The impact of network characteristics on the diffusion of innovations

Renana Peres
School of Business Administration
Hebrew University of Jerusalem, Jerusalem, Israel 91905
peresren@huji.ac.il

August 2013

The impact of network characteristics on the diffusion of innovations

Abstract:

This paper studies the influence of network topology on the speed and reach of new product diffusion. While previous research has focused on comparing network types, this paper explores explicitly the relationship between topology and measurements of diffusion effectiveness. We study simultaneously the effect of three network metrics: the average degree, the relative degree of social hubs (i.e., the ratio of the average degree of highly-connected individuals to the average degree of the entire population), and the clustering coefficient. A novel network-generation procedure based on random graphs with a planted partition is used to generate 160 networks with a wide range of values for these topological metrics. Using an agent-based model, we simulate diffusion on these networks and check the dependence of the net present value (NPV) of the number of adopters over time on the network metrics. We find that the network metric that most influences diffusion is the relative degree of social hubs. This result emphasizes the importance of strong hubs. The effect of the average degree is positive but weaker. The clustering coefficient has a negative impact on diffusion, a finding that contributes to the ongoing controversy on the benefits and disadvantages of transitivity. These results hold for both monopolistic and duopolistic markets.

Keywords: clustering, average degree, degree distribution, agent-based models, diffusion of innovations, word of mouth, diffusion of innovations.

1 Introduction

The social influence processes that take place in a given network are shaped and affected by the network's topological characteristics. In this paper, we study how the topological or structural characteristics of a social network influence new-product diffusion in that network, in terms of speed of diffusion and the number of adopters.

Classical works in diffusion, focusing on the flow of information among individuals, assumed a fully connected market [1]. More recently, literature has begun to acknowledge the role of network topology in social influence processes, exploring diffusion in topologies such as small world networks [2] and scale-free networks [3]. Empirical studies have explored how diffusion is influenced by aspects of network structure, including weak, long-distance ties [4] and the existence of social hubs [5], and have examined how network structure affects the performance of marketing strategies such as new-product seeding [6,7].

Despite this interest, to our knowledge, the direct impact of topological network metrics on diffusion has not been studied. Specifically, there has not been a systematic assessment, carried out across multiple networks, testing the *simultaneous* and *direct* impact of multiple topological metrics on the magnitude and speed of diffusion. Although some comparative studies have been conducted, most of them compare network types (e.g.[8]), referring to the comprehensive set of properties characterizing each network rather than isolating the specific role of each structural dimension.

In this paper, we conduct a methodological investigation of the impact of network topology on the diffusion of a new product to the market. Specifically, we focus on the following structural metrics: the average degree, the relative degree of social hubs (i.e., the ratio between the average degree of the most-connected nodes and the overall average degree), and the clustering coefficient. We apply a graph-theory procedure called random graphs with a planted partition, which has so far not been used in network research, and use it to generate 160 networks, with a large range of values of the investigated metrics. We conduct agent-based simulations of new product diffusion in these networks, both in a monopoly and under duopolistic competition. We test the relative influence of each structural metric on the effectiveness of diffusion, measured as the Net Present Value (NPV) of the number of adopters. Our main findings are:

1. Among the investigated metrics, the relative degree of hubs has the strongest positive impact on diffusion. This result is interesting in light of the controversy on the contribution of social hubs [9,10]. Our results indicate that the effect of hubs is stronger than the effect of the overall average degree, which is also significant and positive, but weaker.
2. The clustering coefficient has a negative impact on diffusion. This result is in line with previous works comparing network types; however, it isolates the role of clustering from that of other topological network metrics. In addition, this finding contributes to an ongoing discussion on the benefits and disadvantages of transitivity (that is, the likelihood that the other nodes connected with a node are also connected to one another), of which clustering is a measure [11], implying the possible drawbacks of transitivity in the context of diffusion.

This paper offers three main contributions: First, it measures the *direct impact* of structural parameters on diffusion. We vary three major network metrics, and run a multivariate regression to explore simultaneously their *relative* roles. Second, we use a network generation method that has not been used so far in network research, to create a set of networks with a wide range of values for the various metrics we examine, without the need to use networks of different types. Third, the agent-based simulation enables us to represent real-life diffusion scenarios by (i) considering both a monopoly and a competitive market; (ii) using a cascade agent activation model [12]; this is the individual-level analog to the Bass diffusion model [1] , which is the standard model used to describe diffusion processes; and (iii) evaluating the effectiveness of the diffusion process by measuring the NPV of the number of adopters, which reflects both the reach of the diffusion process and its speed. Previous studies used such simulations, but did not focus on isolating the roles of specific topological metrics in the diffusion process.

The rest of this paper is organized as follows: In section 2 we review the literature and describe the topological metrics we use and their anticipated influence. Section 3 describes the network generation procedure. Section 4 describes the agent-based model. The results and conclusions are presented in sections 5 and 6, respectively.

2 Network structure and diffusion propagation

Social influence processes are strongly affected by the topology of the network in which they take place. Thus far, most studies focusing on the relationships between network topologies and social influence have been non-comparative in nature, with each paper focusing on a single network type. To the extent that comparison was done, it focused on comparing performance across different types of networks, and was mostly theoretic without measuring actual diffusion or information flow (see [13,14]; see [8] for an exception of an empirical study). For example, in small world networks, which are "rewired" lattice graphs in which several lattice ties are replaced with random connections, information flow and influence processes are expected to be more rapid than in regular lattice graphs, due to the "shortcuts" between nodes[1] [14 15]. Likewise, information flow in fully random graphs is expected to be rapid [16,17].

These studies provide important insights with regard to the influence of network structure on information flow. However, a specific network type usually binds several topological dimensions: For example, small-world networks combine both high clustering and short path length; most random graphs have a Poissonian degree distribution (with some exceptions; see [16]). In scale-free networks the clustering depends on the network size [13]. Thus, it is hard to isolate the relative role of each structural dimension on the basis of an overall comparison of different types of graphs. For example, if diffusion in a lattice graph is slower than in a random graph, is this because of the lower clustering of the random graph, or perhaps the variability in the degree of the nodes? Do the hubs in scale-free networks generate faster diffusion in comparison with other graphs, although the average degree in scale-free networks is very low? These questions are still unanswered.

---

[1] The network literature uses a variety of terms to describe network members and their connections. We use the term "network member" when talking about the real individuals in the network, "node" – to describe this person in the theoretical context, and "agent" when speaking about the agent-based model. In all contexts, we refer to the connections among network members, nodes, or agents as "ties".

The goal of this paper is to study systematically the influence of structural factors on the speed of diffusion. We focus on three structural dimensions: average degree, relative degree of social hubs, and clustering coefficient.

2.1 Average degree

The degree of a node is the number of ties it has with other nodes. The average degree of a network is the average of the degrees of all the nodes in the network, and can be considered as a metric of the network's level of connectivity. Empirical comparisons of the average degrees of real-life networks indicate values ranging from ~7 in some neural networks to <113 in a networks of film actors [18,19]. For theoretical networks, the average degree can often be derived from the networks' creation procedure: For example, for a regular lattice or for a Watts-Strogatz network, where all nodes have an equal number of *k* ties, the average degree is *k*, and a random Erdos-Renyi graph of *N* nodes with a tie probability *p* will have an average degree of $N \cdot p$.

2.2 Relative degree of social hubs

Social hubs are nodes with a high degree in relation to the degrees of other nodes. Not all networks have social hubs. A regular lattice does not have any. In random graphs and small world networks, the degree distribution is Poissonian [13], and social hubs do not differ much from the other members of the network. Social hubs are most prevalent in scale-free networks, where degree distribution follows a power law, with most of the nodes having a small number of ties, and a small number of nodes having an extremely high degree [13,20].

The contribution of social hubs to diffusion processes is a subject of ongoing debate. Some studies argue that hubs enhance the spread of information [21] and the speed of diffusion [6], and hence it is beneficial to target these individuals when attempting to introduce new products or ideas into a network [7]. Other studies claim that the role of hubs is complicated and depends on the level of contagion in the system [22], as well as on these individuals' inherent propensity to adopt [5]. Watts and Dodds (2007) suggest that in many cases, the large mass of less-connected units in a network determines the speed and magnitude of social influence [9].

Here, we contribute to this discussion by systematically varying the level of the relative degree of hubs across the networks we generate, and testing the influence of this variable, while controlling for other topological metrics. We measure the relative degree of social hubs as the *degree ratio* of the average degree of the top 10% most connected nodes to the overall average degree.

2.3 Clustering Coefficient

The clustering coefficient serves as a measure of a network's transitivity, that is, the likelihood that if nodes A and B are connected to each other, and nodes B and C are connected to each other, then nodes A and C are also connected. In other words, the clustering coefficient indicates the likelihood that a person in a given network is friends with the friends of his or her friends. Clustering can be measured either locally for each node—counting the number of its ties, and seeing how many of them are connected to each other—or globally, counting the numbers of "triangles" relative to the number of "open triples", where an open triple is a single node with ties to two other nodes. For example, if nodes *a*, *b* and *c* are all connected to each other, they form a single triangle, and three open triples: *abc*, *bac*, *acb*. Herein we use Newman's (2003) definition of a global clustering coefficient [19]:

(1) $C = \frac{3 X number\ of\ triangles\ in\ the\ network}{number\ of\ open\ triples}$.

Clustering is strongly dependent on network type, and can also vary to some extent across networks of the same type. For a one-dimensional lattice arranged in form of a ring, where each node is connected to its *k* nearest neighbors, the clustering coefficient is $C = \frac{3(k-2)}{4(k-1)}$, which approaches $C \sim 3/4$ for large values of *k*. In a random graph, clustering tends to be much lower, given by the approximation of *k/N*, where *N* is the network size, and *k* is the average degree of the network. The clustering coefficient of a small-world Watts-Strogatz network is between these values [14] and depends on the value of *p*, the rewiring probability of the regular lattice $C \sim \frac{3(k-1)}{2(2k-1)}(1-p)^3$ [23]. In scale-free networks generated by the Barabasi-Albert procedure, clustering is higher than that in a random graph, but lower than that in a small world graph, and is given approximately by $C \sim N^{-0.75}$ [13]. The strong dependence of clustering on network type can be an obstacle for systematically testing how clustering affects diffusion, since it is hard to determine whether differences result from the clustering or from other characteristics that are typical of specific network types, such as level of randomness, path length, etc.

Network theory is split with regard to whether transitivity is beneficial for network processes. In some contexts, a node, that is, an individual in the network, is better off not closing a triad, but rather forming ties with a new node; in other contexts, such as when redundancy and multiple points of influence are needed, individuals gain from closing a triad and increasing the network's transitivity (see [11] for review). Here, we shed some light as to the role of transitivity in the context of diffusion, by varying the clustering coefficient, and testing directly its effect on diffusion, controlling for other topological metrics.

Network literature suggests additional network metrics that might contribute to diffusion processes, including average path length, diameter, and density. We focus on the three discussed above because they represent independent topological network dimensions, and since they are global and can be pre-determined as input for the network generation procedure suggested below.

3 Generating the networks

To test how network metrics influence diffusion, it is necessary to generate a large number of networks, with a wide range of values for each metric under investigation. Since the standard network types (random, lattice, small-world, etc.) largely dictate dependencies among the network metrics, we introduce here a new network generation method, based on merging several random graphs, which can generate networks with a wide range of values for the average degree, relative degree of social hubs, and clustering coefficient. This method, termed "random graphs with a planted partition", is drawn from graph theory [24,25][2] and to the best of our knowledge has not yet been used for simulating social networks.

---

[2] We thank Michael Krivelevich for several useful discussions on this topic.

Assume a market of size N. For simplicity, ties are symmetric, i.e., if node a is connected to b, then b is also connected to a. The N nodes are organized into three separate bins (denoted bins 1, 2 and 3) of sizes N_1, N_2, and N_3, respectively. The probability that a customer from bin i and a customer from bin j are connected is p_{ij}. For three bins, there are six probabilities: p_{11}, p_{22}, p_{33}, p_{12}, p_{13}, p_{23}, as illustrated in Figure 1. If all these probabilities are the same, the graph is a standard random graph. Manipulating the probabilities allows one to create random networks with different values for the average degree, degree ratio and clustering as follows:

The average degree of the nodes in bin i is given by $D_i = p_{ii} \cdot (N_i - 1) + p_{ij} \cdot N_j + p_{ik} \cdot N_k$, where i,j,k=1,2,3 and $i \neq j \neq k$. The term $p_{ii} \cdot (N_i - 1)$ is the average number of ties that the node has with other nodes in bin i (assuming a node cannot connect to itself), and the two other terms are the number of ties that the node has with nodes in bins j and k, respectively. Hence, the overall average degree is $D = \sum_i (D_i \cdot N_i / N)$. If we arbitrarily define bin 1 as the bin of hubs and set p_{11} to be higher than the other probabilities, then we can define the relative degree of hubs as D_1/D.

Figure 1: A random graph with a planted partition

To calculate the clustering coefficient, we need to calculate the number of triangles and the number of open triples. The number of triangles is given by the following expression:

$$(2)\quad Triangles = \sum_{i=1}^{3} \binom{N_i}{3} p_{ii}^3 + \sum_{i=1; j,k \neq i}^{3} \binom{N_i}{2} p_{ii} \left(p_{ij}^2 N_j + p_{ik}^2 N_k\right) + N_1 N_2 N_3 p_{12} p_{13} p_{23}$$

The first sum is the number of triangles in which all nodes belong to the same bin. The second sum is the number of triangles in which two nodes are in the same bin and one is in a different bin: If $i = 1$, for example, we first choose a pair of nodes in bin 1 (there are $\binom{N_1}{2}$ such pairs); their probability to be connected is p_{11}. Taking randomly a node in bin 2 (there are N_2 such nodes), the probability that both nodes of bin 1's pair will be connected to it is p_{12}^2; the same logic applies for bin 3. The third term is the number of triangles in which each node is in a different bin.

The number of open triples is calculated in a similar way. A unit in bin i can be in an open triple with either two other units in i, two units in j, or two units in k, or with one in i and one in j, one in i and one in k, or one in j and one in k. Since there are N_i units in bin i, the number of open triples that contain a node from bin i is given by:

(3)

$$OpTr_i = N_i\left(\binom{N_i-1}{2}p_{ii}^2 + \binom{N_j}{2}p_{ij}^2 + \binom{N_k}{2}p_{ik}^2 + (N_i-1)N_j p_{ii}p_{ij} + (N_i-1)N_k p_{ii}p_{ik} + N_j N_k p_{ij}p_{ik}\right)$$

The clustering coefficient is then calculated as $C = \frac{3X\ Triangles}{OpTr_1+OpTr_2+OpTr_3}$.

Given values of the average degree, relative degree of hubs, and clustering coefficient, we can calculate the probabilities that satisfy these values and then use a random number generator to create the networks corresponding to these probabilities. Note that this procedure does not guarantee a solution for every degree/ratio/clustering combination, or that an obtained solution is unique (since we have 6 probabilities and 3 equations). However, since the goal is to generate networks with a desired structure and not to obtain a specific network, this procedure is adequate for our purposes.

This procedure was used to generate 160 networks, using the following values: $N = 1000$, $N_1 = 100$, $N_2 = 450$, and $N_3 = 450$. The distribution of parameters across the networks is described in Figure 2. As Figure 2 illustrates, the average degree ranges from 2 to 49 (average is 24.5), the clustering coefficient ranges from 0.01 to 0.48 with an average of 0.21, and the relative degree of hubs (i.e., the ratio between the average degree of the 10% most connected nodes and the overall average degree) ranges from 1.09 to 6.7 with an average of 3.96. Figure 3 illustrates the network structure of a network with an average degree of 7.5, relative degree of hubs of 6.67, and a clustering coefficient of 0.45. The probabilities are: $p_{11} = 0.49$, $p_{22} = 0.0001$, $p_{33} = 0.012$, $p_{12} = 0.013$, $p_{13} = 0.013$, $p_{23} = 0$.

Figure 2: The distribution of parameters across the 160 networks

Figure 3: An example of a random graph with a planted partition. Average degree = 7.5, relative degree of hubs = 6.67, clustering coefficient 0.45. $p_{11} = 0.49$, $p_{22} = 0.0001$, $p_{33} = 0.012$, $p_{12} = 0.013$, $p_{13} = 0.013$, $p_{23} = 0$.

We next use these networks as the basis for an agent-based diffusion model.

4 An agent-based model for simulating diffusion

An agent-based model is used here to describe the diffusion process, in line with the agent-based econophysics approach [26]. Agent-based models are increasingly used for describing complex economic systems, and have been extensively used to describe diffusion of innovations [27,28]. Here, a network with a pre-specified structure is created through the random-graph-with-a-planted-partition generation process described above, and an agent-based model is used to simulate the process of new product diffusion in this network. The model is composed of 1000 agents, and is used to describe the diffusion of both a monopolistic firm with no competition and a duopoly with two competing firms.

4.1 Adoption Probabilities

For each network, we simulate the diffusion of the adoption of a new product. Assume first a monopolistic case. Time is discrete, the market starts with all agents at state "0", and when an agent adopts, its state changes to "1". We assume that adoption is uni-directional, that is, an agent can convert from 0 to 1, but cannot dis-adopt and convert from 1 to 0. The transition rule is based on classical diffusion theory, which suggests that the adoption decision is a result of the combined influence of two factors: ***external influence***, represented by the probability δ that an agent will be influenced by sales people, advertising, promotions, and other marketing efforts[3]; and ***internal influence***, which refers to the influence of all means of social interaction such as word of mouth, or imitation. We denote by q_i the susceptibility of agent i to the internal influence of a single other agent, i.e., the probability that a given agent in a given time period will convince agent i to adopt.

The agent activation rule used here is based on a competing risk, or a cascade, approach, where at time t, each prior adopter connected to agent i independently tries to convince i to adopt. Thus, the discrete-time hazard of i to adopt is 1 minus the probability that all these adopters, as well as the advertising efforts, failed the task: $P_i(t) = 1 - (1-\delta)(1-q_i)^{S_i(t)}$, where $S_i(t)$ is the number of adopters in i's personal social network [29].

Note that this formulation is not the only means of describing contagion in agent-based models. Other models, such as those based on the Ising model analogue, have used a threshold-based approach, where the agent changes states when a certain threshold of utility is reached [6]. However, the competing-risk formulation is more appropriate for modeling new product diffusion, since it converges to the Bass model as the discrete time interval reduces to zero [30]. In addition, this approach considers the overall influence from all agents connected to the potential adopter i, unlike other models, which choose randomly the agent that influences i [22].

Libai, Muller and Peres [18] extended this monopolistic model to describe adoption in a competitive scenario, as follows: if two firms, A and B, compete in the market, then an agent can be in one of three states "0", "A" or "B". Each of the competing firms is assigned its own values for external influence, δ_{iA} and δ_{iB}, and for the internal influence of a single agent, q_{iA} and q_{iB}.

---

[3] In the diffusion literature δ is often denoted as p. Here, p is used to represent probabilities.

Adopters of A and B each independently influence a potential adopter *i* to adopt their respective firms. The probability of *i* being successfully convinced by at least one adopter of A or B is given by:

$P_i^A(t) = 1-(1-\delta_{iA})(1-q_{iA})^{S_i^A(t)}$; $P_i^B(t) = 1-(1-\delta_{iB})(1-q_{iB})^{S_i^B(t)}$,

Where S_i^A and S_i^B denote all agents in *i*'s personal social network who have adopted either A or B. Now, in a discrete time period *t*, there could be one of three scenarios: **a)** Agent *i* is convinced to adopt from A but is not convinced to adopt from B. The probability of this happening is $P_i^A(1-P_i^B)$. **b)** Agent *i* is convinced to adopt from B only. The probability of this is $P_i^B(1-P_i^A)$. **c)** Agent *i* is persuaded to adopt from both A and B, but as per the model assumptions has to choose one of the two. The probability of this happening is $P_i^A P_i^B$, and the agent adopts from A rather than B according to the ratio of the probabilities, $\lambda_{iA} = \frac{P_i^A}{P_i^A+P_i^B}$. The probabilities of *i* actually adopting from firm A, adopting from firm B, or not adopting from either are, respectively:

$$P_i(adopt\ A) = P_i^A(1-P_i^B) + \lambda_{iA} P_i^A P_i^B$$

(4)

$$P_i(adopt\ B) = P_i^B(1-P_i^A) + \lambda_{iA} P_i^B P_i^A \quad P_i(adopt\ none) = (1-P_i^A)(1-P_i^B)$$

where $\lambda_{iA} = \frac{P_i^A}{P_i^A+P_i^B}$, $\lambda_{iB} = 1-\lambda_{iA}$

In the simulation, for each agent in each period, the adoption probability is realized by drawing a random number from a uniform distribution and comparing it to adoption probabilities P_i(*adopt A*) and P_i(*adopt B*). If this number is between 0 and P_i(*adopt A*), A is adopted; if it is between P_i(*adopt A*) and P_i(*adopt A*) + P_i(*adopt B*), B is adopted; and if it is between P_i(*adopt A*) + P_i(*adopt B*) and 1 there is no adoption (since the number is randomly drawn, the order A, B does not matter).

A single run of the simulation (namely, an adoption process starting from zero with a given network structure, δ and *q*) ends after 30 time periods, which is consistent with common practice in similar models [29]. Given the parameter values used here, the 30 time periods are such that most of the market has adopted by that time.

For simplicity, δ_{iA} and δ_{iB} are assumed to be equal for firms A and B, and identical across agents. To take into account the possible heterogeneous nature of customer propensity to be affected by others, the value of q is assumed to be normally distributed throughout the network. For robustness, we also examined cases in which q was distributed in a power law distribution with the power-law exponent parameter simulated in the commonly used range of 2-3. We also looked at a uniform distribution in which the range was plus minus the standard deviation used in the Normal distribution analysis. The results reported next are robust to the specification of q.

The values of δ and q were chosen to be consistent with previous research regarding the ranges of these parameter values in diffusion models and in agent-based models (e.g.[19]). The parameter δ was assigned the following values: δ = 0.001, 0.005, 0.01, 0.05, 0.1. The values of q differ across networks: In networks with high average degrees, our preliminary simulations show that the interesting dynamics occur for lower values of q, (since the combination of high q values and high degree generates almost instantaneous diffusion). Therefore, two sets of q values were used: For the 80 networks with lower average degrees, we used a distribution with an average q of 0.005, 0.01, 0.02, 0.03, 0.04, and for the 80 networks with higher average degrees, we used distributions with an average q of 0.08, 0.1, 0.12, 0.16. The experimental design was a full factorial experiment, where for each network we tested all combinations of the different values of δ and q (from the appropriate set).

For each of the 160 networks, we ran the simulation 6000 times: we ran the simulation for each combination of δ and q as elaborated above, for both monopolistic and duopoly scenarios, and, to control for random effects, for each parameter set we ran 120 simulations. Simulations were conducted using C++ code. The overall runtime on an Intel 3.1 GHz i5-2400 core processor, 16GB RAM was 47 days.

4.2 The time value of diffusion

An effective diffusion process is quick, and it affects a large number of network members. To measure the effectiveness of diffusion, we evaluated the NPV of the number of adopters, that is, the discounted sum of the number of adopters. Thus, for a given firm μ the NPV is $\sum_{t=1}^{T} S_\mu(t)/(1+d)^t$, where $S_\mu(t)$ is the number of agents who adopted from firm μ (μ=A,B) at time period t. Note that in our simulations, the firms are symmetric, so $S_A(t)=S_B(t)$. T is the total time horizon, and d is the discount factor. In the simulation, a standard discount factor of 10% is used.

5 Results

To estimate the impact of network topology on diffusion, we examined how the NPV of the number of adopters is dependent on the various topological metrics and diffusion parameters. Figure 4 comprises three graphs, illustrating, respectively, the dependence of the *ln_NPV* on each of the three topological metrics we used. For illustration purposes we present values for δ=0.005 and q=002, averaging the values of each remaining variable across all simulation runs and across all the networks. The *ln* was used to enable a comparison to be made between the coefficients of the monopoly and duopoly cases.

To evaluate the relative role of each of the topological metrics in the diffusion process, we ran a multivariate regression, where *ln_NPV* was regressed simultaneously against the topological metrics and diffusion parameters. The regression data were pooled over the 160 networks. The explanatory variables were the average degree (*Degree*), the relative degree of hubs (*Rel_hubs*), and the clustering coefficient (*Cluster*). The diffusion parameters δ and q were mean-centralized. The resultant regression equation was the following:

(5) $\ln\_NPV = \alpha_0 + \alpha_1 Degree + \alpha_2 \mathrm{Re}l\_hubs + \alpha_3 Cluster + \alpha_4 \delta + \alpha_5 q + \varepsilon$.

Each data point in the regression was the average over the 120 runs with the same parameter combination. The estimation was performed separately for the monopoly and duopoly conditions. For the duopoly case the NPV of one of the firms (firm A) was used; since both firms are identical, the choice of A or B is equivalent.

Figure 4: Network performance as a function of topological metrics, for monopoly and duopoly. $\delta = 0.005$, $q = 0.02$, averaged on all other parameters and on all networks.

The results are displayed in Table 1. The table presents the regression coefficients for the monopoly and duopoly cases. In both cases, the topological metric that has the strongest impact on diffusion is the relative degree of hubs. That is, the higher the ratio between the average degree of the top 10% most connected agents and the overall average degree, the more effective the diffusion process. The effect of the average degree is also positive, but smaller. This result contributes to the controversy as to the importance of social hubs to the diffusion process: When controlling for diffusion parameters and other network metrics, the impact of the average degree of social hubs is positive and strong, even more than that of the overall average degree of the network.

The effect of the clustering coefficient is negative, indicating that global clustering has a negative effect on diffusion. This finding contributes to the discussion on the role of transitivity in network processes [11]: While transitivity might have benefits, in the context of diffusion its impact is negative. By varying the clustering coefficient independently of other network metrics, and testing it simultaneously with the other metrics, we have managed to isolate its effect and assess its relative role in diffusion.

Interestingly, these results are robust to the competitive market structure and are the same for both monopolistic and duopolistic markets. All results are significant with $p < 0.001$.

Table 1: Estimation results in monopolistic and duopolistic markets

6 Discussion and Conclusions

This paper focuses on the effects of average degree, relative degree of hubs and clustering coefficient on the diffusion of a new product. We introduce a network generation procedure, based on random graphs with a planted partition, to generate 160 networks encompassing a large range of values for the parameters under investigation. Using agent-based models, we simulate diffusion processes on these networks, for monopolistic and duopolistic markets.

By directly manipulating the topological metrics and measuring their impact simultaneously, we can *isolate the relative influence* of each metric. We find that the relative degree of hubs, as well as the average degree, have strong and positive effects on diffusion. The clustering coefficient, however, has a negative impact on diffusion: the higher the level of global clustering, the weaker the diffusion process.

These findings shed light on the ways in which underlying network topologies influence diffusion, and they contribute to two ongoing discussions in the network literature. The first result emphasizes the importance of hubs, i.e., nodes whose degrees are relatively high compared with the degrees of the rest of the population. Specifically, our analysis suggests that a degree distribution that tends towards a power-law is likely to be associated with a more effective diffusion process (in terms of NPV) compared with a more uniform degree distribution. This result demonstrates that the role of social hubs in diffusion processes is intricate and depends on the aspect of diffusion that is being tested: While the relative degree of hubs might be less important when measuring the length of individual cascades of influence (the length of a single information, or influence chain) [9], it is important to the overall effectiveness of diffusion.

Our finding regarding the negative impact of clustering contributes two interesting insights to the ongoing discussion on the pros and cons of network transitivity [11]: First, the result sheds light on the role of transitivity, in isolation from other network metrics. While most other studies have compared network types, thereby considering clustering in combination with other network metrics, here we measure the direct effect of clustering on diffusion. Second, our result demonstrates the drawbacks of transitivity in the context of diffusion. It seems that the redundancies generated by high clustering impede diffusion.

Note that although this paper studies a wide range of parameter values, its findings cannot automatically be generalized to all network types. Since the network creation procedure is based on random graphs, the degree distribution is a Poissonian mix function, and is different from some degree distributions found in real networks, such as the power law distribution. The results are not expected to be fundamentally different for such networks; however, further research is needed to verify this empirically.

7 Acknowledgements

This paper was supported by the Israel Science Foundation, The Israeli Association for Internet Study, and the Kmart fund. My thanks to Michael Krivelevich for discussion on random graphs with a planted partition. I also thank Zeev Rudnick, Barak Libai and Eitan Muller. This work is supported by the Israeli Internet Association, and the Kmart foundation.

8 References

---

[1] F.M. Bass, A new product growth model for consumer durables, Manage. Sci. 15 (1969) 215-227.

[2] J. Goldenberg, B. Libai, E. Muller, Talk of the network: A complex systems look at the underlying process of word-of-mouth, Market. Lett. 12 (2001) 211-223.

[3] M. Barthelemy, A. Barrat, R. Pastor-Satorras, A. Vespignani, Velocity and hierarchical spread of epidemic outbreaks in scale-free networks, Phys. Rev. Lett. 92 (2004) 178701-1-4.

[4] M. Granovetter, The strength of weak ties, Am. J. Sociol, 78 (1973) 1360-1380.

[5] J. Goldenberg, S. Han, D. Lehmann, J. Hong, The role of hubs in the adoption process, J. Marketing, 73 (2009) 1-13.

[6] C.E Laciana, S.L. Rovere, Ising-like agent-based technology diffusion model: Adoption patterns vs. seeding strategies, Physica A, 390 (2011) 1139-1149.

[7] F. Stonedahl, W. Rand, U. Wilensky, Evolving viral marketing strategies, in: Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation, ACM, New York, 2010, pp. 1195-1202.

[8] M. Uchida, S. Shirayama, Influence of a network structure on the network effect in the communication service market, Physica A, 387 (2008) 5303-5310.

[9] D. J. Watts, P.S. Dodds, Influentials, networks, and public opinion formation, J. Consum. Res. 34 (2007) 441-458.

[10] K. Keller, J. Berry, The Influentials: One American in Ten Tells the Other Nine How to Vote, Where to Eat, and What to Buy, The Free Press, New York, 2003.

[11] M.J. Burger, V. Buskens, Social context and network formation: An experimental study, Soc. Networks 31 (2009) 63–75.

[12] J. Leskovec, L.A. Adamic, B.A. Huberman, The dynamics of viral marketing, ACM TWEB 1 (2007) 5-51.

[13] R. Albert, A.L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys. 74 (2002) 47–97.

[14] D. J. Watts, S. H. Strogatz, Collective dynamics of ‘small-world’ networks, Nature 393 (1998) 440-442.

[15] M. E. J. Newman, D. J. Watts, Scaling and percolation in the small-world network model, Phys. Rev. E, 60.6 (1999)7332-7342.

[16] M.E.J. Newman, D.J. Watts, S.H. Strogatz, Random graph models of social networks, Proc. Natl. Acad. Sci. U.S.A 99 Suppl 1(2002) 2566-2572.

[17] P. Erdős, A. Rényi, On random graphs, Publicationes Mathematicae Debrecen 6 (1959) 290-297.

[18] B. Libai, E. Muller, R. Peres, Decomposing the value of word-of-mouth seeding programs: Acceleration vs. expansion, J. Market. Res. 50 (2013) 161-176.

[19] M.E.J. Newman, The structure and function of complex networks, SIAM R 45 (2003) 167–256.

[20] M. Faloutsos, P. Faloutsos, C. Faloutsos, On power-law relationships of the internet topology, ACM SIGCOMM Comp. Comm. R. 29 (1999) 251-262.

[21] M.A. Janssen, W. Jager, Simulating market dynamics: Interactions between consumer psychology and social networks, Artif. Life 9 (2003) 343-356.

[22] L. Deng, Y. Liu, F. Xiong, An opinion diffusion model with clustered early adopters, Physica A 392 (2013) 3546–3554.

[23] A. Barrat, M. Weigt, On the properties of small-world network models, Eur. Phys. J. B 13 (2000) 547-560.

[24] S. Fortunato, Community detection in graphs, Phys. Rep. 486 (2010) 75-174.

[25] A. Condon, R.M. Karp, Algorithms for graph partitioning on the planted partition model, Random Struct. Algor.,18 (2001) 116-140.

[26] C. Schinckus, Between complexity of modelling and modelling of complexity: An essay on econophysics, Physica A 392 (2013) 3654-3665.

[27] M. Haenlein, B. Libai, Targeting revenue leaders for new products, J. Marketing 77 (2013) 1-55.

[28] R. Garcia, Uses of agent-based modeling in innovation/new product development research, J. Prod. Innovat. Manag., 22 (2005) 380-398.

[29] J. Goldenberg, B. Libai, E. Muller, Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata, AMS Rev. 9 (2001) 1-18.

[30] J. Goldenberg, O. Lowengart, D. Shapira (2009), Zooming in: Self emergence movements in new-product growth, Market. Sci. 28 (2) 274-292.