

# Visual Inference for Goodness-of-Fit for A Network Model

Sam Tyner \*

Department of Statistics, Iowa State University  
and

Heike Hofmann

Department of Statistics, Iowa State University

September 30, 2017

## **Abstract**

The text of your abstract. 200 or fewer words.

*Keywords:* statistics, social networks, more

---

\*The authors gratefully acknowledge ...

# 1 Introduction

Talk about goodness-of-fit

Talk about political networks

## 2 Data

The data we use are of collaboration networks in the United States Senate during the 111<sup>th</sup> through 114<sup>th</sup> Congresses, overlapping with Barack Obama’s presidency. These senates began on January 6, 2009 and ended on January 3, 2017.<sup>1</sup> In this network, ties are directed from senator  $i$  to senator  $j$  when senator  $i$  signs on as a cosponsor to the bill that senator  $j$  authored. There are (somewhat surprisingly) many hundreds of ties between senators when they are connected in this way, so we simplify the network by computing a single value for each senator-senator collaboration called the *weighted propensity to cosponsor* (WPC). This value is defined in Gross et al. (2008) as

$$WPC_{ij} = \frac{\sum_{k=1}^{n_j} \frac{Y_{ij(k)}}{c_{j(k)}}}{\sum_{k=1}^{n_j} \frac{1}{c_{j(k)}}} \quad (1)$$

where  $n_j$  is the number of bills in a congressional session authored by senator  $j$ ,  $c_{j(k)}$  is the number of cosponsors on senator  $j$ ’s  $k^{th}$  bill, where  $k \in \{1, \dots, n_j\}$ , and  $Y_{ij(k)}$  is a binary variable that is 1 if senator  $i$  cosponsored senator  $j$ ’s  $k^{th}$  bill, and is 0 otherwise. This measure ranges in value from 0 to 1, where  $WPC_{ij} = 1$  if senator  $i$  is a cosponsor on every one of senator  $j$ ’s bills and  $WPC_{ij} = 0$  if senator  $i$  is never a cosponsor any of senator  $j$ ’s bills.

## 3 Model Definition and Choice

To model family that we apply to the Senate data is the Stochastic Actor-Oriented Model (SAOM) family. This family of models incorporates both network structure and node-level

---

<sup>1</sup>Details of how this data can be downloaded are provided by Franois Briatte at <https://github.com/briatte/congress>

information to describe how a network changes over time. This model was first introduced by Snijders in 1996 (Snijders 1996). The two titular properties of SAOMs, stochasticity and actor-orientation, are crucial to understanding networks as they exist naturally: social networks are ever-changing as relationships decay or grow in seemingly random ways, and most actors in them have characteristics that could affect how they change their ties to other nodes in the network.

In cosponsorship networks, it is well known that party affiliation and reciprocity of relationships are major influences on structure (Nils Ringe and Cho 2016).

### 3.1 Choosing appropriate models

In order to determine which models to explore for application of our novel goodness-of-fit testing method, we first fit many possible models and selected a few significant effects. To determine the significance of the effects, we followed this procedure:

1. Define the simple effects structure of the data: the rate parameters and the outdegree and reciprocity parameters.
2. Add each possible effect one-at-a-time to the model structure, as determined by the effects documentation function.
3. Fit each model to the data and check for convergence.
  - (a) If the model converged, move to 4.
  - (b) If the model did not converge, use the previous fitted values as starting values and repeat 5 times or until convergence, whichever comes first.
4. Test the added parameter for significance using the Wald-type test.
5. Report out the estimate of the additional parameter, its standard error, Wald  $p$  value, and convergence criterion.

After completing the procedure for all model effects, we selected effects to include that converged and had a Wald  $p$ -value of less than 0.10, and seemed to have a reasonable interpretation for our data.

The most significant effect was the jumping transitive triplet (JTT) parameter for the party covariate, which was estimated to be about -6 with a standard error of 0.11. This effect relies on the number of transitive closures formed between two senators from different parties. Due to the divisive nature of American politics, the party effect is dominant. Another significant effect was the same JTT parameter for the sex covariate, with an estimate of about 3 with a standard error of 0.89. Another significant effect was the covariate-related similarity score-weighted transitive triplets for the number of bills authored by a senator. This effect was estimated at about 10 with standard error of 3.9, and the high positive effect suggests senators tend to collaborate with other senators who author about the same number of bills they do. Finally, we found the same party transitive triplet effect was also significant, with a fitted value of 1.3 and standard error of 0.7.

List of effects:

1. jtt - party
2. jtt - sex
3. similiarity trans trip - bills
4. same tt - party

## 4 Simulation Procedure

1. Fit each model 1,000 times and save the parameter estimates, standard errors, and overall convergence criterion.
2. From the models that converged, compute the mean of the 1,000 estimates of each parameter.
3. Use the mean values from step 2 as the parameter values in the model and simulate 100 observations from the model. These simulated networks will become the "true data."
4. To each simulation, refit the model form from which it was simulated.

5. Repeat steps 4-5 for various permutations of model forms and parameter values; e.g. double the parameter values, quadruple the parameter values, change the signs on the parameters, etc.
6. Simulate additional "alternative data" from the model in step 3 and the fitted model in step 4.
7. Place the appropriate node-link diagram of each "true data" observation in a lineup with  $M - 1$  ( $M=12?$   $M=9?$ ) "alternative data" node-link diagrams.
8. Show the lineups to participants, with no participant seeing the same "true data" or "alternative data" plot twice.

## 5 Exploratory plots

## References

- Gross, J. H., Kirkland, J. H. & Shalizi, C. R. (2008), 'Cosponsorship in the U.S. senate: A multilevel two-mode approach to detecting subtle social predictors of legislative support', *Unpublished Manuscript*.
- URL:** [http://www.latinodecisions.com/files/4013/3840/2978/Gross-Kirkland-Shalizi\\_Multilevel-Cosponsorship\\_PolAnlys-submission.pdf](http://www.latinodecisions.com/files/4013/3840/2978/Gross-Kirkland-Shalizi_Multilevel-Cosponsorship_PolAnlys-submission.pdf)
- Nils Ringe, J. N. V. & Cho, W. T. (2016), *The Oxford Handbook of Political Networks*, Oxford University Press, chapter Legislative Networks.
- Snijders, T. A. B. (1996), 'Stochastic actor-oriented models for network change', *Journal of Mathematical Sociology* **21**(1-2), 149–172.
- URL:** <http://www.scopus.com/inward/record.url?eid=2-s2.0-0000413317&partnerID=40&md5=d981a59ed505ebc7fe083b42a8b9a179>