Network Analysis Using a Local Structure Graph Model: Application to Alliance Formation

Olga Chyzh* and Mark S. Kaiser†

July 14, 2016

Abstract

We introduce a local structure graph model (LSGM)—a class of statistical estimators that allows for modeling a number of theoretical processes among network edges. LSGMs are especially relevant to political science, as a large number of political networks are outcomes of edge- rather than node-level political processes. Formation of political coalitions and voting blocks, balancing and bandwagoning, policy learning, imitation, diffusion, and tipping-point dynamics and cascade effects are all processes, theoretical understanding of which require focusing on relationships among edges within networks, rather than nodes. Central theories of coalition formation, for example, emphasize that edges form in response to formation of other edges, both within ideological proximity (bandwagoning) and at the opposite end of the ideological spectrum (balancing). Such processes, however, are not easily modeled using the traditional network approach of treating actors (legislators or international states) as network nodes and alliances among them as edges. Any network processes, in which network edges form in response to formation or characteristics of other edges are best modeled using a “second-degree” network, in which an edge (e.g., an alliance) is thought of as a node, and a relationship among edges (e.g., belonging to the same neighborhood) are treated as edges. We demonstrate the properties of LSGMs using Monte Carlo simulations and an empirical application to the international alliance network.

*Assistant Professor of Political Science and Statistics, Iowa State University, http://www.olgachyzh.com/
†Professor of Statistics, Iowa State University.
1 Introduction

Many network processes that are of interest to political science research—such as formation of coalitions and voting blocks, balancing and bandwagoning, policy learning, imitation, diffusion, and tipping-point dynamics and cascade effects—require focusing on relationships among edges within networks, rather than nodes. Central theories behind coalition formation, for example, emphasize that edges form in response to formation of other edges. Voting blocs in legislators form to balance opposing blocs. International alliances frequently form with the goal of balancing against a similar actions of a political rival: e.g., the Soviet bloc formed the Warsaw Pact in response to the US and its allies forming the North Atlantic Treaty Organization (NATO) in the aftermath of World War II.

Such substantively important processes, however, are not easily modeled using the traditional network approach of treating actors (legislators or international states) as network nodes and alliances among them as edges. Any network processes, in which network edges form in response to formation or characteristics of other edges are best modeled using a “second-degree” network, in which an edge (e.g., an alliance) is thought of as a node, and a relationship among edges (e.g., belonging to the same neighborhood) are treated as edges.\(^1\)

We demonstrate that the theoretical processes of interest may be estimated using an adaptation of a class of network/spatial models known in statistics as local structure graph models (LSGMs), that fall within the broader class of Markov random field models (Casleton, Nordman, and Kaiser 2016). To formulate an LSGM, one must first specify a set of full conditional distributions for each potential edge in the network, i.e. the distribution of the presence/absence of an edge given the outcomes for all potential edges and a set of exogenous covariates. Thus, each conditional distribution is specified in terms of a neighborhood

\(^1\) Despite similarity in terminology, the problem we address is not the same as modeling second-degree links of the type “if A is connected to B and B is connected to C, then A and C share a second degree connection.” Instead, we highlight the theoretical advantages and propose a estimation framework for modeling network connectivities among edges, rather than connectivities among nodes.
structure that explicitly identifies the degree of “local” dependency between each edge and each other edge. Such a conditional distribution allows for a derivation of a global or joint probability model for the network. We use Monte Carlo simulations to demonstrate that the model recovers unbiased and efficient parameter estimates that are also easily interpretable. Finally, we supplement simulations with an application to the formation of the international alliance network among international states between 1947-2001.

2 Conceptualizing Edge Formation as Spatial Diffusion

Network analysis has found wide application in all sub-fields of political science. Scholars have applied network tools to derive insights on the functioning of legislatures (Cho and Fowler 2010), diffusion of policies (Desmarais, Harden, and Boehmke 2015), and conflict and cooperation among international states (Kinne 2013). Network analysis has informed both theory-building and inference. Thus, some research employs network game theory and agent-based modeling with the goal of accurately simulating networks of interest and studying their properties (Jackson 2008; Chyzh 2016a; Gallop 2016; Maoz and Joyce 2016; Siegel 2009). Other research focuses on development and application of network-informed probabilistic estimators that would allow for deriving statistical inferences. Such probabilistic network modeling may be further classified into exponential random graph models (ERGMs), latent space models (LSMs), and spatial autoregressions (SARs).

ERGMs account for network dependencies via the inclusion of covariates that correspond to specific topological features of the graph (e.g., reciprocity, triads, 2-stars). These network topologies, also known as Markovian features, are defined as counts of all elements of a certain class weighed as a proportion of the total count that could potentially form in the given graph. The parameters associated with such covariates will then inform us of the prevalence of each type of element in the observed realization of the graph Wasserman and Faust (1994);
As a result of their formulation, ERGMs provide the most leverage when the network dependencies are treated as a nuisance or when the researcher is interested in modeling the types of dependencies that are easily translated into particular Markovian features. While some theoretical processes are easily modeled via Markovian features, a number of important political dynamics, such as coalition-building, balancing and bandwagoning, or spatial diffusion are less amenable to the ERGMs framework.

Coalition-building (e.g., alliance coalitions among international states, voting coalitions in legislatures, building voter support for a political candidate), for example, is often theorized as an outcome of two processes—balancing or formation of blocks at the opposite sides of ideological spectrum (Morrow 1991; Kedar 2005)—and bandwagoning or the tendency of weaker players to align with the expected winner (Sweeney and Fritz 2004; Hassell 2016). If we conceptualize an alliance between two states or two legislators as an edge in a network, then the balancing theory predicts that the realization of any given edge is conditional on realization of edges located on the opposite side of the ideological spectrum: i.e. the competing coalitions form in response to one another. The bandwagoning theory, in contrast, predicts that edge formation will cluster within ideological space: i.e. formation of a voting alliance will likely trigger additional allies to jump on-board.

While one may, of course, try to imagine possible ways to model coalition formation within the ERGM set-up (e.g., using triads or two-stars), any Markovian features would get at the theorized processes only indirectly. A coefficient on triads, for example, would capture the general tendency of two actors forming an alliance, given that they are both allied with the same third actor, yet it would provide no information regarding the distance or proximity of such actors on the ideological spectrum. Importantly, while the theorized coalition-building process takes place among pairs of edges (i.e. edges form in response to other edges), most of the commonly used Markovian topologies are constructed at the nodal
LSMs, in turn, allow for a hierarchical approach to modeling network data, in which dependence is accounted for by considering the relevant node-specific latent variables that are sources of non-independence, such as group membership or position within social space (Hoff and Ward 2004; Minhas, Hoff, and Ward 2016). Just like with ERGMs, LSMs traditionally focus on sources of dependence associated with nodes, whereas many important applications require modeling dependence among edges.

The third type of network modeling—SARs—account for non-independence among observations by including lag structures which measure theorized sources of connectivity (Anselin 2013; Besag 1974; Hays, Kachi, and Franzese 2010). While existing social science applications of spatial autoregression are also limited to modeling diffusion among nodes, rather than edges, we propose an extension that would allow for precisely that.

To illustrate and compare a traditional network model that focuses on nodes to a second-degree network that focuses on edges, Subfigure 1a of Figure 1 shows a randomly generated network of five nodes, while Subfigure 1b shows a second degree version of the same network. Note that in Subfigure 1b, edges from Subfigure 1a are depicted as nodes and relationships level of analysis.
Figure 2: Visualizing International Alliance Formation as a First Degree Network

(a) 1947

(b) 1949

(c) 1951

(d) 1955

Notes: Alliance data are obtained from the Correlates of War Project (Gibler 2009). Nodes represent international states that formed any alliance in a given year.

between them (whether they share a common node) as edges. For example, node 13 in Subfigure 1b corresponds to the edge between nodes 1 and 3 in Subfigure 1a, and edges 13 and 32 (conceptualized as nodes in Subfigure 1b) are connected by an edge because these edges share a common node (node 3).

Figures 2 and 3 demonstrate the same idea applied to the network of international alliances. Figure 2 presents a traditional visualization of the alliance network with interna-
Figure 3: Visualizing International Alliance Formation as a Second Degree Network

![Visualizations for different years](image)

Figure 3, in contrast, visualizes international alliance networks during the same slices in time as a second degree network or a network, in which alliances are treated as nodes that

tional states as nodes and actors of forming an alliance as edges at various slices in time. Importantly, while these visualization provide information related to alliance density and clustering, it is very difficult to judge from these visualizations whether alliances form as a result of balancing or bandwagoning by international states—a question of central theoretical importance within the literature.

Figure 3, in contrast, visualizes international alliance networks during the same slices in time as a second degree network or a network, in which alliances are treated as nodes that
are aligned in space based on their ideological distance from one another (a continuous conceptualization of a relationship/edge).\textsuperscript{2} This second-degree visualization of the international alliance network provides much more information regarding the balancing and bandwagoning processes that lead to the formation of the alliance network. In particular, Figure 3 allows for assessing three important patterns related to alliance formation, all of which obscured by a first-degree network visualization.

First, we see that international alliances tend to form between ideologically similar rather than different states—most alliances cluster close to the diagonal of the graph (a line at a 45\(^\circ\) angle would represent the location of all alliance partners with identical ideal scores) rather than in the areas off the diagonal. While this pattern is expected, it is nonetheless useful to be able to confirm this intuition by visualizing the data in a relevant way. Second, Figure 3 highlights clustering in ideological space within the same temporal period, which is consistent with bandwagoning rather than balancing theory of alliance formation. Third, an examination of consecutive temporal slices implies that balancing may happen over time: while the US and its allies form a large number of alliances in the first three temporal periods shown, while the Soviet block seems to only have responded 7 years later.\textsuperscript{3} In the next section, we present a statistical approach that allows to model this and similar processes in a controlled interpretable way using a extension of an LSGM model.

\textsuperscript{2}To calculate ideological distance between alliances, we utilized ideal point scores, developed by Bailey, Strezhnev, and Voeten (2015). Bailey, Strezhnev, and Voeten (2015) use UN voting data to align all international states on a standardized normal scale (from about -3 to +3) between 1947-2012, where higher scores are associated with the US and its allies and lower scores are attributed to Russia/Soviet Union block. In Figure 3, each alliance partner ideal point score provides one of the alliance coordinates in a two dimensional space, and ideological distance is calculated as the Euclidean distance between alliances.

\textsuperscript{3}Years not shown either experienced no alliance formation activity (1948, 1950) or were similar to the years shown, i.e. the US and its allies formed a small number of additional alliances in 1952, 1953, and 1954. The Soviet block did not form any alliances until 1955.
3 A Local Structure Graph Model with Continuous Neighborhoods

As argued above, many political networks, especially those that form as a result of balancing/bandwagoning dynamics, or diffusion/learning processes are best studied with the focus on edges and relations between them rather than the traditional focus on nodes and their connections/edges.

To model such processes, suppose $i$ is an edge in a network of $n$ edges (e.g., an alliance in the example discussed above), $i \in 1, 2, \ldots, n$ so that $i$’s local is denoted as $s_i = (u_i, v_i)$ in Cartesian space. Further, denote the binary random variable, $y(s_i)$, that denotes the presence of absence of an edge, such that:

$$y(s_i) = \begin{cases} 
1 & \text{if } y^* > 0 \\
0 & \text{if } y^* \leq 0 
\end{cases}$$

Next, let us define $i$’s neighbors as $N_i$, so that $y(N_i) = \{y(s_j) : s_j \neq s_i\}$. While edge neighborhoods in LSGMs have been traditionally defined as binary (i.e. two edges either belong to the same neighborhood or not), a continuous conceptualization of neighborhoods adopted here is more theoretically appropriate for social science applications. Hence, we depart from the traditional LSGM formulation in that, rather than defining edges as belonging into a set of discrete neighborhoods, we treat neighborhoods as a continuous space, in which some edges are located in closer proximity than others. Distance between edges is denoted by an $nxn$ matrix $w$, whose $ij$ cell represents the degree of connectivity between edges $i$ and $j$ and 0s on the major diagonal (edges do not belong to their own neighborhood). The connectivity matrix $w$ may represent geographical distance between edges, their ideological similarity, or any other pairwise measure of relationship.
Next, we make a Markov assumption of conditional spatial independence of the form:

\[ f(y(s_i)|y(s_j) : s_j \neq s_i) = f(y(s_i)|y(N_i)) \]  

(1)

Thus, the realization of any given edge \( i \) is dependent on realization of every other edge in its neighborhood \( N_i \), yet conditionally independent of realization of edges in its neighbors’ neighborhoods. Intuitively, this assumption simply means that \( i \) is most heavily affected by its immediate neighbors rather than indirect effects through its neighbors’ neighborhoods.

Consistent with the alliance example above, we assume a binary conditional distribution, which is expressed in exponential form as:

\[ P(Y(s_i) = y(s_i)|y(N_i)) = \exp \left[ y(s_i) A_i(y(N_i)) - B(y(N_i)) \right], \]

(2)

where \( A_i \) is a natural parameter function and \( B_i = \log[1 + \exp(y(s_i)A_i(y(N_i))) \right), \) and \( y(N_i) \) is a vector of values of the binary random variables (edges) of \( i \)’s neighbors. For binary conditional distributions, the natural parameter function takes the form:

\[ A_i(y(N_i)) = \log \left( \frac{\kappa_i}{1-\kappa_i} \right) + \eta w(y(s_j) - \kappa_j), \]

(3)

where \( \log \left( \frac{\kappa_i}{1-\kappa_i} \right) = X_i \beta \), \( X_i \) is a vector of exogenous covariates, \( \beta \) is a vector of estimation parameters, \( w \) is a matrix of spatial connectivities among edges, and \( \eta \) is its parameter.

Parameters where:

\[ p_i = \frac{\exp(A_i(N_i))}{1 + \exp(A_i(N_i))} \]

(4)

The point estimates recovered by maximizing the PL function have been shown to be unbiased and consistent for the general case of Markov random fields models (Casleton,
Nordman, and Kaiser 2016; Guyon 1995). Standard errors may be obtained via bootstrap. In what follows, we use Monte Carlo simulations to demonstrate the properties of the parameter estimates for the special case of the model presented in Equation 2, and follow up with a empirical application to data on international alliance formation.

4 Monte Carlo Simulations

We start by generate information for 100 observations (nodes), \( i = 1, 2, \ldots, n \) with characteristics captured by variable \( X_i \), drawn from a standard normal distribution. We proceed to convert these data to a dyadic format (edges) by pairing each observation with each other observation and omitting self-referencing pairs of the type \( i - i \) for a total of 9900 edges. To generate a meaningful spatial matrix, \( w \), we place each pair on an evenly spaced ten-by-ten grid and calculate the Euclidean distance between the two units making up each edge. Next, we use a Gibbs sample with randomly initialized values to generate random variable, \( Y(s_i) \).

The Gibbs sampler starts with the randomly-generated starting values for the dependent variable and iteratively updates them, one observation at a time, using the specified parameter values, \( \beta_0 = \beta_1 = 1 \) and \( \eta = .05 \), following the data-generating process specified in Equation 2. The Gibbs sampler was run with a burn-in of 20,000, after which sample graphs were retained from 100,000 subsequent rounds with 50 iterations for thinning.

The results of the Monte Carlo simulations are presented in Figure 4. As expected, the 90% confidence intervals, displayed in the figure, converge around the true value of each parameter. The positive coefficient on \( \eta \) indicates the presence of a direct dependence among the realizations of neighboring edges.
To further demonstrate the benefits of LSGM, we apply it to perform an initial test of the balancing and bandwagoning theories of international alliance formation. Referencing the Concert of Europe as a prime example, proponents of the balancing theory argue that states form alliances to counter-balance the power or their adversaries (Waltz 1979; Walt 1985). An alliance formed by one bloc should, therefore, lead to the formation of a counterbloc. The bandwagoning theory, in contrast, emphasizes other causes for alliance formation, such as ideological similarity or policy alignment (Lake 2009; Morrow 1991). As empirical examples, the proponents of the bandwagoning theory point to the continuation and expansion of NATO during the US’ unipolar moment after the end of the Cold War (Mastanduno 1997). An empirical implication of bandwagoning, then, is that alliance formation should cluster within ideological state: once a pair of states forms an alliance, other ideologically similar states may want to jump on the bandwagon and form alliances with the same states.

The LSGM allows for conducting of empirical tests of such theories by modeling alliances as edges that form in response to realizations of other edges. Bandwagoning would expect that edges will cluster within ideological space (and time). Balancing, in contrast, would expect that edges will form at the diametrically opposite parts of the ideological spectrum:
for example, an alliance formation by the US and its ideological partner should trigger Russia or China to seek allies within its own ideological realm.

We test these predictions using international alliance formation data from the Correlates of War Project (Gibler 2009). The dependent variable is a dichotomous measure of whether a pair of states formed any alliances in a given year.\(^4\) Consistent with the above exposition, the unit of analysis is a network edge (formation of an alliance), and the dataset contains observations of all pairs of states that could potentially form an alliance in a given year between 1947-2001.

Spatial connectivity \(w\) between alliance is measured using international state ideal scores based on United Nations General Assembly voting (Bailey, Strezhnev, and Voeten 2015). We treat each potential ally’s ideal score as a coordinate, which allows us to align all potential alliances in a two-dimensional space. Each \(ij\) cell of the \(w\) matrix thus contained a measure of the Euclidean distance between \(i\) and \(j\) in this two-dimensional ideological space. Shorter distances indicate policy similarity while greater distances indicate policy dissimilarity.

Next, we account for possible short-term temporal dependence (i.e., balancing or bandwagoning may take place in the next year) by including a temporal lag weighed by \(w\). The resulting natural parameter function closely reminds Equation 2 plus the temporal term:

\[
A_{it}(y_{it}) = \log \left( \frac{\kappa_{it}}{1 - \kappa_{it}} \right) + \sum_{s_{jt} \in N_{it}^0} \eta w_{ijt}(y(s_{jt}) - \kappa_{jt}) + \sum_{s_{jt} \in N_{it}^1} \alpha w_{ijt}(1 - \kappa_{jt}), \tag{5}
\]

where \(N_{it}^0\) denotes \(i\)'s neighbors in the contemporaneous time period \(t\), or \(N_{it}^0 = \{y_{jt}\}\), \(N_{it}^1\) denotes \(i\)'s neighbors in the previous time period \(t-1\), or \(N_{it}^1 = \{y_{jt} : y_j(t-1) = 1\}\), \(w_{ijt}\) is the \(ij^{th}\) cell of the connectivity matrix \(w\), and \(\alpha\) is the parameter associated with the temporal lag. The weighed temporal lag assesses whether alliances in the current time period form in

---

\(^4\)Operationalizing the dependent variable as alliance formation rather than simply the presence of an alliance allows for a more direct test of the theories, as presumably states balance and bandwagon in response to recent actions of others. See Gibler and Wolford (2006) for a discussion on differences of the two measures.
response to alliances that formed in the previous time period, weighed by their distance to
$i$ in ideological space.\textsuperscript{5} Every additional alliance that formed in the previous year makes a
positive contribution to the natural parameter function, with alliances that are further away
in the ideological space making the largest contribution. As a result, a positive coefficient on
the temporal term would indicate that the probability of an alliance formation increases in
response to the number of alliances that were formed in the preceding year in the opposite
part of the ideological spectrum. A negative coefficient on the temporal term would indicate
that alliances are more likely to form when when most alliances in the previous year were
formed in the same part of the ideological spectrum.

Finally, we control for additional theories of alliance formation by including several ex-
genous edge-level covariates. Consistent with prior research, we expect that pairs of states
are more likely to form a military alliance if they engage in international trade and are jointly
democratic (Lai and Reiter 2000). We also expect that states are more likely to ally if they
are approximately even in terms of military capabilities (Kimball 2006). Data on interna-
tional trade are obtained from the Correlates of War Project (Barbieri, Keshk, and Pollins
2009), data on military symmetry/asymmetry are obtained from Arena (2016). \textit{Military
Power Ratio} is measured as the ratio of the military capabilities of the more powerful state
in a pair of states to the total military capabilities of the pair, or $\frac{\max(m_1,m_2)}{m_1+m_2}$.

The results of the estimation are presented in Table 1. The coefficient on \textit{Spatial Depen-
dence} is positive, indicating that international alliances tend to form at the opposite sides
of the ideological space. This finding is consistent with the balancing model of alliance for-
formation, in which states balance against the growing power of their adversaries, once they
observe alliance formation among ideologically distant states. The temporal lag is negative,

\textsuperscript{5}Since alliance formation is a rare event (alliance formation equals 1 only in approximately 0.24 percent
of 399,993 observations) and are not uniformly distributed across years (instead clustering in a few years),
it is of little to account for time using a conventional temporal lag (i.e., did an alliance between the same
two states form in the previous year?).
indicating that bandwagoning may happen in the next year. The results suggest evidence for both theories: alliances tend to cluster in ideological space (bandwagoning), yet growing coalitions tend to be countered by formation of opposing blocs (balancing).

The coefficients on the control variables are as expected. Military Power Ratio has a negative effect, suggesting that symmetric alliances are more common than asymmetric ones. Dyadic Trade and Joint Democracy have a positive effect, indicating that trade and similar political institutions enhance military cooperation.

6 Conclusion

This paper introduces an LSGM—a statistical estimator that is a great fit for modeling many theoretical processes that are core to political science applications. We demonstrated the desirable asymptotic properties of the estimator using Monte Carlo simulations and provided an illustrative application to modeling the formation of the international alliance network.

More broadly, we emphasized that the traditional network focus on nodes and relationships among them (edges), central to many current network analysis research, fail to model many processes of interest that occur among edges rather than nodes. Modeling such processes requires going beyond the traditional first degree networks to higher degree networks (e.g., dependencies among edges). The LSGM provides a tool for testing for such dependencies in
a controlled interpretable way.

By treating neighborhoods in continuous rather than discrete terms, typical to statistical applications (e.g., Casleton, Nordman, and Kaiser 2016), the paper helps adapt LSGM to social science applications, where the conceptions of space and similarity (ideological, economic, political) are rarely black and white (Beck, Gleditsch, and Beardsley 2006; Chyzh 2016b; Thies, Chyzh, and Nieman 2016).

LSGM has many potential applications beyond the study of international alliances, offering a statistical approach to modeling any processes that take place among network edges rather than nodes. It applies, for instance, to the study of coalition building in legislatures and other domestic and international bureaucracies. The estimator may also be used in research on information diffusion, or tipping-point processes, such as community outreach related to building support for a particular policy.
References


