2008

Influence of Structure Parameters on the Information Diffusion Process in Virtual Networks

Sascha Vitzthum
Emory University, sascha_vitzthum@bus.emory.edu

-Research In Progress-

INFLUENCE OF STRUCTURE PARAMETERS ON THE INFORMATION DIFFUSION PROCESS IN VIRTUAL NETWORKS

Social Aspects of IS

SASCHA VITZTHUM

Goizueta Business School

Emory University

sascha_vitzthum@bus.emory.edu

Abstract

This paper examines the effects of composition, size and structure of virtual networks on the diffusion of information. The research extends the Axelrod cultural model (ACM), an established theory rooted in political science research, to explain information dissemination among social actors in ICT-enabled virtual networks (Axelrod 1997). The investigation is focused on the impact of network structure parameters and gives insight on how a manipulation of those parameters influences the diffusion of information in virtual networks.

Keywords: Axelrod Culture Model, Social Network, Diffusion of Innovation, Dissemination of Information, Network Analysis, Computational Model

INFLUENCE DES PARAMETRES DE STRUCTURE SUR LE PROCESSUS DE DIFFUSION DE L'INFORMATION DANS LES RESEAUX VIRTUELS

Résumé

Cette étude examine les effets de la composition, de la taille et de la structure des réseaux virtuels sur la diffusion de l'information. Nous proposons d'étendre le modèle culturel d'Axelrod (Axelrod 1997) pour expliquer la diffusion de l'information entre les acteurs sociaux dans les réseaux virtuels. L'enquête se concentre sur l'impact des paramètres de structure du réseau et explique comment la manipulation de ces paramètres influence la diffusion de l'information dans les réseaux virtuels.

Introduction

The rise of access to and use of information and communication technologies (ICT), in particular the Internet, has spawned virtual social networks that display unique patterns of information diffusion. New forms of collaboration and knowledge transfer are exemplified through the rise of the open source software development and the increase of virtual communities of practice and online social networks. The extended reach of today's social and professional networks impacts both the diversity of the participants and the structure of the network itself. The pool of participants for a certain network is no longer restricted to a certain locale or country. This *democratization of access* has led to a global pool of potential users. Moreover, the anonymity of the Internet does not allow for discrimination of participants based on origin or educational background, leading to a *democratization of participation* and a broader, more diverse portfolio of opinions and beliefs among Internet users.

The diffusion of information between actors in social networks has been a widely studied topic in the social sciences. Examples of diffusion of information-related research are word of mouth, diffusion of innovation, knowledge diffusion in organizations and dispersion of cultures. Previous research has identified actor relationships, similarity and proximity to influence the dissemination of information through social networks (i.e. Brown and Reingen 1987, Rogers 2003). However, the majority of the previous studies only examined phenomena in small populations that belonged to local networks. The majority of previous research examined networks where agents had limited reach and lacked the diversity of virtual networks.

The main purpose of this research is to acknowledge the changing nature of social networks caused by the increasing access to and use of ICT and to investigate the effects of network composition, size and structure on information diffusion in these virtual networks. The research does so by extending the Axelrod cultural model (ACM), an established theory rooted in political science research that examines information dissemination among actors with limited communication reach (Axelrod 1997). The extension allows an analysis of networks with more diverse actors and random structure, the two main outcomes of the democratization of access and participation. The investigation is focused on the impact of network structure parameters and gives insight on how a manipulation of those parameters can influence the diffusion of information.

Theoretical Background

In his famous work on the diffusion of innovations, Rogers (2003) showed the importance of social structures and communication networks on the diffusion of information. Rogers (2003) generalized that homophily among and communication proximity between actors in a network would increase the likelihood of the diffusion of shared ideas. Homophily in that regard should be understood as shared interests or common beliefs between actors. Communication proximity, however, refers to the directness of the communication. Thus, he proposes that an idea is better communicated directly from person to person instead of through one or multiple intermediaries. As such, if an actor wants to communicate an idea, he should talk in person to as many like-minded actors as possible to propagate it throughout a network. However, the author also recognized that those two factors decrease the chance of novel information being distributed within the communication network, since fewer actors will share novel beliefs that deviate from the homophile beliefs. Thus, novel ideas can only be communicated to a small set of actors, which themselves can have relationships with few actors that are susceptible to the novel idea.

The diffusion of innovation theory has direct implications in the context of virtual networks. First, the increasing diversity of users does not necessarily impact the behaviors and beliefs of the other users. The proverbial behavior of birds of a feather flocking together, where common beliefs are seen as a bonding mechanisms for relationships, will allow only dominant ideas or information to be shared and reinforced (McPherson et al. 2001). Conversely, novel ideas are not shared by the majority of the actors, leading to the creation of niche communities. Second, the structure of the communication networks needs to be investigated. Virtual networks allow for equal participation which is not governed by institutional structure. Thus, initially the communication structure of a virtual network is not predetermined, mirroring a random distribution of relationships (ties) between actors (nodes). A random distribution allows every node to have on average the same amount of direct ties, permitting the dissemination of novel ideas. Since real world data on the development of communication structures comparable networks at different stages of formation have been hard to collect, researchers started to employ computational models to simulate the diffusion process.

Roger's theory was adopted by Axelrod and applied to context of cultural expansion in the realm of political science. Axelrod (1997) was among the first researchers to employ computational modeling to analyze the spread of information. The basic premise of the ACM mirrored Rogers proposition of homophily: actors who share similar cultural attributes, which include language, beliefs attitudes and behaviors, are more likely to interact and further adapt each other's values (Axelrod, 1997). The notion of communication proximity was simplified: instead of recreating a communication structure, Axelrod equated communication proximity with geographical proximity. In his experiment, Axelrod used a spatial lattice of 100 agents with different cultural values (modeled as five digit strings) to set up the simulation. The diffusion mechanism was parsimonious: Every round an "activated" agent would donate one of its traits to one of its four immediate neighbors in the cardinal directions with a probability based on the number of their shared traits. Thus, after a successful donation, the two involved agents share a common trait on an additional feature, making them more similar and increasing the probability of an exchange in the future.

Once agents share the same traits in all five features, they are considered a culture. The main outcome of the ACM is equilibrium where only few stable cultural regions emerge. The equilibrium is achieved when neighboring cultures do not share a common trait on any feature, preventing further exchanges. Axelrod theorizes that his findings mirror the real world where only few cultures exist. He concluded by proposing that adoption of information based on similar beliefs leads to local convergence, but global polarity between cultures. The parsimony of the experiment and the consistency of the results started a stream of research that tested the generalizability of the results by incrementally lifting the restrictions of the initial model.

Several extensions of the ACM are concerned with the increase of interaction ranges and its effect on cultural heterogeneity. Shibanai et al. (2001) investigated the effect of a global mass media and modeled it as a "generalized other," which acts as a direct (fifth) neighbor to each agent. Their experiment showed that a global agent can speed up the convergence of cultures, while at the same time yielding a smaller number of distinct cultures at the end of the simulation. Greig (2002) also investigated global impacts on the number of stable regions. Similar to the original ACM, Greig increased the number of potential neighbors and ran simulations for discrete levels of neighborhood sizes. He replicated the ACM hypothesis that with increasing number of neighbors, the average number of stable regions decreases. He observed that his replication of the original model yielded an average of 4.1 unique cultures. For any bigger neighborhood size, the number of cultures dropped below 1.5. Most of his runs yielded quasi-homogeneity with an average of 99.15% of the population belonging to the dominant culture. Ward (2006) introduced the concept of virtual neighbors to the ACM model. Her model showed that an increase in access to global communication and thus to virtual neighbors, decreases the number of unique cultures over time.

Two general effects of the system parameter changes can be generalized from the ACM research stream: a) In locally restricted communication networks, an increase of communication range leads to fewer distinct cultures, and b) Initial local similarity leads to a higher number of distinct cultures in the system. In other terms, similarity leads to diversity through the creation of boundaries, while range influences the size of the distinct territories. Hence, in a network where only the parameters of geographic proximity are varied, the general outcome of local convergence and global polarization still holds true; only the extend of the outcome changes. However, as argued before, geographic proximity is a special case of communication proximity. Staying true to Rogers' (2003) theory, the ties in a communication network need to be analyzed. Thus, to operationalize the construct of Rogers' theory <u>and</u> to better approximate the random structure of today's virtual networks, the geographic restrictions on the ACM need to be lifted. An analysis based on network level parameters will allow for a better understanding of the information diffusion process in virtual networks.

Research Model

The setup of the ACM on a grid, where communication proximity equals geographical proximity, leads to the diffusion and eventual dominance of a few shared cultures rather than to a system of many diverse and novel cultures. From a network perspective, the spatial grid is an extreme network structure. This section will summarize three network structure parameters which in the original model can be considered extreme, and how these attributes impact the diffusion of information. In particular, the research model investigates the impact on the process variables of the information diffusion: total activations (total number of trait exchanges between agents), the time to convergence (number of rounds taken before completion), and average diffusion velocity (total number of activations / total number of rounds). The research model is summarized in Figure 1.

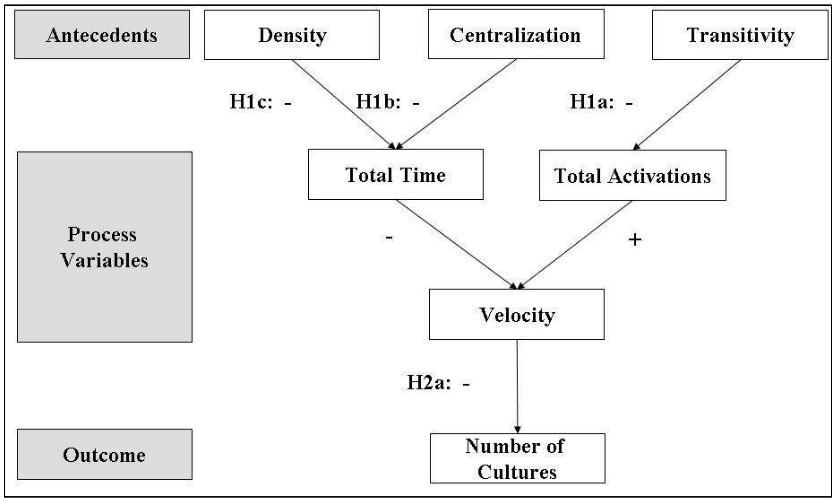

Figure 1: Research model

Transitivity

Tie strength is observed through the transitivity score of the network. Transitivity is defined as the proportion of node triplets in a network that have three direct ties (Wasserman & Faust 1999). High transitivity indicates a high degree of reciprocity. For example, if three nodes have ties with each other, information can freely flow between the actors using direct communication. Since all nodes are connected, the triplet is considered transitive, forming strong ties. However, if there are only ties between nodes A and B (AB) and nodes B and C (BC), no direct communication can take place between nodes A and C. Indirect connections result in low transitivity or weak ties. Generally speaking, a low transitivity indicates the existence of weak ties, whereas a dominance of strong ties results in more transitive networks. The original ACM has transitivity of 0 and thus an abundance of weak ties.

Granovetter (1973) demonstrated that weak ties between different personal networks can enable the diffusion of unique information. Drawing on the theory of weak ties (Granovetter 1973), Rogers suggested that an imbalanced distribution of communication ties within the network can lead to the emergence of more novel ideas. Rogers recognized that the emergence of the Internet increased the availability of personal networks with weak ties (Rogers 2003, Rosen 2000). However, a transitivity of 0 will not be achieved in virtual networks, because it is unlikely in a large network that two actors don't have independent relationships with third actor. As such, the ACM model, which only consists of intransitive triplets, skews the outcomes of the simulation. In particular, recalling the basic exchange mechanism of the ACM, communication between actors is less efficient because information cannot directly be communicated from one particular agent to another, unless the agents are direct neighbors. The lack of strong ties in the ACM yields more exchanges of traits (activations) before the system converges in a quasi-homogenous state. Therefore, the low transitivity skews the number of activations compared to a real world virtual network with at least some strong ties.

> *Hypothesis 1a: The transitivity of the network has a negative impact on the total number of agent activations.*

Degree Centralization

Degree centralization describes the composition of the ties within a network. In particular, it compares the distribution of ties around nodes within a system (Wasserman & Faust 1999). If, for example, one actor has ties to all other actors, while the other actors do not share connections among themselves, the centralization of the network is 1. On the contrary, if actors share the same amount of connections, centralization is 0. Social network analysis literature shows that degree centralization is a good predictor of overall efficiency of information flow (Cook et al. 1983). High network centralization suggests that information is better broadcast through well- connected agents that act as information hubs. Thus, with faster information distribution the system converges faster.

In the original model, the degree centralization is 0. It seems unlikely that in a large virtual network each member has direct contact with exactly the same amount of peers. On the contrary, usually popular members or moderators

in the network have more connections than the average member. As such, the impact of degree centralization is underemphasized in the original model. Accordingly, the original model should converge slower[1] compared to a virtual network with high degree centralization.

> *Hypothesis 1b: Degree centralization of the network has a negative impact on the time of convergence of the systems.*

Density

Network density describes a network as a ratio of actual ties within the system over the maximum potential connections in the systems. Wasserman and Faust (1999, p. 182) argue that density by itself does not sufficiently describe the centralization of a network. However, it can provide useful information about the network structures as long as it is used in conjunction with the aforementioned centralization measures. A density of 1 indicates that each actor has a direct tie to each other actor within the system. Increasing the network density should have a similar effect, increasing network centralization by raising the average probability of communications and, thus, forcing the system to converge faster.

> *Hypothesis 1c: Density of the network has a negative impact on the time of convergence of the systems.*

The first set of hypotheses is distinct from previous research by relating the measures of the network structure to the process variables of the simulation: time to conversion and activations until convergence. In the next section, I will propose a hypothesis that relates the process measures to the final outcomes of the simulation.

In previous research studies (Axelrod 1997, Greig 2002), network density was increased by extending the range of interaction, which led to fewer distinct cultures in the converged systems. However, I argue that this was only an indirect result. As discussed before, network measures influence process variables. Laguna et al. (2003) used a process parameter to model the rate of exchange between agents to predict final outcomes. In the context of this network analysis, convergence velocity is the ratio of total number of agent activations over time of convergence. I propose convergence velocity as predictor of the distribution of cultures. In particular, the results of Laguna et al. (2003) suggest that higher velocity will lead to fewer distinct cultures.

> *Hypothesis 2a: The velocity of convergence will have a negative impact on the number of distinct cultures in the converged state.*

The research model is depicted in Figure 1. In order to investigate this research model, I adjusted the ACM by removing the geographical proximity constraint and introducing a network structure that, along with the original similarity constraint, governs the exchange of cultural features.

The Experiment

The original ACM consisted of 100 agents, with 5 features each having one of 10 traits. An agent's culture is described as a 5 digit string (i.e. 4, 3, 5, 9, 1). Axelrod used a geographical distribution of the agents on a 10-by-10 grid. One of the premises of the theoretical model was that the reach of communication is limited. This condition was implemented in the computational model by allowing agents to only interact with their immediate neighbors in the cardinal directions (up to four).

Once an agent pair is chosen to communicate with each other, the probability of adopting a new common trait for a feature is based on the number of already shared traits. If a pair shares 3 traits, the probability of exchanging a common trait for an additional feature is 60% (3 shared features divided by 5 total features). If no feature is shared, no exchange occurs. As a result, exchanges take place until all agents either share all or no traits. When equilibrium is reached, the average number of cultural regions is recorded.

The median number of stable regions was 3. However, 14% of the runs yielded only one stable region, whereas 10% of the runs resulted in more then six regions. The original paper also suggests that the number of stable regions depends upon the range of interactions or the size of the neighborhood. Axelrod showed that the average number of

[1] Time to convergence is reflected by the number of rounds. Every round consists of 50 total activation attempts (one per agent pair).

stable regions increases as interactions over greater distances occur. For small neighborhoods (4 neighbors), the average number of stable regions was 3.4, whereas large neighborhoods (12 neighbors) yielded an average of 1.5 stable regions.

I programmed two simulations. First, I replicated the ACM to validate the simulation. The simulation of the extension followed the original model with the exception that rather than being connected by geographic proximity, agents are part of an imposed network structure. The structure is generated using a fixed coefficient (.0404) for the network density (which equals the network density in the original model) and is operationalized in a 100x100 sociomatrix with 1 signaling a bidirectional tie and 0 signaling a lack there of. As with the original ACM, the exchange of traits occurs when at least one trait is the same. However, the second condition for an interaction is tie as defined by the network structure. A directional tie has to exist in order to fulfill this second condition.

Overall, 100 runs with different initial network structures and traits are carried out. For all activation cycles the changes in traits and number of cultures as well as the relevant network and process measures are recorded. The replication of the original ACM yielded similar results, with the median of final cultures (3), and the percentage of heterogeneous systems with more than six cultures (10%) being equal and the number of systems with homogenous cultures being slightly higher (14% vs. 16%).

The extension removed the spatial network structure. Bidirectional ties were randomly assigned (seed of 400 ties) over 100 runs, thus producing variations in density, transitivity and centralizations for each run. Table 1 shows the means of the variables of both the base experiment and the extension experiment.

Table 1: Means of Variables

	Transitivity	Centralization	Density	Activations	Time	Velocity	# of Cultures
Original ACM	0	0	0.040	-[2]	-[2]	-[2]	3.2
Base	0	0	0.040	12562	572	22.5	3.63
Extension	0.0415	0.594	0.041	9740	388	25.6	4.17

Two OLS and one Poisson regressions were undertaken to test the hypotheses. The regression results are depicted in Table 2. The results of the regression show that all hypotheses are supported and that the network parameters only indirectly influence the final outcome of the simulation. The parameters influence convergence velocity, which in turn influences the number of final cultures, supporting hypothesis 2a.

Hypothesis 1a that proposed transitivity has a negative effect on total activations is supported. Stronger ties allow for more activations of between actors of similar nature, since similarity increases the chance of activation. Moreover, strong ties foster the dissemination of redundant information, which further leads to a convergence because exchanges between actors stop, once they the have the same traits for all the features

Hypothesis 1b, which stated that degree centralization negatively effects time to convergence, is also supported. If there are popular actors that have more ties than the rest of their peers, they can act as information hubs. As such, if clusters of agents are better-connected than the remaining agents, they act as information bridges, leading to faster convergence and fewer final regions. Lastly, a higher network density, which can be interpreted as a higher connection average by each culture, has the same effect as centralization, since on average agents have higher capabilities of acting as information hubs.

The results show that the geographic proximity is only a limited substitute for the Roger's (2003) theorized communication proximity of agents. The network structure imposed by geographical proximity does not take the influences of transitivity and centralization into account. Per design, only network density indirectly influences the number of final cultures. This explains why the results of previous research (Axelrod 1997, Greig 2002) where an increase in range, which essentially is an increase of the average connections equally across cultures, leads to a decline in final cultures.

---

[2] The original paper does not report the number of activations or the time to converge, thus a velocity could not be calculated.

Table 2: Regression Results

	Time	Activations	Cultures (Poisson)
Intercept	**746.7 (136.5)*** **	**13452.3 (4420.7)** **	**3.24 (0.73)*** **
Transitivity	204.2 (0.248)	**-44785.9 (26671.4)** **	-3.47 (4.52)
Centralization	**-2272.2 (816.1)** **	-31343.2 (26435.1)	2.31 (4.36)
Density	**-5673.7 (3288.4) .**	260.5 (106513.2)	-12.63 (17.88)
Velocity	N/A	N/A	**-0.053 (0.007)*** **
Adjusted R^2	0.085	0.021	N/A
Significance Levels: '***' 0.001, '**' 0.01, '*' 0.05, '.' '0.1			

Conclusion

The main objectives of this research were to acknowledge the changing nature of social networks caused by the increasing access to and use of ICT and to investigate the effects of network composition, size and structure on information diffusion in these virtual networks. Virtual networks are free from geographical restrictions and allow for communication between large communities of users that previously were unable to interact. The research shows that network structure parameters have a distinct, albeit indirect influence, on the outcome of information diffusion processes in virtual networks. Moreover, it emphasizes that geographical proximity- based models only have limited applicability to the mechanisms of information diffusion in virtual social networks.

The findings of the research provide a validation of Roger's initial theory of information diffusion in the context of virtual networks: As long as users connect to each other and have something in common, information can diffuse through networks. Ultimately, the democratization of access also leads to a democratization of information, which is exemplified through collaborative efforts such as Wikipedia or the open text project. However, while ideas of the majority will be shared, the attention should focus on how effectively those minority opinions or as Rogers calls them, novel ideas, are communicated. In that regard, there will be a need in the future to design virtual social networks so novel ideas can also be promoted. By effectively manipulating the network structure parameters, novel ideas can be propagated faster to a broader audience.

There is an increasing interest in studying the complexities of social network mechanisms in IS and management research. Recently social network analysis (SNA) has been used to analyze IS proficiency in organizational units and information seeking, and knowledge management (Kane 2007, Borgatti and Cross 2003, Alavi and Kane 2005). As most networks are enabled by IT artifacts, SNA will become of every increasing importance to IS research. Developing computational models that can validate findings and test the impact of parameters will become an integral part of future research.

This initial computational model has lifted some of the limitations imposed by previous research. However, as with every simulation, it has limitations of its own. In a virtual network, all users do not behave the same way, at the same time or in the same order. There are particularly active communicators that share their thoughts frequently with a variety of users, while there are many users that do not communicate at all. The direction of communication and influence, the difference between the strength of receiving and sending ties, and the dynamic nature of network structures are additional phenomena that occur in virtual networks that could not be captured in this experiment. Nevertheless, as more empirical evidence or theory in respect to those topics are gathered, the simulation can always be extended and refined and eventually become a close approximation of information diffusion processes in virtual networks.

References

Alavi, M., and Kane, G.C. "Casting the Net: A Multimodal Network Perspective on Knowledge Management, "Proceedings of the *26th International Conference on Information Systems*, Las Vegas, NV, 2005.

Axelrod, R. "The Dissemination of Culture: A Model with Local Convergence and Global Polarization," *The Journal of Conflict Resolution* (41:2) 1997, pp 203-226.

Borgatti, S.P., and Cross, R. "A relational view of information seeking and learning in social networks," *Management Science (49:4), Apr 2003b, pp 432-445.*

Brown, J. J. and Reingen, P.H. "Social Ties and Word-of-Mouth Referral Behavior," *Journal of Consumer Research* (14:3), 1987, pp 350-362

Cook, K.S., Emerson, R.M., and Gillmore, M.R. "The Distribution of Power in Exchange Networks: Theory and Experimental Results," *The American Journal of Sociology* (89:2) 1983, pp 275-305.

Granovetter, M.S. "The Strength of Weak Ties," *The American Journal of Sociology* (78:6) 1973, pp 1360-1380.

Greig, J.M. "The End of Geography? Globalization, Communications, and Culture in the International System," *Journal of Conflict Resolution* (46:2) 2002, pp 225-243.

Kane, G.C. "IS Proficiency in Social Networks, "Proceedings of the *28th International Conference on Information Systems,* Montreal, ON, 2007.

Laguna, M.F., Abramson, G., and Zanette, D.H. "Vector opinion dynamics in a model for social influence," *Physica A: Statistical Mechanics and its Applications* (329:3-4) 2003, pp 459-472.

McPherson M., Smith-Lovin L., and Cook J.M. "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology* (27), 2001, pp 415-444

Rogers, E.M. *Diffusion of innovation,* The Free Press: New York, 1995.

Rosen, E. *The Anatomy of Buzz: How to Create Word of Mouth Marketing,* Doubleday/Currency: New York, 2000.

Shibanai, Y., Yasuno, S., and Ishiguro, I. "Effects of Global Information Feedback on Diversity: Extensions to Axelrod's Adaptive Culture Model," *The Journal of Conflict Resolution* (45:1) 2001, pp 80-96.

Ward, Joycelyn, "Digital Disparity and the Dissemination of Culture," Working Paper, Goizueta Business School, Emory University, Atlanta, 2001

Wasserman, S., and Faust, K. *Social Network Analysis: Methods and Applications* Cambridge University Press, 1999.