- Hiltrop, J. M., 1999, The Quest for the Best: Human Resource Practices to Attract and Retain Talent, *European Management Journal*, (17), pp. 422–430.
- Isaac, M., & Senge, P., 1994, *Overcoming limits to learning in computer-based learning environment*, in: Morecroft, J. W., & Sterman, J., (ed.), Modeling for Learning Organizations, Productivity Press (Portland), pp. 270–273.
- Kim, D., 1989, *Learning Laboratories: Designing a Reflective Learning Environment*, in Milling, P., & Zahn, E., (ed), Computer-Based Management of Complex Systems, Springer-Verlag (Berlin).
- Legare, T.L., 1998, The human side of mergers and acquisitions, *Human Resource Planning* (21), pp. 32–41.
- Liker, J., Choi, T., 2004, Building Deep Supplier Relationships, *Harvard Business Review, December*, pp. 104–113
- Locke, W. A., & Latham, G. P., 2002, Building a practically useful theory of goal setting and task motivation: A 35-year odyssey, *American Psychologist*, 57, pp. 705–717.
- Matsui, T., Okada, A., & Inoshita, O., 1983, Mechanism of feedback affecting task performance, *Organizational Behavior and Human Performance*, (31), pp. 114–122.
- Morecroft, J., 1984, Strategy Support Models, *Strategic Management Journal*, (5), pp. 215–229.
- Pfeffer, J., 1994, *Competitive Advantage through People: Unleashing the Power of the Work Force*, Harvard Business School Press (Boston).
- Pfeffer J., 1998, *The human equation*, Harvard Business School Press (Boston).
- Ross, W.T. Jr, Anderson, E., & Weitz, B.A., 1997, The causes and consequences of perceived asymmetry of commitment to the relationship, *Management Science*, 43, pp. 680–704.
- Seijls, G.H., & Latham, G.P., 2001, The effect of learning, outcome, and proximal goals on a moderately complex task, *Journal Organizational Behaviour*, (22), pp. 291–307
- Senge, P., 1990, *The Fifth Discipline: The Art and Practice of the Learning Organizations*, Doubleday (New York).
- Sterman, J., 1994, Learning in and About Complex Systems, *System Dynamics Review*, (10), 2–3
- Sterman, J., 2000, *Business Dynamics. Systems Thinking and Modeling for a Complex World*, Mc Graw Hill (Boston).
- Vennix, J., 1996, *Group Model Building*, Wiley (Chichester).
- Winch, G., 2001, Management of the ‘Skills Inventory’ in times of Major Change, *System Dynamics Review*, (17), pp. 151–159.
- Wright, P. M., & McMahan, G. C., 1992, Theoretical Perspectives for Strategic Human Resource Management, *Journal of Management* (182), pp. 295–320.
- Wright, P., 1998, Human resources-strategy fit: Does it really matter?, *Human Resource Planning*, (21), pp. 56–57.
- Yilmaz, C., Sezen, B., & Turner Kabaday, E., 2004, “Supplier fairness as a mediating factor in the supplier performance–reseller satisfaction relationship”, *Journal of Business Research* (57), pp. 854–863.

Chapter 13

Diffusion and Social Networks: Revisiting Medical Innovation with Agents

Nazmun N. Ratna*, Anne Dray, Pascal Perez, R. Quentin Grafton,

David Newth, Tom Kompas

*Crawford School of Economics and Government
ANU College of Asia and the Pacific
The Australian National University, Canberra, Australia
nazmun.ratna@anu.edu.au

1. Introduction

In this chapter we reanalyze *Medical Innovation* by Coleman, Katz and Menzel (1966), the classic study on diffusion of Tetracycline, which at that time was a newly introduced antibiotic. Their pioneering study elaborated on how different patterns of interpersonal communications can influence the diffusion of a medical innovation in four medical communities in Illinois.

The motivation for our reanalysis is to capture the complex interactions involved in the diffusion process by combining Agent-based Modeling (ABM) and network analysis. Based on the findings in *Medical Innovation*, we develop a diffusion model called *Gammaym*. The topology of networks generated in *Gammaym*, and its evolution, are analyzed to evaluate the network structure influencing the diffusion process.

We describe the original study and the rationale for our study in the following section. Section 3 describes the modeling framework, modeling sequences and methods. Simulation results under different scenarios are analyzed in Section 4. The structure of the social networks depicted in our model, and its evolution are analyzed in Section 5. We explore the significance of types of network integration after normalizing adoption curves in terms of numbers for professionally integrated and isolated doctors in Section 6. The paper concludes with discussions and a review of the implications of the simulation results.

2. Diffusion and Networks

2.1. Diffusion of Tetracycline

Tetracycline was launched in November 1953. The success of its release was uncertain for the pharmaceutical companies because of an already established market for broad-spectrum antibiotics. Four US Midwestern cities: Peoria, Bloomington, Quincy and Galesburg, were selected by Coleman et al. who wanted to set their study in a group of three or four cities in one contiguous area that was not in the shadow of a large medical center, but where the cities differed from each other in terms of available hospital facilities: the number of teaching hospitals and general hospitals and population. The Coleman et al. sample constituted 148 general practitioners, internists, and pediatricians in active practice, of which 126 (85 percent of the sample) were interviewed. In an attempt to evaluate the importance of social networks, each of them was asked about their close associates (e.g., friends, colleagues and advisors) in the medical community. In order to measure the time of adoption, a prescription audit in the local pharmacies was carried out for 125 doctors (121 general practitioners, internists, and pediatricians, and 4 listed as surgeons or proctologists) over a 16-month period following the release of Tetracycline for general sale. The prescriptions of the doctors were audited for three successive days at approximately monthly intervals (Coleman et al. 1966: 194).

Coleman et al. (1966) study identifies two broad categories of variables influencing the diffusion process. The first category represents personal traits or individual variables, affecting individual receptivity. Individual variables are: 1. type of practice, 2. medical background, 3. contacts with out-of-town institutions, 4. media behavior, and 5. orientations and attitudes. The second category defines social variables influencing the adoption process as a result of social or professional ties to other members of the community. The influences of professional networks were evaluated on the basis of four factors that enable exchange of information in professional settings: 1. hospital affiliation, 2. shared office, 3. advice seeking, and 4. discussion. The friendship structure in the original study, emerged from the informal discussion among doctors in a social setting and doctors were asked to name three doctors whom they see most often socially.

The original study revealed strong evidence for social contagion, i.e., doctors' decisions to adopt Tetracycline were strongly influenced by the people they were connected to either socially or professionally. It appeared that the 'integrated' doctors differed very little from their isolated colleagues at the beginning of the introduction of Tetracycline, but their rate of adoption accelerated to produce an increasing gap between 'integrated' and 'isolated' doctors (Coleman et al. 1966).

2.2. Social Contagion vs. Media Exposure

Burt (1987) (1327) used the Coleman et al. study to derive three major conclusions: "1. where contagion occurred, its effect was through structural equivalence"¹ not

¹ "Structural equivalence is the degree that two individuals occupy the same position in a social system" (Valente 1995:55). A structural equivalence model of diffusion postulates that individuals are

cohesion; 2. regardless of contagion, adoption was strongly determined by a physician's personal preferences, but these preferences did not dampen or enhance contagion; 3. there was no evidence of a physician's network position influencing adoption when contagion was properly specified in terms of structural equivalence".

Strang and Tuma (1993) apply an event-history framework on the Coleman et al. (1966) data to analyse a class of diffusion models incorporating spatial and temporal heterogeneity. Their result contradicts Burt, as their models indicate that, "Contagion in medical innovation is not a simple product of structural equivalence. Cohesive ties based on advice giving and discussion also contribute to diffusion, as do structures of similarity in physicians' orientation towards their work" (Strang and Tuma 1993: 638).

Valente (1995) tests his threshold/critical mass (T/CM) model on *Medical Innovation* data. The T/CM model requires a threshold measure and a determination of the critical mass, "a system level measure of the minimum number of participants needed to sustain a diffusion process (Valente 1995: 79). The study indicates that the opinion leaders, who have greater exposure to external influence, play a dominant role in the diffusion process. Another study by Van Den Bulte and Lilien (2001) provides strong support for external influence in the diffusion of tetracycline by incorporating a data set on advertisement volume. As the empirical results were unable to detect statistically significant contagion effect once advertising effect is controlled for, the authors conclude that the data do not show that diffusion was driven by contagion operating over social networks, and that earlier analyses confounded the contextual effect of the aggressive marketing effort by the pharmaceuticals with social contagion (Van den Bulte and Lilien 2001).

2.3. Rationale for Gammamym

We reanalyze the *Medical Innovation* study by applying ABM for two principal reasons: the non-inclusion of dynamics of network structure and complexity in any of the previous studies, and a limited exploration of the impact of external influence or media exposure from previous reanalyses of the Coleman et al. data.

The major limitation of all the previous studies is their static exposition of network structure, which falls short of representing the evolving process. These dynamics can be described by the *dynamics of the network*, or the evolving structure of the network itself such as the making and breaking of network ties. The other set of dynamics refers to *dynamics on the network*, created due to the actions taken by individuals. The outcome of their individual actions is influenced by what their neighbors are doing and, therefore, the structure of the network. In the real world, both kinds of dynamics exist simultaneously (Watts 2003). Until now, however, neither *Medical Innovation* study nor the subsequent reanalyses of the original dataset have incorporated any of those dynamic features. Our study therefore complements the extant work.

Our work also makes a contribution in that the complexity generated in the diffusion process has not been examined by any previous studies. In Coleman et al. (1966) study, the extent of influence was evaluated for pairs of individuals. Individual network (discussion, friendship or advice) was, therefore, perceived as a set of

discrete/disjointed pairs. Given the existence of overlapping networks and consequent influences on doctors' adoption decisions, the complexity of actual events can not be captured by pair analysis. Indeed, the authors themselves considered that analyzing in pairs was the major limitation of the study². ABM enables us to address this limitation by considering the whole network as a unit of analysis.

Apart from the study by Van den Bulte and Lilien (2001), external influence in the diffusion process has also not been analyzed rigorously for the *Medical Innovation* data set. External influence is created from different forms of advertising and intensive coverage by representatives from the drug companies. In the *Medical Innovation* study, doctors acknowledged a 'detailman' or a pharmaceutical representative, and also direct mail from the drug companies, as the major sources of initial information about Tetracycline. The original study however did not contain any information on marketing effort by drug companies. An attempt to collect data on marketing effort was constrained by the fact that back issues of the *Journal of the American Medical Association* (JAMA) did not have any advertising supplement, as they were removed before binding for storage (Van Den Bulte and Lilien 1999). In our study, we examine the impact of external or media influence by performing sensitivity analysis. The simulation results allow us to explore the actions of the pharmaceutical companies, which has not been examined in a MAS (Multi-Agent System) setting before.

3. Modeling Framework

Using the Smalltalk programming language, we develop *Gammamym* with the CORMAS platform under VisualWorks environment. The codes and the detailed UML diagrams for *Gammamym* are available in the CORMAS website: <http://cormas.cinrad.fr/en/applica/gammamym.htm>.

First, we discuss the spatial representation and the rationale for specified attributes in *Gammamym* for each objects/agents with reference to the original study. The decision making process by the social agents is then described and the section concludes with a discussion on modeling sequence.

3.1. Spatial Representation and Passive Objects

Based on the distribution of office partnership and number of hospitals in the original study, we represent the medical community in an 8 X 8 spatial grid. The unit cell captures three different locations for professional interactions: Hospitals, Practices and a Conference Center, which are created as passive objects. In the original study, on average 47 percent of doctors were alone in an office, 20 percent were in clinics,

17 percent were working with two colleagues and 15 percent were sharing an office with one colleague. We captured the categories of office partnership into three practice types³: *Private* (alone in office), *Center* (shared office with two partners) and *Clinic* (working with four colleagues). The doctors in the *Gammamym* model are thus distributed among 46 private, 11 centers and 4 clinics. *Gammamym* specifies two hospitals, as all the cities, on average, have two hospitals. Conference center, the third passive entity, provides the context in which a much larger group of professionals can interact with each other. Although the original study did not specify a conference center as such, our rationale for this specification is to signify the importance of contact with out-of-town institutions (Coleman et al. 1966: 44).

Our model, *Gammamym*, has sixty-one practices, two hospitals and one conference center, randomly located over an 8 X 8 spatial grid. Random allocation of these practices over the grid reflects that spatial representation is not sensitive to distance. In other words, the doctors' decision to go to hospitals or the conference center does not depend on the location of their practices.

3.2. Social Agents

Gammamym depicts two kinds of social agents - *Doctor* and *Laboratory*. Initially located in their respective practices, *Gammamym* creates 99 doctors, the principal agents in the diffusion process. A laboratory, on the other hand, influences doctors' adoption decisions by sending information through multiple channels: medical representatives, journals and commercial flyers.

Located and Communicating Agents: Doctors

In *Gammamym*, doctors are specified with the attributes generating network effects only. Individual traits have impacts on the adoption decision. Nevertheless, we opt for this simplification on the basis of the correlation coefficient estimated in *Medical Innovation*. Four network variables, shared office, advice seeking, discussion and friendship, showed a stronger association to the date of first use of Tetracycline than any other individual variables, with the single exception of total volume of prescriptions for the class of drugs. Holding the volume of prescriptions constant, the association between integration and adoption increased (Coleman et al. 1966: 92). Thus, we explore if all the doctors are homogenous in terms of their individual attributes or 'degree of predisposition to adoption' (Burt 1987), to what extent does

³ Elaborating on significance of office partnerships to generate network effects, the researchers in the original study specified its three broad categories (Coleman et al. 1966: 73): alone in office, shared office, clinic. Shared office was further classified on the basis of number of colleagues being one, two, and three or more. Shared office for three or more colleagues was not incorporated as only 3 doctors in City B was identified for this category. We specified 5 doctors to practise in a clinic based on the representation of office partnerships in City D (Coleman et al. 1966: Figure 15). A majority of doctors in City A were practising alone (53 percent), in contrast to which 50 percent of doctors were affiliated with clinics in City C. None of the doctors in City B was practicing in a clinic. See *Medical innovation* (P 73-74) for

² "To analyse pairs of individuals instead of single individuals may seem like only a very modest step towards the analysis of networks of social relations. It would be more satisfactory, and truer to the complexity of actual events, if it were possible to use longer chains and more ramified systems of social relations as the units of analysis. But so little developed are the methods for the analysis of social processes, that it seemed best to be content with the analysis

integration matter for adoption decision? In our model, a random friendship network, and professional networks are created through discussions with office colleagues, or through hospital visits or conference attendance, or all of the above.

Constrained by the unavailability of the original dataset, we opt for a random friendship network for *Gammamym*. The doctors are, thus, initialized with random number of friends and *counter for friends*⁴, both ranging from 0–3. In the original study, indices of similarity⁵ of pairs were analyzed in order to identify the attributes the doctors share with the colleagues they choose as friends, discussion partners or advisors. We treat the indices with reservation, because of their limited statistical relevance with only 111 friendship pairs (Coleman et al. 1966: 143). The contacts with friends are not spatially defined, as *Medical Innovation* contains no information in this regard. Each doctor communicates with his/her friend(s) and evaluates mean adoption rate when the counter for friend is 4.

Professional interactions are spatially defined to signify discussion networks⁶. We do this to signify the importance of tacit knowledge or non-codified knowledge, which requires face-to-face contacts for its transmission (Appold 2004; Balconi, Breschi and Lissoni 2004; Becker and Dietz 2004; Douriax 2003; Fritsch and Franke 2004; Grafton, Knowles and Owen 2004; Howells 2002). The doctors, therefore, consider the others as discussion partners if they are situated in the same cell. After each visit to hospitals or conference center, doctors return to their practices. Office partnership is, therefore, of central importance for networking as the doctors in *Gammamym* made most of their interactions in their practices.

Communicating Agent: Laboratory

In *Gammamym*, we incorporate one pharmaceutical company as a communicating social agent, termed as the Laboratory (LAB from here on). The LAB adopts a mixed marketing strategy with three different channels to send information about the new drug:

- i. Detailman (pharmaceutical representative) visiting practices;
- ii. Flyers, available at the conferences;
- iii. Journals, sent to doctors' practices.

⁴ An attribute for the doctors to specify the frequency of interactions with their friends.

⁵ The index varies from a value of +1, where all doctors choose others like themselves, to -1, where doctors choose only others who are unlike themselves. Based on the indices of similarity (in parenthesis for each factor), the order of importance for choosing friends was: religion (0.403), professional age (0.195), place of growing up (0.146) and medical school (0.167).

⁶ Our intention to distinguish among discussion and advice networks was constrained, primarily, by unavailability of original data. The criteria for advice networks, was further complicated by the fact that none of the common background factors seemed to work for choosing the advisor, as the indices of similarity for religion, professional age, place of growing up and medical school are respectively 0.109, 0.042, 0.077, 0.018 (Coleman 1966:143). Interestingly enough, analyses by Burt (1987) and Stang and Tuma (1993), despite using the original

57 percent of the doctors asked to reconstruct their stages of diffusion, identified detailman as the first source of information. Hence, we opt for a blanket exposure of all doctors to detailman. In *Gammamym* the detailman visits all the doctors at their practices. At each time step the LAB keeps records of the practices visited by the detailman and randomly chooses one of the remaining practices or available practices, only if at least one of the doctors is present at the practice. As the time step signifies one week, so the detailman is able to contact all the doctors during his visit.

In the original study, drug house mails or direct mail advertising were identified as the second major source of information. Assuming similar influences by direct mail advertising by the pharmaceutical companies and journal advertisements, we specify journals as the second option for the LAB. The LAB issues journals only when the number of newly adopted doctors in the previous time step, i.e., the last increment, is less than half of the average number of adopted doctors. Journals are sent to all the practices and thereby ensure a blanket exposure to all doctors.

From the perspective of the LAB, inclusion of flyers is crucial as it adds another dimension to the marketing mix by targeting a large group of doctors at the same time. To avoid the notion of blanket exposure to all doctors, we specify the criterion that the LAB sends flyers based on the number of previous conference participants. Receiving information from a flyer is therefore conditional upon the number of available flyers, and not all the doctors attending the conference receive the flyers.

3.3. Decision-making Process

Doctors' decisions to adopt a new drug involve interdependent local interactions among different entities in *Gammamym*. Diffusion scholars have long recognized that an individual's decision about adoption is a process that occurs over time, consisting of several stages (Coleman et al. 1966; Rogers 1995). We specify five stages of adoption: i. Awareness or knowledge, ii. Interest, iii. Evaluation/mental trial, iv. Trial, and v. Adoption/acceptance. In our model *readiness* is specified as the attribute signifying the above stages of adoption. All doctors are initialized with *readiness* number of 4 that is decremented when they receive an alert from different sources.

In *Gammamym*, the doctors are homogenous in terms of individual predisposition to knowledge. Thus, the criterion for alert generation is same for all doctors irrespective of their adoption stages or sources. For doctors working in centers and clinics, they evaluate the mean adoption rate of their office colleagues while situating at their practices. Otherwise (during visits to the hospitals or conference center), each doctor randomly chooses five of their acquaintances and evaluates their mean adoption rate. Thus, discussions with other doctors, either friends or colleagues at practices, conferences, or hospitals generate an alert when the mean adoption rate is 0.50 or above. In the case of the LAB, on the other hand, an alert is created each time a doctor receives information from the detailman, flyers or journals.

Doctors' *readiness* is gradually reduced with alerts from all the aforementioned sources i.e. doctors move to the next stage of adoption when alert is greater than or equal to one. When the readiness reaches zero (Adoption/acceptance stage), doctors

3.4. Modeling Sequence

Gammamym is divided into four phases: i) managing professional interactions; ii) external influence; iii) decision making process; and iv) networks formation. At each time step, *Gammamym* resets the attributes of the practices. Thus, the doctors are at their respective practices at the beginning of each simulation.

Phase I entails the methods for doctors' professional interactions. Primarily, the doctors interact with the office colleagues at their practices. Hospitals are another location for professional interactions, where they have their monthly visits. The doctors are initiated with an attribute, counter for hospital, ranging from 0 to 3. The doctors randomly choose one of the hospitals and move to the designated cell when the counter for hospital is 4. The third location for information exchange is the conference centre. The invitations to conferences are sent randomly to 30 doctors. Doctors receive the invitation and move to the conference centre if they are available at their practices at the time step when the invitations are sent. Phase II and Phase III are discussed earlier in Section 3.2 and Section 3.3 respectively. Phase II depicts the mixed marketing strategy adopted by the LAB. Phase III is the decision-making process based on readiness. Phase IV constitutes the methods for network formation.

Gammamym generates 137 matrices, the network matrices for professional and friendship ties at each time step, and the adoption matrix at the end of simulation. Network matrices contain the number of interpersonal interactions for each doctor at each time step. The adoption matrix, on the other hand, specifies the adoption status (adoption vs. non-adoption) at each time step for all the doctors. These matrices are the links for our novel approach of integrated modeling as the analysis on network topology and evolution of networks in Section 5 is based on these matrices.

4. Simulation Results

In this section our discussion traces the shape of the cumulative diffusion curves under three scenarios. The three scenarios are specified to evaluate the degree of influence by different factors in the diffusion process:

1. Baseline Scenario; with one 'seed' or initial adopter, one detailman and one journal;
2. Heavy Media Scenario; with one initial adopter but different degrees of external influence, by increasing the number of detailman to five and the number of journals to four;
3. Integration Scenario; one initial adopter, without any external influence from the LAB.

All three scenarios have been run over 68 weeks or 17 months, which was the time length for the original study. As several random functions are included in the algorithm, each scenario is repeated 100 times in order to estimate the output's variability. For each of the cases, the seed or the innovator is chosen among the doctors who are

practicing at centers, i.e., doctors who have two colleagues⁷. The cumulative diffusion curves (CDC), representing the percentage of adopting doctors at each time step are shown in Figure 1. All three curves are derived after averaging over 100 simulations.

Baseline scenario (Figure 1) with one innovator and one detailman generates a logistic or S-shaped curve, similar to those found in cases of mixed influence diffusion models (Ryan and Gross 1943; Mahajan and Peterson 1985; Rogers 1995; Valente 1995). In this scenario, an initial phase of slow diffusion exists until the first inflection point at the 24-time step where 23 percent of doctors have adopted the new drug. Thereafter, the rate of adoption speeds up as more doctors are exposed to someone who has already adopted. The rate of adoption then gradually begins to level off as fewer doctors remain in the population who are yet to adopt. Overall, 92 percent of doctors adopted the new drug by the end of 48 weeks, or within 12 months. With one innovator in the system, diffusion seeds through the system as the doctors' adoption decisions are simultaneously influenced by their exposure to the new drug through actions by the LAB, as well as interpersonal communication.

Figure 1. Cumulative diffusion curves for three scenarios: Baseline, Heavy Media and Integration

⁷ We explored with initial adopters with different levels of social and professional integration. The cumulative diffusion curves for each of the cases represent the same process with point of inflection at 36th time step and levelling off at 50th time step irrespective of the integration

The steepest diffusion curve (Figure 1) represents a heavy media scenario, where 50 percent of the population adopts the new drug at the end of 12 weeks. The rate of diffusion increases up to 16 time steps and decreases afterwards as only 18 percent of doctors at that time have failed to adopt and remain unaffected. At the 25 time step, the CDC levels off as all the doctors have adopted the new drug.

The integration scenario represents an extremely slow diffusion process (Figure 1). As the only means to have an alert is to be in contact with the initial adopter, only 18 percent of the population adopts the new innovation at the end of 68 time steps.

5. Network Analysis

In this section we will examine: 1. the topology of the interaction networks that are generated by the agent based simulations; 2. how the networks evolve over time; and 3. the way the uptake evolves across these networks.

5.1. Network Topology

We first calculate the degree distribution from the interaction matrices generated at each time step from four possible alternatives: i. regular lattices; ii. random graphs (Erdős and Rényi 1959); iii. small world networks (Watts and Strogatz 1998); and iv. scale-free networks (Barabási 2002). Degree is defined as the number of edges connected to a given node. The results were averaged over 100 replications of the simulation. Figure 2 compares the degree distribution resulting from *Gammaym* and an Erdős-Rényi (1959) random graph. From this figure we can see the networks created from the simulation are indistinguishable from a random graph. We also perform a χ^2 test against an Erdős-Rényi random graph with the same number of nodes and edges. There was no statistically significant difference between the degree distributions of *Gammaym* and the random graph when $\alpha=0.01$. We do not claim that the structures of the network in the four medical communities in 1950's Illinois were random, but that the structure of the network developed in *Gammaym* appears to be random.

Many social systems display two statistical properties; first they display a high degree of clustering when compared to random graphs; second they have an average shortest path-length similar to that found in random graphs. These two properties are known as the small-world properties. To determine if the interaction networks generated by *Gammaym* have small world properties, we calculate the clustering coefficient and path-length statistics.

The clustering coefficient (CC) is a measure of the number of friends of friends type relationships found within the network. For a given node N_i with k_i neighbors, E_i is defined to be number of links between the k_i neighbors. The clustering coefficient is the ratio between the number of links that exist between neighbors (E_i) and the potential number of links $k_i(k_i - 1)/2$ between the neighbors.

Figure 2. Degree distribution of the interaction networks in *Gammaym* and Erdős-Rényi random graph

The average clustering coefficient is:

$$CC = \frac{1}{N} \sum_{i=1}^N \frac{2E_i}{k_i(k_i - 1)}$$

The average shortest path length (PL) is defined as:

$$PL = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=i+1}^N PL_{mn}(i, j)$$

PL_{mn} is the minimum distance between nodes i and j .

A network is said to have small world properties if, compared to an Erdős-Rényi random graph, the following conditions hold: $PL \approx PL_{rand}$ and $CC \gg CC_{rand}$. The comparison between the interaction networks generated by the simulation model, and an ensemble of random graphs reveals that $PL_{model} = 1.98$ and $PL_{rand} = 1.99$ also $CC_{model} = 0.172$ and $CC_{rand} = 1.68$. This shows that the networks depicted in *Gammaym* simulation do not have the small-world properties.

5.2. Evolution of Networks

In this section we analyze the evolution of networks in *Gammamym*. As agents interact with each other, new connections form between them, while others are reinforced. Figure 3 shows the evolution of connectivity between agents within the simulation. The connectivity is defined as $L/(N(N-1))$, where L is the total number of interactions within the system and N is the total number of agents within the simulation. Connectivity values are averaged over an ensemble of 100 runs. From Figure 3, it can be seen that the connectivity within the system grows as a function of $n \log n$ over time. An interesting property in Figure 3 is the sudden jumps in connectivity associated with conference events. The most notable of these occurs at time step 55 and appears to capture an element of many real world 'networking' events, where people come together to exchange ideas and experiences.

Initially the agents are only connected to doctors that they directly interact with, within their practices, defined as *contact groups*. Hence, the initial state of the system consists of a set of disconnected contact groups. This means that the initial flow of information within the system is limited to only small isolated pockets of interactions. However, as connectivity increases, the size and the nature of these contact groups changes.

In order to study the nature and structure of the contact groups, we define a *cluster* as a set of doctors belonging to the same contact group. We explore the structure and evolution of the professional network and then the combination of the professional and friendship networks⁸ jointly, over the 68 weeks. At each time step within the simulation

Figure 3. Evolution of connectivity within professional and friendship networks

we count the number of clusters of agents and calculate the average cluster size and the average shortest path length between agents within the system (Figure 4). All statistics are averaged over an ensemble of 100 runs. Figure 5 is a graphical illustration of network at the beginning i.e. time step = 0 (Figure 5 (A, C)) and after the formation of the giant cluster at time step 10 (Figure 5 (B, D)).

In case of evolution of professional networks, we see that initially the simulation contains approximately 61 clusters (Figure 4 A), as defined by 61 practices over the spatial grid (Section 3.1). As new interactions occur, the number of clusters quickly decays. The average cluster size quickly explodes as the agents become consumed by the giant cluster (Figure 4 B). The system begins to behave as one giant cluster after about 7–10 time steps (Figure 5 B). We note that the average shortest path length between nodes initially increases rapidly (Figure 4 C), as more and more nodes become connected to the giant cluster. When the system becomes connected, such that there exists a path between all agents, the shortest path-length between any two nodes is, on average, relatively long. This is because of the sparsity of interaction matrices and the network containing many long paths. However, as the system increases in connectivity the path-lengths become smaller until there is approximately two degrees of separation between any two agents within the system (Figure 4 C). In summary, the system initially consists of a number of disconnected contact groups or clusters, but quickly evolves to form a single connected component or cluster while the size and structure of these clusters define how far and how quickly Tetraacycline is adopted.

Figure 4 (D–F) represents evolution of clusters over time, when both professional and friendship networks are considered. Given the static nature of the friendship networks, the evolution for both the networks essentially signifies a similar process

Figure 4. Network Statistics for professional networks- (A) Number of clusters; (B) Average cluster size; (C) Average shortest path length; Network statistics for professional and friendship networks- (D) Number of clusters; (E) Average cluster size; (F) Average shortest path length

⁸ In *Gammamym* the specification for interactions with friends (see Section 1.3.2) implies that the

Figure 5. Formation of giant cluster: (A) Professional networks at $t=0$; (B) Professional networks at $t=10$; (C) Professional and friendship networks at $t=0$; (D) Professional and friendship networks at $t=10$

revealed in Figure 4(A-C) in case of professional network only. Noticeable, however, is the strong coalescent effect of the friendship relationships that drive the 61 initial clusters for professional networks down to approximately 3.5 clusters when friendship and professional networks are considered jointly (Figure 4 D). We also note that, in Figure 4 (F) average shortest path length between nodes falls over the time period. The evolution of network in this case, starts with a few giant clusters with long paths (Figure 5 C). As the connectivity increases through time the shortest path length, on average, becomes smaller. At the end of 68 time steps, the degrees of separation between any two agents are marginally smaller than in case of professional networks only (Figure 4 F).

5.3. Evolution of Uptake

To analyze the differences in the speed of diffusion depicted in Section 1.4, we look at how Tetracycline uptake evolves under three scenarios. We, therefore, define an *uptake cluster* as a set of agents who belong to the same contact group and each has already adopted Tetracycline. In this context the uptake cluster can be thought of as a ‘cluster’ commonly encountered in percolation studies (Stauffer 1979). Figure 6 shows how the uptake of Tetracycline evolves through time

Figure 6. Uptake clusters for three scenarios: (A) Number of clusters within the system; (B) Maximum cluster size; (C) Minimum cluster size; (D) Average cluster size; and (E) Standard deviation of cluster size

In the base scenario, starting from one seed (innovator) the average number of uptake clusters increases up to 1.6 at time step 15 (Figure 6 A). Then, it decreases to one giant uptake cluster (time step 30) as the size of the existing clusters increases gradually before merging. We also observe that the number of uptake clusters explodes rapidly under the heavy media scenario (Figure 6 A). Compared to the base scenario, much faster dynamics is displayed with the giant uptake cluster forming at time step 15 with heavy media exposure. At time step 20, 100% of the doctors have adopted the drug and the size of the uptake cluster equals the size of the social network. The integration scenario displays a very slow process of diffusion. On average, only one uptake cluster forms during the 68 time steps, and its maximum size barely reaches 20 by the end of the simulation (Figure 6 B). Thus, despite a very efficient deployment of the giant group, the diffusion struggles to spread across the whole network without external influence.

6. Integration and Diffusion

In this section we investigate: integration into which network matters more, for whom and at what stage. The previous sections analyzed the diffusion process for all doctors under different degrees of external influence. Given the existence of overlapping networks, this section attempts to disentangle the impacts of professional

model with uniform distribution of agents in terms of their professional integration. In this case, 101 doctors are initialized as: 33 in 33 private practices, 33 in 11 centers and 35 in 7 clinics. To distinguish between impacts of social and professional relationships, we select the most connected and least connected groups for comparison: i. both socially and professionally integrated, ii. integrated professionally, but socially isolated, iii. isolated professionally, but socially integrated, and iv. both professionally and socially isolated.

Given three types of office partnership and four types of friendship status, a range of definitions for professional or social integration or isolation is comprehensible. First, we define professionally integrated doctors as the ones with maximum number of colleague or have four colleagues in a clinic. Socially integrated doctors have maximum number of friends or three friends as specified in *Gammamym*. Isolates are the ones practicing alone or having no friends. The diffusion curves (Figure 7A) for the above groups of doctors are derived after averaging over 100 simulations for the benchmark scenario (with one initial adopter and one detainee).

Another set of adoption curves is derived by relaxing the definition of integration to friendship networks. The group of socially integrated doctors now includes doctors with two or more friends; and the socially isolated doctors are the ones with one or no friends (Figure 7B).

Normalized adoption curves, in both cases indicate that although underlying networks evolve in predictable ways, as has been analyzed in Section 5, the uptake is a function of initial starting conditions. Initial conditions, however, are determined by level of social or professional integration. Diffusion proceeds exactly the same way, for all groups of doctors, up to the formation of the giant cluster and after that, integration to different networks generates different impacts.

Figure 7A. Impact of integration on diffusion

Figure 7B. Impact of integration on diffusion

Our results show that integration is crucial for initial uptake as the number of doctors, who are both socially and professionally integrated, is higher than for the rest of the groups. Though the speed of diffusion is not significantly different from the rest of the groups, diffusion among the least connected doctors, both socially and professionally isolated, is the slowest at the initial stage (Figure 7(A,B)). It is important to note that the simulation results for the two groups of doctors, where their integration status in two different networks is mutually reinforcing each other, replicate the findings in the original study.

Our results are rather interesting for two groups of doctors with contrasting integration status: integrated professionally, but socially isolated; and isolated professionally, but socially integrated. The degree of professional integration has a stronger impact. The similarity of adoption rates for group one (doctors with four colleagues and two or more friends) and group two (doctors with four colleagues, but with one or no friends) implies that, even if doctors are socially isolated, professional interactions enable them to adopt quickly (Figure 7(A,B)). Nevertheless, integration to a friendship network is more significant for doctors who are professionally isolated. Given the differences in the adoption rates between group three and group four doctors, we can imply that, to some extent, integration to friendship networks can make up for professional isolation. In other words, even if a doctor is practicing alone, social connections can work as a source of information.

7. Concluding Remarks

We develop an agent-based model, called *Gammamym*, to analyze the diffusion process. This is inspired by the classic *Medical Innovation* study on the adoption of

the limited availability of proper techniques/methods during the 1950s, the original study focused on interpersonal influence for pairs of individuals. In our study, we overcome this limitation of addressing complexity and dynamics of actual adoption using agent-based modeling. Our integrated modeling with *Gammamym* reveals original results within the existing literature of diffusion research, and also complements the extant work on *Medical Innovation*.

Our approach of integrated modeling enabled us to look at network properties such as connectivity and degree distribution, among others, in terms of the interaction matrices generated by the agent-based model. On the basis of these properties we determine that the interaction networks depicted in *Gammamym* are random graphs. Complexity of the diffusion process is explained by analyzing evolution of networks or dynamics on the networks. We find that initially the system consists of a number of disconnected components and quickly evolves, after 7–10 time steps, to form a single connected component. The analysis of network topology also indicates that underlying networks evolve in predictable ways, and the uptake is a function of the initial starting conditions.

Analysis of the evolution of uptake or adoption of Tetracycline enables us to disentangle the extent of different factors affecting adoption. Despite stressing the complementarity between network theory and diffusion research, a large body of diffusion literature has so far failed to examine the dynamic structures of the interpersonal networks and their evolutions over the diffusion process. Our model shows that although the media does not influence the network structure, it does have a major impact in accelerating the diffusion process. Under heavy media exposure undertaken by the pharmaceutical company to increase sales of Tetracycline, the average size of uptake group with agents who have adopted the new drug rises faster than otherwise. Moreover, all the agents adopt the new drug within 25 time steps, much earlier than with a baseline scenario of much less media exposure. We validate the previous evidence of a dominant media influence by comparing the speed of diffusion for three scenarios: baseline, heavy media and integration. The dominant role of media also suggests there is a trade-off between the efforts expended in promoting a new drug, and the speed of the diffusion process.

The comparison of the cumulative diffusion curves of *Gammamym* with those of *Medical Innovation*, is largely restricted because of unavailability of the original dataset. Nevertheless, our exploration with normalized adoption curves for groups of doctors, based on different degrees of social and professional integration, reveals that after the formation of the joint cluster, the initial uptake is dependent on the level of integration. Despite the predictability of evolution of networks, as suggested by network topology, this signifies the importance of social or professional ties for adoption decisions, as was outlined in the original study. Overall, our results provide a step towards understanding the complexity and dynamics of the diffusion process in the presence of professional and social networks.

References

- Appold S J, 2004, "Research Parks and the Location of Industrial Research Laboratories: An Analysis of the Effectiveness of a Policy Intervention." *Research Policy*, 33, pp.225–243.
- Balconi M, Breschi S and Lissoni F, 2004, "Networks of Inventors and the Role of Academia: An Exploration of Italian Patent Data." *Research Policy*, 33, pp.127–145.
- Barabási A L, 2002, *Linked: The New Science of Networks*, Cambridge, Massachusetts.
- Becker W and Dietz J, 2004, "R&D Cooperation and Innovation Activities of Firms: Evidence for the German Manufacturing Industry." *Research Policy*, 33, pp.209–223.
- Burt R S, 1987, "Social Contagion and Innovation: Cohesion Versus Structural Equivalence." *American Journal of Sociology*, 92, pp.1287–1335.
- Doutirax J, 2003, "University-Industry Linkages and the Development of Knowledge Clusters in Canada." *Local Economy*, 18(1), pp.63–79.
- Erdős P and Rényi A, 1959, "On Random Graphs." *Publicationes Mathematicae Universitatis Debrecenensis*, 6, pp.290–297.
- Fritsch M and Franke G, 2004, "Innovation, Regional Knowledge Spillovers and R&D Cooperation." *Research Policy*, 33, pp.245–255.
- Grafton R Q, Knowles S and Owen D P, 2004, "Total Factor Productivity, Per Capita Income and Social Divergence." *Economic Record*, 80(250), pp.302–313.
- Howells J R L, 2002, "Tacit Knowledge, Innovation and Economic Geography." *Urban Studies*, 39(5–6), pp.871–884.
- Mahajan V and Peterson R A, 1985, *Models of Innovation Diffusion*, Newbury Park, CA: Sage.
- Rogers E M, 1995, *Diffusion of Innovations*, New York: The Free Press.
- Ryan B and Gross N C, 1943, "The Diffusion of Hybrid Seed Corn in Two Iowa Communities." *Rural Sociology*, 8(1), pp.15–24.
- Stauffer D, 1979, "Percolation." *Physics Reports*, 54, pp.1–74.
- Strang D and Tuma N B, 1993, "Spatial and Temporal Heterogeneity in Diffusion." *The American Journal of Sociology*, 99(3), pp.614–639.
- Valente T W, 1995, *Network Models of the Diffusion of Innovations*, Cresskill, New Jersey: Hampton Press, Inc.
- Van Den Bulte C and Lilien G, 1999, "A Two-Stage Model of Innovation Adoption with Partial Observability: Model Development and Application", Working Paper: 1–44, *Institute for the Study of Business Markets, The Pennsylvania State University*.
- Van Den Bulte C and Lilien G, 2001, "Medical Innovation Revisited: Social Contagion Versus Marketing Effort." *American Journal of Sociology*, 106(5), pp.1409–35.
- Watts D J, 2003, *Six Degrees: The Science of a Connected Age*, New York: W. W. Norton Company.
- Watts D J and Strogatz S H, 1998, "Collective Dynamics of 'Small-World' Networks." *Nature*, 393, 440–442.