Bilinear Mixed-Effects Models for Affiliation Networks

Yanan Jia\textsuperscript{*} \hspace{1cm} Catherine A. Calder\textsuperscript{*}†

June 24, 2014

Abstract

An affiliation network is a particular kind of two-mode social network that consists of a set of ‘actors’ and a set of ‘events’ where ties indicate an actor’s participation in an event. While event affiliations are fundamental in defining the social identity of individuals, statistical methods for studying affiliation networks are less well developed than methods for studying one-mode, or actor-actor, networks. One way to analyze affiliation networks is to consider one-mode network matrices that are derived from an affiliation network, but this approach may lead to the loss of important structural features of the data. The most comprehensive approach is to study both actors and events simultaneously. In this paper, we extend the bilinear mixed-effects model developed for one-mode networks to affiliation networks by considering dependence patterns in the interactions between actors and events. We describe a Markov chain Monte Carlo algorithm for Bayesian inference and illustrate the proposed methodology through an analysis of extracurricular activity participation data.

**Keywords:** Bayesian modeling, generalized linear model, social networks, Markov chain Monte Carlo (MCMC)

1 Introduction

In typical statistical analyses, the primary goal is learning about properties of individual units. When the property of interest involves interactions between multiple units rather than

\textsuperscript{*}Department of Statistics, The Ohio State University, Columbus, OH, USA

\textsuperscript{†}Email: calder@stat.osu.edu
properties of the individual units themselves, the units can be considered a network. Network data are widely used to represent relational information among interacting units. Units are referred to as nodes in a network, and relationships between the nodes are represented by ties/edges. Pairs of nodes, that may be either linked or not, are called dyads in a network. Networks can be represented by matrices and mathematical graphs.

We use the term mode to differentiate sets of distinct nodes in a network. The most common type of network is a one-mode network in which all units are of the same type. A typical example is a friendship network where all nodes are individuals or actors and ties between all actors are well defined. The ties in a one-mode network may be directed or undirected. For directed ties, we can denote one-mode networks with $n$ nodes as an $(n, n)$ adjacency matrix $Y$. In most cases, the entries of $Y$ are binary, where $y_{ij} = 1$ if there is a tie from node $i$ to node $j$, and $y_{ij} = 0$ otherwise. It is also possible to have (non-negative) integer-valued network datasets where the entries $y_{ij}$ are the number of ties from node $i$ to node $j$. In the case of undirected ties, the adjacency matrix will be symmetric so that $y_{ij} = y_{ji}$.

Two-mode networks contain relational information about two distinct sets of entities, specifically about ties between nodes of different modes. Two-mode networks can capture more relational structure than the standard one-mode representation of such data and are a natural representations for relational data involving affiliations between sets of entities. The term “affiliation” usually refers to membership or participation data. Arguably, the most well known affiliation dataset is the “Southern Women” network collected by Davis and Gardner (1941), which consists of attendance records at various social events in a small southern town. This dataset is an affiliation network since the ties represent affiliations between a set of actors (women), denoted by $A$, with a set of events (social events), denoted by $E$. Affiliation networks, such as the Southern Women network, allow us to study the dual perspectives of actors and events where connections among members of one of the modes are based on linkages established through the second mode (i.e., women are connected because they attend the same social events and social events are connected through the women that participate in them).
An affiliation network can be denoted by an $n^a \times n^e$ affiliation matrix $Y = \{y_{ik}\}$, which records the affiliation of each actor with each event, where rows index actors and columns index events, and $n^a$ and $n^e$ are the total number of actors and events, respectively. The entries of this matrix, $y_{ik}$, can be binary variables or non-negative integer-valued variables. If actor $i$ is affiliated with event $k$, then $y_{ik} \geq 1$ and $y_{ik} = 0$ otherwise, where $i = 1, 2, \ldots, n^a$ and $k = 1, 2, \ldots, n^e$. Each row of $Y$ describes an actor's affiliation with the events. Similarly, each column of $Y$ describes the membership of an event. An affiliation network can also be represented by a bipartite graph, or a graph in which the nodes can be partitioned into two subsets corresponding to the distinct modes, and all lines are between pairs of nodes belonging to the different modes. For affiliation networks, since actors are affiliated with events, and events have actors as members, all lines in the bipartite graph are between nodes representing actors and nodes representing events.

Statistical methods for one-mode networks are fairly well developed. The exponential random graph model (ERGM) is one of the most popular methods for analyzing networks (Frank and Strauss, 1986; Wasserman and Pattison, 1996; Pattison and Wasserman, 1999; Robins et al., 1999). Although ERGMs are useful for modeling global network characteristics, they are known to possess some undesirable properties. Robins et al. (1999) and Handcock et al. (2003) discussed these challenges associated with ERGMs: the intractability of the normalizing constant in the likelihood function of ERGMs, model degeneracy, and no natural mechanism for including covariate information in the model. For these reasons, alternative models built on latent variables have attracted considerable attention recently. These models include mixed-effects models (van Duijn et al., 2009; Zijlstra et al., 2009; Hoff, 2003, 2005, 2009), the stochastic blockmodel (Wang and Wong, 1987; Snijders and Nowicki, 1997; Snijders, 1997), and latent space models (Hoff et al., 2002). All of these latent variable models assume conditional independence of the probability of ties between dyads. That is, the elements of $Y$ are independent conditional on latent variables $Z$: $P_{\theta}(Y = y | Z = z) = \prod_{i,j} p_{\theta}(Y_{ij} = y_{ij} | Z = z)$. Conditional independence does not imply that latent variable models cannot capture network dependencies of interest. Indeed, some of the more sophisticated latent variable models make
clever use of latent structures to capture types of dependence. The conditional independence of edges implies that model degeneracy is not an issue, which facilitates the construction of models. In addition, the conditional independence of tie probabilities leads to computational advantages in model fitting (Hunter et al., 2012). A latent variable model that we build on in this paper is the bilinear mixed-effects model proposed by Hoff (2005), which is an extension of latent space models for one-mode networks. This model uses interacting latent variables to capture certain types of higher-order dependence patterns often present in social networks.

While there is a rich literature of statistical methods for one-mode networks, methods for two-mode networks are limited. One approach known as the “conversion’,’ or projection method (Newman, 2001), relies on the two one-mode networks that can be derived from an affiliation network: $Y Y'$ is the one-mode network for actors and $Y' Y$ is the one-mode network for events. However, information is lost by converting an affiliation network into two one-mode networks. For instance, if we use binary matrices to represent the one-mode networks, then we lose information about both the number and the properties of the shared partners of the other set. Alternatively, we can build models for two-mode networks to analyze both actors and events simultaneously. Wang (2009) extended ERGMs to the two-mode situation. However, these models suffer from the limitation of the one-mode ERGMs described above.

In this paper, we build on ideas from Hoff (2005) and extend the bilinear mixed-effects models to the two-mode settings. Our model differs from the one-mode bilinear model in that the bilinear effect for a pair $(i, k)$ in our model is the inner product of unobserved characteristic vectors specific to actor $i$ and event $k$, respectively. These bilinear effects can capture certain forms of dependence in a two-mode social network as we discuss below. Then we develop a Markov chain Monte Carlo (MCMC) algorithm providing Bayesian inference.

This paper is organized as follows. In next section, we discuss basic models for affiliation network data. Section 3 introduces types of dependence often seen in two-mode network datasets and argues that the basic models are not sufficiently able to capture these dependencies. In Section 4 we add a bilinear effect to the basic models we discussed previously.
Section 5 describes parameter estimation through an MCMC algorithm. An application of our proposed methodology to student extracurricular activity participation is presented in Section 6. We conclude in Section 7 with a discussion of some directions for future research.

2 Basic Models for Affiliation Network Data

The data we model in this article consist of an $n^a \times n^e$ sociomatrix $Y$, with entries $y_{ik}$ denoting the value of the relation between actor $i$ and event $k$ and additional covariate information associated with actors, events, and dyads.

2.1 Fixed Effects Model

Since most affiliation network data, $y_{ik}$, are binary or (non-negative) integer valued, we specify models using the standard generalized linear model framework. We let

$$Pr(Y = y|\beta) = \prod_{i=1}^{n^a} \prod_{k=1}^{n^e} Pr(Y_{ik} = y_{ik}|\beta),$$

where each component of $Y$ follows an exponential family distribution. We relate $\mu_{ik} \equiv E(Y_{ik}|\beta)$ to a set of covariate variables $x_{ik}$ via a link function denoted by $g(\cdot)$:

$$\theta_{ik} = g(\mu_{ik}) = \beta' x_{ik},$$

where $\beta$ is a $r$-dimensional vector of unknown regression coefficients. We decompose $x_{ik}$ into $x_{ik} = (x_{ik}^d, x_{i}^a, x_{k}^e)$, where $x_{ik}^d$ is an $r^d$-dimensional covariate vector associated with (actor $i$, event $k$) dyad, $x_{i}^a$ is an $r^a$-dimensional covariate vector associated with actor $i$, $x_{k}^e$ is an $r^e$-dimensional covariate vector associated with event $k$, implying $r^d + r^a + r^e = r$. The model can then be rewritten as

$$\theta_{ik} = g(\mu_{ik}) = \beta_d' x_{ik}^d + \beta_a' x_{i}^a + \beta_e' x_{k}^e.$$  (1)
where $\beta_d, \beta_a,$ and $\beta_e$ are vectors of unknown regression coefficients with dimension $r_d, r_a$ and $r_e$, respectively. The affiliation network data are measured on a set of actors and a set of events. Since actors and events comprise multiple dyads, the observations $y_{ik}$'s are likely not (conditionally) independent given the regression coefficients, and we need a model which can capture dependence induced by the shared actors and events making up the dyads.

### 2.2 Mixed Effects Models

Since subsets of observations are taken on the same actors or events, we consider mixed models including actor and event random effects of the form

$$
\theta_{ik} = g(\mu_{ik}) = \beta_d'x_{ik}^d + \beta_a'x_{i}^a + \beta_e'e_{k}^e + a_i + e_k,
$$

where $\mu_{ik} \equiv E(Y_{ik}|\theta_{ik})$, and $a_i$ and $e_k$ represent the actor and event random effects, respectively. For discrete data subject to overdispersion (Poisson, binomial), an observation level residual is also present, so that,

$$
\theta_{ik} = g(\mu_{ik}) = \beta_d'x_{ik}^d + \beta_a'x_{i}^a + \beta_e'e_{k}^e + a_i + e_k + \gamma_{ik},
$$

with $\gamma_{ik}$ usually taken as i.i.d. errors \cite{Congdon2010}. We can interpret the $\gamma_{ik}$s as dyad random effects.

The observations $\{Y_{ik} : i = 1, \ldots, n^a, k = 1, \ldots, n^e\}$ are modeled as conditionally independent given the random effects, denoted by $a = (a_1, \ldots, a_{n^a})'$, $e = (e_1, \ldots, e_{n^e})'$, and $\gamma = \text{vec}(\Gamma)$ for the $n^a \times n^e$ matrix $\Gamma$ with elements $\gamma_{ik}$ for $i = 1, \ldots, n^a$ and $k = 1, \ldots, n^e$. That is,

$$
Pr(Y = y|\beta, a, e, \gamma) = \prod_{i=1}^{n^a} \prod_{k=1}^{n^e} Pr(Y_{ik} = y_{ik}|\beta, a_i, e_k, \gamma_{ik}).
$$

We take the different types of random effects to be mutually independent and Gaussian with mean zero and variances $\sigma_a^2, \sigma_e^2$ and $\sigma_\gamma^2$, respectively:

$$
a | \sigma_a^2 \sim \text{MVN}(0, \sigma_a^2 I_{n^a \times n^a}),
$$
\[
\epsilon \mid \sigma_e^2 \sim \text{MVN}(0, \sigma_e^2 I_{n \times n^e}),
\]

and

\[
\gamma \mid \sigma_\gamma^2 \sim \text{MVN}(0, \sigma_\gamma^2 I_{n^\gamma \times n^\gamma}),
\]

where \( I_{n \times n} \) denotes the \( n \)-dimensional identity matrix and \( n^\gamma = n^a \times n^e \). For the model given by (2), we can treat \( \sigma_\gamma^2 \) as a constant equal to zero in the model given by (3). Therefore, we refer to the model given by (3) as the generalized linear mixed effects model for affiliation networks.

Letting \( \epsilon_{ik} \) denote the \((i, k)\) dyad random effect (i.e., \( \epsilon_{ik} = a_i + e_k + \gamma_{ik} \)) and marginalizing over the \( a_i \)'s, \( e_k \)'s, and \( \gamma_{ik} \)'s, it follows that

\[
\begin{align*}
E(\epsilon_{ik}^2) &= \sigma_a^2 + \sigma_e^2 + \sigma_\gamma^2 \\
E(\epsilon_{ik}\epsilon_{jk}) &= \sigma_e^2 \\
E(\epsilon_{ik}\epsilon_{il}) &= 0 \\
E(\epsilon_{ik}\epsilon_{il}\epsilon_{jl}) &= \sigma_a^4 + \sigma_e^4
\end{align*}
\]

where \( \sigma_a^2 \) and \( \sigma_e^2 \) capture the components of the total variation in the \( \epsilon_{ik} \)'s explained by dyads containing the same actor or event, respectively. As we will discuss in the next section, the restriction this model places on the fourth-order (i.e., \( E(\epsilon_{ik}\epsilon_{il}\epsilon_{jk}\epsilon_{jl}) \)) moments is limiting and makes this mixed model unable to capture the types of dependences frequently encountered in real affiliation networks.

### 3 Dependence Patterns in Affiliation Networks

Network data differ from other types of dependent data in that ties often tend to be transitive, balanced, and clusterable (Wasserman and Faust, 1994). In one-mode friendship networks, we often see patterns that indicate “a friend of a friend is a friend,” a statement that translates to properties of sets of three dyads (triangles). In particular, this pattern is called transitivity. Balance is a generalized version of transitivity defined for signed relationship of the type \( y_{ij} \) is positive if there is a tie between nodes \( i, j \) and negative otherwise. A triad
\(i, j, k\) is said to be balanced if \(y_{ij} \times y_{jk} \times y_{ki} > 0\). Clusterability is a generalization of the concept of balance. A triad is clusterable if it is either balanced or the pairwise relationships within the triad are all negative.

Here we extend these definitions to the two-mode setting. We say a set of one pair of actors \(i, j\) and one pair of events \(k, l\) form a cycle, and offer the following definitions.

**Definition 3.1** For affiliation relations, a cycle \(i, j\) and \(k, l\) is transitive if \(y_{ik} > 0\), \(y_{il} > 0\), and \(y_{jk} > 0\), then \(y_{jl} > 0\).

Transitivity implies that if actors \(i\) and \(j\) attend event \(k\) together and actor \(i\) attends another event \(l\), then we expect actor \(j\) also to attend event \(l\).

**Definition 3.2** For general affiliation signed relations, a cycle \(i, j\) and \(k, l\) is said to be balanced if \(y_{ik} \times y_{il} \times y_{jk} \times y_{jl} > 0\).

Since the number of edges between a cycle is an even number, we note that balance and clusterability are identical conceptually in the two-mode setting.

For general signed relations among units, many theories of social systems suggest that the relationships within a cycle tend to be balanced. For example, if \(y_{ik} = 1\) and \(y_{jk} = 1\), which means the relationships between actor \(i\) and event \(k\) and between actor \(j\) and event \(k\) are positive, then it is more likely that either both \(y_{il} = 1\) and \(y_{jl} = 1\) or both \(y_{il} = 0\) and \(y_{jl} = 0\).

In other words, if actors \(i\) and \(j\) both participate in event \(k\), then they are likely to either both participate in event \(l\) or both not participate in event \(l\). In real affiliation networks, we expect to see more evidence of balance than we would expect if the presence of ties is completely random. This translates into the presence of particular balanced patterns among cycles, which are illustrated in Figure 1. As discussed by Wang (2009), these configurations, 0-path \((L_0)\), 2-path \((L_2)\), actor 2-stars \((S_{A2})\), event 2-stars \((S_{E2})\), and four-cycles \((C_4)\), shown in Figure 1 are the balanced cycles often seen in affiliation networks. Four-cycles \((C_4)\) are the simplest local closures representing the strength of ties when two-mode networks are
Figure 1: Balanced Cycles. A cycle is defined by \(\{i, j\} \in A\) and \(\{j, k\} \in E\). Solid lines connecting actors and events denote ties, and dashed lines denote the absence of ties.

projected onto the two one-mode networks. The number of \(C_4\) cycles including actors \(i\) and \(j\) is the number of events the two actors share and thus captures the presence of transitivity in affiliation networks.

Returning to the generalized linear mixed-effects model discussion in Section 2.2, we consider the Pearson residuals,

\[
\hat{\xi}_{ik} = \frac{y_{ik} - \hat{\mu}_{ik}}{SE(\hat{\mu}_{ik})},
\]

from fitting the generalized linear mixed model to an observed affiliation network. We say the residuals \(\{\hat{\xi}_{ik}, \hat{\xi}_{il}, \hat{\xi}_{jk}, \hat{\xi}_{jl}\}\) form a cycle, and let \(\hat{S}_4\) denote the set of all residual cycles. In addition, let \(\hat{C}_4\) be the subset of the \(\binom{n_a}{2}\binom{n_e}{2}\) elements of \(\hat{S}_4\) such that \(\hat{\xi}_{ik} \times \hat{\xi}_{il} \times \hat{\xi}_{jk} \times \hat{\xi}_{jl} > 0\). Then the proportion of observed positive residual cycles,

\[
\hat{\pi} \equiv \frac{|\hat{C}_4|}{|\hat{S}_4|},
\]

where \(|\cdot|\) denotes cardinality, should be equal to the expected proportion of positive residual cycles, \(\pi_0\), if the ties are independent. By considering the balanced cycles shown in Figure 1 it can be shown that \(\pi_0 = p^4 + (1 - p)^4 + 6p^2(1 - p)^2\), where \(p\) is the proportion of positive residuals. If \(\hat{\pi}\) is significantly larger than \(\pi_0\), then we have more balanced residual cycles than we would expect under independence, indicating lack of fit of the generalized linear mixed model. Section 6 illustrates the use of this model diagnostic on a real affiliation dataset.
4  Bilinear Mixed-Effects Models

In order to allow for more transitivity and balance beyond generalized linear mixed effects model for affiliation networks, we add a bilinear random effect to the model given by (3). As with Hoff (2005)’s bilinear mixed effects model for one-mode network data, this addition enables us to capture the expected balance tendencies in two-mode network relations.

Unlike the one-mode bilinear model where the bilinear term is comprised of the inner product of a single random vector, our bilinear term is the inner product \( u_i'v_k \), where \( u_i \) and \( v_k \) are two independent \( t \times 1 \) random vectors that corresponds to unobserved characteristics specific to actor \( i \) and event \( k \), respectively. If we add this inner product of latent variables \( u_i \) and \( v_k \) to (3), we have

\[
\epsilon_{ik} = a_i + e_k + \gamma_{ik} + \varepsilon_{ik},
\]

where the random effects \( a_i, e_k \) and \( \gamma_{ik} \) are still taken to be multivariate normal with means and covariances are as given in Section 2.2. The set of bilinear term \( \{\varepsilon_{ik} = u_i'v_k, i = 1, \ldots, n^a, k = 1, \ldots, n^e\} \) allows us to capture balance. To see this, consider the case in which \( t = 1 \), where \( t \) is the dimension of the \( u_i \) and \( v_k \) vectors. In this case, the \( \varepsilon_{ik} \)’s correspond to the residuals from the version of the model without the bilinear term. Since \( \varepsilon_{ik} \times \varepsilon_{il} \times \varepsilon_{jk} \times \varepsilon_{jl} = (u_i u_j v_k v_l)^2 \geq 0 \), the bilinear term can be seen to capture positive residual cycles. Of course in a real dataset, we do not expect networks to be completely balanced. By taking \( t > 1 \), the bilinear term captures the balance tendencies of real networks without forcing every residual cycle to be positive.

We assume the \( u_i \)’s and \( v_k \)’s are mutually independent and follow \( t \)-dimensional multivariate normal distributions so that

\[
\begin{align*}
\mathbf{u}_i | \Sigma_u & \sim \text{MVN}(0, \Sigma_u) \\
\mathbf{v}_k | \Sigma_v & \sim \text{MVN}(0, \Sigma_v).
\end{align*}
\]

In addition, we assume \( u_i \perp u_j \) for \( \{i, j = 1, \ldots, n^a : i \neq j\} \) and \( v_k \perp v_l \) for \( \{k, l = 1, \ldots, n^e : k \neq l\} \). It follows that \( \varepsilon_{ik} \)’s have moments

\[
E(\varepsilon_{ik}) = 0
\]
\[ E(\varepsilon_{ik}^2) = \text{trace}(\Sigma_u \Sigma_v) \]
\[ E(\varepsilon_{ik} \varepsilon_{jk} \varepsilon_{jl} \varepsilon_{il}) = \text{trace}(\Sigma_u \Sigma_v \Sigma_u \Sigma_v). \]

The other second, third, and fourth order moments are all equal to zero. For simplicity, we assume \( \Sigma_u = \sigma_u^2 I_{t \times t}, \) \( \Sigma_v = \sigma_v^2 I_{t \times t}. \) In this case, the moments of bilinear term become
\[ E(\varepsilon_{ik}^2) = t \sigma_u^2 \sigma_v^2, \]
\[ E(\varepsilon_{ik} \varepsilon_{jk} \varepsilon_{jl} \varepsilon_{il}) = t \sigma_u^4 \sigma_v^4. \]

The nonzero moments of the \( \epsilon_{ik}s \) are now
\[ E(\epsilon_{ik} \epsilon_{il}) = \sigma_a^2 \]
\[ E(\epsilon_{ik} \epsilon_{jk}) = \sigma_e^2 \]
\[ E(\epsilon_{ik}) = \sigma_a^2 + \sigma_e^2 + \sigma_r^2 + t \sigma_u^2 \sigma_v^2 \]
\[ E(\epsilon_{ik} \epsilon_{jk} \epsilon_{jl} \epsilon_{il}) = \sigma_a^4 + \sigma_e^4 + t \sigma_u^4 \sigma_v^4. \]

The other moments of the \( \epsilon_{ik}s \) remain equal to zero.

The bilinear effect \( \varepsilon_{ik} = u'_i v_k \) can be interpreted as a mean-zero random effect that is able to capture particular types of dependence in affiliation network data. If we consider the pair of vectors \((u_i, u_j)\) (or the pair \((v_k, v_l)\)) and they have similar direction and magnitude, then \( u'_i v_k \) and \( u'_j v_k \) (or \( u'_i v_k \) and \( u'_i v_l \)) will not be too different. A probability measure over these unobserved characteristics induces a model in which the presence of a tie between an actor and an event is dependent on the presence of other ties. Relations modeled as such are probabilistically balanced.

5 Parameter Estimation

The parameters we want to estimate are \( \{\beta_d, \beta_a, \beta_e, \sigma_a^2, \sigma_e^2, \sigma_r^2, \sigma_u^2, \sigma_v^2\} \). Following Hoff (2005), we work with the following representation of our model:
\[ \theta_{ik} = \beta_d x_{ik}^d + (\beta_a x_{ik}^a + a_i) + (\beta_e x_{ik}^e + e_k) + \gamma_{ik} + u'_i v_k \]
\[ = \beta_d x_{ik}^d + \mu^a_i + \mu^e_k + \gamma_{ik} + u'_i v_k, \quad (4) \]
where \( \mu^a_i = \beta'_a x_{ik}^a + a_i \) and \( \mu^e_k = \beta'_e x_{ik}^e + e_k \) can be viewed as actor and event specific effects, respectively. We then define \( z_{ik} = \theta_{ik} - u'_i v_k = \beta_d x_{ik}^d + \mu^a_i + \mu^e_k + \gamma_{ik} \), and let \( z = \text{vec}(Z) \),
where $Z$ is the $n_a \times n_e$ matrix with elements $z_{ik}$ for $i = 1, \ldots, n_a$ and $k = 1, \ldots, n_e$. We take $\theta$ to be the $n_a \times n_e$ matrix with elements $\theta_{ik}$, and let $u$ be a $n_a \times t$ matrix with rows $u_i$ for $i = 1, \ldots, n_a$ and $v$ be a $n_e \times t$ matrix with rows $v_i$ for $i = 1, \ldots, n_e$. Then we can write

$$z = \text{vec}(\theta - uv') = X_d \begin{pmatrix} \beta_d \\ \mu_a \\ \mu_e \end{pmatrix} + \gamma$$

where $X_d$ is the appropriate design matrix constructed using (4) and $\gamma$ is a vector with dimension $n_\gamma$ as described in Section 2.2. From (5), it is clear that conditional on the $\theta$’s, $u$’s and $v$’s, the other parameters can be sampled using a standard Bayesian normal-theory regression approach.

Our general Gibbs sampler given below is analogous to Hoff (2005)'s algorithm, except for the sampling of the bilinear effects:

1. Sample linear effects:
   - Sample $\beta_d, \mu_a, \mu_e | \beta_a, \beta_e, \sigma_a^2, \sigma_e^2, \sigma_\gamma^2, \theta, u, v$ (linear regression)
   - Sample $\beta_a, \beta_e | \mu_a, \mu_e, \sigma_a^2, \sigma_e^2, \sigma_\gamma^2$ (linear regression)
   - Sample $\sigma_a^2, \sigma_e^2$, and $\sigma_\gamma^2$ from their full conditionals

2. Sample bilinear effects:
   - For $i = 1, \ldots, n_a$ sample $u_i | u_{-i}, v, \theta, \beta_d, \mu_a, \mu_e, \sigma_a^2$ (linear regression)
   - For $k = 1, \ldots, n_e$ sample $v_k | v_{-k}, u, \theta, \beta_d, \mu_a, \mu_e, \sigma_v^2$ (linear regression)
   - Sample $\sigma_a^2$ and $\sigma_v^2$ from their full conditionals

3. Update $\theta$: For actor $i$ and event $k$
   - Propose $\theta_{ik}^* \sim N(\beta_d'X_{ik} + \mu_a^2 + \mu_e^2 + u_i'v_k, \sigma_\gamma^2)$
   - Accept $\theta_{ik}^*$ with probability $[p(y_{ik}|\theta_{ik})/p(y_{ijk}|\theta_{ik})] \wedge 1$

The full conditional distributions of $\beta_d, \mu_a, \mu_e, \beta_a, \beta_e, \sigma_a^2, \sigma_e^2, \sigma_\gamma^2, \sigma_u^2, \sigma_v^2, u_i$ and $v_k$ are given in Appendix A.
For binary affiliation network cases, we can use the same algorithm as above with $\sigma^2_\gamma$ set to a fixed value since over-dispersion is not appropriate for generalized linear models for binary responses. We use this algorithm to fit the model to a binary affiliation network dataset in Section 6 with $\sigma^2_\gamma = 1$.

In addition to the parameters, the dimension of the latent bilinear effects, $t$, is unknown. Choice of $t$ will generally depend on the goal of the analysis. If we want to visualize the bilinear terms in order to understand latent structure in an affiliation network, we can simply choose $t = 1, 2, \text{ or } 3$. If the goal is prediction, we can examine higher dimensions and compare models using the Deviance Information Criterion (DIC: Spiegelhalter et al. 2002), or perhaps formally include $t$ in the model space and employ a reversible jump MCMC algorithm for model fitting (Green 1995). In Section 6, we try different value for $t$ and report the corresponding DIC. Lastly, if we are interested in whether the model captures particular features of the observed network (Wasserman and Faust 1994) or in examining particular aspects of lack-of-fit, we can evaluate the model with posterior predictive checks (Besag 2000).

6 Application

6.1 Data Description

To illustrate our proposed bilinear mixed-effects model, we consider a binary affiliation network of student participation in extracurricular activities (McFarland 1999). McFarland considers the network over three years, but we focus on the network as it exists in 1996. The 1996 network consists of 1295 students and 91 student organizations, in which participation is recorded along with individual-level attributes of grade, gender, and race. Among the 1295 students, 373 have missing values. We remove these students from our illustrative analysis, and randomly select $n^a = 300$ students. We combine similar clubs together (see Appendix B) and the newly constructed network has $n^e = 37$ activities (events). We use the size of the
clubs from the full dataset as the event-specific covariates and the gender of students as the actor specific covariates (coded as 1 for girls and 0 for boys). The data used in our analysis are shown in Figure 2.

6.2 Evidence of Higher-Order Dependence

Before we fit the bilinear logistic regression model, we check the balance in the data by fitting the basic logistic mixed-effects model given by (2) and examining the dependence patterns in residuals as described in Section 3. The residuals have a mean around zero (0.018) and standard derivation of 0.22. For this real dataset, the fraction of residuals that are positive is \( p = 0.0705 \) (distribution is highly right skewed) and the proportion of positive residual cycles, \( \hat{\pi} \), is 0.8842. Under independence of the ties, the expected proportion of positive residual cycles is \( \pi_0 = p^4 + (1 - p)^4 + 6p^2(1 - p)^2 = 0.7724 \). If we randomly generate a \( n_a \times n_e \) matrix with 7.05 percent positive values (+1) and 92.95 percent negative values (-1) 100 times, the upper bound is 0.8697. Therefore, we conclude that the observed proportion of positive residual cycles is significantly greater than expected under independence (p-value \( \approx 0 \)).

6.3 Priors

Prior distributions for the random effect variances (\( \sigma_a^2, \sigma_e^2, \sigma_u^2, \) and \( \sigma_v^2 \)) are taken to be independent and distributed as IG(1, 1), where IG\((a, b)\) denotes the inverse gamma distribution with shape \( a > 0 \) and scale \( b > 0 \). The priors for \( \beta \) are normally distributed \( \beta \sim \text{MVN}(0, I_{r \times r}) \). The variance of the prior distribution of \( \beta \) is small since we are in the logistic regression setting.
Figure 2: Illustration of the one-mode network of extracurricular activities obtained by projecting the two-mode network described in Section 6.1. The blue circles represent the $n^e = 37$ activities (events), where the size of circles corresponds to the total number participants in the activity. If at least one student participates in a pair of activities, a green line is drawn connecting the activities. The line thickness corresponds to the number of students who participate in both activities. This figure was constructed using functions in the igraph R package (Csardi and Nepusz, 2006).
Table 1: DIC values for models with varying dimension of the components of the bilinear term.

| Parameters | $\beta_d$ | $\beta_a$ | $\beta_e$ | $\alpha_e$ | $\sigma_u$ | $\sigma_v$ |
|------------|--------|--------|--------|---------|--------|--------|
| Mean       | -5.111 | 0.399  | 0.052  | 0.215   | 0.321  | 2.238  |
| SD         | 0.204  | 0.140  | 0.006  | 0.061   | 0.118  | 1.032  | 0.301  |

Table 2: Posterior means and standard deviations of model parameters when $t = 3$.

6.4 Results

The MCMC algorithm described in Section 5 was run for 100,000 iterations for values of $t = 0$ (no bilinear term), 1, 2, and 3. Trace plots suggest that the Markov chain reach its stationary distribution well before 30,000 iterations, so we conservatively base our inferences on the last 30,000 iterations. The DIC corresponding to different values of $t$ are listed in Table 1. In terms of the DIC criterion, the fit improves as the dimension of the components of the bilinear terms increases. Based on the DICs, we choose to report inferences on the $t = 3$ model.

Table 2 provides the posterior mean and standard deviations for all scalar model parameters when $t = 3$. Since $E[\beta_a|Y] = 0.399$, the odds of a girl attending extracurricular activities is $e^{0.399} = 1.490$ higher than the odds of boys attending, with 95 percent credible interval (1.126, 1.972). As expected, the relationship between club size and the expected log odds of participation is also positive with $E[\beta_e|Y] = 0.052$, implying an estimated odds ratio $e^{0.052} = 1.053$ with 95 percent credible interval (1.041, 1.066).

Next, we analyze the posterior distribution of the latent vectors $u$ and $v$. The dimension of the bilinear term $uv'$ is $n_a \times n_e$. To reduce the dimensionality, we performed a principal
components analysis of the posterior mean of the bilinear term. The first two principal components preserve 84.3 percent of the total variance, so we can summarize our inferences using a biplot (Gabriel, 1971), which is shown in Figure 3. Here, the arrows represent the loadings of each of the activities onto the first two principal components and the coordinates of the points are the corresponding scores for each actor. To facilitate the interpretation, the arrows are colored based on the assigned categories of activities listed in Appendix B (Note that these categories were not used in the model fitting.) Points corresponding to female students are colored pink and points corresponding to male students are colored blue.

From Figure 3 we can see that generally the arrows corresponding to activities in the same categories tend to point in similar directions. For example, we see that Newspaper and Yearbook, the two activities in the News category, both point to the bottom right capturing the apparent tendency of students to either participate in both or neither of these activities. On the other hand, there are examples where the arrows do not cluster based on category. For example, in Music category, Orchestra and Choir point in the opposite direction from Band indicating a lack of overlap in participants in these groups. From the distribution of the blue and pink points, we see that males dominate the upper left quadrant of the biplot, which makes intuitive sense since the loading vectors of the boys-only sports (Football, Baseball, and Wrestling) are oriented in this direction. We can also see that girls tend to be more active in the Service, News, and Cheer categories.

7 Discussion

This article has extended the bilinear mixed-effects model to the two-mode setting. This model is based on a generalized linear mixed-effects model with the addition of an inner product of two latent characteristic vectors. As we have shown, it is able to capture the types of dependence patterns (balance) usually seen in affiliation network. Our model improves on existing ad hoc methods for analyzing affiliation networks in that we coherently capture the uncertainty in our inferences on the latent structure of a network. Visualizing this
Figure 3: Biplots of the components of the bilinear term – see Section 6.4 for a description. The biplot on the left contains the subset of arrows corresponding to activities with estimated $v$'s with first or second component greater than 0.15 in absolute value. The biplot on the right contains the arrows for the remaining estimated $v$'s.
uncertainty is a challenging task, and we plan to work on this important task in future work. As is often the case for fitting complex Bayesian models using MCMC, computation can be challenging when the sample size (number of actors and events) is large. Accordingly, future work will seek to exploit computational tricks to take advantage of sparsity in affiliation networks, as well as explore approximation techniques.

Lastly, the latent structure of our model provides a natural mechanism for incorporating grouping structure in either the actors or events. For example, in the extracurricular activity example, we could have built in the expected similarity between activities in the same category by placing a hierarchical prior on the activity random effects and the components of the \( v \)'s. In this way, we would be able to explain the dependences in affiliation network both within and across different groups of events.

Acknowledgements

The authors thank Professor Chris Browning, Department of Sociology at The Ohio State University, for introducing them to affiliation networks. Support for this work was provided by grants from the National Science Foundation (NSF DMS-1209161), the National Institutes of Health (NIH R01DA032371), the William T. Grant Foundation, and The Ohio State University Institute for Population Research.

References

Besag, J. (2000). *Markov Chain Monte Carlo for Statistical Inference*. University of Washington, Center for Statistics and the Social Sciences.

Congdon, P. D. (2010). *Applied Bayesian Hierarchical Methods*. Chapman & Hall/CRC, Boca Raton, FL.
Csardi, G. and Nepusz, T. (2006). “The igraph software package for complex network research.” *InterJournal, Complex Systems*, 1695.

Davis, A. and Gardner, B. B. (1941). *Deep South*. University of Chicago Press, Chicago, IL.

Frank, O. and Strauss, D. (1986). “Markov graphs.” *Journal of the American Statistical Association*, 81, 832–842.

Gabriel, K. R. (1971). “The biplot graphic display of matrices with application to principal component analysis.” *Biometrika*, 58, 3, 453–467.

Green, P. J. (1995). “Reversible jump Markov chain Monte Carlo computation and Bayesian model determination.” *Biometrika*, 82, 4, 711–732.

Handcock, M. S., Robins, G., Snijders, T., Moody, J., and Besag, J. (2003). “Assessing degeneracy in statistical models of social networks.” *Journal of the American Statistical Association*, 76, 33–50.

Hoff, P. D. (2003). “Random effects models for network data.” In *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, 303–312.

— (2005). “Bilinear mixed-effects models for dyadic data.” *Journal of the American Statistical Association*, 100, 469, 286–295.

— (2009). “A hierarchical eigenmodel for pooled covariance estimation.” *Journal of the Royal Statistical Society, Series B*, 71, 971–992.

Hoff, P. D., Raftery, A. E., and Handcock, M. S. (2002). “Latent space approaches to social network analysis.” *Journal of the American Statistical Association*, 97, 1090–1098.

Hunter, D. R., Krivitsky, P. N., and Schweinberger, M. (2012). “Computational Statistical Methods for Social Network Models.” *Journal of Computational and Graphical Statistics*, 21, 856–882.
McFarland, D. A. (1999). “Organized Behavior in Social Systems: A Study of Student Engagement and Resistance in High Schools.” Ph.D. thesis, University of Chicago, Chicago, IL.

Newman, M. E. J. (2001). “Scientific collaboration networks. I. Network construction and fundamental results.” Physical Review, 64, 016131, 1–8.

Pattison, P. and Wasserman, S. (1999). “Logit models and logistic regressions for social networks: II. Multivariate relations.” British Journal of Mathematical and Statistical Psychology, 52, 169–193.

Robins, G., Pattison, P., and Wasserman, S. (1999). “Logit models and logistic regressions for social networks: III. Valued Relations.” Psychometrika, 64, 371–394.

Snijders, T. A. B. (1997). “The statistical evaluation of social network dynamics.” Sociological Methodology, 31, 361–395.

Snijders, T. A. B. and Nowicki, K. (1997). “Estimation and prediction for stochastic block-models for graphs with latent block structure.” Journal of Classification, 14, 75–100.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). “Bayesian measures of model complexity and fit.” Journal of the Royal Statistical Society, Series B, 64, 4, 583–639.

van Duijn, M. A., Gile, K. J., and Handcock, M. S. (2009). “A framework for the comparison of maximum pseudo likelihood and maximum likelihood estimation of exponential family random graph models.” Social Networks, 31, 52–62.

Wang, P. (2009). “Exponential random graph models for affiliation networks.” Social Networks, 31, 12–25.

Wang, Y. J. and Wong, G. Y. (1987). “On the geometry of discrete exponential random families with application to exponential random graph models.” Journal of the American Statistical Association, 82, 397, 8–19.
Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambrige University Press, Cambrige.

Wasserman, S. and Pattison, P. (1996). “Logit models and logistic regression for social networks: I. An introduction to Markov graphs and p*.” *Psychometrika*, 61, 401–425.

Zijlstra, B. J. H., van Duijn, M. A. J., and Snijders, T. A. B. (2009). “MCMC estimation for the p2 network regression model with crossed random effects.” *British Journal of Mathematical and Statistical Psychology*, 62, 143–166.
APPENDICES

A Full Conditional Distributions

Full conditional distribution of \((\beta_d, \mu^a, \mu^e)\)

The full conditional distribution of \((\beta_d, \mu^a, \mu^e)\) is proportional to

\[ p(z|\beta_d, \mu^a, \mu^e, \sigma_\gamma^2) \propto p(\mu^a, \mu^e|\beta_a, \beta_e, \sigma_a^2, \sigma_e^2) \times p(\beta_d). \]

Assume the prior distribution of \(\beta_d\) follows a multivariate normal distribution

\[ \beta_d \sim MVN(\mu_{\beta_d}, \Sigma_{\beta_d}). \]

We already know that \(\mu_i^a = \beta'_a x_i^a + a_i\) and \(\mu_k^e = \beta'_e x_k^e + e_k\). So, we have

\[ \mu^a, \mu^e|\beta_a, \beta_e, \sigma_a^2, \sigma_e^2 \sim MVN(X_{ae}\beta_{ae}, \Sigma_{ae}), \]

where \(X_{ae}\) is a \((n^a + n^e) \times 2\) matrix, \(\beta_{ae} = (\beta_a, \beta_e)'\), and \(\Sigma_{ae} = \begin{pmatrix} \sigma_a^2 I_{n^a} & 0 \\ 0 & \sigma_e^2 I_{n^e} \end{pmatrix} \).

Since \(\beta_d, \mu^a, \mu^e\) are independent and Gaussian, we can rewrite their joint distribution as

\[ \beta_d, \mu^a, \mu^e|\beta_a, \beta_e, \sigma_a^2, \sigma_e^2 \sim MVN \begin{pmatrix} \mu_{\beta_d} \\ X_{ae}\beta_{ae} \end{pmatrix}, \begin{pmatrix} \Sigma_{\beta_d} & 0 \\ 0 & \Sigma_{ae} \end{pmatrix} \].

Let \(z_{ik} = \theta_{ik} - u'_i v_k = \beta'_d x_{ik}^d + \mu_i^a + \mu_k^e + \gamma_{ik}\) so

\[ z|\beta_d, \mu^a, \mu^e, \sigma_\gamma^2 \sim MVN \begin{pmatrix} \beta_d \\ \mu^a \\ \mu^e \end{pmatrix}, \sigma_\gamma^2 I_{n \gamma} \] \]

It follows that the full conditional distribution \((\beta_d, \mu^a, \mu^e)\) is multivariate normal with the following mean and covariance:
\[
\Sigma = \left( \begin{array}{cc}
\Sigma_{\beta_d}^{-1} & 0 \\
0 & \Sigma_{ae}^{-1}
\end{array} \right) + X_D^T X_D / \sigma_\gamma^2
\]

\[
\mu = \Sigma \left( \begin{array}{c}
\Sigma_{\beta_d}^{-1} \beta_d \\
\Sigma_{ae}^{-1} X_{ae} \beta_{ae}
\end{array} \right) + X_D^T z / \sigma_\gamma^2
\]

**Full conditional distribution of** \((\beta_a, \beta_e)\)

The full conditional distribution of \((\beta_a, \beta_e)\) is proportional to \(p(\mu^a, \mu^e | \beta_a, \beta_e, \sigma_a^2, \sigma_e^2) \times p(\beta_a, \beta_e)\). Assume the combined regression parameter has a multivariate normal prior: \((\beta_a, \beta_e) \sim MVN(\mu_{ae}, \Sigma_{ae})\).

Therefore, the full conditional is a multivariate normal distribution with the following mean and variance:

\[
\Sigma = \left( X_{ae}^T \Sigma_{ae}^{-1} X_{ae} + \Sigma_{\beta ae}^{-1} \right)^{-1}
\]

\[
\mu = \Sigma \left( X_{ae}^T \Sigma_{ae}^{-1} \mu^a \right) + \Sigma_{\beta ae}^{-1} \mu_{ae}
\]

**Full conditional distribution of** \(\sigma_a^2, \sigma_e^2\)

We restrict \(\Sigma_a = \sigma_a^2 I_{n^a \times n^a}\) and \(\Sigma_e = \sigma_e^2 I_{n^e \times n^e}\). For \(\sigma_a^2 \sim IG(\alpha_{a1}, \alpha_{a2})\), and \(\sigma_e^2 \sim IG(\alpha_{e1}, \alpha_{e2})\), the full conditionals are independent and

\[
\sigma_a^2 | \mu^a \sim IG(n^a/2 + \alpha_{a1}, \alpha_{a2} + (\mu^a - X_a \beta_a)'(\mu^a - X_a \beta_a)/2)
\]

and

\[
\sigma_e^2 | \mu^e \sim IG(n^e/2 + \alpha_{e1}, \alpha_{e2} + (\mu^e - X_e \beta_e)'(\mu^e - X_e \beta_e)/2).
\]
Full conditional distribution of $\sigma_\gamma^2$

We restrict $\Sigma_\gamma = \sigma_\gamma^2 I_{n^\gamma \times n^\gamma}$. Using prior distribution of $\sigma_\gamma^2 \sim IG(\alpha_1, \alpha_2)$, the full conditional distribution of $\sigma_\gamma^2$ is

$$
\sigma_\gamma^2 | \beta_d, \mu^a, \mu^e, z \sim IG \left( \alpha_1 + n^\gamma / 2, \alpha_2 + \left[ z - X_D \left( \begin{array}{c} \beta_d \\ \mu^a \\ \mu^e \end{array} \right) \right]' \left[ z - X_D \left( \begin{array}{c} \beta_d \\ \mu^a \\ \mu^e \end{array} \right) \right] / 2 \right)
$$

Full conditional distribution of $u_i$

Let $\theta_{ik} = \beta_d^i x_{ik}^d + \mu^a_i + \mu^e_k + \gamma_{ik} + u_i^v v_k$, as before, and $\hat{\theta}_{ik} = E(\theta_{ik}|\beta_d, \mu^a, \mu^e, x_{ik}) = \beta_d^i x_{ik}^d + \mu^a_i + \mu^e_k$. Then $\epsilon_{ik} = \theta_{ik} - \hat{\theta}_{ik} = v_k' u_i + \gamma_{i}/2$.

Considering the full conditional of $u_i$, we have

$$
\begin{pmatrix}
\epsilon_{i,1}^u \\
\epsilon_{i,2}^u \\
\vdots \\
\epsilon_{i,n^e}^u \\
\end{pmatrix} =
\begin{pmatrix}
v_1 \\
v_2 \\
\vdots \\
v_{n^e} \\
\end{pmatrix} u_i +
\begin{pmatrix}
\gamma_{i,1} \\
\gamma_{i,2} \\
\vdots \\
\gamma_{i,n^e} \\
\end{pmatrix},
$$

so $e_i^u | v, u_i, \sigma_\gamma^2 \sim MVN(\mathbf{v} u_i, \sigma_\gamma^2 I_{n^e})$. Therefore, sampling $u_i$ from its full conditional is equivalent to a Bayesian linear regression problem. Assuming the $u_i$'s are a priori independent and $u_i \sim MVN(0, \Sigma_u)$, the full conditional of $u_i$ is multivariate normal with

$$
\Sigma = (\Sigma_u^{-1} + \mathbf{v}' \mathbf{v} / \sigma_\gamma^2)^{-1}
$$

and

$$
\mu = \Sigma \mathbf{v} e_i^u / \sigma_\gamma^2.
$$
Full conditional distribution of $v_k$

Similar to the derivation of the full conditional distribution of $v_i$, we have $e_{ik} = \theta_{ik} - \hat{\theta}_{ik} = u'v_k + \gamma_i/2$.

Considering the full conditional of $v_k$, we have

$$
\begin{pmatrix}
    e_{1,k} \\
    e_{2,k} \\
    \vdots \\
    e_{n_a,k}
\end{pmatrix}
= 
\begin{pmatrix}
    u_1 \\
    u_2 \\
    \vdots \\
    u_{n_a}
\end{pmatrix}v_k + \begin{pmatrix}
    \gamma_{1,k} \\
    \gamma_{2,k} \\
    \vdots \\
    \gamma_{n_a,k}
\end{pmatrix}
$$

$e_k^v \sim \text{MVN}(uv_k, \sigma^2_v I_{n_a})$. Therefore, sample $v_k$ from its full conditional is also equivalent to a Bayesian linear regression problem. Assuming the $v_k$'s are a priori independent and each $v_k \sim \text{MVN}(0, \Sigma_v)$, the full conditional of $v_k$ is multivariate normal with

$$
\Sigma = (\Sigma_v^{-1} + u'u/\sigma^2_v)^{-1}
$$

and

$$
\mu = \Sigma u'e_k^v / \sigma^2_v.
$$

Full conditional distribution of $\sigma^2_u, \sigma^2_v$

We restrict $\Sigma_u = \sigma^2_u I_{t \times t}$ and $\Sigma_v = \sigma^2_v I_{n_a \times n_a}$ and let $\sigma^2_u \sim \text{IG}(\alpha_{u1}, \alpha_{u2})$ and $\sigma^2_v \sim \text{IG}(\alpha_{v1}, \alpha_{v2})$.

Then the full conditionals are

$$
\sigma^2_u | u \sim \text{IG}(n^u t / 2 + \alpha_{u1}, \alpha_{u2} + \text{trace}(u'u) / 2)
$$

and

$$
\sigma^2_v | v \sim \text{IG}(n^v t / 2 + \alpha_{v1}, \alpha_{v2} + \text{trace}(v'v) / 2).
$$
B Data Processing

In our analysis of McFarland’s (1999) data for year 1996, we collapsed several extracurricular activity categories as described below. For example, our club “Pep” includes both members of “Pep.Club” and “Pep.Club.Officers.” In addition, we grouped the activities into one of eight types labeled as 1-8 below. These groups were not used in fitting our model, but were helpful in interpreting our fitted model.

1. Language
   (a) Asian
   (b) Spanish includes Hispanic.Club, Spanish.Club, Spanish.Club..high.,Spanish.NHS
   (c) Latin
   (d) French includes French.Club..low., French.Club..high., French.NHS
   (e) German includes German.Club, German.NHS

2. Academic Competition
   (a) Debate
   (b) Forensics includes Forensics, Forensics..National.Forensics.League.
   (c) Chess
   (d) Science.Olympiad
   (e) Quiz.Bowl
   (f) Academic.Decathalon

3. News
   (a) Newspaper
   (b) Yearbook includes Yearbook.Contributors, Yearbook.Staff
4. Cheer

   (a) *Pep* includes Pep.Club, Pep.Club.Officers
   (b) *Drill*
   (c) *Cheer* includes Cheerleaders..8th, Cheerleaders..9th, Cheerleaders..Spirit.Squad, Cheerleaders..JV, Cheerleaders..V

5. Service

   (a) *National Honor Society*
   (b) *Drunk.Driving* includes Drunk.Driving, Drunk.Driving.Officers
   (c) *Key*

6. Art/Theater

   (a) *Art*
   (b) *Theatre*
   (c) *Thespian*

7. Music

   (a) *Band* includes Band..8th, Band..Marching..Symphonic., Band..Jazz
   (b) *Orchestra* includes Orchestra..8th, Orchestra..Full.Concert, Orchestra..Symphonic
   (c) *Choir* includes Choir..treble, Choir..concert, Choir..women.s.ensemble, Choir..a.capella, Choir..chamber.singers, Choir..vocal.ensemble..4.women., Choir..barbershop.quartet..4.men.

8. Sports

   (a) *Football* includes Football..8th, Football..9th, Football..V
   (b) *Soccer*
   (c) *Volleyball* includes Volleyball..8th, Volleyball..9th, Volleyball..JV, Volleyball..V
(d) *Basketball* includes Basketball..boys.8th, Basketball..boys.9th, Basketball..boys.JV, Basketball..boys.V, Basketball..girls.8th, Basketball..girls.9th, Basketball..girls.JV, Basketball..girls.V

(e) *Baseball* includes Baseball..JV..10th., Baseball..V

(f) *Softball* includes Softball..JV..10th., Softball..V

(g) *Cross.Country* includes Cross.Country..boys.8th, Cross.Country..girls.8th, Cross.Country..boys.V, Cross.Country..girls.V

(h) *Golf*

(i) *Swim* includes Swim...Dive.Team..boys, Swim...Dive.Team..girls

(j) *Tennis* includes Tennis..boys.V, Tennis..girls.V

(k) *Track* includes Track..boys.8th, Track..girls.8th, Track..boys.V, Track..girls.V

(l) *Wrestling* includes Wrestling..boys.8th, Wrestling..V