ERGMito

Statistical Models for Small Team Social Networks

George G Vega Yon Aileen Dinkjian Sarah Hamm-Alvarez Kayla de la Haye

University of Southern California
Keck School of Medicine

June, 2020

Keck School of Medicine of USC

What makes a successful team?

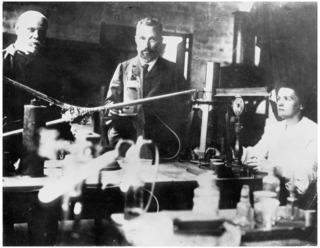


Figure 1 Nobel in Physics 1903: Henri Becquerel, Pierre Curie, and Marie Curie

Figure 2 Nobel in Economics 2019: Abhijit Banerjee, Esther Duflo, and Michael Kremer

Figure 3 US 2019 Women's Soccer team

What makes a successful team?

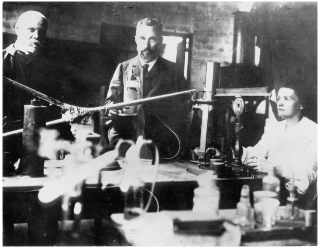


Figure 1 Nobel in Physics 1903: Henri Becquerel, Pierre Curie, and Marie Curie

Figure 2 Nobel in Economics 2019: Abhijit Banerjee, Esther Duflo, and Michael Kremer

Figure 3 US 2019 Women's Soccer team

A lot depends on human relations...

What makes a successful team?

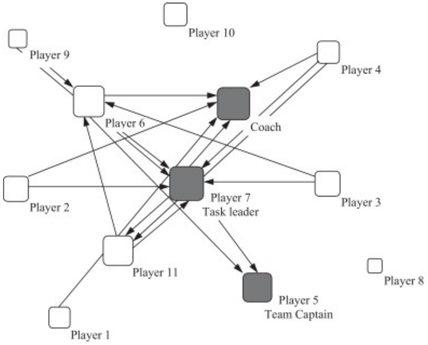


Figure 4 A volleyball team leadership network (Fransen et al., 2015)

What makes a successful team?

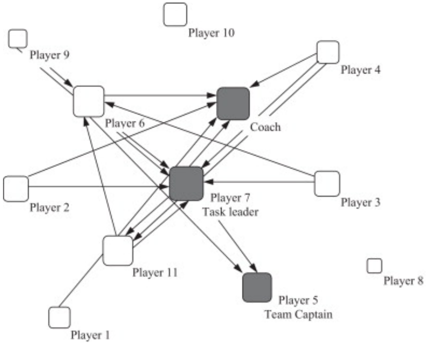
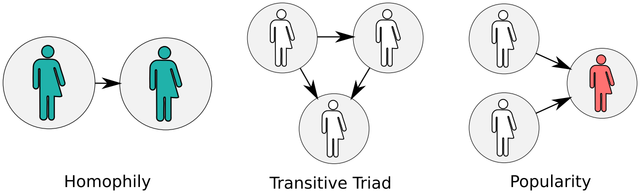


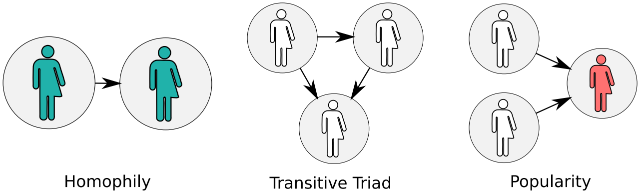
Figure 4 A volleyball team leadership network (Fransen et al., 2015)

What processess govern the formation of social networks?

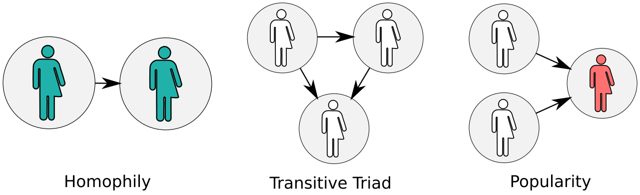
- Is it *homophily*, *social balance* (transitivity), *popularity*?



- Is it *homophily*, *social balance* (transitivity), *popularity*?
- Exponential Family Random Graph Models, aka ERGMs, can help us.



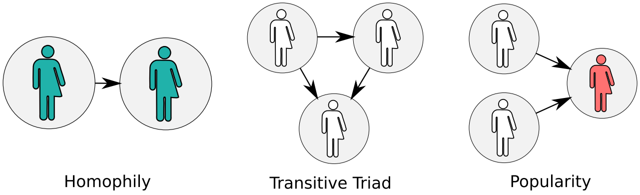
- Is it *homophily*, *social balance* (transitivity), *popularity*?
- Exponential Family Random Graph Models, aka ERGMs, can help us.
- In simple terms, ERGMs allow us to *do*



Keck School of Medicine of USC

- Is it *homophily*, *social balance* (transitivity), *popularity*?
- Exponential Family Random Graph Models, aka ERGMs, can help us.
- In simple terms, ERGMs allow us to *do*

  statistical inference on what network patterns/structures/motifs govern social networks



Fairly complex models for studying networks

A vector of model parameters

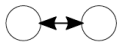
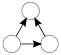
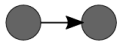
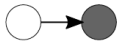
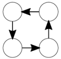
A vector of sufficient statistics

$$\Pr\left(\mathbf{Y} = \mathbf{y} \mid \theta, \mathbf{X}\right) = \frac{\exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}, \mathbf{X}\right)\right\}}{\sum_{\mathbf{y}' \in \mathcal{Y}} \exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}', \mathbf{X}\right)\right\}}, \quad \forall \mathbf{y} \in \mathcal{Y}$$

Observed data

The normalizing constant

All possible networks

Representation	Description
	Mutual Ties (Reciprocity) $\sum_{i \neq j} y_{ij} y_{ji}$
	Transitive Triad (Balance) $\sum_{i \neq j \neq k} y_{ij} y_{jk} y_{ik}$
	Homophily $\sum_{i \neq j} y_{ij} \mathbf{1}\left(x_i = x_j\right)$
	Attribute-receiver effect $\sum_{i \neq j} y_{ij} x_j$
	Four Cycle $\sum_{i \neq j \neq k \neq l} y_{ij} y_{jk} y_{kl} y_{li}$

Exponential Random Graph Models

Fairly complex models for studying networks

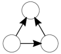
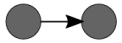
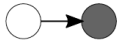
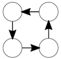
A vector of model parameters

A vector of sufficient statistics

$$\Pr\left(\mathbf{Y} = \underbrace{\mathbf{y} \mid \theta, \mathbf{X}}_{\text{Observed data}}\right) = \frac{\exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}, \mathbf{X}\right)\right\}}{\sum_{\mathbf{y}' \in \mathcal{Y}} \exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}', \mathbf{X}\right)\right\}}, \quad \forall \mathbf{y} \in \mathcal{Y}$$

The normalizing constant

All possible networks

Representation	Description
	Mutual Ties (Reciprocity) $\sum_{i \neq j} y_{ij} y_{ji}$
	Transitive Triad (Balance) $\sum_{i \neq j \neq k} y_{ij} y_{jk} y_{ik}$
	Homophily $\sum_{i \neq j} y_{ij} \mathbf{1}\left(x_i = x_j\right)$
	Attribute-receiver effect $\sum_{i \neq j} y_{ij} x_j$
	Four Cycle $\sum_{i \neq j \neq k \neq l} y_{ij} y_{jk} y_{kl} y_{li}$

We will be using a novel extension that focuses on small networks (**ERGMitos**)

Keck School of Medicine of USC

A vector of model parameters

A vector of sufficient statistics

$$\Pr\left(\underbrace{\mathbf{Y} = \mathbf{y} \mid \theta, \mathbf{X}}_{\text{Observed data}}\right) = \frac{\exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}, \mathbf{X}\right)\right\}}{\sum_{\mathbf{y}' \in \mathcal{Y}} \exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}', \mathbf{X}\right)\right\}}, \quad \forall \mathbf{y} \in \mathcal{Y}$$

The normalizing constant

All possible networks

ERGMs for small networks

A vector of model parameters

A vector of sufficient statistics

$$\Pr\left(\mathbf{Y} = \underbrace{\mathbf{y} \mid \theta, \mathbf{X}}_{\text{Observed data}}\right) = \frac{\exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}, \mathbf{X}\right)\right\}}{\underbrace{\sum_{\mathbf{y}' \in \mathcal{Y}} \exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}', \mathbf{X}\right)\right\}}_{\text{The normalizing constant}}}, \quad \forall \mathbf{y} \in \underbrace{\mathcal{Y}}_{\text{All possible networks}}$$

Small means that we can calculate the normalizing constant exactly, and thus:

ERGMs for small networks

A vector of model parameters

A vector of sufficient statistics

$$\Pr\left(\mathbf{Y}=\mathbf{y}\mid\theta,\mathbf{X}\right)=\frac{\exp\left\{\theta^{\mathbf{t}}s\left(\mathbf{y},\mathbf{X}\right)\right\}}{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left\{\theta^{\mathbf{t}}s\left(\mathbf{y}',\mathbf{X}\right)\right\}},\quad\forall\mathbf{y}\in\mathcal{Y}$$

Observed data

The normalizing constant

All possible networks

Small means that we can calculate the normalizing constant exactly, and thus:

- Improved accuracy (bias),

Keck School of Medicine of USC

A vector of model parameters — A vector of sufficient statistics

$$\Pr\left(\mathbf{Y}=\mathbf{y}\mid\theta,\mathbf{X}\right)=\frac{\exp\left\{\theta^{\mathbf{t}}s\left(\mathbf{y},\mathbf{X}\right)\right\}}{\sum_{\mathbf{y}'\in\mathcal{Y}}\exp\left\{\theta^{\mathbf{t}}s\left(\mathbf{y}',\mathbf{X}\right)\right\}},\quad\forall\mathbf{y}\in\mathcal{Y}$$

Observed data

The normalizing constant

All possible networks

Small means that we can calculate the normalizing constant exactly, and thus:

- Improved accuracy (bias),
- Smaller Type I Error rates,

ERGMs for small networks

A vector of model parameters

A vector of sufficient statistics

$$\Pr\left(\mathbf{Y} = \underbrace{\mathbf{y} \mid \theta, \mathbf{X}}_{\text{Observed data}}\right) = \frac{\exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}, \mathbf{X}\right)\right\}}{\sum_{\mathbf{y}' \in \mathcal{Y}} \exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}', \mathbf{X}\right)\right\}}, \quad \forall \mathbf{y} \in \mathcal{Y}$$

The normalizing constant

All possible networks

Small means that we can calculate the normalizing constant exactly, and thus:

- Improved accuracy (bias),
- Smaller Type I Error rates,
- Faster estimation (about 20 times faster)

ERGMs for small networks

A vector of model parameters

A vector of sufficient statistics

$$\Pr\left(\mathbf{Y} = \mathbf{y} \mid \theta, \mathbf{X}\right) = \frac{\exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}, \mathbf{X}\right)\right\}}{\sum_{\mathbf{y}' \in \mathcal{Y}} \exp\left\{\theta^{\mathbf{t}} s\left(\mathbf{y}', \mathbf{X}\right)\right\}}, \quad \forall \mathbf{y} \in \mathcal{Y}$$

Observed data

The normalizing constant

All possible networks

Small means that we can calculate the normalizing constant exactly, and thus:

- Improved accuracy (bias),
- Smaller Type I Error rates,
- Faster estimation (about 20 times faster)
- A world of possibilities...

(Vega Yon, Salughter and de la Haye (2019))

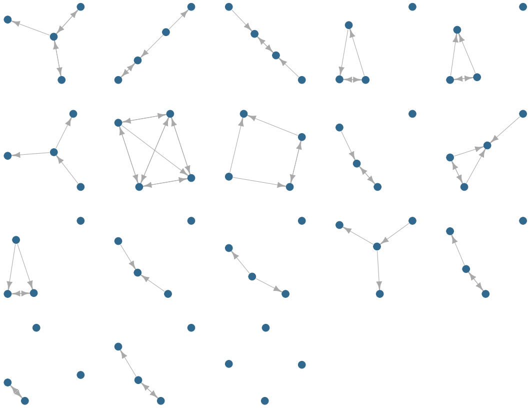


Figure 5 Advice seeking networks observed in an experiment.

Some results on the emergence of advice-seeking networks in small teams

1. Higher Social Perceptiveness *means* more connections.
2. The friend of my friend is my friend.

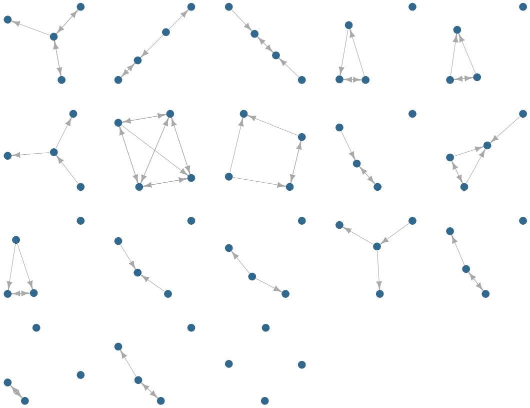


Figure 5 Advice seeking networks observed in an experiment.

Some results on the emergence of advice-seeking networks in small teams

1. Higher Social Perceptiveness *means* more connections.
2. The friend of my friend is my friend.
3. Women ask for advice.

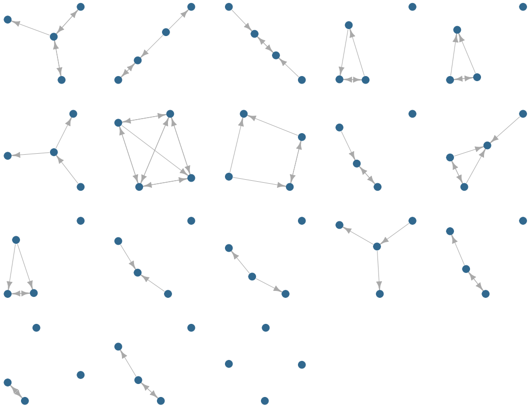


Figure 5 Advice seeking networks observed in an experiment.

Concluding remarks

- The whole is not the sum of the parts

Concluding remarks

- The whole is not the sum of the parts
- Understanding local-macro dependencies is key.

Concluding remarks

- The whole is not the sum of the parts
- Understanding local-macro dependencies is key.
- We can use ERGMitos to elucidate the processes that govern small networks.

Concluding remarks

- The whole is not the sum of the parts
- Understanding local-macro dependencies is key.
- We can use ERGMitos to elucidate the processes that govern small networks.
- Insight into **team networks** is valuable for understanding team **process and performance**, and improving **team design**

ERGMito

Statistical Models for Small Team Social Networks

George G Vega Yon Aileen Dinkjian Sarah Hamm-Alvarez Kayla de la Haye

University of Southern California
Keck School of Medicine

June, 2020

Keck School of Medicine of USC

Thanks!