Modeling Interstate Alliances with Random Graphs

David J. Marchette,[1*] Carey E. Priebe[2]

[1]Naval Surface Warfare Center,
17320 Dahlgren Rd, Dahlgren, VA 22448, USA
[2]Applied and Mathematical Statistics,
Johns Hopkins University, Baltimore, MD 21218, USA

[*]To whom correspondence should be addressed; E-mail: dmarchette@gmail.com.

A new model of random graphs, the random dot product graph (RDPG) is described. This model is well suited to social networks, since it defines the edges in the graph in terms of a vector of "attributes". The model is illustrated through application to a time series of graphs defined by the alliances between nation states.

We consider a time series of graphs defined in terms of alliances between countries. We investigate data representing alliances between a total of 214 nations collected from 1816 to 2000 (*1*). The data are available at `cow2.la.psu.edu`. For each pair of nations, alliance is coded as in Table 1. There are some missing values in the interstate alliance data, and in this study we treat these as missing edges. Various methods for imputing the missing values could be considered instead. While the edges are colored by alliance type (see Table 1), we will consider only the simplified graph with binary edges: existence or absence of an alliance.

Table 1: Alliance codes in the alliance dataset.

0 or NA	No alliance	
1	Defense pact	intervene militarily if partner attacked
2	Neutrality	remain militarily neutral if partner attacked
3	Nonaggression pact	consultation and/or cooperation in a crisis

We will construct an alliance graph for each year. Thus the vertices of the graph will consist of the 214 nations, and there will be an edge between two vertices if the corresponding nations had an alliance during the year. This produces a time series of graphs. We will describe a new model of random graphs and apply this model to the time series.

We do not believe that most interesting random graphs have independent edges, and so we seek a model that relaxes this requirement. A simple model, which has some interesting (and possibly relevant) properties is the random dot product model.

A random dot product graph (RDGP) is a random graph model containing Erdös-Renyí random graphs (*2*) as a sub-model, in which each vertex v is assigned a vector $x_v \in \mathbb{R}^d$. The probability of an edge from v to w is a function of the dot product of the vectors:

$$p_{ij} = f(x_v' x_w).$$

In this paper we will set f to be a simple threshold:

$$f(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 \leq x \leq 1 \\ 1 & x \geq 1 \end{cases}$$

The vectors x_v are fixed, and new graphs are drawn from the collection of all graphs on n vertices according to the edge probabilities defined above.

It should be noted that in addition to being a generalization of the Erdös-Renyí random graph, the RDPG is a generalization of random intersection graphs (*3*) and a sub-model of latent position models (*4*).

The motivation for these come from social networks applications. It seems reasonable to assume that individuals have a collection of attributes which (to a large extent) define the probabilities of connection with others. For example, in a friendship graph the attributes might correspond to interest areas and personality, and the probability of an edge (friendship) between two individuals is largely driven by the overlap of their interests and the compatibility of their personalities. This is admittedly a simplification. It is not clear that in real problems such attribute vectors can be reliably identified. However, the model is a reasonable approximation, and one can proceed to fit the model without assuming any particular interpretation to the vectors. This will be our philosophy in this work.

One reason for choosing the RDPG over the distance model described in (*4*) is parsimony for a certain class of graphs. We are interested in finding groupings within the vertices, and the RDPG allows us to define these groups as being those which share the same vector. While (with suitable assumptions on the probabilities) the two models can be used to define the same random graphs, it turns out that the dimension d of the vectors needed in the distance model can be much larger for certain cases than is necessary for the dot product model. As we will see, a dimension of $d = 2$ is adequate for the problem we consider.

Although computationally challenging, fitting the vectors to a given graph or set of graphs is relatively straightforward. Scheinerman (*5*) gives a linear algebra method that tries to minimize the Frobenious norm for the edge probabilities, and maximal likelihood is straight forward due to the fact that the probabilities are conditionally independent, given the vectors. Thus, the likelihood is:

$$L = \prod_{i \neq j} \left(x_i' x_j\right)^{a_{ij}} \left(1 - x_i' x_j\right)^{1-a_{ij}} . \tag{1}$$

For a graph $G = (V, E)$, a partition is a collection of subsets of vertices $P = \{P_1, \ldots, P_k\}$ such that $P_i \cap P_j = \emptyset$ for $i \neq j$ and $\cup P_i = V$. We will assume the number of partitions K is known a priori. In a social network context, these groups might be club membership, interest groups, religious affiliation, or some unobserved grouping that one would like to discover. Thus, we are provided with a set of partition labels $\mathcal{L}$, and seek a map $h : V \to \mathcal{L}$. We can rewrite Equation (1) as

$$L = \prod_{i \neq j} \left(\left(\sum_{k=1}^{K} \pi_{i,k} y_k \right)' \left(\sum_{k=1}^{K} \pi_{j,k} y_k \right) \right)^{a_{ij}} \left(1 - \left(\sum_{k=1}^{K} \pi_{i,k} y_k \right)' \left(\sum_{k=1}^{K} \pi_{j,k} y_k \right) \right)^{1-a_{ij}} \tag{2}$$

This admits an EM algorithm:

1. E-step: fix $Y = \{y_i\}$ and choose $P = \{p_{i,k}\}$ to maximize Equation (2).

2. M-step: fix P and choose Y to maximize Equation (2).

3. Repeat until convergence.

By then assigning to each vertex the vector with the largest p, we have an algorithm for fitting the K vectors and making the vertex assignment.

We have several model selection problems in this work. First, one must decide on the dimensionality d of the attribute vectors x_i. In a particular application, this should be in part driven by scientific considerations. One may hypothesize that there are two main factors defining the relationships, and thus choose $d = 2$ and observe the model fit. It should be noted that we want to pick d as small as possible, due to the additional variance in the estimators as the dimension increases. For the purpose of this work we will set $d = 2$. Investigation of higher values of d showed no significant improvement in the models.

Another problem is the choice of K. Many papers have been written about model selection similar to this one in various guises, but the basic idea comes down to choosing an appropriate penalty for model complexity. We will run several values of K and plot the likelihood as a function of K. Note that for $K = 1$ the graph is the Erdös-Renyí random graph and the maximum likelihood estimate is a straightforward calculation.

To analyze the alliance data, we use a moving window of width 20 years, during which the adjacency matrices of the graphs are averaged to estimate the probabilities which are then fed to the model. The step size is 1 so the windows overlap. This windowing stabilizes the fitting algorithm, at the cost of potentially averaging out short term changes. As will be seen, interesting structure and changes are still detected.

The basic statistics of the graphs are depicted in Fig. 1A. Some structure is immediate, such as the changes in density at the years 1887 and 1937. The number of alliances, as well as the number of countries engaging in alliances, is increasing in time, while the density of edges in the graph tends to decrease after a peak in the early 1900s. Fig. 1B. shows several RDPG models fit to the data. The y-axis is the log likelihood scaled by the number of possible edges in the graph. This allows the comparison of the models across time by removing the variability caused by the different number of non-isolated vertices. As can be seen, the Erdös-Renyí model is a poor fit to this data, indicating that there is indeed interesting structure in the alliances. The green curve shows the model in which each country has its own vector. This is the least parsimonious model, and the most difficult to interpret. As can be seen, while the $K = 2, 3$ models are not as good a fit, they are significantly better than the Erdös-Renyí model, and are much more easily interpreted than the general model. Note that there is little difference between the $K = 2, 3$ models until the spike which occurs in 1936, after which the $k = 3$ model tend to do better. This is also the point at which the two models seem to start to become less accurate than the general model. The degree distributions for the graphs are depicted in Fig. 1.C. This can be considered an ensemble of distributions, one for each graph, with the size of the dot indicating the number of graphs that have that value of the distribution. This shows that these graphs tend not to follow a power law distribution.

We investigate the groups in Fig. 2. In Fig. 2A. we show the groups for the $k = 3$ model. Here we are depicting the stability of the groups, indicating how well the group at time t matches the group at time $t - 1$. We have indicated several interesting years on the plot, as well as anotating a few of the countries or groups of countries ("Old Europe" consists of nation states that no longer exist, such as Austria-Hungary and Hesse Grand Ducal). Fig. 2B. depicts the same plot for the $K = 2$ model. These make an interesting contrast. Note that prior to 1936, the $K = 2$ model was much more stable, indicating that (since the log likelihood for the two models are essentially the same) this is the better model during this time period (which agress with the Occam principle). After 1936 the $K = 2$ model is much less stable than the larger model, indicating that the groups are not as well defined in the smaller model. In 1850 the groups consisted of the nations that later formed the Austro-Hungarian empire in one group, the other nations in the other, while in 1900 the groups consist of Austria-Hungary, Germany, Italy, Romania and United Kingdom in one, Argentina, Bolivia, Brazil, China, France, Japan Korea, Peru, Portugal, Russia, Spain, Sweden, Uruguay, Yugoslavia in the other. After 1936, there is a transition period (also called World War II), after which the groups (in the $K = 3$ model) are fairly stable as: the US and its alies, the Soviet Union and its alies, and all others. This is depicted in Fig 3. in which we show the edge probabilities for these three groups. Interestingly enough, in 1982 the groups change so that Russia and the US are in the same group. It is only this year in which this happens. The US alies have strong ties through 1980 and after 1983, while the Russia alies have very week ties in the period from 1962-1981 and after 1983.

This analysis shows that the RDPG model is a powerfull tool for modeling social relationships. By comparing different models, we can learn about the dynamics of the different groups, and better understand the underlying structure of the relationships.

References

1. D. M. Gibler, M. Sarkees, *Journal of Peace Research* **41**, 211 (2004).

2. B. Bollobás, *Random Graphs* (Cambridge University Press, Cambridge, 2001).

3. M. Karonski, K. Singer, E. Scheinerman, *Combinatorics, Probability and Computing* **8**, 131 (1999).

4. P. D. Hoff, A. E. Raftery, M. S. Handcock, *JASA* **97**, 1090 (2002).

5. E. Schenerman, Random dot product graphs (2005). Talk given at IPAM, www.ipam.ucla.edu/schedule.aspx?pc=gss2005.

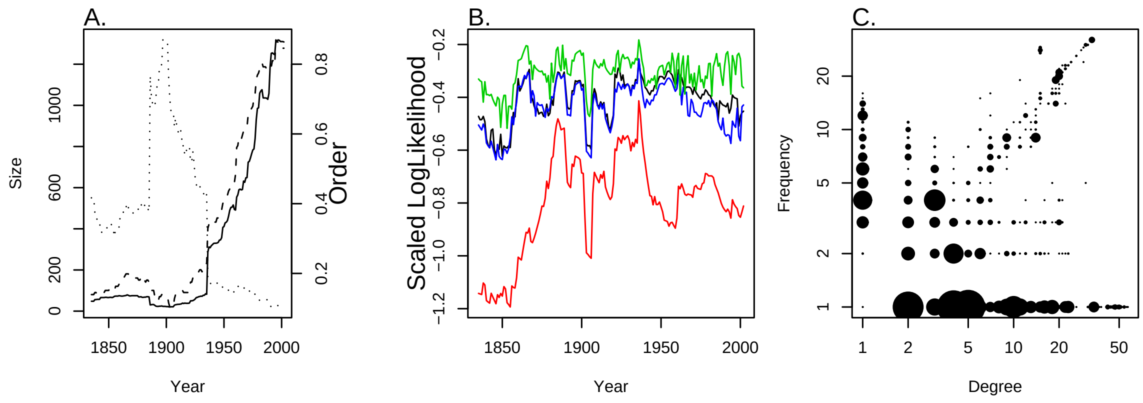
A.
Size
Order
Year
B.
Scaled LogLikelihood
Year
C.
Frequency
Degree

Fig. 1.

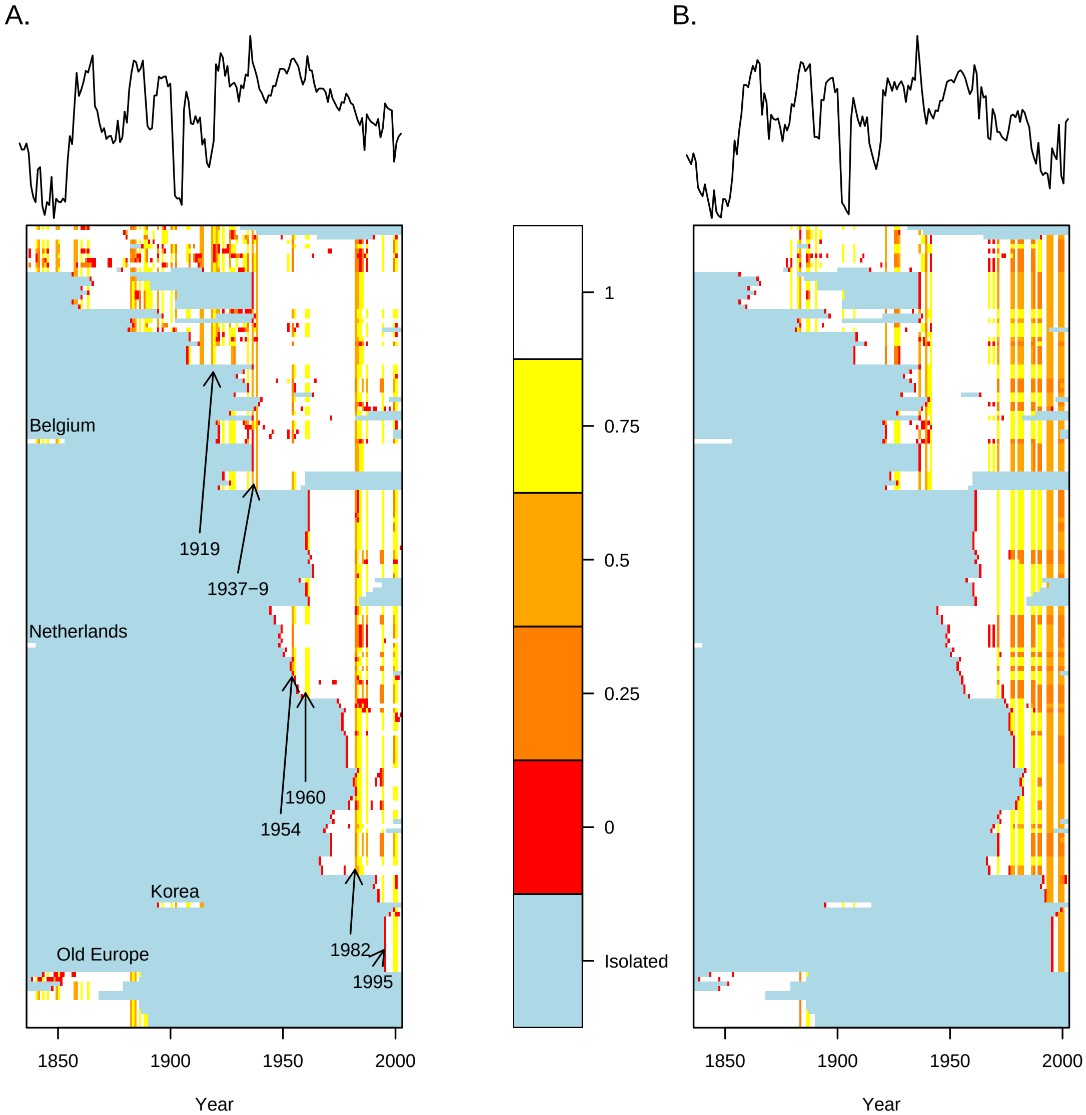
A.
B.
Belgium
Netherlands
Korea
Old Europe
1919
1937−9
1954
1960
1982
1995
1
0.75
0.5
0.25
0
Isolated
1850
1900
1950
2000
Year

Fig. 2.

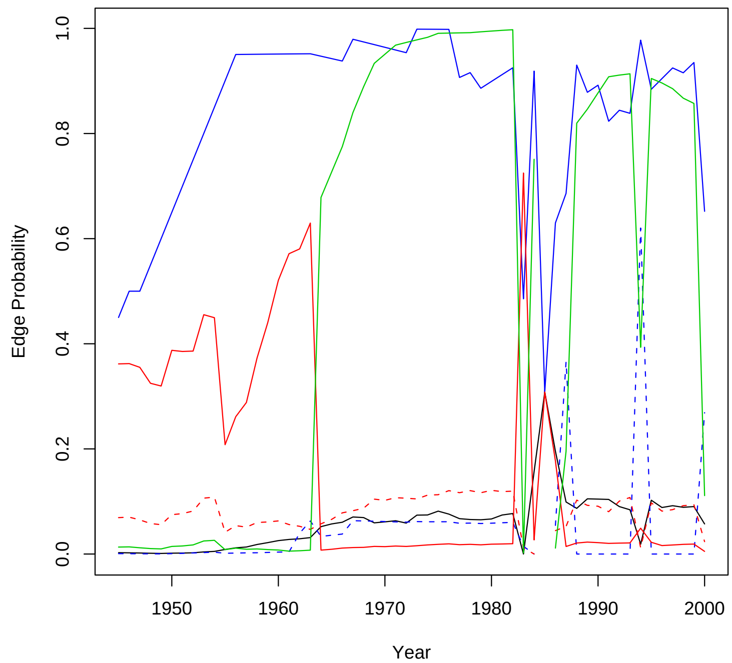

Fig. 3.

Fig. 1.(A) Sizes (solid), orders (dashed) and density (dotted, on a scale from 0.0 to 1) of the graphs defined by the alliances. The large jump in density occurs at 1887 and the drop is at 1937. (B) Scaled loglikelihood values for several values of K. For each graph, the loglikelihood is scaled by the number of possible edges: $|V|(|V|-1)/2$. The green curve corresponds to the maximal model: allowing a separate vector for each vertex. The red curve corresponds to the Erdös-Renyí random graph, which is the $K = 1$ model; the blue curve corresponds to the $K = 2$ model; the black curve corresponds to the $K = 3$ model. (C) Degree distributions of the graphs. For each graph, the degree distribution is plotted, with the size of the dot indicating the number of graphs overplotting that value.

Fig. 2.(A) Group-membership plot for the $K = 3$ model. The x-axis corresponds to time, the y-axis to country. For each year and each country, the color corresponds to the amount of overlap

between the country's group in the current year as compared to that of the previous year: let G_t correspond to the set of countries in the group associated with country c in year t. Then the overlap is defined as: $|G_t(c) \cap G_{t-1}(c)|/|G_t(c) \cup G_{t-1}(c)|$, The scaled loglikelihood is plotted above the image. (B) Group-membership plot for the $K = 2$ model.

Fig. 3. Edge probabilities for the groups containing the US and Rusia between the years 1945 and 2000 for the $K = 3$ model. For each year the vector associated with the United States is identified, as is that associated with Russia, and the third vector (Other) associated with non-isolated vertices. Plotted are the edge probabilities within and between the different groups: US-Russia is in black; within-group US is solid blue, within-group Russia is solid red; within-group Other is solid green; US-Other is dotted blue; Russia-Other is dotted red.