Network Modeling

Temporal Exponential Random Graph Models

Viviana Amati   Jürgen Lerner

Dept. Computer & Information Science
University of Konstanz

Winter 2015/2016
(last updated: December 14, 2015)

Longitudinal network data.

Let $G^{(1)}, \ldots, G^{(T)}$ be $T > 1$ graphs that represent the state of **the same network** at T different points in time $t_1, \ldots, t_T$.

Without any assumptions the joint probability of the T graphs is

$$P(G^{(1)}, \ldots, G^{(T)}) \;=\; \prod_{h=1}^{T} P(G^{(h)} | G^{(h-1)}, \ldots, G^{(1)}) \;.$$

For any timepoint t_h with $1 < h \leq T$ we are interested in the conditional distribution

$$P(G^{(h)} | G^{(h-1)}, \ldots, G^{(1)}) \;,$$

that is the distribution of the network at time t_h conditional on all previous observations.

Longitudinal network data: Markov assumption.

Normally we make the (first-order) **Markov assumption** stating that for any timepoint t_h with $1 < h \leq T$ it is

$$P(G^{(h)} | G^{(h-1)}, \ldots, G^{(1)}) = P(G^{(h)} | G^{(h-1)}) \ .$$

This can be formulated informally as:

> *Given the present, the future is independent of the past.*

Or more simplistic:

> *Same present implies same future (even if the past was different).*

For $k \geq 1$, the k**-order Markov assumption** states that for any timepoint t_h with $k < h \leq T$ it is

$$P(G^{(h)} | G^{(h-1)}, \ldots, G^{(1)}) = P(G^{(h)} | G^{(h-1)}, \ldots, G^{(h-k)}) \ .$$

Temporal ERGM (TERGM).

A TERGM specifies the conditional probability of the graph at time t_h, given the graph at the preceeding timepoint as

$$P_\theta(G^{(h)}|G^{(h-1)}) = \frac{1}{\kappa(\theta)} \exp\left(\sum_{i=1}^{k} \theta_i \cdot s_i(G^{(h)}, G^{(h-1)})\right)$$

with

- $s_i \colon \mathcal{G} \times \mathcal{G} \to \mathbb{R}$ for $i = 1, \ldots, k$ (*statistics*);
- $\theta_i \in \mathbb{R}$ for $i = 1, \ldots, k$ (*parameters*); $\theta = (\theta_1, \ldots, \theta_k)$;
- *normalizing constant* κ defined by

$$\kappa(\theta) = \sum_{G' \in \mathcal{G}} \exp\left(\sum_{i=1}^{k} \theta_i \cdot s_i(G', G^{(h-1)})\right) .$$

Note: the function $P_\theta(G) = P_\theta(G|G^{(h-1)})$ defines a (non-temporal) ERGM on $\mathcal{G}$.

The remainder of these slides discusses the relation between TERGM **statistics** and the **network effects** modeled by them.

Statistics of TERGMs: first distinction.

In general, TERGM statistics $s_i(G^{(h)}, G^{(h-1)})$ are functions of both, the preceeding network $G^{(h-1)}$ and the network $G^{(h)}$ of the current time point. For instance,

- the **number of persistent ties**, i. e., ties of $G^{(h)}$ that were also ties in $G^{(h-1)}$;
- the **number of new ties**, i. e., ties of $G^{(h)}$ that were non-ties in $G^{(h-1)}$.

A TERGM statistic $s_i(G^{(h)}, G^{(h-1)})$ can also be independent of $G^{(h-1)}$, that is, be a function of only $G^{(h)}$. For instance,

- the **number of ties**, i. e., ties of $G^{(h)}$ (irrespective of whether they were ties in $G^{(h-1)}$ or not).

Note that at most two of these three statistics can be used together; which ones are used changes the interpretation of positive vs. negative parameters.

Statistics of TERGMs: a remark.

A TERGM statistic $s_i(G^{(h)}, G^{(h-1)})$ that is independent of $G^{(h)}$, i. e., that is only a function of $G^{(h-1)}$, cannot be used.

Since the conditioning network $G^{(h-1)}$ is fixed for the whole distribution

$$P_\theta(G^{(h)}|G^{(h-1)}) = \frac{1}{\kappa(\theta)} \exp\left(\sum_{i=1}^{k} \theta_i \cdot s_i(G^{(h)}, G^{(h-1)})\right) \quad ,$$

such a statistic would be constant for all networks $G^{(h)}$. (Thus the associated parameter could not be estimated.)

Homophily: tie existence, formation, or duration.

Assume nodes have a `gender` covariate.

A reasonable minimal TERGM to test for homophily would include the four following statistics.

1. The **number of ties**, controling for the density of $G^{(h)}$.
2. The **number of persistent ties**, controling for the inertia of ties when going from $G^{(h-1)}$ to $G^{(h)}$.
3. The **number of same-gender ties**, assessing homophily in $G^{(h)}$.
4. The **number of persistent same-gender ties**, assessing whether same-gender ties tend to last longer than mixed-gender ties.

Exchanging **“new”** with **“persistent”** in statistics (2) and (4) would test whether same-gender ties are more likely to be created than mixed-gender ties.

Dyadic dependence vs. lagged dyadic dependence.

We represent a graph $G^{(h)} = (V, E^{(h)})$ by its adjacency matrix

$$(y^{(h)}_{u,v})_{u,v=1,\ldots,|V|} \quad \text{where} \quad y^{(h)}_{u,v} = 1 \Leftrightarrow (u,v) \in E^{(h)} \quad .$$

Can define two variants of a **reciprocity** statistic:

$$\begin{aligned}
\texttt{recip}(G^{(h)}, G^{(h-1)}) &= \sum_{u \neq v} y^{(h)}_{u,v} \cdot y^{(h)}_{v,u} \\
\texttt{lagged-recip}(G^{(h)}, G^{(h-1)}) &= \sum_{u \neq v} y^{(h)}_{u,v} \cdot y^{(h-1)}_{v,u}
\end{aligned}$$

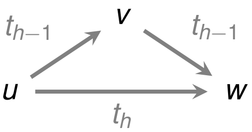
Note: a count in `lagged-recip` does not imply that there ever was a mutual tie at the same point in time.

A count in `lagged-recip` implies that (v, u) gets reciprocated by the reverse tie (u, v) **at the next time point.**

Dyadic dependence vs. lagged dyadic dependence.

Can define two variants of a **transitivity** statistic:

$$\begin{aligned}
\texttt{trans}(G^{(h)}, G^{(h-1)}) &= \sum_{u,v,w} y_{u,v}^{(h)} \cdot y_{v,w}^{(h)} \cdot y_{u,w}^{(h)} \\
\texttt{lagged-trans}(G^{(h)}, G^{(h-1)}) &= \sum_{u,v,w} y_{u,v}^{(h-1)} \cdot y_{v,w}^{(h-1)} \cdot y_{u,w}^{(h)}
\end{aligned}$$



A count in `lagged-trans` implies that the

- 2-path from u over v to w at time t_{h-1}
- gets closed by (u, w) at time t_h.

Characterization of fully independent TERGMs.

If all statistics are sums of products of $(y_{u,v}^{(h)})$ and $(y_{u,v}^{(h-1)})$,

and if all products contain at most **one entry from** $(y_{u,v}^{(h)})$,

then the TERGM is **fully independent.**

For instance, `lagged-trans` leaves dyads independent

$$\begin{aligned}
\texttt{lagged-trans}(G^{(h)}, G^{(h-1)}) &= \sum_{u,v,w} y_{u,v}^{(h-1)} \cdot y_{v,w}^{(h-1)} \cdot y_{u,w}^{(h)} \\
\texttt{trans}(G^{(h)}, G^{(h-1)}) &= \sum_{u,v,w} y_{u,v}^{(h)} \cdot y_{v,w}^{(h)} \cdot y_{u,w}^{(h)}
\end{aligned}$$

but `trans` implies non-independence among incident dyads.

Separable temporal ERGMs.

(A specific sub-class of TERGMs.)

Existence, formation, and dissolution.

The **existence** of ties is the result of two sub-processes:

1. the **formation** process, explaining the rate at which new ties are formed; and
2. the **dissolution** process, explaining how long existent ties tend to last until they get dissolved.

It is plausible that formation and dissolution are shaped by different social mechanisms.

For instance, it could be that *homophily*

1. explains the formation of ties (e. g., by providing opportunities for friendship creation);
2. has no influence on the dissolution of ties (once a tie is established, actor differences might be less important).

A non-temporal ERGM necessarily confounds formation and dissolution; some temporal ERGMs too.

Separable temporal ERGMs (STERGMs).

An STERGM makes the **assumption that formation and dissolution are conditionally independent, given the preceeding network**.

This excludes, for instance, “contracts” of the form

> *I create a tie with you if you break up your tie with that other person.*

The separability assumption would also exclude modeling of marriage ties if monogamy is the norm
(unless time lags are so short that divorce and re-marriage never occurs in the same time-interval).

Formation network and dissolution network.

Let $G^{(h)} = (V, E^{(h)})$ and $G^{(h-1)} = (V, E^{(h-1)})$ be two consecutive observations of a longitudinal network.

- The **formation network** $G^{+} = (V, E^{+})$ has edge set
$$E^{+} = E^{(h)} \cup E^{(h-1)} \quad .$$
- The **dissolution network** $G^{-} = (V, E^{-})$ has edge set
$$E^{-} = E^{(h)} \cap E^{(h-1)} \quad .$$

$G^{(h)}$ can be reconstructed from $G^{(h-1)}$, G^{+} and G^{-} by

$$E^{(h)} = E^{+} \setminus (E^{(h-1)} \setminus E^{-}) = E^{-} \cup (E^{+} \setminus E^{(h-1)}) \quad .$$

Thus, modeling $G^{(h)}$ given $G^{(h-1)}$ is equivalent to modeling G^{+} and G^{-} given $G^{(h-1)}$.

Formation model and dissolution model.

Modeling $G^{(h)}$ given $G^{(h-1)}$ is equivalent to modeling G^{+} and G^{-} given $G^{(h-1)}$; that is

$$P(G^{(h)}|G^{(h-1)}) = P(G^{+}, G^{-}|G^{(h-1)}) \ .$$

By the **separability assumption** we have

$$P(G^{+}, G^{-}|G^{(h-1)}) = P(G^{+}|G^{(h-1)}) \cdot P(G^{-}|G^{(h-1)}) \ .$$

The two models on the right-hand side are called the **formation model** and the **dissolution model**, respectively.

Separable temporal ERGMs.

Assumption: formation and dissolution are independent:

$$P(G^{(h)}|G^{(h-1)}) = P(G^{+}|G^{(h-1)}) \cdot P(G^{-}|G^{(h-1)}) \ .$$

Specifically for TERGMs we have

$$\begin{aligned}
P_{\theta}(G^{(h)}|G^{(h-1)}) &= \frac{1}{\kappa(\theta)} \cdot \exp\left(\sum_{i=1}^{k} \theta_i \cdot s_i(G^{(h)}, G^{(h-1)})\right) \\
P_{\theta^+}(G^{+}|G^{(h-1)}) &= \frac{1}{\kappa^+(\theta^+)} \cdot \exp\left(\sum_{i=1}^{k^+} \theta_i^+ \cdot s_i^+(G^{+}, G^{(h-1)})\right) \\
P_{\theta^-}(G^{-}|G^{(h-1)}) &= \frac{1}{\kappa^-(\theta^-)} \cdot \exp\left(\sum_{i=1}^{k^-} \theta_i^- \cdot s_i^-(G^{-}, G^{(h-1)})\right)
\end{aligned}$$

Note that the normalizing constants are different.

Formation model of an STERGM.

The formation network has edge set $E^+ = E^{(h)} \cup E^{(h-1)}$.

Since $G^{(h-1)}$ is fixed, G^+ is a random draw from

$$\mathcal{G}^+(G^{(h-1)}) = \{G = (V, E) \in \mathcal{G}\, ;\; E^{(h-1)} \subseteq E\} \; .$$

This implies the normalizing constant of the formation model:

$$\begin{aligned} P_{\theta^+}(G^+ | G^{(h-1)}) &= \frac{1}{\kappa^+(\theta^+)} \cdot \exp\left(\sum_{i=1}^{k^+} \theta_i^+ \cdot s_i^+(G^+, G^{(h-1)})\right) \\ \kappa^+(\theta^+) &= \sum_{G \in \mathcal{G}^+(G^{(h-1)})} \exp\left(\sum_{i=1}^{k^+} \theta_i^+ \cdot s_i^+(G, G^{(h-1)})\right) \end{aligned}$$

Formation model explains: **which ties are added to $E^{(h-1)}$?**

Dissolution model of an STERGM.

The dissolution network has edge set $E^{-} = E^{(h)} \cap E^{(h-1)}$.

Since $G^{(h-1)}$ is fixed, G^{-} is a random draw from

$$\mathcal{G}^{-}(G^{(h-1)}) = \{G = (V, E) \in \mathcal{G}\,;\; E \subseteq E^{(h-1)}\} \; .$$

This implies the normalizing constant of the dissolution model:

$$\begin{aligned} P_{\theta^{-}}(G^{-}|G^{(h-1)}) &= \frac{1}{\kappa^{-}(\theta^{-})} \cdot \exp\left(\sum_{i=1}^{k^{-}} \theta_i^{-} \cdot s_i^{-}(G^{-}, G^{(h-1)})\right) \\ \kappa^{-}(\theta^{-}) &= \sum_{G \in \mathcal{G}^{-}(G^{(h-1)})} \exp\left(\sum_{i=1}^{k^{-}} \theta_i^{-} \cdot s_i^{-}(G, G^{(h-1)})\right) \end{aligned}$$

Dissolution model: **which ties are removed from $E^{(h-1)}$?**

Interpretation of parameters in STERGMs.

Example: let θ be the parameter of the `mutual` statistic.

The **formation model** explains the creation of new ties.

- A positive (negative) θ implies that reciprocated ties are created with higher (lower) probability.
- The formation model does not say anything about those dyads that have been edges in the previous network.

The **dissolution model** explains the persistence of old ties.

- A positive θ implies that reciprocated ties are dissolved with lower(!) probability;
  that is, they are persistent with higher probability;
  that is, they tend to have longer duration.
- A negative θ implies that reciprocated ties are dissolved with higher(!) probability;
  that is, they are persistent with lower probability;
  that is, they tend to have shorter duration.
- The dissolution model does not say anything about those dyads that have been non-edges in the previous network.

Fitting STERGMs with the `tergm` package.

General form of the `tergm` function:

```
tergm( dyn.net,
       formation = ~<stat1> + <stat2> +...,
       dissolution = ~<stat1'> + <stat2'> +...,
       estimate = ...)
```

`dyn.net` is a list of networks or a `networkDynamic` object.

The `formation` and `dissolution` formulas can include different terms.

`estimate` gives the estimation method; in our case it is normally `CMLE` (conditional maximum likelihood estimation).

Modeling the four waves of the Knecht Data.

Model assumes that parameters are *homogeneous over time*.

statistics	formation	dissolution
edges	-3.16(0.24)***	-1.06(0.35)**
mutual	+1.54(0.24)***	+2.78(0.43)***
cyclicalties	-0.35(0.12)**	-0.98(0.20)***
transitiveties	+0.80(0.19)***	+1.06(0.22)***
nodematch.gender	+0.71(0.18)***	+0.46(0.26).
absdiff.delinq	-0.43(0.12)***	-0.12(0.15)

Note: delinquency explains the formation but not the dissolution of ties.

Transitive and cyclical ties suggest a latent *hierarchical* structure.