A Network-Based Covariate Augmented Factorization Approach for Modeling Facebook Common Knowledge Experiments

Hao He[1], Xueying Liu[1], Neil Kattampallil[2], Vicki Lancaster[2], Gizem Korkmaz[3], Chris J.Kuhlman[4], and Xinwei Deng[1]

[1] Department of Statistics, Virginia Tech, Blacksburg, VA, USA
haoh@vt.edu, xliu96@vt.edu, xdeng@vt.edu
[2] Biocomplexity Institute, Universiy of Virginia, Arlington, VA, USA
neilalex525@gmail.com, val7zv@virginia.edu
[3] Statistics and Data Science, Westat, Rockville, MD, USA
GizemKorkmaz@westat.com
[4] Advanced Research Computing, Virginia Tech, Blacksburg, VA, USA
ckuhlman@vt.edu

Abstract. Common knowledge (CK) human subject experiments on a simulated environment of the Facebook social media platform are of great interest to understand how agents (players) reason and coordinate actions based on information collected over a network. However, outcomes from CK experiments often deviate from the outcomes of the theoretical CK model. It is important to model and analyze such deviations. Due to the prohibitive cost of conducting CK experiments for every player under all combinations of experimental factors, each player typically participates in only a subset of experiments. In this work, we propose a covariate-augmented factorization approach (CAFA) to model discrepancies between CK experimental data and theoretical CK predictions. The proposed CAFA considers latent factors to factorize the discrepancy matrix with respect to players and CK experiments, augmented by experimental covariates as additional regressors. The numerical analysis on real experimental data shows that the proposed CAFA not only achieves high prediction accuracy but also provides insight on important experimental factors affecting the discrepancy.

Keywords: Common knowledge · Human subjects games · Matrix completion · Logistic factorization · Group-LASSO

1 Introduction

Background and Theoretical Model. *Common knowledge (CK)* is a phenomenon that is generated among a collection of individuals where the members all know information about themselves and each other that enables them to reason about not only what they will do, but also about what all other people in the collection will do. Specifically, in a social network $G(V,E)$ with node set V and edge set E, CK considers information $\mathbb{I}$ such that each member u_i of a group $D \subseteq V$ knows $\mathbb{I}$, each u_i knows that every other member u_j $(i \neq j)$ knows $\mathbb{I}$, each u_i knows that every other member u_j $(i \neq j)$ knows

that u_i knows $\mathbb{I}$, and so on, ad infinitum [9]. In this setting, D is the collection of people that share CK. The CK phenomenon has been identified and observed by social scientists and psychologists (e.g., [2, 14]). Controlled laboratory experiments on human subjects in a Facebook-like [8] and other [14] settings have demonstrated that groups can produce CK.

There are many applications for CK. In socio-economic situations, applications of CK include forming teams in business and sports [3], recruiting individuals for groups [10], and establishing social norms [13]. It is of great interest to quantitatively model the phenomenon of CK for prediction and inference of future events. To our best knowledge, there are only two *quantitative* models of CK: a model of where people interact face-to-face [2] and a model of humans interacting on Facebook [9]. In this work, we focus on the model of CK that is constructed from human actions on Facebook. We denote this model as the *theoretical* CK on Facebook model, whose details are given in [9], and are too lengthy to provide here.

Common Knowledge-Based Collective Action Experiments. A collective action game was devised and implemented so that players could interact through web browsers [8]. The game mimics important features of the Facebook social media platform wherein players share information via their walls or timelines. Since Facebook users and their communications can be represented as a graph, nodes of a biclique in the Facebook network represent a group of users that share CK through the information on their walls [9]. Games were played with human subjects (agents) in order to understand whether they use CK to reason and coordinate actions. Collective action is represented by the Granovetter model [4], where each player is assigned a threshold θ. An ego player requires θ other members of the CK set to participate before—or at the *same* time as—the ego player activates/participates. (For definiteness, the context of the game is whether to participate in a protest event.) In [4], each player has sufficient information to decide whether to participate, but in this game, because of CK, each player has sufficient information to determine what *every* player in the CK set will decide because player thresholds are part of information $\mathbb{I}$ that is CK. This leads to an interesting situation where multiple players may decide to participate *simultaneously* because they know their thresholds *will be* met because they can *predict* what other players will decide. The game setups specifically include graphs that contain bicliques, often in the form of star (sub)graphs, so that CK is present. Given information about (local) graph structure and player thresholds, each player in each game decides whether to participate. These decisions can be compared to results from a theoretical model [9]. Game details are in [8].

Motivation and Contribution. Outcomes from CK experiments often deviate from outcomes of the theoretical CK model. Figure 1 illustrates a comparison between the predictions from the theoretical model and the outcomes from experimental data. For the theoretical model (left panel), the predictions of player decisions are deterministic and are solely dependent on the game setting. Real participants, however, display multiple sources of uncertainty in their decisions. First, for a given game setting, different players may behave differently (different risk tolerance). Second, even the same player may change their decision when encountering an identical setting in another game.

Therefore, it is beneficial to incorporate such uncertainty in modeling player behaviors through a probabilistic adjustment.

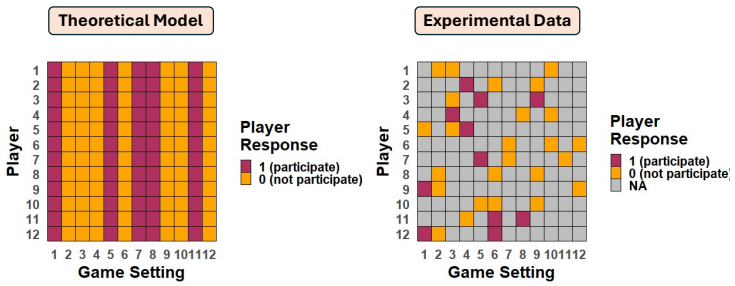

Fig. 1: Illustration of a comparison between the theoretical model and experimental data for the CK experiment. Only the first 12 of the 390 game players are shown on the y-axis and only 12 of the 450 games settings are shown on the x-axis. "NA" means that player did not participate in that game.

In practice, however, experiments are costly to conduct. As illustrated in the right panel of Figure 1, if each row represents a player, each column an experimental setting, and each entry a binary decision, the resulting matrix is sparse (most of the entries are "NA" because each player only plays a small subset of experimental settings). One of the challenges is to complete this sparse matrix so that we can bridge the gap between the theoretical predictions and the limited, heterogeneous experimental observations.

To address these challenges, this work focuses on modeling the discrepancy between the heterogeneous experimental observations and the deterministic outcomes from the theoretical model. Specifically, we make several main contributions as follows. **First**, we formulate the analysis of the CK experiment as the quantification of the discrepancy between experimental data and the theoretical model in Section 2. The theoretical model can provide deterministic predictions for each player under every experimental setting. The experimental data capture players' heterogeneity to some extent. The proposed framework considers the modeling of discrepancy as a calibration for the theoretical model to realize the merits of both theoretical rationale and empirical evidence. **Second**, to quantify the discrepancy under the limited experimental data, we formulate the learning problem as a sparse matrix completion problem, i.e., imputing the missing entries in a sparse matrix of discrepancy with respect to players and CK experiments. Thus, it can borrow strength from both between-player dependency and between-experiments dependency to enable accurate prediction of players' actions for unobserved experimental settings. **Third**, we propose a network-based covariate augmented factorization approach (CAFA) to impute the missing entries (i.e., player's action at unobserved experimental settings) with high accuracy via the accommodation of network-based covariates in the experimental setting. The proposed method builds on early works in logistic matrix factorization [6], extends that framework to exploit side information [11], and applies structured penalty (i.e., group LASSO) [12] to shrink whole blocks of coefficients. To our knowledge, no previous study has simultaneously combined structured penalty and logistic matrix factorization with side information.

2 Experiments and Problem Formulation

Online experiments were conducted within a game platform that was built to simulate decision-making within groups on Facebook. Constructs such as friends and the networks they induce, walls (timelines), and reading information on these walls were included. After reading walls and before making a decision, each player had the option of sending a message to any number of the players on shared walls as to whether she intended to participate or not. After all such messages were sent and received, players made their binary decisions whether to participate (state 1) or not (state 0). These are one-shot games and are described in more detail elsewhere [8]. Experimental results can then be compared to predictions from the theoretical model [9].

Each player decides whether to participate under different experimental conditions. Each such condition is a combination of the following six factors. The first factor is *Network structure (NS)* with three levels: **star**, **clique** and **circle**. The structures are shown in Figure 2. The second factor is *Communication type (CT)*. It has three levels: **none**, **bilateral** and **wall**. None means no messaging between players; bilateral means each participant may only message with its neighbors; wall means each participant may message with its neighbors and its neighbors' neighbors. One of two pre-determined messages may be sent: "I will participate" and "I will not participate." The third factor is *Network knowledge (NK)*. It has two levels: **global** and **local**. Global means all the identity and communication link information are given to all participants (e.g., who the participants are and who they are connected to); local means that participants only know the identity and communication link information of the neighbors that they can see on different walls. The fourth factor is *Player threshold (PTH)* with two levels **high (3)** and **low (1)**. A player's threshold θ represents the integer number of other people that need to participate in order for this player to transition from state 0 to 1. High means $\theta = 3$; low means $\theta = 1$. The fifth factor is *Number of high thresholds (NTH). It has 5 levels*: 0, 2, 3, 4, 5, representing the number of players with **high** thresholds in the group of one game. The sixth factor is *Player positions* with five levels: 1, 2, 3, 4, 5. A value of $i \in \{1,\ldots,5\}$ means the player is at position u_i in Figure 2.

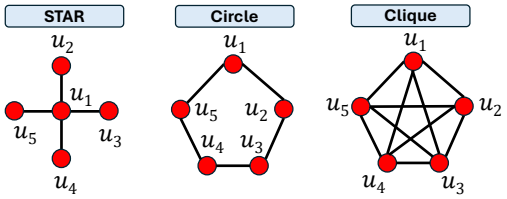

Fig. 2: Network structure (NS) of star (left), circle (middle), and clique (right).

The total number of combinations of the factors NK, CT, NS, NTH, and POS is $(2 \times 3 \times 3 \times 5 \times 5 = 450)$, since PTH is determined by NTH and POS. In this work, we are specifically focusing on the *theoretical* CK on Facebook model, which means CT = *wall* (because the shared information $\mathbb{I}$ arises from Facebook walls) and NK = *local*,

leaving us ($3 \times 5 \times 5 = 75$) possible experimental conditions. The experimental data studied in this work includes 675 observations played by 45 players. Each player plays 15 out of 75 experimental conditions (the data is sparse with 20% of entries observed). The matrix completion problem is then modeled as follows. Let $i \in P$ be a player, and let $j \in C$ be a game (experimental) setting (note that here each j corresponds to an assignment of values to experimental factors listed above), where P is the set of all players and C is the set of all game settings. We model:

$$y_{e_j}^{(i)} \;=\; y_{t_j} \;+\; \delta_j^{(i)} \pmod 2, \tag{1}$$

where $y_{e_j}^{(i)} \in \{0,1\}$ is the binary decision of player i under the game setting of j, $y_{t_j} \in \{0,1\}$ represents the binary prediction from the theoretical CK model of the player decision, and $\delta_j^{(i)} \in \{0,1\}$ is the binary adjustment capturing the discrepancy between theoretical model and the experiment for player i in game setting j. Note that, the $\delta_j^{(i)}$ accommodates the effects of messaging from the experiments, which the theoretical model does not address. The focus of this work (i.e., the target value in the matrix completion problem), is to model $\delta_j^{(i)}$.

3 Methodology

Covariate-Augmented Factorization (CAFA) with Group-LASSO Penalty. The binary data completion problem is modeled using the following covariate-augmented logistic factorization. Let N be the number of players (rows) and M be the number of experimental conditions (columns). Each condition $j \in \{1,\ldots,M\}$ corresponds to a covariate vector $\mathbf{x}_j \in \mathbb{R}^p$, where p refers to the number of the one-hot encoded experimental factors (i.e., dummy variables). That is, for each categorical factor, we create a one-hot design matrix: every level is represented by a binary indicator column that takes the value 1 if the observation belongs to that level and 0 otherwise. Each $\mathbf{x}_j$ represents a combination of the features. We denote the number of covariate groups (subsets of collections of one-hot encoded factors) by S. Each subset $s \in \{1,\ldots,S\}$ corresponds to all features (dummy variables) associated with one experimental factor.

We introduce two latent matrices for players and experimental conditions: $U = \big(\mathbf{u}_1,\ldots,\mathbf{u}_N\big)^\top \in \mathbb{R}^{N\times K}$; $V = \big(\mathbf{v}_1,\ldots,\mathbf{v}_M\big)^\top \in \mathbb{R}^{M\times K}$. We also introduce a coefficient matrix: $B = (\beta_1,\ldots,\beta_p)^T = \big(B^{(1)},\ldots,B^{(S)}\big)^T \in \mathbb{R}^{p\times K}$, which maps the covariate vector $\mathbf{x}_j$ into a K-dimensional vector $B^T\mathbf{x}_j$ (K is the dimension of the latent space). Here, for a subset $s \in \{1,\ldots,S\}$, the sub-matrix $B^{(s)}$ of B corresponds to the coefficients of the dummy variables. Then we define $\mathbf{w}_j \;=\; \mathbf{v}_j + B^T\mathbf{x}_j \in \mathbb{R}^K$. We model the probability of $\delta_j^{(i)} = 1$ (i.e., there is a difference between the model's prediction and the experimental observation for player i under condition j) via the logistic function:

$$P\big(\delta_j^{(i)} = 1 \mid \mathbf{u}_i, \mathbf{w}_j\big) \;=\; g\big(\mathbf{u}_i^\top \mathbf{w}_j\big) \;=\; \frac{1}{1+\exp\big(-\mathbf{u}_i^\top \mathbf{w}_j\big)}\,. \tag{2}$$

Here, $\mathbf{u}_i^\top \mathbf{w}_j = \mathbf{u}_i^\top\big(\mathbf{v}_j + B^T\mathbf{x}_j\big)$. To enable structured variable selection on the covariates, we impose a group-penalty on B. For each sub-matrix $B^{(s)}$, let $\|B^{(s)}\|_2$ be the

Euclidean norm. Then the penalty is $\mathscr{P}(B) = \lambda \sum_{s=1}^{S} \left\| B^{(s)} \right\|_2$, where $\lambda > 0$. To address numerical stability and to prevent overfitting—especially given the sparsity in our observations—small ℓ_2 penalties are added on U and V [7].

Let $\Omega \subseteq \{1,\ldots,N\} \times \{1,\ldots,M\}$ be the set of observed (i,j) pairs, and let $\delta_j^{(i)} \in \{0,1\}$ be the corresponding binary outcomes. The final penalized objective function is:

$$\min_{U,V,B} f(U,V,B) = -\ell(U,V,B) + \lambda \sum_{s=1}^{S} \|B^{(s)}\|_2 + \frac{\rho_1}{2}\|U\|_F^2 + \frac{\rho_2}{2}\|V\|_F^2 , \qquad (3)$$

with $\rho_1, \rho_2 > 0$ chosen to be small constants. The ridge terms $\frac{\rho_1}{2}\|U\|_F^2$ and $\frac{\rho_2}{2}\|V\|_F^2$ ($\|A\|_F^2 = \sum_{i,j} a_{ij}^2$ is the squared Frobenius norm of a matrix) are introduced solely to stabilize the scale of the latent factors and to make each block update strictly convex; any choice $\rho_1, \rho_2 > 0$ fulfils these purposes (see, e.g., [7]). Accordingly we set $\rho_1 = \rho_2 = 0.01$, following the common "one-percent" ridge used in matrix-factorization practice. By minimizing f, we determine the parameters in U, V, and B, and use the estimates to recover the predicted probability of $P(\delta_j^{(i)} = 1)$ in Equation (2), and predict $\delta_j^{(i)}$ in Equation (1): $\delta_j^{(i)} = 1$ if $P(\delta_j^{(i)} = 1) > 0.5$, and $\delta_j^{(i)} = 0$ otherwise.

Parameter Estimation. We adopt a block coordinate descent approach [12] by updating U, V, and B iteratively. The update for each of U and V is a standard logistic regression problem with quadratic regularizer. Such problems are efficiently solved using Newton's method, as detailed in [5]. With U and V fixed, the term $\mathbf{u}_i^\top \mathbf{v}_j$ may be omitted in the objective function. Then consider the subproblem for B:

$$\min_B - \sum_{(i,j)\in\Omega} \left[\delta_j^{(i)} \log g\left(\mathbf{u}_i^\top (B^T \mathbf{x}_j)\right) + (1-\delta_j^{(i)}) \log\left(1 - g\left(\mathbf{u}_i^\top (B^T \mathbf{x}_j)\right)\right) \right] + \lambda \sum_{s=1}^{S} \|B^{(s)}\|_2.$$

This is a group-penalized logistic regression problem, which can be efficiently solved via proximal gradient methods [15]. In each update, compute the gradient of the smooth part and then apply the group soft-thresholding operator on $B^{(s)}$ [15]. Each iteration guarantees a decrease in the objective function $f(U,V,B)$ in Equation (3). Convergence to a stationary point follows from standard results in block coordinate descent with proximal updates, provided the appropriate smoothness and boundedness conditions are met [12].

4 Results of Case Study

In this section, we present a cross-validated comparison between the proposed method and two benchmark methods: Single Value Decomposition (SVD) [7] and Logistic Matrix Factorization (LMF) without covariate information [6] (i.e., using only two latent matrices U and V to model the probability $P(\delta_j^{(i)} = 1 | u_i, v_j)$ in Equation (2), without X and B, via w). Since SVD is not directly applicable to binary data, the data is converted from $\{0,1\}$ to $\{-1,1\}$, and 0 is considered as the threshold for classification. Then, denote the filled (predicted) value by SVD as $\hat{m}_{ij}$, it is linearly rescaled to $0 \le \hat{p}_{ij} \le 1$:

$\hat{p}_{ij} = 0.5 + \frac{1}{2}\frac{\hat{m}_{ij}}{m_{\max}}$ if $\hat{m}_{ij} > 0$; $\hat{p}_{ij} = 0.5 - \frac{1}{2}\frac{\hat{m}_{ij}}{m_{\min}}$ if $\hat{m}_{ij} \leq 0$, where $m_{\max} = \max_{(i,j)} \hat{m}_{ij}$, and $m_{\min} = \min_{(i,j)} \hat{m}_{ij}$. In this way, the three models can be compared using metrics like the area under curve (AUC) for the ROC curve [1]. For each model, 20 iterations of random player-wise training-testing splits are conducted to measure the variance. The choice of $K = 3$ is determined based on cross-validation.

The left panel of Figure 3 displays the mean AUC value over 20 iterations for each model. In all training-testing split scenarios, the proposed method outperforms the other two benchmark methods. The two-sample t-tests for mean AUC value between CAFA and LMF yield p-values of $0.070, 0.001, 0, 0$ for training proportions of 20%, 40%, 60%, 80%, respectively. Here we only display the comparison between CAFA and LMF because the differences between CAFA and SVD are obvious for training proportions larger than 20%. The results suggest statistically significant differences in the performance measures between CAFA and the other two models for almost all training-testing split scenarios (the only exception is when training data is extremely limited (20%)). The rationale behind the results is that, the XB component in the proposed method helps better capture the information of the covariates that constitute each experiment setting combination. It is noteworthy that when using 80% of the available data (about only $80\% \times 20\% = 16\%$ of all entries of the matrix), the mean AUC is 0.895, indicating that the proposed method is capable of accurately completing the sparse matrix.

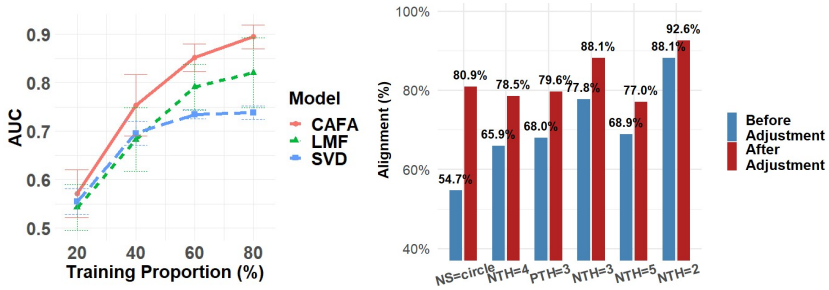

Fig. 3: The left panel shows the mean AUC for each model under four scenarios of training-testing split. The error bars calculated over 20 iterations are included to quantify the variance. The right panel displays a comparison of alignment percentages before and after including $\delta_j^{(i)}$ in Equation (1) for selected factor levels using CAFA. From left to right, the change is in descending order.

The right panel of Figure 3 shows the alignment percentage plots constructed based on 5-fold cross-validation (i.e., partition data into 5 pieces, and treat each piece as testing data). It indicates that $\delta_j^{(i)}$ effectively improves the theoretical model to capture players' behavior in practice, especially at factor levels where the theoretical model originally struggled. For example, the left three pairs of bars (NS = *circle*, NTH = 4 and PTH = 3), the alignment percentages increase from 54.7% to 80.9%, from 65.9% to 78.5%, and from 68% to 79.6% (percentage increases are, respectively, 48%, 19%, and 17%). Other factor levels are handpicked to show moderate increase in the alignment

percentages after adjustment. The only cases where the alignment percentages decrease are NTH = 0 (from 99.3% to 97%), NS = *clique* (from 98.7% to 94.2%), and NS = *star* (from 86.7% to 84.9%), and these are the cases where theoretical models originally performed well.

5 Summary

In this work, we take advantage of both the strength of the underlying game-theoretic model and the rich patterns of the experimental data to more effectively characterize human behavior in common knowledge settings. Specifically, we propose a Covariate-Augmented Factorization approach (CAFA) to perform logistic matrix completion with covariate information on sparse data. Our findings reveal that the CAFA is capable of imputing the sparse matrix with high accuracy, and it can automatically identify important experimental factors for quantifying the discrepancy.

References

1. Bradley, A.P.: The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognition **30**(7), 1145–1159 (1997). https://doi.org/10.1016/S0031-3203(96)00142-2
2. Chwe, M.S.Y.: Structure and strategy in collective action. Am. J. of Soc. **105**, 128–156 (1999)
3. DeChurch, L.A., Mesmer-Magnus, J.R.: The cognitive underpinnings of effective teamwork: A meta-analysis. Journal of Applied Psychology **95**(1), 32–53 (2010)
4. Granovetter, M.: Threshold models of collective behavior. American Journal of Sociology **83**(6), 1420–1443 (1978)
5. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2nd edn. (2009)
6. Johnson, C.: Logistic matrix factorization for implicit feedback data. In: Advances in Neural Information Processing Systems. pp. 1–9 (2014)
7. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. Computer **42**(8), 30–37 (2009)
8. Korkmaz, G., Capra, M., Kraig, A., et al.: Coordination and common knowledge on communication networks. In: AAMAS (2018)
9. Korkmaz, G., Kuhlman, C.J., Marathe, A., et al.: Collective action through common knowledge using a Facebook model. In: AAMAS (2014)
10. Kozlowski, S.W., Ilgen, D.R.: Enhancing the effectiveness of work groups and teams. Psychological Science in the Public Interest **7**(3), 77–124 (2006)
11. Lee, A., Kim, B.: Online matrix completion with side information. In: Proceedings of the International Conference on Machine Learning. pp. 234–243 (2019)
12. Meier, L., van de Geer, S., Bühlmann, P.: The group lasso for logistic regression. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **70**(1), 53–71 (2008)
13. Roberts, D.: The Pueblo Revolt: The Secret Rebellion That Drove the Spaniards Out of the Southwest. Simon & Schuster (2004)
14. Thomas, K.A., DeScioli, P., et al.: The psychology of coordination and common knowledge. J. of Personality and Social Psychology **107**, 657–676 (2014)
15. Yuan, M., Lin, Y.: Model selection and estimation in regression with grouped variables. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **68**(1), 49–67 (2006)