|}{N}.$$

3. Once the degree of an initiator becomes at least 1, we call the subgraph which consists of an initiator and its neighbours, a *group*. Then, any participating node can get attached to any node (member) with probability *p* in any group depending on its closeness to a group where closeness is defined as the minimum spatial distance among all spatial distances of members in a group.

4. The process continues until no isolated participating node exists in the system.

For example, consider the phenomena of group formation of participants having similar research interest in a workshop organized in an institute. Here, we assume that the nodes represent researchers in the institute and *V* is the set of all researchers among them who are participating in the workshop. Thus, for the network formation, we assume $N \gg |V|$ is the total number of researchers available in the institute. We assume that the workshop schedule includes *k* parallel sessions focused on different sub areas of research in each such session. The speakers in each session at a given instant can be considered as the initiators for the resultant network formation, in which, the closeness of a researcher (node) can be defined by considering his research interest or social acquaintances with fellow researchers. The resultant network can be thought of as an outcome of the crystallization process.

We simulate social networks using the above mentioned model by setting $N = 27 \times 10^3$, $|V| = 2 \times 10^3$ and we select 10 random initiators for the generation of the resultant network, for example, see Fig. 5.5 where the size of the nodes are proportional to its degrees and $C_n = 6$.

A network generated by setting 2 intiators with $|V| = 2 \times 10^3$ and $N = 27 \times 10^3$ is provided in Fig. 5.6.

5.2.1 Community structures in the resultant social network

Real world network inherits community structure or modularity. In the context of social structure, community is a group of similar people who have more connections inside the community and less interactive to the rest of the world. Newman defined that community is a group of nodes which has more number of links than expected in random networks [Newman and Girvan, 2004]. To measure the quality of community structure inside a network, modularity index *Q* is given by Newman which is defined as

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta_{c_i c_j}$$

where *m* is the number of edges in the network, k_i and k_j are the degrees of the nodes *i* and *j* which belong to communities c_i and c_j respectively, and δ is the Kronecker delta function.

In order to investigate the existence of community structures in the resultant network due to the proposed model, we proceed as follows. We generate the networks for different values of $|V|$, in particular, we consider $|V| = t \times 10^2$ where $t = 2 : 10$ and $N = 2 \times 10^3$ is fixed. We observe that as $|V|$ (the concentration of solution) increases, the number of modules increases and the modularity index decreases. See in Figs. 5.7(a) and 5.7(b). We use Louvain algorithm for the detection of the community structures [Blondel *et al.*, 2008]. It is needless to mention that this phenomena is common in various social networks, for example, [Girvan and Newman, 2002].
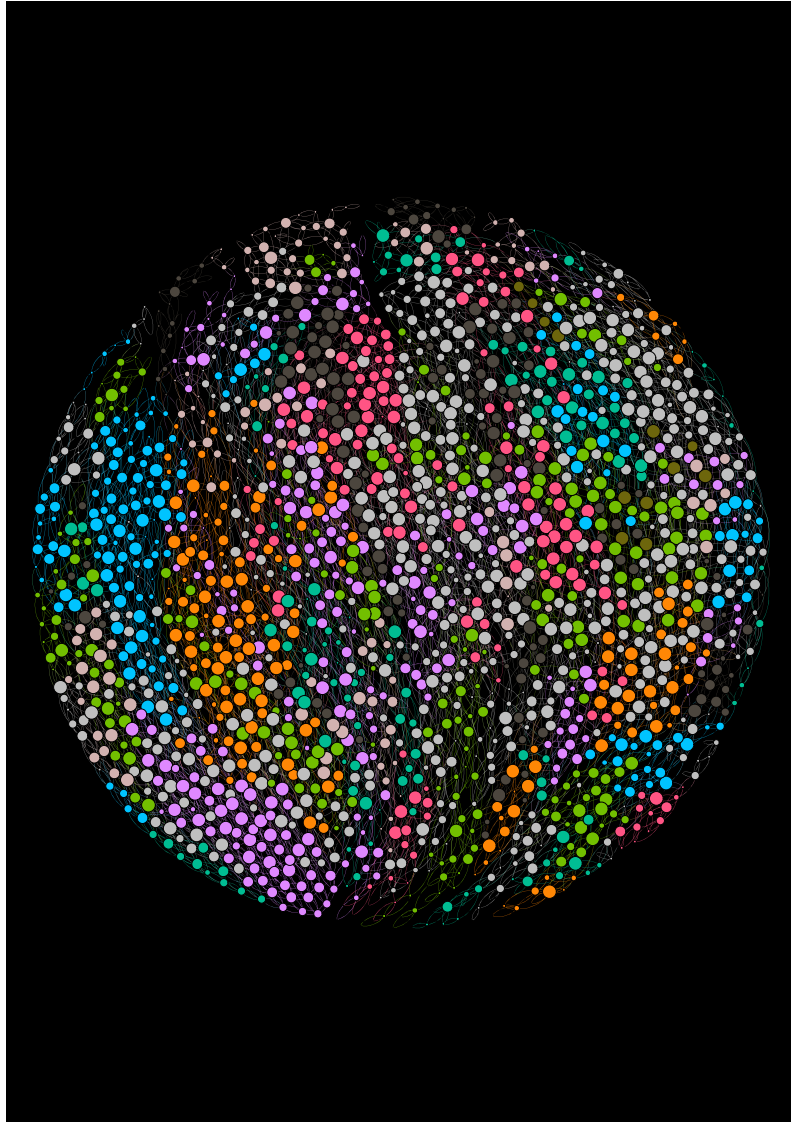
Figure 5.5 : A network generated by using the proposed model. The clusters of nodes with same color denote the communities
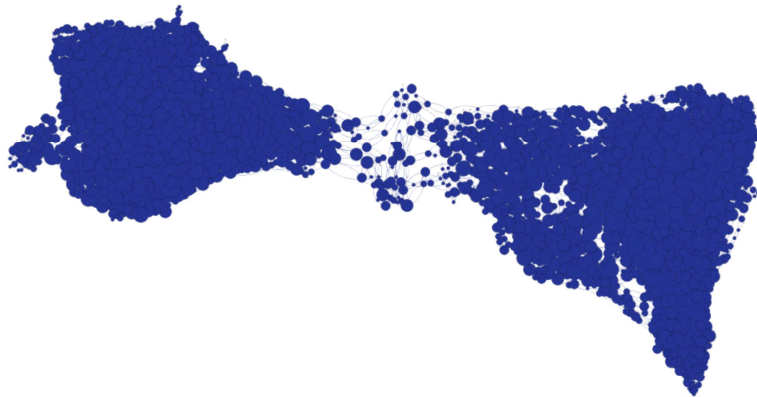
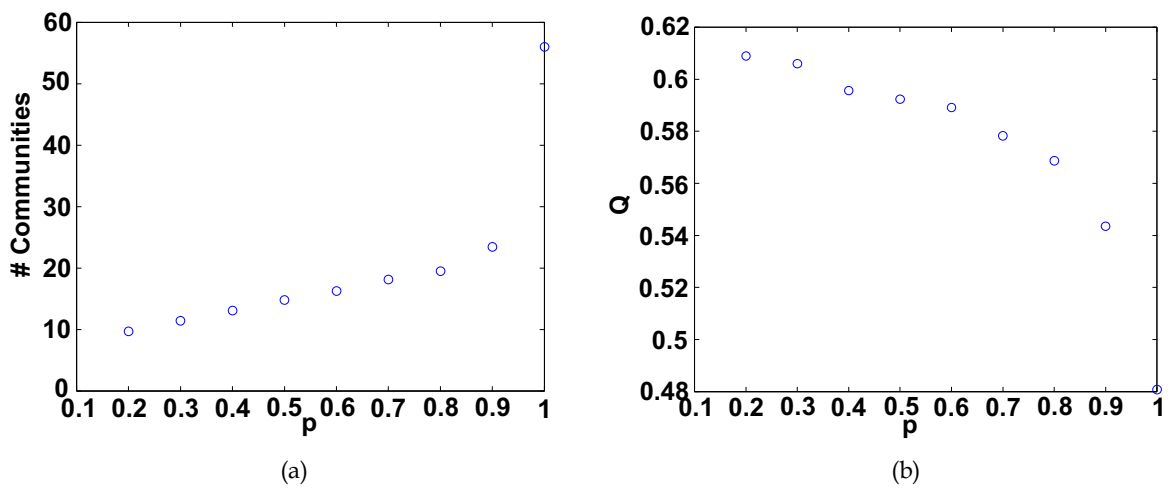**Figure 5.6 :** A network generated by crystallization process.

(a)

(b)

Figure 5.7 : In (a) Number of communities with respect to concentration. In (b) Modularity index Q with respect to concentration.

5.3 CONCLUSION

In this chapter we investigate crystallization in the perspective of network formation process. The basic rules of the network formation in the context of crystallization are defined that are able to explain the nucleation process as a network formation phenomena. A similar approach is used to define a social network model.

…