Quantifying Social Influence in Social Networks with Stochastic Block Influence Model

Abstract

Understanding the effect of social influence has received increasing attention in computer science, economics, and social science in recent years. In this paper, we introduce a generative model, called the Stochastic Block Influence Model (SBIM), to study the effect of social influence on adoption behaviors, where people make decisions based on their preferences and the social influence exerted by neighbors. Inspired by the stochastic block model, we assume that social influence varies across and within social blocks, and we infer two types of relationships: the interactions and the strength of influence among social blocks. We test our model on the adoption of microfinance in an Indian village, which shows that our method outperforms the baseline model. By analyzing the block-to-block social influence matrix, we find that social blocks with overlapping characteristics are associated with positive influence, while blocks with a lack of overlap are associated with negative influence. Moreover, the less diverse social blocks exert a negative influence on more diverse ones in general. The model and these findings have significant implications for information diffusion, viral marketing, and technology adoptions.

Introduction

We now live in an increasingly connected environment (Travers and Milgram 1967; Backstrom et al. 2012). The inter-connections among people not only foster information diffusion but also enable the inter-dependencies in the decision-making among peers. The understanding and modeling of how the hidden social influence changes people's decision-making are, therefore, essential and critical for many practical applications, such as viral marketing, political campaigns, and large-scale health-related behavioral change (Fowler and Christakis 2008; Pan, Altshuler, and Pentland 2012; Leng et al. 2018). There have been theories aiming to model social influence operating on simple model assumptions, such as independent cascade (Kempe, Kleinberg, and Tardos 2003) and complex contagion (Centola and Macy 2007). However, they fail to capture the rational decision-making of individuals.

Even though social networks are usually observed, the hidden social influences are always hard to infer from observational data. There are different lines of research in economics and computer science focusing on the quantification of social influence. In economics, researchers use causal inference to distinguish social influence from homophily - a common phenomenon in social networks - using strategies such as propensity score matching (Aral, Muchnik, and Sundararajan 2009), behavioral matching (Leng et al. 2018) and regression adjustment (Angrist 2014). Relevant studies in computer science focus on building a generative cascade and maximizing the likelihood of such a cascade. Gomez-Rodriguez assumes the existence of an unobserved hidden information pathway and maximizes the likelihood of such cascade via stochastic convex optimization (Gomez-Rodriguez, Leskovec, and Krause 2012; Gomez Rodriguez, Leskovec, and Schölkopf 2013). Myers models social influence with internal exposure (within the network) and external exposure to information (from an external event) with a probabilistic model and fits parameters by maximizing the log-likelihood (Myers, Zhu, and Leskovec 2012). Yang develops a linear influence model where the influence functions of individual nodes govern the diffusion through the network (Yang and Leskovec 2010).

Modeling the underlying relational information among agents by segmenting them into blocks has raised considerable interest recently in machine learning (Abbe 2017). Operating under the similar assumption that hidden social groups of agents determine their decision-making, we develop a generative model to infer two types of relationships - social interaction and social influence - based on the observed network structure and the adoption behavior. This is one of the main contributions and advantage of our model. Moreover, rather than assuming that social influence always increases individuals' likelihood of adoption (Kempe, Kleinberg, and Tardos 2003; Centola and Macy 2007; Leng et al. 2018), we empirically learn the effect and flexibly allow it to be negative. Our method is built upon the Stochastic Block Model (SBM), which provides a fertile building block to study more complex social processes. Moreover, the analysis on the inferred social block, block-to-block interaction and block-to-block social influence reveal many interesting stories that are otherwise unobservable. We experiment on the diffusion of microfinance data in an Indian village and analyze the social influence between the learned social blocks. In general, we find that the less diverse so-

cial blocks exert a negative influence on more diverse social blocks. Moreover, social blocks with overlapping professions, gender distributions, and education levels have a positive influence on one another; while social blocks with lower education levels have a negative influence on higher-educated blocks. Furthermore, social blocks with a gender imbalance cause a block to be more open to negative influence from gender-balanced blocks.

The remaining sections are organized as follows. We first briefly discuss the Stochastic Block Model. After that, we describe the new Stochastic Block Influence Model (SBIM). Next, we describe our experiment setting and analyze the results. Lastly, we conclude the paper and state future works.

Background and related work

The Stochastic Block Model is a statistical inference model on networks and

SBM builds upon the Erdos-Renyi random graph model ($G(n,p)$), but generalizes with higher intra-cluster connectivity and lower inter-cluster connections. Assume a random graph $G(V,E)$ that is partitioned into C disjoint blocks $(V_1,...,V_C)$, and the proportion of nodes in each block c is π_c, and $\sum_{c=1}^{C}\pi_c = 1$. Let matrix $\mathbf{B} \in \mathbb{R}^{C\times C}$ denotes the inter-block and intra-block connection probability. We therefore have the probability of a link between v_i and v_j between two separate blocks V_k and V_l as $Pr((v_i,v_j) \in E|v_i \in V_k, v_j \in V_l) = p_{ij}$. Let $\mathbf{m}_i$ be the block assignment of individual i. The mixed membership of individuals can be modeled simply letting $\mathbf{m}_i \in \mathbb{R}^C$, and $\sum_{j=1}^{C} m_{ij} = 1$. Together, we combine the social block vector of all individuals in matrix $\mathbf{M} \in \mathbb{R}^{|V|\times C}$.

Model

Our focus is to study the block-to-block unobserved social influence on individuals' decision-making process and network formation. We desire a model with the following properties:

1. The network formation and social influence are determined by individuals' hidden blocks.
2. The model has the flexibility to leverage both the observed network structure and the adoption of behavior to infer the block assignment.

Stochastic block model provides a framework to learn hidden block structure in our study (Airoldi et al. 2008). Rather than fixing the membership of each individual, we relax the constraint and allow individuals to be assigned to a mixture of social blocks. The main innovation of our model is to jointly infer the block-to-block social influence matrix and block-to-block interaction matrix. Therefore, the block assignment influences not only the presence/absence of a link but also the adoption behaviors, which extends the traditional SBM by studying these two types of relationships.

First, we set down the notations. We have N individuals and K social blocks in the network $\mathcal{G} = (V,E)$. We observe the adjacency matrix representing the connectivity between two individual pairs ($\mathbf{A} \in \mathbb{R}^{N\times N}$) and the binary vector representing individuals' adoption behaviors ($\mathbf{y} \in \mathbb{R}^N$). Let $X_i \in \mathbb{R}^D$ be the characteristics of i and the node is associated with a vector of block membership $\mathbf{m}_i \in \mathbb{R}^C$. D is the dimension of the sociodemographic covariates. For example, an individual has multiple identities, such as having a profession in agriculture, being non-native to the village, and having a low level of education. Therefore, he is part of several social blocks: the farmer, non-native, and low-education social blocks.

To generate the network, the model considers each pair of nodes according to the block assignment and block-to-block interaction $\mathbf{B} \in \mathbb{R}^{C\times C}$. For each pair $\{v_i, v_j\}$, depending on the mixed block assignment of v_i and v_j, $\tilde{A}_{ij}$ is generated according to $\mathbf{B}$. In particular,

$$Pr\big(\tilde{A}_{ij} = 1|\mathbf{m}_i, \mathbf{m}_j\big) = \mathbf{m}_i\mathbf{B}\mathbf{m}_j. \tag{1}$$

To generate the decision-making, the model considers each node and its neighbors according to the block assignments and the adjacency matrix, as well as the block-to-block influence matrix $\mathbf{F} \in \mathbb{R}^{C\times C}$. For each node v_i and its neighbors $G(v_i)$, we aggregate the influence from his/her neighbors by summing up the influence according to the pair-wise influence. As an example of this method, let us consider an individual v_i who is a low-educated farmer from a low caste. To make a decision, this farmer would take into account the influence of his/her neighbors in the same block, likely other individuals working in agriculture and of similar castes. Furthermore, this farmer would also be influenced by other blocks in different manners: for instance, this farmer is part of a gender-imbalanced block and a low-educated block, which is highly influenced by high-educated and high-caste block. Thus, the farmer's decision is made through a combination of his demographic factors, influence from his neighbors, and influence from other blocks. The adoption decision of an individual can then be formulated by,

$$y_i = \text{logit}\big(\beta\mathbf{X}_i + \sum_{j\in G(i)} \big(\mathbf{m}_i F_{ij} \mathbf{m}_j\big) \circ \mathbf{A}\big), \tag{2}$$

where $\circ$ is the element-wise matrix multiplication. The first term $\beta\mathbf{X}_i$ measures the adoption decision conditioned on one's sociodemographics if there is no influence, and the second term aggregates over the influence from i's neighbors. The element-wise product ensures that only direct neighbors exert influence, which can be adapted to incorporate higher-order influences.

We described the probability that governs the agents' adoption behaviors and pair-wise connections. The graphical representation of our model is shown in Figure 1. For the full network, the model assumes the following generative process:

1. For each node $v_i \in V$, draw a C-dimensional mixed membership vector $\mathbf{m}_i \sim \text{Dirichlet}(\mathbf{c}_i) \in \mathbb{R}^C$.
2. Draw block-to-block connectivity matrix $\mathbf{B}_{ij} \sim \text{Beta}(\alpha_{ij}, \beta_{ij}) \in \mathbb{R}^{C\times C}$.
3. Draw block-to-block influence matrix $\mathbf{F} \sim \mathcal{N}(\mu_{ij}, \sigma_{ij}) \in \mathbb{R}^{C\times C}$.

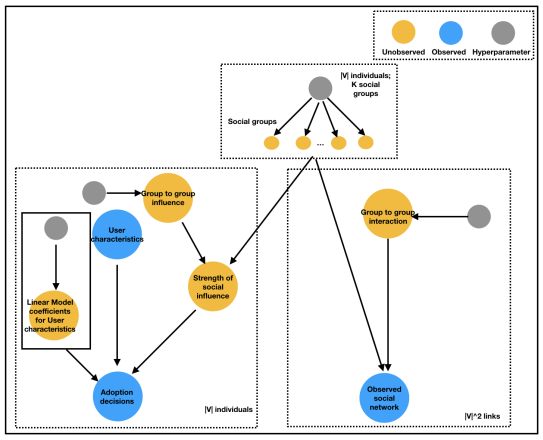


Figure 1: Graphical representation of the Stochastic Block Influence Model (SBIM). We use three types of nodes colored by yellow, blue, and gray, to denote observed variables, unobserved variables, and hyperparameters. We assume people's decision-making process is jointly determined by their sociodemographic information and the social influence exerted by their neighbors. Social influence is determined by the individual's social block, as well as his neighbor's social block. The same social block also generates the network.

4. Draw coefficients of the individual attributes $\beta \sim \mathcal{N}(\boldsymbol{\mu}_b, \boldsymbol{\sigma}_b) \in \mathbb{R}^{1\times D}$.
5. Draw the connection between each pair of nodes v_i and v_j, according to Equation (1).
6. Draw the adoption decision of the individual according to Equation (2).

For abbreviation, we denote Z as the hidden variables, $Z = \{\boldsymbol{b}, \mathbf{m}, \mathbf{B}, \mathbf{F}\}$ and θ as the set of hyperparameters, $\theta = \{\boldsymbol{\mu}_b, \boldsymbol{\sigma}_b, \mathbf{c}, \mu_{F_{ij}}, \sigma_{F_{ij}}, \boldsymbol{\alpha}, \boldsymbol{\beta}\}$

This generative process defines a joint probability distribution over the N node-wise membership matrix $\mathbf{M}$, the block-to-block interaction matrix $\mathbf{B}$, the block-to-block influence matrix $\mathbf{F}$, attributes coefficients $\boldsymbol{\beta}$, observed network $\mathbf{A}$, and the observed adoption decision $\mathbf{y}$.

Given the observed network and adoption behavior, the model defines a posterior distribution - the conditional distribution of hidden block structure and relationships - that decomposes the agents into C overlapping blocks. The posterior will place a higher probability on the configurations of community membership that describe densely connected communities as well as stronger (positive or negative) influences. We present a visualization in Figure 3a and 3b which illustrates that the posterior superimposes a block structure on the original network.

Inference As with many hierarchical Bayesian models, the posterior described above is intractable. We, therefore, use the Markov Chain Monte Carlo (MCMC) algorithm as an approximate statistical inference method to estimate the parameters. MCMC draws a series of correlated samples that converge in distribution to the target distribution and is generally asymptotically unbiased.

Gibbs sampling and Metropolis-Hastings methods converge slowly to the target distribution as they explore the parameter space by random walk (Hoffman and Gelman 2011). Hamiltonian Monte Carlo (HMC) suppresses the random walk behaviors with an auxiliary variable that transforms the problem with sampling to a target distribution into simulating Hamiltonian dynamics. However, HMC requires computing the gradient of the log-posterior, which is tedious in our model, and it requires a reasonable specification of the step size and a number of steps, which would otherwise result in a substantial drop in efficiency (Hoffman and Gelman 2014)

Therefore, we apply a No-U-Turn Sampler (NUTS), a variant to HMC method, to eliminate the need for choosing the number of steps by automatically adapting the step size. Specifically, NUTS builds a set of candidate points that spans the target distribution recursively and stops automatically when it starts to double back and retrace its steps (Hoffman and Gelman 2014). We use the NUTS algorithm implemented in Python PyMC3 (Salvatier, Wiecki, and Fonnesbeck 2016).

Experiments

We study the adoption of microfinance collected by JPAL in an Indian village (Banerjee et al. 2013). We study village indexed by 64 with 257 villagers. The data includes not only the relationships between households, but also a variety of amenities of households, including roofing materials, type of latrine, quality of access to electric power, and so on. The amenities are used as the predictors ($\mathbf{X}$). The outcome variable is the adoption decision on microfinance.

Baselines. We use the Random Forest with covariates ($\mathbf{X}$) and the hidden community learned by spectral clustering on the adjacency matrix. We use the social blocks learned by spectral clustering as a piece of additional information supplementary to $\mathbf{X}$, and then perform random forest. Spectral clustering uses the second smallest eigenvector of the graph Laplacian as the semi-optimal partition (Ng, Jordan, and Weiss 2002).

.

Model Training and Evaluation. To train our model and evaluate the performance for a particular C, the number of social blocks, we cross-validate by randomly splitting the data into 75% training samples and 25% test samples. We repeat this process by ten times. With NUTS, we obtain the point estimates for all latent variables in Z. We then rerun our model (as previously described) with all latent variables fixed to the estimates on the test dataset. This step returns the predicted adoption probability for each villager in the test data. Some critical hyperparameters for NUTS are the number of burn-in samples, the number of samples after burn-in, the target acceptance probability and the number of chains. For all of our NUTS sampling runs, we burn 3,000 samples

to ensure that MCMC mostly converge to the actual posterior distribution. The number of samples after burn-in is 500; usually, only less than ten samples (among the 500) are diverging. Next, we select the target acceptance probability to be 0.8. At the end of each run, we average across the 500 samples to derive point estimates for all latent variables.

We evaluate our method using the AUC since our data is imbalanced. AUC is the area under the Receiver-Operating-Characteristics curve, which is plotted by the false positive rate and correct positive rate for different thresholds. We define a loss metric during the training period to select the best configurations. It is formulated by the negative of the standard improvement measure, which is the absolute improvement in performance normalized by the room for improvement. This metric is formulated by $L = \frac{\text{Baseline}_{\text{Test AUC}} - \text{SBIM}_{\text{Test AUC}}}{1 - \text{Baseline}_{\text{Test AUC}}}$, where the Baseline Test AUC is obtained from the random forest and the SBIM Test AUC is from our SBIM, both evaluated on test sets.

This measure captures the improvement of our method compared to the baseline. Since we have a small test set, a randomly-drawn test set may be harder to predict than others. Measuring the relative improvement ensures that the composition of the test set does not bias the performance.

Hyperparameter Tuning. Our model has six hyperparameters. Since the parameter space is large, we adapt a bandit-based approach to tune the parameters developed in called Hyperband (Li et al. 2017). The Hyperband algorithm adaptively searches for configurations and speeds up the process by adaptive resource allocation and early-stopping. Our adaptation of this algorithm allows each configuration tested to run with full resource due to the sampling used in our methodology, allowing NUTS to run consistently across all configurations.

The ranges from which these hyperparameters were sampled are as follows: $\mu_b \in [-2, 2]$, $\sigma_b \in [-0.1, 1]$, $\mu_b \in [-2, 2]$, $\boldsymbol{c} \in [0.5, 1.5]$, $\mu_{\boldsymbol{F}} \in [-6, 6]$, and $\sigma_{\boldsymbol{F}} \in [0.1, 3]$. We let $\boldsymbol{\alpha} = \boldsymbol{\beta} = 2$ to allow a reasonable and non-skewed prior.

Performance. We first need to choose the optimal number of social blocks before we evaluate the performance. To do so, we first tune the model with different hyperparameters for each C we consider and then calculate the average loss. When testing $C \in \{2, 6, 10, 14\}$, we observe that there is a negative parabolic trend with the loss peaking at its lowest at $C = 10$ blocks, so we use this number of social blocks for further evaluation.

We compare the performance of our model with the baseline in Table 1. We observe that our method outperforms random forest in the test set by 13.8% by the improvement metric, which is the negative of the loss formulated previously, and the baseline random forest model overfits the training set.

Table 1: Model and Baseline Performance

	Mean	Standard Deviation
Baseline Train AUC	0.901	0.010
Baseline Test AUC	0.610	0.095
SBIM Train AUC	0.805	0.022
SBIM Test AUC	0.664	0.062

Discussion

Block Type and Interaction Matrix. As stated in the model, the operating assumption is that there may be several subgroups of people who influence each other whose relationships are hidden and using the SBIM allows us to capture these dynamics. A motivating example of this would be in a village; there could be several types of individuals of different professions, education levels, and a variety of other demographic features. Take one individual, who is a low-educated member of a lower caste. Within a potentially diverse community, it is vital to capture who influences this individual - it could be influenced from a set of people in the community of a similar class or a set of people in the community with similar professions. The block aspect of the model allows us to capture these critical, hidden relationships. SBIM model allows us to infer block membership information for each individual.

We present the size of each social block in Figure 2. Social block two is larger than the other blocks, and the sizes of the rest are similar. It is socially realistic that many individuals are grouped in a block of larger size. We represent the adjacency matrix sorted by this inferred block index from smallest to largest block in Figure 3a. We see that there are many links within all of the blocks along the diagonal, demonstrating that the block model is meaningful and captures more links within than across blocks. The largest block, furthest along the diagonal, has fewer linked individuals, which is reasonable since it is the largest group and we do not expect all individuals in the most significant blocks to be linked solely due to size. We use the standard normalized entropy to measure the diversity of different attributes. Normalized entropy is a metric used to capture the number of types of characteristics within each category while accounting for the frequency of each entity type within a category. It can be formulated by, $D = -\frac{\sum_{i=1}^{k} p_i \log(p_i)}{\sum_{i=1}^{k} \frac{1}{n_i} \log(\frac{1}{n_i})}$, where k refers to the number of types within a category, p_i refers to the probability of each type i, f refers to the number of occurrences of each type n_i.

The gender ratio (R) is measured within a block and is formulated by $R = \frac{r_m}{r_f}$, where r_m and r_f refer to the number of occurrences of males and females respectively. Thus, since R is the ratio of males to females in a block and a high or low gender ratio, both correspond to a high gender imbalance.

These metrics for diversity and gender ratio are used in to evaluate block characteristics for a specific example in Figure 5 and Table 2, in addition to being used to evaluate the group attributes that are associated with different types of influence in Table 3.

We can associate the sociodemographic characteristics belonging to the individuals in a block to generalize block type as consisting of characteristics such as high or low

class, homogeneous or diverse, skilled or less educated, as depicted in Table 2. In this example, each block is associated with a qualitative type, and the attributes within that block leading to such characterization are described. Lower or higher class blocks are designated by caste composition, education levels, and profession types. Homogeneous or diverse blocks are designated by some profession composition, caste types, mother tongue language composition, gender imbalance, and fraction of village natives.

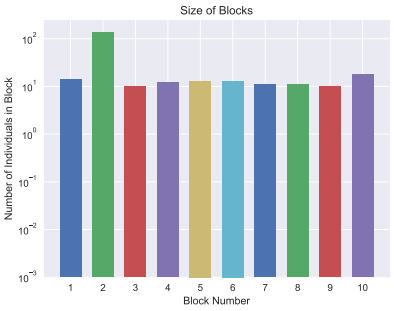
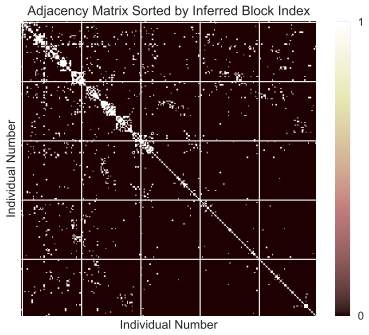


Figure 2: Size of each social block

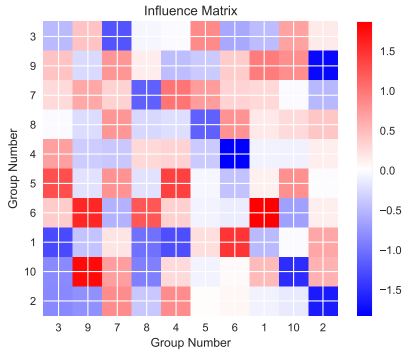
Influence Matrix and Attributes. The block-to-block influence, sorted by increasing block size, is displayed in Figure 3b, where the strength of social influence, allowed to be either positive or negative, is shown. We can see some blocks influence other blocks ranging from strongly negative influence, no influence, to strongly positive influence.

The total influence into and out of each block is depicted in Figure 4, which allows us to evaluate what type of influence a block receives and outputs (net positive, negative, or neutral). For example, we can see diverse, low-class block 5 and senior, low-class block 6 output high levels of positive influence, and diverse, middle-class block 8 receives a net high level of negative influence. We observe that some blocks have a stronger outgoing influence than other blocks and can perceive of these as positive and negative influence leaders. Similar reasoning applies to characterize blocks that receive a high level of influence as follower blocks, furthermore observing the difference in net incoming and outgoing influence within each block as relating to its role in the block-to-block network. We refer to this to interpret different dynamics between social blocks, in addition to then pairing this information with demographic information to make further evaluations about block characteristics associated with different types of influence.

In Figure 5, a subset of the sociodemographic features are displayed for each block, where the network of blocks is connected with varying degrees of influence between them. For example, we can see that lower median-age block four negatively influences the older median-age block 6. The equal gender ratio block 10 positively influences the similarly equal gender ratio block 9. Block 10 influences block 9, where both blocks have similarly high caste diversity. Highly language diverse block 6 positively influences



(a) Interaction matrix.



(b) Influence matrix.

Figure 3: Interaction matrix and influence matrix.

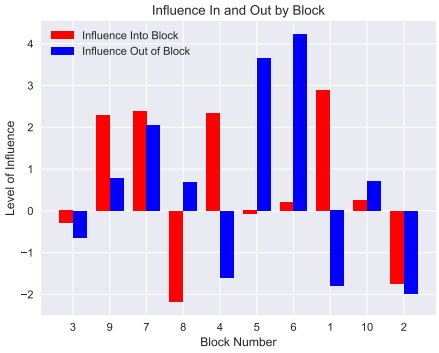


Figure 4: Net Influence In and Out for Each Block

low language diverse block 1. Lower professionally diverse block 1 negatively influences higher professionally diverse block 3.

By analyzing several examples in this manner using block characteristic composition and observing the types and patterns of influence, several general trends arise, as depicted in Table 3. The block attributes most frequently associated with different types of influence are summarized into key trends. Positive influence occurs when two blocks have overlap

Table 2: Block Characteristics Example. The majority refers to the largest subset. Disadvantaged caste refers to lower castes, including the castes OBC (Other Backward Class) and Scheduled castes. Higher education refers to having education levels at P.U.C (pre-university course) and having a "degree or above" designation. Moderate and lower education levels include all levels below this, where moderate levels have more S.S.L.C. (Secondary School Leaving Certificate) levels, and P.U.C. levels and lower levels have mostly primary school education levels.

Block	Block Type	Attributes
1	Homogeneous, low-class	only one disadvantaged caste and one language spoken
		low profession diversity and education levels
2	Diverse, skilled, highly-educated	several different castes from many levels
		diverse languages and diverse, high-skilled professions.
3	Senior, low-class	majority disadvantaged caste
		majority low skill-level professions in agriculture
4	Young, low-class	younger average age, gender imbalanced block
		majority lowest caste members, mostly natives
		higher education
5	Diverse, low-class	diverse number of disadvantaged castes
		moderate language diversity, moderate education
		most jobs in agriculture
6	Senior, low-class	older average age,diverse in low castes
		two languages spoken,,very low education
		lower-class professions
7	Homogeneous, low-class	gender imbalanced, mostly disadvantaged caste
		one language majority
		majority professions in agriculture and sericulture
8	Diverse, middle-class	mostly one language
		caste diverse but mostly lower castes
		diverse professions
9	Diverse, highly-educated, low-class	disadvantaged caste majority
		diverse jobs, higher-class professions (teacher, priest)
		high education level, diverse languages
10	Homogeneous, low-class	gender-balanced
		majority disadvantaged caste, only one language spoken
		majority of professions in agriculture and sericulture

Table 3: Block Attributes Associated with Different Types of Influence. Positive and negative influence refers to influence type from one block to another block. Self-influence refers to positive influence within the same block. Overlap refers to overlapping categories.

Attribute	Positive Influence	Negative Influence	Positive Self-Influence
Gender	Similar gender distribution	Gender-imbalanced block is more open to negative influence from gender-balanced block	Large gender imbalance
Caste	Overlapping majority castes	Lack of overlap in caste composition	Majority village natives
Profession	Profession overlap, in specialty jobs specifically. Large professions diversity	Professionally diverse block receives negative influence from a less professionally diverse block. Lack of profession overlap causes a negative influence.	High job diversity and higher-skilled jobs
Education	Large higher education level overlap	Higher educated block receives negative influence from less educated block	Higher education level
Language	Overlapping language	Lack of overlap in language causes negative influence	Language diversity
Age	None	Older average age block can receive negative influence from younger average age block	Younger average age

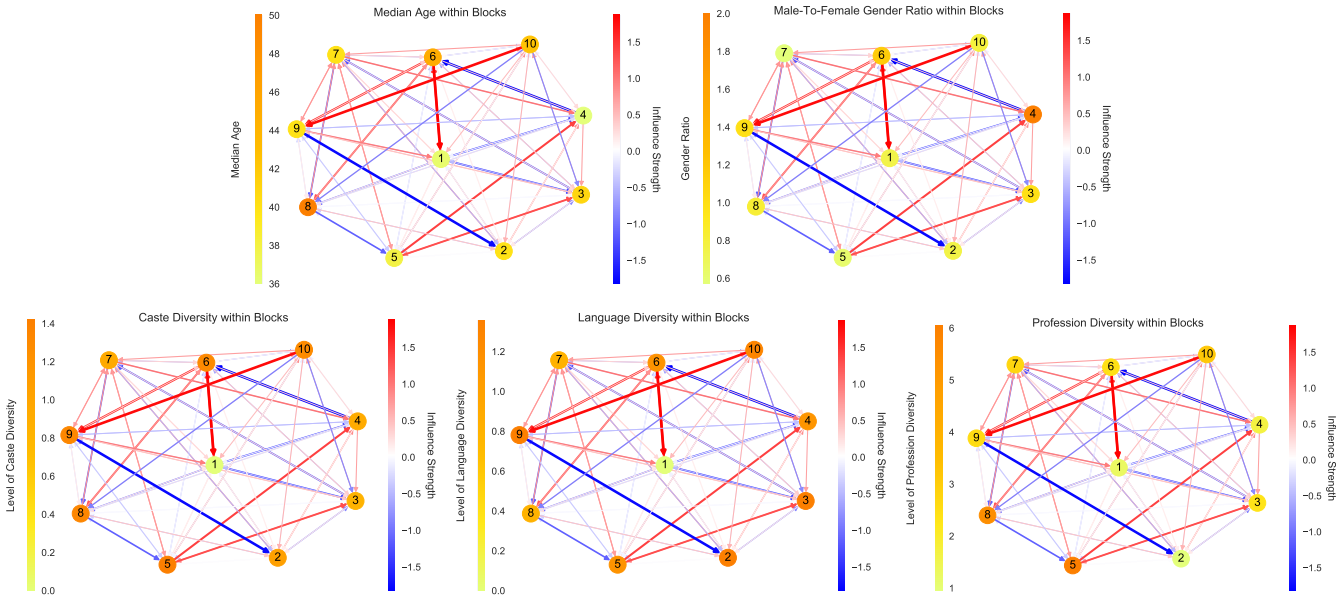


Figure 5: Sociodemographic analysis of each social block and the social influence across social blocks. Each node represents a social block corresponding to the index shown in the previous section. The directed links represent the strength of social influence varying from strong negative (light blue) to strong positive (light red). Colors of the node represent the sociodemographic characteristics within that social block. We display a subset of characteristics, including median age, gender ratio, caste diversity, language diversity, and profession diversity within each block.

in the following characteristics: gender distribution, majority castes, professions, and high profession diversity level, high education and high-skilled jobs, and mother tongue languages. Negative influence frequently occurs when two blocks have a lack of overlap in the following characteristics: gender distribution, caste composition, profession diversity level, education levels, and average age. Furthermore, the direction of negative influence is most frequently observed from a low-class block to a high-class block. Additionally, we frequently observe positive self-influence, which is from a block to itself, and this occurs when a block is younger, highly-educated, having high job diversity and holding higher-skilled jobs, having high language diversity, having a large gender imbalance, and having a large number of village natives.

These trends, when paired with block type characterizations, lead to interesting associations, such as block-to-block perceptions of lower or higher class with influence. Blocks of the higher class designation more frequently received negative influence from lower class blocks. Blocks of similar classes, especially higher classes, had more frequent positive influence. High-class blocks also had more frequent positive self-influence.

Conclusion

In this paper, we propose a generative Stochastic Block Influence Model which infers two types of hidden relationships: block-to-block interaction and block-to-block influence on decision making. Our model not only jointly infers these two types of social relationships, but also allows the inference of negative social influence, which is common in practical cases but has been ignored by the literature. In the adoption of microfinance example we present, the inferred block-to-block influence offers insights into how different social blocks exert influence on individuals' decision-making.

Extending to a diverse range of behaviors and studying if the influence matrix varies across different decision-making with our model is an interesting topic for future work. Moreover, the NUTS sampling used in this study is relatively inefficient. Developing a scalable inference method also points out a future direction.

References

Abbe, E. 2017. Community detection and stochastic block models: recent developments. *arXiv preprint arXiv:1703.10146*.

Airoldi, E. M.; Blei, D. M.; Fienberg, S. E.; and Xing, E. P. 2008. Mixed membership stochastic blockmodels. *Journal of Machine Learning Research* 9(Sep):1981–2014.

Angrist, J. D. 2014. The perils of peer effects. *Labour Economics* 30:98–108.

Aral, S.; Muchnik, L.; and Sundararajan, A. 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences* 106(51):21544–21549.

Backstrom, L.; Boldi, P.; Rosa, M.; Ugander, J.; and Vigna, S. 2012. Four degrees of separation. In *Proceedings of the 4th Annual ACM Web Science Conference*, 33–42. ACM.

Banerjee, A.; Chandrasekhar, A. G.; Duflo, E.; and Jackson, M. O. 2013. The diffusion of microfinance. *Science* 341(6144):1236498.

Centola, D., and Macy, M. 2007. Complex contagions and the weakness of long ties. *American journal of Sociology* 113(3):702–734.

Fowler, J. H., and Christakis, N. A. 2008. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the framingham heart study. *Bmj* 337:a2338.

Gomez-Rodriguez, M.; Leskovec, J.; and Krause, A. 2012. Inferring networks of diffusion and influence. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 5(4):21.

Gomez Rodriguez, M.; Leskovec, J.; and Schölkopf, B. 2013. Structure and dynamics of information pathways in online media. In *Proceedings of the sixth ACM international conference on Web search and data mining*, 23–32. ACM.

Hoffman, M. D., and Gelman, A. 2011. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *ArXiv e-prints*.

Hoffman, M. D., and Gelman, A. 2014. The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research* 15(1):1593–1623.

Kempe, D.; Kleinberg, J.; and Tardos, É. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 137–146. ACM.

Leng, Y.; Dong, X.; Moro, E.; et al. 2018. The rippling effect of social influence via phone communication network. In *Complex Spreading Phenomena in Social Systems*. Springer. 323–333.

Li, L.; Jamieson, K.; DeSalvo, G.; Rostamizadeh, A.; and Talwalkar, A. 2017. Hyperband: A novel bandit-based approach to hyperparameter optimization. *The Journal of Machine Learning Research* 18(1):6765–6816.

Myers, S. A.; Zhu, C.; and Leskovec, J. 2012. Information diffusion and external influence in networks. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 33–41. ACM.

Ng, A. Y.; Jordan, M. I.; and Weiss, Y. 2002. On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*, 849–856.

Pan, W.; Altshuler, Y.; and Pentland, A. 2012. Decoding social influence and the wisdom of the crowd in financial trading network. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*, 203–209. IEEE.

Salvatier, J.; Wiecki, T. V.; and Fonnesbeck, C. 2016. Probabilistic programming in python using pymc3. *PeerJ Computer Science* 2:e55.

Travers, J., and Milgram, S. 1967. The small world problem.

Yang, J., and Leskovec, J. 2010. Modeling information diffusion in implicit networks. In *Data Mining (ICDM), 2010 IEEE 10th International Conference on*, 599–608. IEEE.