Comparing Peer Influences in Large Social Networks - An Empirical Study on CRBT

Bin Zhang†, Paul Pavlou†, Ramayya Krishnan‡, David Krackhardt‡

†: Temple University ‡: Carnegie Mellon University

Abstract

Cohesion and structural equivalence are two competing network models to explain diffusion of innovation. The dispute of which model plays a more influential role has not been resolved. This paper attempts to reconcile this problem in a large network setting – adoption of Caller Ring Back Tones (CRBT) in a cellular telephone conversation network. Since this societal scale networks is very large, we use a novel technique to extract multiple densely connected and self-contained subpopulations from the network. We found subpopulation size in such million-node network only falls in two levels, 200 and 500, in the extraction step. Using a new auto-probit model with network terms, we then compare the competing influences of cohesion and structural equivalence on each of the subpopulation extracted. Finally we use meta-analysis to summarize the estimated parameters from all subpopulations. The results show CRBT adoption is affected by both cohesion and structural equivalence. The size and direction of network influence both can change with the size of group. Structural equivalence has a negative effect on adoption when group size is at about 200, and has a positive effect when group size is at about 500. The effect of cohesion, on the other hand, is consistent.

1 Introduction

The debate about which network influence, cohesion or structural equivalence, plays a more influential role in technology diffusion is still inconclusive. When forming opinions or making decisions, people usually use someone they know or someone in their social network as their frame of reference, taking their opinions into account. This progress, in which an actor adapts his behaviors to those of alters in his social network, is known as contagion or social influence (Duncan et al., 1968; Leenders, 1997). There are two social network models, cohesion and structural equivalence, to explain contagion. Cohesion is made through communication, which is direct contact between actor and alter; while structural equivalence is created through comparison, which occurs when an actor competes with other alters who he considers in a similar social position to him in the network. Both models have been used to explain the progress of contagion. Coleman et al. (1966) studied diffusion of medical innovation and found medical doctors adopted new technology at the early stage because of cohesion. Burt (1987) reanalyzed Coleman et al's data and concluded that contagion does not happen through cohesion but rather through structural equivalence. Since then both camps, cohesion (Rogers and Kincaid, 1981; Harkola and Greve, 1995) and structural equivalence (Strang and Tuma, 1993; Van den Bulte and Lilien, 2001), have found quantitative evidence to support their claims. However, most of these analyses used network autocorrelation models that only include one network term at a time, and compare the term sizes in two independent models. The assumption made by such method is that only one network term would be significant. Since then, some quantitative methods for social networks have been developed to work in the situation where both terms are significant and direct comparison of these two influence sizes can be made. For example, Doreian's two regimes of network effect autocorrelation model (1989). However, it does not support dichotomous response variable. Actually, quite often in social sciences, the behaviors need to be analyzed are binary, for example, whether adopt a new technology or not. But no model is available if we want to compare multiple network terms at the same time and the response variable is dichotomous. Furthermore, the only data being used for the comparison is Coleman et al.'s Medical Innovation data, which only has a node size of 125. Whether network influence shows similar mechanism in large social networks has not been studied yet. Furthermore people argue that Coleman et al.'s data cannot represent adoption of information technologies.

In this paper, we apply a newly designed statistical model, multi-network auto-probit (mNAP) Zhang et al., to investigate the effects of both direct influence (cohesion) and competition (structural equivalence) on Caller Ring Back Tone (CRBT) adoption, which is represented by binary variable, within cellular phone communication social networks. CRBT is becoming one of the most attractive mobile content with a projected revenue of $4.7 Billion in United States by the end of 2012[1]. The penetration rate of CRBT is even larger in Asia and Africa - 95% of the market for digital music in Indonesia comes from CRBT[2]. CRBT replaces plain ring-back tones with music for a caller to hear as he/she waits for the receiver to answer. Unlike other instances of large social networks, which are often extracted from online networking sites, the interaction between a pair of caller and callee entails a stronger notion of intent-to-communicate, thus cellular phone call network is a better approximation of individual's real social network. It is important to understand how the dynamics of adoption are likely to unfold within the underlying phone call social network: whether individual adopts a new technology because of direct contact with adopters, or because of competition with them. However, how new products spread within such networks has not been well studied.

Our data set is from one of the largest cellular phone services in Asia. There are over one million subscribers and over one billion phone call records in our data after preprocessing. Given the size of our data, it is impossible to analyze the whole data size using the current computing power available. Hence in order to investigate CRBT diffusion, we need to use subpopulation to decrease heterogeneity and also make analysis computationally tractable. The subpopulations extracted using Transitive Clustering and Pruning (T-CLAP) (Zhang et al., 2011) are relatively isolated and

[1]Broadcast Music Inc, 2011

[2]Indonesia Finance Today, 03/15/2011

densely connected subnetworks so they can show stronger network influence on CRBT diffusion. We also used mNAP model (Zhang et al., 2011) to analyze the data since no current model supports dichotomous response variable and multiple network terms.

2 Theory

2.1 Cohesion Model

The cohesion and structural equivalence models are two competing network influence models driving diffusion (Valente, 2005). In the cohesion model, a focal person's adoption is influenced by his neighbor who he directly connects to. Those connections are communications between actors. A focal person could be informed by, persuaded by, or receive suggestion from the people in his network. The most famous piece of work about using cohesion model to explain technology diffusion might be of Coleman et al. (1966). They found that medical doctors adopted new technology because of directed ties with adopters. Rogers and Kincaid (1981) also investigated cohesion's effect on innovation diffusion. Different from Coleman et al., they used personal network density as the measure of cohesion. Their result is similar to that of Coleman et al. though, they found that personal network density is positively related to the adoption. The conclusion from the literature suggests that people notice and understand new product through discussion and observation with those who are in their social network (Harkola and Greve, 1995). In this context, cohesion assumes callers who make phone calls to each other will hear the called party's CRBT thus more likely to buy that ring-back tone or get interested in CRBT and eventually adopt the technology. Thus a focal person's direct ties to adopters influences his decision about adoption. So in the hypothesis below we want to test whether the average probability of CRBT adoption by people whom a focal person calls influences the probability of adoption by that person.

Hypothesis 1: cohesion and CRBT diffusion

The probability of CRBT purchased by a focal person is positively related to the ratio of CRBT adopters among the focal person's neighbors.

2.2 Structural equivalence Model

Structural equivalence is also known as role equivalence model. It is a positional model (Burkhardt, 1994). An ego is structurally equivalent to an alter if they connect to the same others. Structural equivalence model describes the competition between ego and alter that have same position in the social network. "Structural equivalence model were developed ... explicitly as a vehicle for describing the structure of role relations defining statuses across multiple networks." (Burt, 1987) For example, a medical doctor wants to maintain an image of innovativeness. After another doctor who he/she shares common friends or advisee with adopts a new technology, the doctor believes the adoption of a new technology will enhance his/her innovative reputation and effective power in the social network, so he wants to adopt before others who are in the same position as him.

In our research, structural equivalence refers to actors in the network having same or similar pattern of relations with others actors. In the structural equivalence model, "the trigger to ego's adoption is adoption by the people with whom he jointly occupies a position in the social structure." (Burt, 1987) In addition to occurring when people adopt the behaviors, attitudes, and beliefs of those with whom they interact, similarity in decision may occur when ego and alters connect to the same others (Coleman et al., 1966; Burt, 1982, 1987). The degree to which a focal individual and another person interact with the same others reflects the extent to which the focal and the other are structurally equivalent. Structural equivalence is a measure of the extent to which individuals communicate with the same other people. Thus, two individuals may be structurally equivalent even if they never communicate with one another. Ego would infer judgement of alters that have the same position in the influential flow of the network, and in order not to lose its influential power, ego would eventually adopt as well.

Burt (1987) reanalyzed the Medical Innovation data and observed that structural equivalent alters of adopters are more likely to adopt. Burt concluded that the effect of contagion was through structural equivalence instead of cohesion. Strang and Tuma (1993) found strong influence from doctors that are structural equivalent and little influence from cohesion. Burkhardt (1994) also compared two effects, cohesion and structural equivalence, with regard to users' attitude, self-efficacy belief and frequency of use of computers. He found "when people evaluate their own personal skills or self images, they rely on those close to them; when they determine job-related attitudes, they are more likely to rely on role (structural) equivalents." Van den Bulte and Lilien (2001) used proportion of exact alters' matches as their measure when reanalyzing the data set again. Their results show, without considering marketing effort, both network influence terms are significant, with structural equivalence being more significant.

We conclude from the literature that people in the same position in a social network will "use each other as a frame of reference for subjective judgments and so make similar judgments even if they have no direct communication with each other" (Harkola and Greve, 1995). In this context, structural equivalence is defined as the Euclidean distance between two callers. The more common friends two callers share, the smaller the Euclidean distance between them. However, in order to ensure the parameter of structural equivalence have a positive relationship with high structural equivalence, we use the inverse of structural equivalence. Our hypothesis about the structural equivalence effect on the diffusion is described as below:

Hypothesis 2: structural equivalence and CRBTs diffusion

The probability of CRBT purchased by a focal person is positively related to the extent to which structural equivalent alters have purchased CRBTs.

Comparing Cohesion and structural equivalence is important because they represent two different target scheme. Suppose we want to convert a focal person's (actor's) adoption in the network shown in Figure 1. If the adoption is driven by cohesion, then we need to influence individuals with many ties such as A, because he will influence all of his neighbors (represented by dashed arrow) eventually. However if adoption is mainly driven by structural equivalence, we need to influence individuals such as C, because he in the same social position as many individuals such as B, D, E and Actor, and can eventually affect them.

[Figure 1 about here.]

Network is usually consists of clusters, which are subnetworks that are densely connected. As the size of the population increases, the variation of the cluster size also likely to increase. According to Dunbar's number, the number of stable social relationship that an individual can maintain is between 100 and 230 Dunbar (1992). In a small network, the strength of relationship between individuals tends to be larger, thus the influence size will be larger too. Likewise, in a large network, the influence size between neighbors tends to be smaller. In a network driven by structural equivalence, the larger the network, the more likely a focal person has a structural equivalent alter. Therefore our hypotheses about the the size of networks on both network influence terms are described as below:

Hypothesis 3: network size and cohesion

The strength of cohesion's impact on probability of CRBT purchased by a focal person is negatively related to the size of the network that the focal person belongs to.

Hypothesis 4: Network size and structural equivalence

The strength of structural equivalence's impact on probability of CRBT purchased by a focal person is positively related to the size of the network that the focal person belongs to.

3 Model

So far most of the empirical models only accommodate one network influence term, and compare the coefficients from two models using Q-test (Leenders, 2002). Doreian (1989) proposed two regimes of network effect autocorrelation model, but it only support continuous response variable. Fujimoto and Valente (2011) directly put network autocorrelation terms at the right hand side of logistic regression. This model is called as Quick and Dirty (QAD) by Doreian (1982), because it does not satisfy the assumption of logistics regression – the observations are not independent. So the estimation results are biased. Another model that is close to what we need is Yang and Allenby's hierarchical Bayesian autoregressive mixture model (2003). It supports categorical dependent variable and multiple networks. But it still not sufficient, because all weights of component networks must be at the same side, and statistically significant at the same time. In summary, none of the current models meet our requirement, thus we use a newly developed model – mNAP (Zhang et al., 2013) to study network influence for the same group of users on CRBT diffusion. mNAP model is implemented using Bayesian methods and the estimates are generated through Markov Chain Monte Carlo routine. It takes both interdependence of actors and their observable attributes such as demographics into consideration. Such interdependence is described by weight matrices. Examples of such matrices are the adjacency matrix and the inverse of Euclidean distance between two individuals common friends (Burt, 1987). The model is described as below:

$$\mathbf{y} = \mathbb{I}(\mathbf{z} > 0)$$

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\theta} + \boldsymbol{\epsilon} \quad \boldsymbol{\epsilon} \sim \mathrm{Normal}_n(0, I_n)$$

$$\boldsymbol{\theta} = \rho_1 \mathbf{W}_1 \boldsymbol{\theta} + \rho_2 \mathbf{W}_2 \boldsymbol{\theta} + \mathbf{u} \quad \mathbf{u} \sim \mathrm{Normal}_n(0, \sigma^2 I_n)$$

$\mathbf{y}$ is the vector of observed binary choices, whether a caller purchased CRBT or not. It is an indicator function of the latent preference of individuals, $\mathbf{z}$. $\mathbf{z}$ could be represented as a function of both exogenous covariates $\mathbf{X}$ and autocorrelation term $\boldsymbol{\theta}$. $\boldsymbol{\theta}$ is responsible for those nonzero covariances in the $\mathbf{z}$. $\boldsymbol{\theta}$ can be described as the sum of product between network structures and observed choice $\mathbf{Wy}$. $\mathbf{W}_i$ represents the two network structures. $\mathbf{W}_1$ is the weighted adjacency matrix for cohesion; $\mathbf{W}_2$ is the matrix for structural equivalence; ρ_1 and ρ_2 are the parameters of two network influence terms respectively.

Euclidean distance is used to measure structural equivalence. In a directed network with unweighted edges the Euclidean distance, d_{ij}, between two actors i and j is the sum of squared difference between i and j's adjacency vectors. Since larger d represents less structurally equivalence between between node i and j, we use the inverse of d_{ij} plus one: $s_{ij} = \frac{1}{d_{ij}+1}$. The new measure has a positive relationship with structural equivalence.

4 Data

CRBT has become more and more popular among cellular phone users globally. It enables a called party customize what a calling party will hear as he/she waits for the called party to answer. For example, a CRBT subscriber selects a popular song as his ring-back tone. When someone calls, the caller will not hear the standard plain ring-back tone but instead will hear the song until the called party (the subscriber) answers the phone or the mailbox takes over. CRBT replaces standard ring-back tones with any tune the subscriber chooses, such as a song or a joke. With the ability to set up personalized ring-back tones, subscribers can instantly express their own individuality. They also make a fashion statement by allowing other callers to hear their own personalized CRBT. Their self-satisfaction can also be fulfilled if purchasers believe others will enjoy their CRBT.

Our data were obtained from a large Asian telecommunications company (source and raw data confidential). We have cellular phone call records and CRBT purchase records over a three-month period, and phone account holders' demographic information such as age and gender. Since the phone call conversation network is directed, asymmetry can exist between callers. We restrict the data to reciprocal calls because symmetric connections imply equal and stable connections while an asymmetric connection indicates an unstable relationship (Hanneman and Riddle, 2005). We define reciprocity for dyads (A, B) as the condition in which A calls B and B calls A in the same calendar month. We interpret reciprocity as an increase probability that the two parties are acquaintances. Thus we further constrain our analysis to include only the data that involve reciprocal dyads. Constrained by these requirements, the size of our phone call record becomes to about 197 million calls from 1.4 million customers. This network not only is too large to analyze, but also contains many clusters that contain different influence sizes in them, therefore it should be analyzed by using multiple subpopulations and meta-analysis.

[Table 1 about here.]

A detailed description of the preprocessed data is listed in Table 1. The dependent variable was measured as a binary variable, indicating whether a caller downloaded CRBT in a three-month period or not. The independent variables included in our models are gender, which is the gender of the cellular phone account holder, and age of the account holder. We also include the outdegree of the caller, which is the number of unique users a subscriber calling to, to observe the exogenous effect of number of neighbors. Cohesion is defined as callers who make phone calls to each other (0 or 1). Since the number of people a caller calls are drastically different, we normalize the cohesion matrix by dividing each row by the total number of adopters, to make the matrix element to be the percentage of adoption among neighbors.

5 Analysis

Before addressing the question of network influence over the diffusion of CRBT technology, we face another challenging problem: finding an ideal subset of the network. Our network is large, a network with 1.3 million of users. With this wealth of interconnectedness and access to data, have come two major challenges. The first and more important is with statistics. Heterogeneity across subsets of the population increases with the size of the population. Assuming that we have enough computing power, analysis of the whole population would confound across these distinct effects existing in different subsets. The second challenge is computational cost. Most of the social network analysis packages do not scale well. That is, for some class of questions and analytic routines, our standardized desktop systems are not able to finish analyzing large networks in a realistic time frame. Memory-wise desktops do not have enough resources to accommodate a structure of a large network. One solution to this problem is using subpopulations of a smaller size that is computable within the restrictions of memory size in a realistic time frame. We want subpopulations to have the following favorable characteristics: first, they have high density within the network. Internal density shows strong connections among actors in the group, so content induction is more likely to happen. Second, these networks still have variation in connections. Third, relatively few ties from within the community to the external network. We want to avoid 'boundary leakage' (nodes have more edges to the external network than the internal network, hence contaminating the structure of the extracted networks). Fourth all these subpopulations should have relatively small node sizes so the estimation of our model can be finished in reasonable amount of time. Since no algorithm existing to provide an ideal balance of quality and speed, we used T-CLAP algorithm (Zhang et al., 2011) to identify dense and relatively independent subpopulations. In a recent study, T-CLAP outperforms leading algorithms in community detection such as Infomap, which can also be used for subpopulation extraction. Since the algorithm does not set strict threshold of the subpopulation size. Instead, it returns subpopulation with local maximal I-E ratio, a measure for cluster quality. The size of returned subpopulation is close to its true value. We extracted more than 100 subpopulations, which covers more than 50% nodes in the whole population. Interestingly, the size of returned subpopulation can be categorized as just two groups, one is about 200 and the other is about 500. The results of estimation of all of them follows the same pattern in each size level, so we randomly pick five from 200 level and five from 500 level. This pattern also suggests that in the cellular phone call network, the number of contact an individual could manage is larger than Dunbar's number. The size, I-E ratio, where higher indicates denser and more cohesive, and density of all 10 subpopulations is listed in Table 2.

[Table 2 about here.]

The descriptive statistics of independent variables for each subpopulation is listed in Table 3.

[Table 3 about here.]

We present the results of network autocorrelation model for all the subpopulations, where the dependent variable is the whether an individual adopt CRBT or not. Table 4 present the results for 10 subpopulations. We find that cohesion effect is consistent across all subpopulations. The result shows that cohesion effects from all subpopulations are significant at 0.05 level, the size ranges from 0.056 to 0.075. Such result confirms our Hypothesis 1. It shows that callers receive strong influence through direct connections to alters in the same group who have already adopted. An explanation is: if a caller calls more CRBT subscribers, he gets exposure to more ring-back tones, and is more likely to hear ring tones interest him, thus he is more likely to buy ring tones too. Finally, the effects of outdegree are positive and significant at 0.01 level, which suggest if a caller calls to more people, the probability of purchasing a CRBT increases. Such result is also consistent with the finding for the Hypothesis 1.

We also observed significant effect of structural equivalence, which suggest the adoption of CRBT is impacted by both cohesion and structural equivalence. For the structural equivalence model, caller evaluate alters who are in the same social position of a phone call network as him. Same social position in a network means people are in the same kinds of relations, with the same kinds of people. Interestingly the effect size of structural equivalence varies with the size of subpopulation. When the subpopulations are at the size of about 200, the effect of structural equivalence is significant and negative, meaning individuals with more adopters having the same social position are less likely to adopt. The result seems absurd at the first glance, but actually is reasonable. One explanation is that in a smaller group, people weight individuality higher than being fashion or being different. When the subpopulations are at the size of about 500, the structural equivalence effect is significant at positive value. One explanation is that in a cellular phone call social network, parties who call each other are likely to be friends or belong to same group under some relationships. The enthusiasm of showing others about his adoption of frontier fashion and individuality is higher. The satisfaction of letting friends appreciate his fashion taste or simply an interesting tone is also higher. Motivated by this thought, a perceived competition is created among these subscribers. Ego will know about an alter he does not necessary call to has adopted ring-back tone through common friends they both call. The more ring tones those alters bought, the more CRBT the ego will adopt. The results also suggest that network influence size might be correlated with the subpopulation size, hence meta analysis to summarize the estimated parameters for cohesion and structural equivalence based on subpopulation size is required.

In general coefficient for cohesion has higher significance level and size than that of structural equivalence, which suggests cohesion effect is more influential on CRBT adoption than structural equivalence. Cohesion has a significant effect on a caller's decision on adoption of CRBT – adoption is influence by adopters among neighbors. When the group size is large, people tend to imitate competitors. When the group size is small, people tend to differentiate themselves from others. Based on the results, we should use different trigger schemes for groups at different sizes. If we want to trigger more adoptions, for smaller groups, we only need to concentrate on individuals with many connections; while for larger groups, we should not only concentrate on users who are popular, but also these who are in the similar positions in the group.

[Table 4 about here.]

We then break down our subpopulations to two groups based on their sizes. One group for subpopulations is at size of about 200, the other group for subpopulations is at size of about 500. Through the comparison of the two pooled mean effect, we could determine whether influence size of cohesion and structural equivalence varies with the size of network, thus address Hypothesis 3 and 4. The summarized cohesion effect in networks of 200-level is in Table 5.

[Table 5 about here.]

In the smaller network, cohesion has a pooled mean at 0.063, with a 95% confidence interval at (0.027, 0.099). The larger network, at about 500, has a pooled mean cohesion effect at 0.048, with a confidence interval at (0.023, 0.089). The results show the cohesion effect is stronger in smaller subpopulations. The result of meta-analysis confirmed our Hypothesis 3 that in smaller group individuals has stronger influence to each other.

[Table 6 about here.]

For the five networks at the size of about 200, the summarized structural equivalence effect is statistically significant with a size of -0.0067, and a 95% confidence interval of $(-0.011, -0.0025)$ (Table 7). This result confirms that when the network size is smaller, individuals in the network tend to make different choice from people who are at the same social position.

[Table 7 about here.]

When the network size increases to 500, the summarized structural equivalence effect is statistically significant at 0.0092, with a confidence interval of $(-0.011, -0.0025)$ (Table 8). The positive effect size suggests that in a large network, at about 500, individuals tend to imitate other people who are at the same social position. Combining the results in the last two tables, we conclude that Hypothesis 4 is confirmed. Furthermore, the direction of influence from structural equivalence are different when network size varies. This is a new finding in such kind of research. Such result can only be gotten using subpopulations at different size and a model that supports multiple network autocorrelation terms.

[Table 8 about here.]

5.1 Robustness Check

Significant network autocorrelation term does not necessary mean a model with two terms is better. We use Deviance Information Criterion (DIC) to measure the goodness-of-fit of our model. It is similar to model selection criteria AIC and BIC, and calculated from MCMC draws. The definition of DIC is shown as below. Models with the lowest value are the best model.

$$DIC = 2\bar{D} - D(\bar{\theta})$$

where $\bar{D} = \mathbb{E}[-2\log(p(y|\theta))]$, $D(\bar{\theta}) = -2\log(p(y|\bar{\theta}))$.

As the results shown in the Figure 2, when we have network influence terms, the model is better. The model with cohesion only is better that the one with structural equivalence only. The best model, the one with the lowest DIC our model, is the one with both cohesion and structural equivalence. Thus we should include two network terms in the model. The goodness-of-fit test confirmed that CRBT adoption is influenced by the joint effects of cohesion and structural equivalence.

[Figure 2 about here.]

6 Conclusion

The debate among researchers about two classes of network models, cohesion and structural equivalence, and the impact on diffusion in social networks still persists. However, other than Coleman's classical Medical Innovation data, few new data sets have been used to address this research question. Reconciling these findings is very important because each model represents different adoption triggering strategy. Our study is one of the very few to investigate multiple network influence terms on diffusion, particularly in large network context. One major challenge in analysis comes from the size of the data. We used an innovative algorithm T-CLAP to extract subpopulations from the global network. Given the fact that more and more data sets are in societal scale, this method could have wide application. We find that in a large network whose node size is at million level, group size only falls in two levels, 200 and 500. Using the subpopulations extracted, we analyze the effects of cohesion and structural equivalence's on adoption of CRBT by using mNAP model. Our results show that the adoption of CRBT is consistently correlated with cohesion. When the size of subpopulation is small (at the level of 200), CRBT adoption is negatively correlated with structural equivalence; when the size of subpopulation is large, at the level of 500, adoption is positively correlated with structural equivalence. Between the two network influences, cohesion has a more significant impact. So the strength of communication or connection still has stronger influence on adoption.

References

Burkhardt, M. E. (1994). Social interaction effects following a technological change: A longitudinal investigation. *Academy of Management Journal*, 37(4):869–898.

Burt, R. S. (1982). *Toward a Structural Theory of Action : Network Models of Social Structure, Perception, and Action*. Quantitative Studies in Social Relations. New York: Academic Press.

Burt, R. S. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology*, 92(6):1287.

Coleman, J. S., Katz, E., and Menzel, H. (1966). *Medical innovation: A diffusion study*. Bobbs-Merrill Co.

Doreian, P. (1989). *Two Regimes of Network Effects Autocorrelation*. The Small World. Ablex Publishing.

Dunbar, R. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, 20:469–493.

Duncan, O. D., Haller, A. O., and Portes, A. (1968). Peer influences on aspirations: A reinterpretation. *The American Journal of Sociology*, 74(2):119–137.

Hanneman, R. and Riddle, M. (2005). *Introduction to social network methods*. Online.

Harkola, J. and Greve, A. (1995). Diffusion of technology: Cohesion or structural equivalence? In *Academy of Management Best Papers Proceedings*, pages 422–426, Vancouver, Canada.

Leenders, R. T. (1997). *Longitudinal behavior of network structure and actor atributes: modeling interdependence of contagion and selection.*, chapter Evolution of Social Networks. Gordon and Breach, New York.

Leenders, R. T. (2002). Modeling social influence through network autocorrelation: constructing the weight matrix. *Social Networks*, 24:21–47.

Rogers, E. M. and Kincaid, L. D. (1981). *Communication Networks: Toward a New Paradigm for Research*. Free Press.

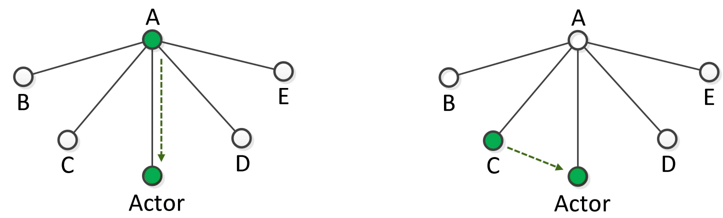
Strang, D. and Tuma, N. B. (1993). Spatial and temporal heterogeneity in diffusion. *The American Journal of Sociology*, 99(3):614–639.

Valente, T. W. (2005). Models and methods for innovation diffusion. In Carrington, P. J., Scott, J., and Wasserman, S., editors, *Models and Methods in Social Network Analysis*. Cambridge University Press.

Van den Bulte, C. and Lilien, G. L. (2001). Medical innovation revisited: Social contagion versus marketing effort. *American Journal of Sociology*, 106(5):1409–1435.

Zhang, B., Krackhardt, D., Krishnan, R., and Doreian, P. (2011). An effective and efficient sub-population extraction method in large social networks. *Proceedings of International Conference on Information Systems*.

Zhang, B., Thomas, A. C., Doreian, P., Krackhardt, D., and Krishnan, R. (2013). Contrasting multiple social network autocorrelations for binary outcomes, with applications to technology adoption. *ACM Transactions on Management Information Systems*.



(a) Cohesion model: actor's adoption is influenced by direct tie from A

(b) Structural equivalecne model: actor's adoption is influenced by C, which is in the same social position

Figure 1: Cohesion model vs structural equivalence model

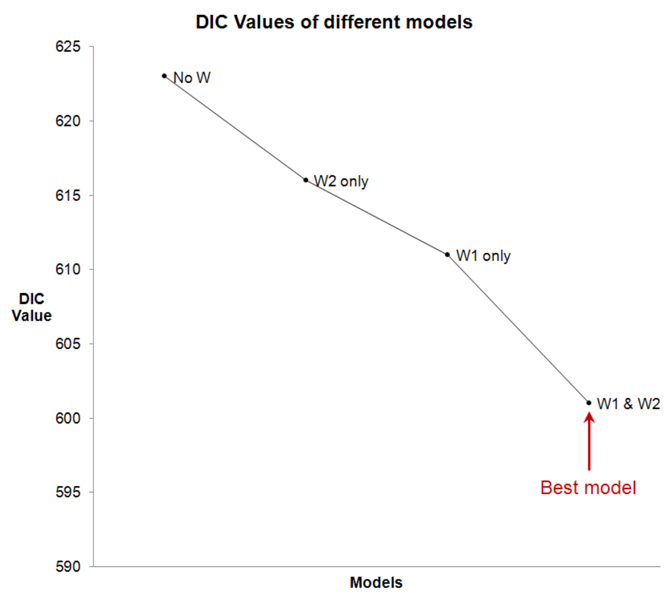


Figure 2: DIC of four different models

Table 1: Variable Description

Variable	**Description**
$\mathbf{y}$	Dependent variable, whether caller purchased CRBT (binary)
Gender	Gender of cellular phone account holder
Age	Reported age of cellular phone account holder
Degree	Number of callees that a caller has
$\mathbf{W}_1$	cohesion effect matrix, normalized
$\mathbf{W}_2$	structural equivalence effect matrix

Table 2: Extracted subpopulations structure characteristics

Subpopulation	n	I-E **ratio**	**Density**
1	191	0.83	0.13
2	202	0.43	0.10
3	213	0.48	0.039
4	238	0.82	0.030
5	263	0.82	0.029
6	465	0.43	0.075
7	485	0.82	0.033
8	553	0.82	0.029
9	563	0.43	0.021
10	597	0.82	0.082

Table 3: Descriptive statistics of independent variables

	Subpopulations									
Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
n	191	202	213	238	263	465	485	553	563	597
Gender	0.19	0.12	0.033	0.11	0.065	0.57	0.12	0.27	0.12	0.050
0=male	(0.50)	(0.45)	(0.18)	(0.40)	(0.29)	(0.89)	(0.41)	(0.62)	(0.42)	(0.25)
Age	45	42	38	42	42	48	45	40	42	36
	(11)	(12)	(11)	(11)	(12)	(12)	(13)	(13)	(13)	(12)
Degree	13.6	10.0	3.6	4.0	4.5	18.1	9.3	9.8	5.3	3.4
	(14.9)	(10.6)	(5.1)	(5.2)	(5.6)	(18.6)	(13.6)	(15.5)	(6.7)	(1.8)

Table 4: Results of analysis using mNAP model, cohesion vs structural equivalence

Variable	**Subpopulations (1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
Gender	0.38	-0.20	-0.034	0.27	2.0	-0.11	1.4^{***}	0.69	0.094	0.90
	(0.51)	(3.3)	(0.33)	(0.51)	(0.12)	(0.31)	(0.66)	(0.42)	(0.46)	(0.87)
Age	0.023	-0.026	0.053^{**}	0.024	$0.028^{\dagger}$	$0.0086^{\dagger}$	0.064^{**}	0.37^{**}	0.065	-0.008
	(0.018)	(0.017)	(0.014)	(0.015)	(0.017)	(0.021)	(0.014)	(0.013)	(0.040)	(0.017)
Degree	-0.0014	0.040^{**}	0.030^{**}	0.0043	0.069^{**}	0.040^{**}	0.042^{**}	0.023^{**}	0.043^{**}	0.024^{**}
	(0.0088)	(0.0098)	(0.0095)	(0.020)	(0.020)	(0.022)	(0.019)	(0.033)	(0.044)	(0.0070)
Cohesion	0.07^{**}	0.06^{**}	0.065^{*}	0.056^{*}	0.061^{**}	0.053^{*}	0.043^{**}	0.045^{**}	0.053^{*}	0.046^{**}
	(0.022)	(0.022)	(0.026)	(0.070)	(0.019)	(0.022)	(0.015)	(0.015)	(0.022)	(0.014)
Struct.	$-0.0059^{\dagger}$	-0.0098^{*}	-0.010^{*}	-0.0079^{*}	-0.0015^{*}	0.0066^{*}	0.0098^{*}	0.0010^{*}	0.034^{*}	0.0049^{*}
equiv.	(0.0031)	(0.0041)	(0.0045)	(0.0035)	(0.0008)	(0.0030)	(0.0041)	(0.0006)	(0.015)	(0.0024)

**: $p < 0.01$, *: $p < 0.05$, †: $p < 0.10$

Table 5: Meta-analysis for cohesion effect, subpopulations size ≈ 200 pooled

Subpopulation	**Cohesion**	**95% C.I. Lower**	**Upper**	**Weight**
1	0.07	0.026	0.11	1.1
2	0.06	0.026	0.14	0.15
3	0.065	0.0034	0.13	0.52
4	0.056	0.01	0.10	0.66
5	0.061	0.037	0.085	2.6

Summary effect = 0.063, 95% C.I. = (0.027, 0.099)

Table 6: Meta-analysis for cohesion effect, subpopulations size ≈ 500 pooled

Subpopulation	**Cohesion**	**95% C.I. Lower**	**Upper**	**Weight**
6	0.053	0.021	0.085	0.81
7	0.043	0.039	0.15	0.91
8	0.045	0.016	0.075	1.0
9	0.053	0.011	0.094	1.1
10	0.046	0.030	0.062	1.2

Summary effect = 0.048, 95% C.I. = (0.023, 0.089)

Table 7: Meta-analysis for structural equivalence, subpopulations size $\approx$ 200 pooled

Subpopulation	Structural Equivalence	95% C.I. Lower	Upper	Weight
1	−0.0059	−0.0096	−0.0022	1.0
2	−0.0098	−0.014	−0.0059	0.82
3	−0.010	−0.017	−0.0034	0.90
4	−0.0079	−0.014	−0.0017	1.1
5	−0.0015	−0.0026	−0.00043	1.2
Summary effect = −0.0067, 95% C.I. = (−0.011, −0.0025)				

Table 8: Meta-analysis for structural equivalence, subpopulations size $\approx$ 500 pooled

Subpopulation	Structural Equivalence	95% C.I. Lower	Upper	Weight
6	0.0066	0.0043	0.0089	0.63
7	0.0098	0.0028	0.017	1.1
8	0.001	0.0004	0.0016	1.1
9	0.034	0.019	0.049	0.97
10	0.0049	0.0032	0.0066	1.2
Summary effect = 0.0092, 95% C.I. = (0.0048, 0.014)				