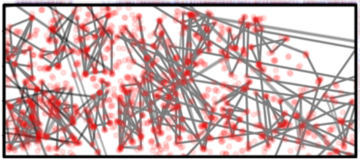
Temporal Exponential Random Graphs with Vertex Dynamics

Zack W Almquist

University of Minnesota

June 15th, 2018
EPIC - SNA
Columbia University



Road Map

1) Overview of Population/Network Processes
2) Examples
3) A Model for Dynamic Networks with Population Processes
4) Empirical Example
5) Further Reading

Who Shows up to the Party?

Who Shows up to the Party?

Dynamic Networks with Population Processes

Dynamic Populations

- Educational settings
  - Classrooms
  - Yearly turnover
- Human societies
  - Communication
  - Interaction
- Demography
  - Forecasting
  - Birth/Death/Mobility
- Organizations
  - Org. Demography
  - Org. Ecology

- e.g., Coleman (1961); Hodgkinson (1985); Feldman (1972)
- e.g., Mayhew and Levinger (1976); Chase (1980)
- e.g., Lee (1987); Lee and Carter (1992)
- e.g., Hannan and Freeman (1977, 1984); Anderton et al. (1983); Romo and Schwartz (1995)

Dynamic Networks

- Educational settings
  - Friendship over time
  - Influence networks (e.g., smoking)
  - e.g., Newcomb (1961); Nordlie (1958); Mercken et al. (2010)
- Human societies
  - Communication
  - Interaction
  - e.g., van de Rijt (2011); Doğan et al. (2009); Buskens and van de Rijt (2008)
- Demography
  - Sexual contact networks
  - Kinship networks
  - e.g., Entwisle et al. (2007); Fischer (1982); Boyd (1989)
- Organizations
  - Collaboration
  - Composition
  - e.g., Carley (1999); Butts (2009); Mintz and Schwartz (1981)

Dynamic Networks with Dynamic Populations

- Sexual contact networks
  - HIV Transmission
  - e.g., Add Health, Uganda
  - e.g., Bearman et al. (2004); Morris and Kretzschmar (1997)
- Organizational collaboration during disasters
  - Disaster response
  - e.g., Hurricane Katrina, 9/11
  - e.g., Butts et al. (2012); Butts (2008)
- Informal groups
  - Interaction
  - e.g., movements, gangs
  - e.g., Freeman et al. (1988); Freeman (1992)

Overview of Classic Population and Network Processes

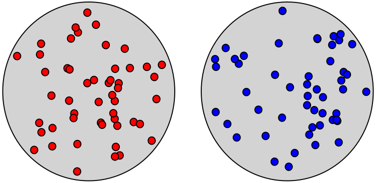
This work looks to combine demographic and network processes, in order to better understand their interaction on social phenomena (e.g., organizational collaboration, sexual contact networks, informal group processes . . .)

Overview of Classic Population and Network Processes

1) Population Processes
2) Network Processes

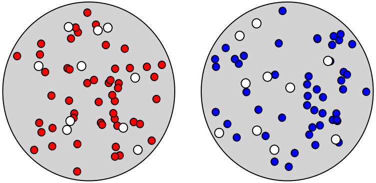
Population Processes

- Fertility
  - Entry
- Mortality
  - Exit
- Migration
  - Mobility



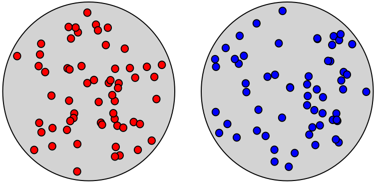
Population Processes

- Fertility
  - Entry
- Mortality
  - Exit
- Migration
  - Mobility



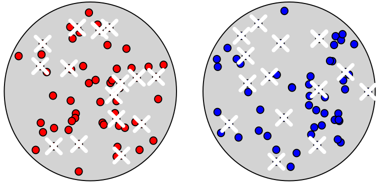
Population Processes

- Fertility
  - Entry
- Mortality
  - X Exit
- Migration
  - Mobility



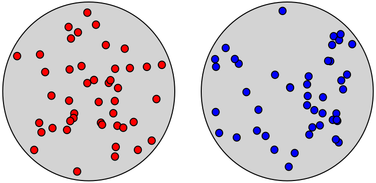
Population Processes

- Fertility
  - Entry
- Mortality
  - X Exit
- Migration
  - Mobility



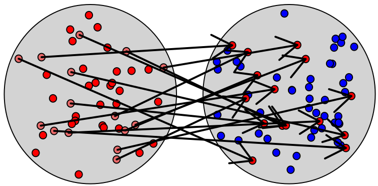
Population Processes

- Fertility
  - Entry
- Mortality
  - Exit
- Migration
  - Mobility



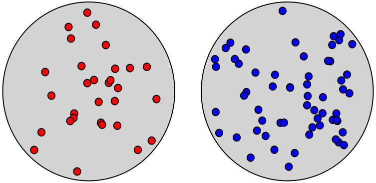
Population Processes

- Fertility
  - Entry
- Mortality
  - Exit
- Migration
  - Mobility



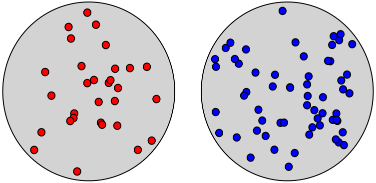
Population Processes

- Fertility
  - Entry
- Mortality
  - Exit
- Migration
  - Mobility



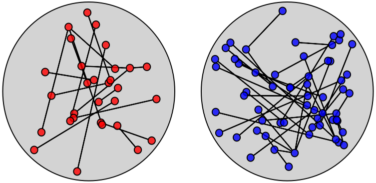
Network Processes

- Tie Formation
  - Adding ties
- Tie Dissolution
  - Removing ties
- Structural Patterns
  - e.g., Triadic closure



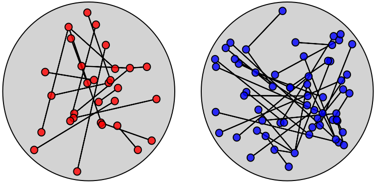
Network Processes

- Tie Formation
  - Adding ties
- Tie Dissolution
  - Removing ties
- Structural Patterns
  - e.g., Triadic closure



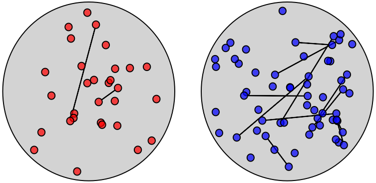
Network Processes

- Tie Formation
  - Adding ties
- Tie Dissolution
  - Removing ties
- Structural Patterns
  - e.g., Triadic closure



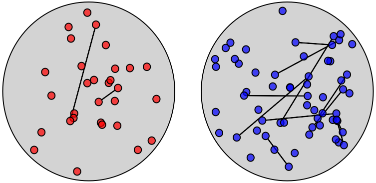
Network Processes

- Tie Formation
  - Adding ties
- Tie Dissolution
  - Removing ties
- Structural Patterns
  - e.g., Triadic closure



Network Processes

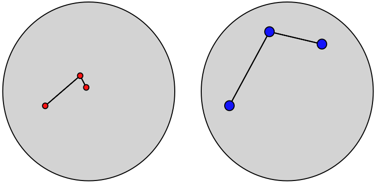
- Tie Formation
  - Adding ties
- Tie Dissolution
  - Removing ties
- Structural Patterns
  - e.g., Triadic closure



Classic Network Analysis

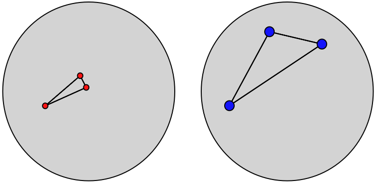
Network Processes

- Tie Formation
  - Adding ties
- Tie Dissolution
  - Removing ties
- Structural Patterns
  - e.g., Triadic closure



Network Processes

- Tie Formation
  - Adding ties
- Tie Dissolution
  - Removing ties
- Structural Patterns
  - e.g., Triadic closure



Road Map

1) Overview of Population/Network Processes
2) Examples
3) A Model for Dynamic Networks with Population Processes
4) Empirical Example
5) Further Reading

But What Happens When These Processes Interact?

- Sexual contact networks
  - Disease spreads via contact
  - This spread in turn influences the population process
  - This in turn affects the network
- Organizational response to disasters
  - Mass convergence
  - Emergent coordination tasks
  - Different organizations show up on different days
- Informal groups
  - Communication/interaction
  - Information passing
  - Different individuals on different days

Road Map

1) Overview of Population/Network Processes
2) Examples
3) A Model for Dynamic Networks with Population Processes
4) Empirical Example
5) Further Reading

Modeling Dynamic Networks with Populations Dynamics

- Empirical motivation
- Model based motivation
- Statistical models for networks
- A statistical model for dynamic networks with population dynamics

Classic Approach

Treat Networks as fixed populations

- Organizational Collaboration
  - Treat every organization as appearing on every day

Org. Collaboration

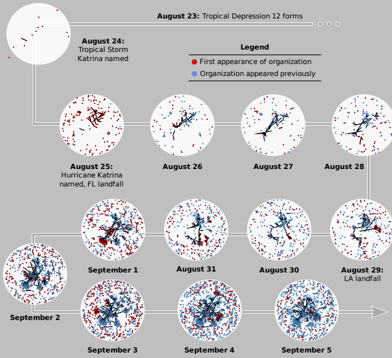


Figure from: Butts et al. (2012)

Population Dynamics in Networks

Why is This a Problem?

- Major problems for realistic models: Population size and composition dynamics
  - Little work in the network methods literature; prior work limited to very stylized cases
- Extreme problem for phenomena like mass convergence in disasters, conversation dynamics, disease spread, etc.

Org. Collaboration

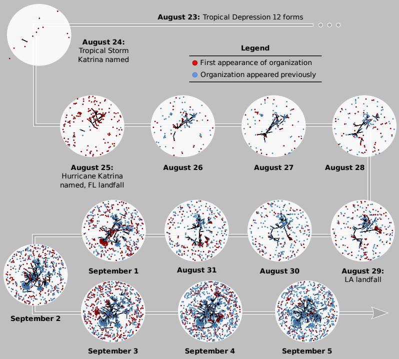
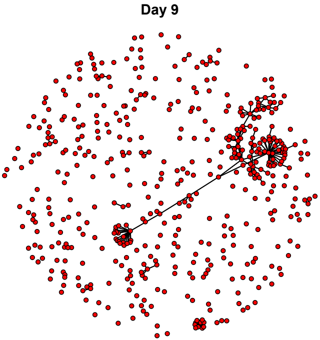


Figure from: Butts et al. (2012)

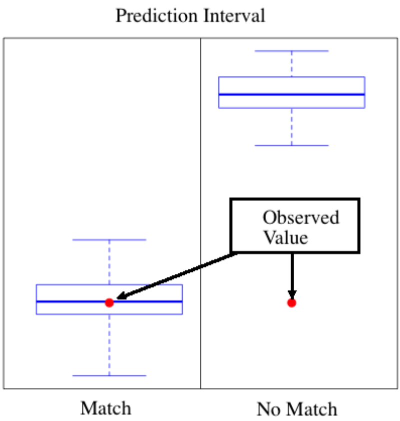
What Happens When the Population Process is Ignored?



What Happens When the Population Process is Ignored?

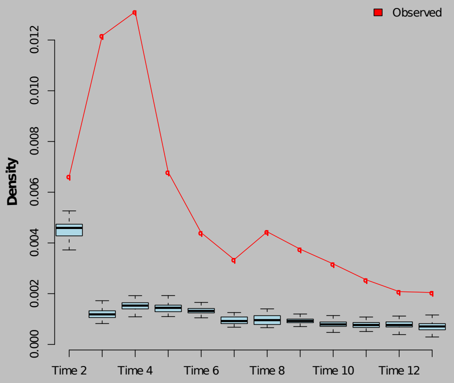
Aside

- Density $= \frac{\text{\# Observed Edges}}{\text{\# Possible Edges}}$
- Prediction Interval (boxplot)



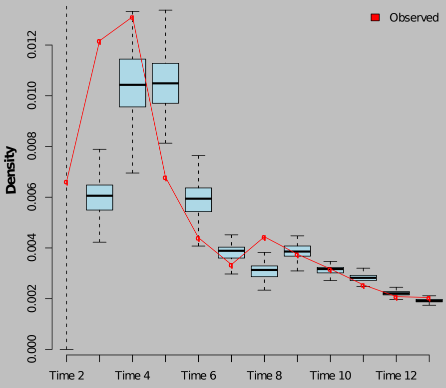
What Happens When the Population Process is Ignored?

Naive Population Model



What Happens When the Population Process is Ignored?

Perfect Population: Who Shows Up

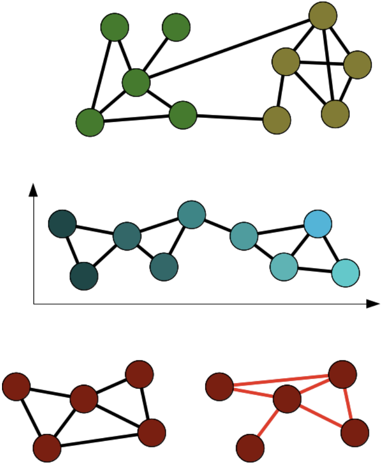


Network Modeling: Predict the formation and structure of social networks

- Many examples
  - Conditional uniform graphs, Bernoulli graphs (baseline models)
  - Holland and Leinhardt's p_1 (within dyad dependence)
  - Degree distribution models, growth models, etc.
- Exponential Random Graph Models
  - Draw on theory of statistical exponential families
  - A way of representing and working with new and existing models

Initial Intuition: Factors in Tie Formation

- All ties are not equally probable
  - Chance of an (i, j) edge may depend on properties of i and j
  - Can also depend on other (i, j) relationships
- Some examples:
  - Homophily
  - Propinquity
  - Multiplexity



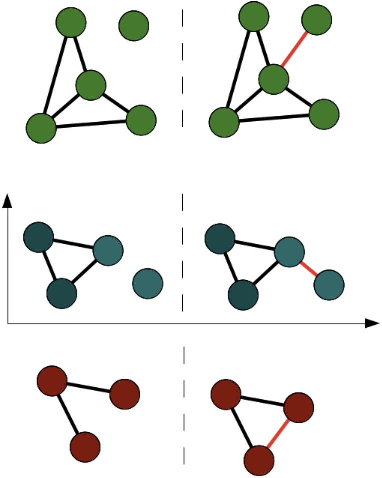
Exponential Random Graph Model (ERGM)

$$\Pr(G = g \mid X) = \frac{\exp\left(\theta^T s(g, X)\right)}{\sum_{g' \in \mathcal{G}} \exp\left(\theta^T s(g', X)\right)}$$

- G is a random network
- g belongs to the support, $\mathcal{G}$
- θ is a vector of parameters
- $s(g)$ is a known vector of graph statistics on g

Initial Intuition: Factors in Tie Formation over Time

- All ties are not equally probable
  - Chance of an (i, j) edge may depend on properties of i, j and t
  - Can also depend on other (i, j, t) relationships
- Some examples:
  - Homophily
  - Propinquity
  - Triadic/clustering effects
    - Shared partner effects



Temporal Exponential Random Graph Model (TERGM)

$$\Pr(G_t = g_t \mid G_{t-1}, \dots, G_{t-k}, X) = \frac{\exp\left(\theta^T s(g_t, G_{t-1}, \dots, G_{t-k}, X)\right)}{\sum_{g' \in \mathcal{G}} \exp\left(\theta^T s(g_t', G_{t-1}, \dots, G_{t-k}, X)\right)}$$

- G_t is independent of $G_1, \dots, G_{t-2}$ given G_{t-1} (Hanneke et al., 2010)
- Arbitrary lags $(G_{t-1}, \dots, G_{t-k})$
- g_t belongs to the support, $\mathcal{G}$
- s here is a vector of real-valued sufficient statistics
- θ is a vector of parameters
- Notice that the denominator is intractable

Temporal Exponential Random Graph Model (TERGM)
(e.g., Robins et al., 1999; Hanneke and Xing, 2007)

$\Pr(G_t | G_{t-1}, \dots, G_{t-k})$

$\Pr(V_t | Z_{t-1}, \dots Z_{t-k}) \Pr(Y_t | V_t, Z_{t-1}, \dots, Z_{t-k})$ $\quad$ $\Pr(Y_t | Y_{t-1}, \dots, Y_{t-k})$

Dynamic Network logistic-Regression (DNR) w/ Vertex Dynamics (Almquist and Butts, 2012) $\quad$ Classic DNR

Temporal Exponential Random Graph Model (TERGM)
(e.g., Robins et al., 1999; Hanneke and Xing, 2007)

$$\Pr(G_t | G_{t-1}, \ldots, G_{t-k})$$

$$\Pr(V_t | Z_{t-1}, \ldots Z_{t-k}) \Pr(Y_t | V_t, Z_{t-1}, \ldots, Z_{t-k}) \qquad \Pr(Y_t | Y_{t-1}, \ldots, Y_{t-k})$$

Dynamic Network logistic-Regression (DNR)
w/ Vertex Dynamics (Almquist and Butts, 2012)

Classic DNR

Key Insight

- Comes from separating the population process (the vertex set) from the network process ...

Dynamic Network logistic-Regression w/ Vertex Dynamics

- Note that this can be recast as a logistic regression problem
- Formally,

$$\begin{aligned}
&\Pr(V_t \mid Z_{t-1}, \ldots, Z_{t-k}, X) \\
&\qquad = \prod_{i=1}^{n} B\left(\mathbb{I}(v_i \in V_t) \middle| \operatorname{logit}^{-1}\left(\psi^T w(i, Z_{i-1}, \ldots, Z_{i-k}, X)\right)\right) \\
&\Pr(Y_t \mid V_t, Z_{t-1}, \ldots, Z_{t-k}, X) \\
&\qquad = \prod_{(i,j)\in V_t \times V_t}^{n} B\left(Y_{ij,t} \middle| \operatorname{logit}^{-1}\left(\theta^T u(i, j, V_t, Z_{i-1}, \ldots, Z_{i-k}, X)\right)\right), \\
&\text{where } Z_{t-k} = (Y_{t-k}, V_{t-k})
\end{aligned}$$

A flexible framework (TERGM subfamily) that readily scales to thousands of vertices/time points

A Dynamic Network Model with Population Dynamics

Dynamic Network logistic-Regression w/ Vertex Dynamics

- Note that this can be recast as a logistic regression problem
- Formally,

$$\log\left(\frac{\Pr(\mathbb{I}(V_i \in V_t) = 1)}{\Pr(\mathbb{I}(V_i \in V_t) = 0)}\right) = \psi^T w_i(Z_{t-1}, \ldots, Z_{t-k}, X)$$

$$\log\left(\frac{\Pr(Y_{ij,t} = 1)}{\Pr(Y_{ij,t} = 0)}\right) = \theta^T u_{ij}(V_t, Z_{t-1}, \ldots, Z_{t-k}, X)$$

A flexible framework (TERGM subfamily) that readily scales to thousands of vertices/time points

Model Adequacy Assessment for DNR w/ Vertex Dynamics

Model Assessment by Prediction

i) Model Selection

- e.g., BIC, AIC ...

ii) Simulation analysis of graph statistics (a.k.a. Graph Level Indices (GLI); Hunter et al. (2008))

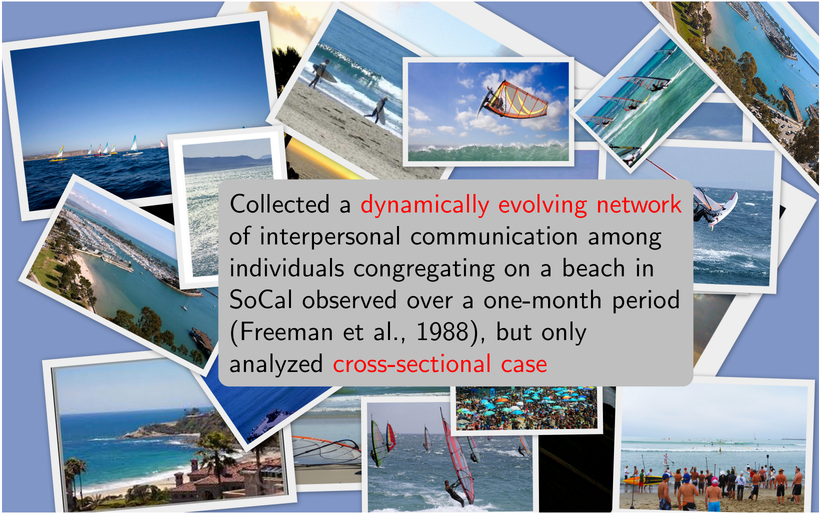
- Simulate one-step-ahead prediction for each time point and compare against observed value

Road Map

1) Overview of Population/Network Processes
2) Examples
3) A Model for Dynamic Networks with Population Processes
4) **Empirical Example**
5) **Further Reading**

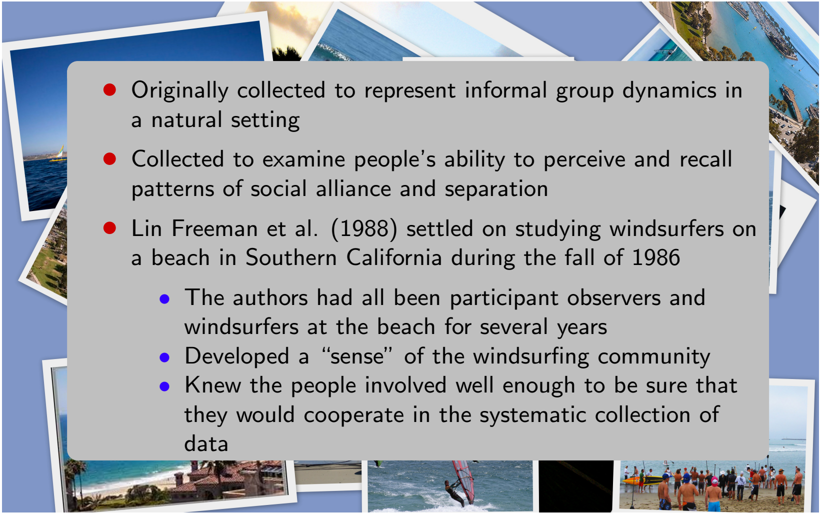

Going to the Beach!

Going to the Beach!



Collected a dynamically evolving network of interpersonal communication among individuals congregating on a beach in SoCal observed over a one-month period (Freeman et al., 1988), but only analyzed cross-sectional case

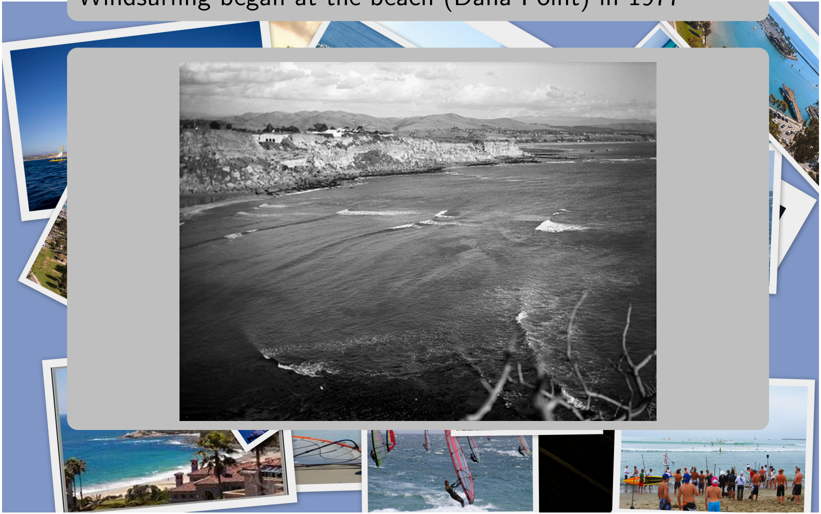
The Data: Social Relationships on the Beach



- Originally collected to represent informal group dynamics in a natural setting
- Collected to examine people's ability to perceive and recall patterns of social alliance and separation
- Lin Freeman et al. (1988) settled on studying windsurfers on a beach in Southern California during the fall of 1986
  - The authors had all been participant observers and windsurfers at the beach for several years
  - Developed a “sense” of the windsurfing community
  - Knew the people involved well enough to be sure that they would cooperate in the systematic collection of data

The Data: Social Relationships on the Beach

Windsurfing began at the beach (Dana Point) in 1977



The Data: Social Relationships on the Beach

Started slow, but gradually, as more people got into the sport, a group of "regulars" was established

The Data: Social Relationships on the Beach

This produced insulated group dynamics and a “community” bound together by a common activity

The Data: Social Relationships on the Beach

By 1983, stable membership of about 25 to 30 (Group1)

The Data: Social Relationships on the Beach

That same year, windsurfing experienced an upsurge of interest and beginners took it up in increasing numbers. The established windsurfers began to refuse to welcome new windsurfers

The Data: Social Relationships on the Beach

At first, the new wave of beginners were social isolates. Then, gradually, as their numbers grew, these newcomers began to drift together to form a second informal group (Group2)

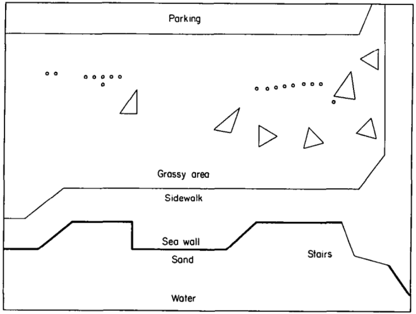


Fig. 1. Location of old group and new group on the beach at noon on 27 September

"Figure 1 shows where people (and their sails) were located on the beach at noon on 27 September and illustrates this [two groups] separation" (Freeman et al., 1988)

Members of each group seemed, to some degree, to limit their interaction to fellow group members

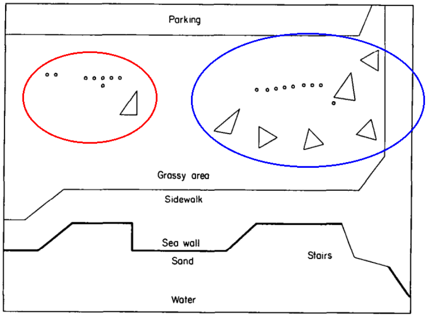
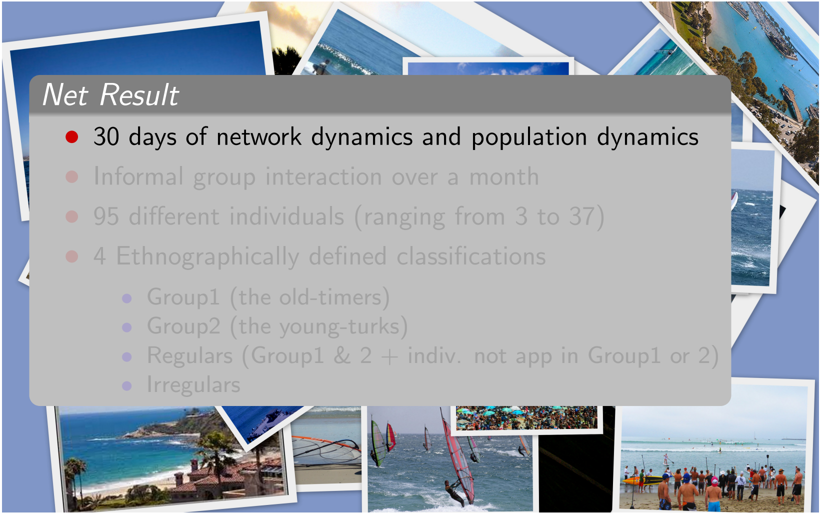


Fig. 1. Location of old group and new group on the beach at noon on 27 September

"Figure 1 shows where people (and their sails) were located on the beach at noon on 27 September and illustrates this [two groups] separation" (Freeman et al., 1988)



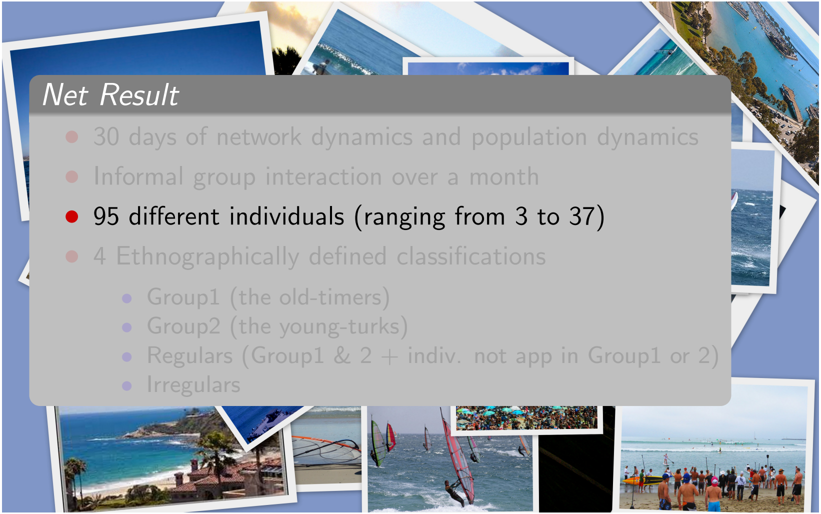
Net Result

- 30 days of network dynamics and population dynamics
- Informal group interaction over a month
- 95 different individuals (ranging from 3 to 37)
- 4 Ethnographically defined classifications
    - Group1 (the old-timers)
    - Group2 (the young-turks)
    - Regulars (Group1 & 2 + indiv. not app in Group1 or 2)
    - Irregulars

The Data: Social Relationships on the Beach

Net Result

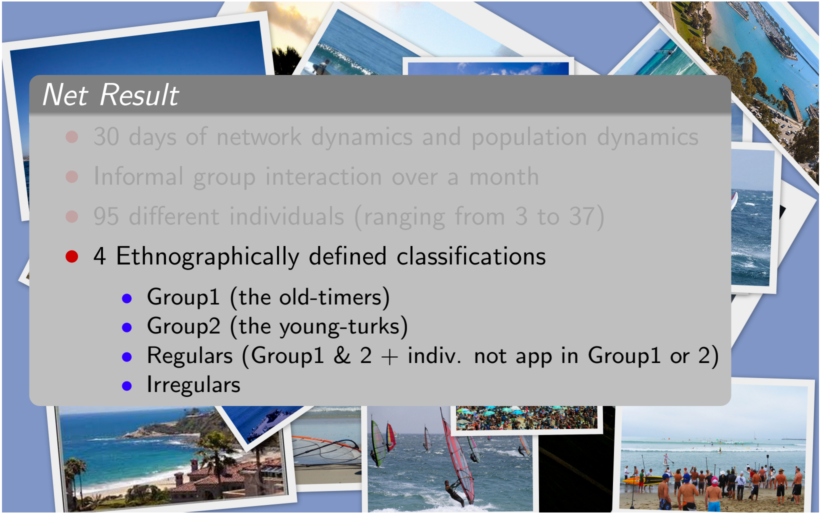
- 30 days of network dynamics and population dynamics
- Informal group interaction over a month
- 95 different individuals (ranging from 3 to 37)
- 4 Ethnographically defined classifications
  - Group1 (the old-timers)
  - Group2 (the young-turks)
  - Regulars (Group1 & 2 + indiv. not app in Group1 or 2)
  - Irregulars



Net Result

- 30 days of network dynamics and population dynamics
- Informal group interaction over a month
- 95 different individuals (ranging from 3 to 37)
- 4 Ethnographically defined classifications
  - Group1 (the old-timers)
  - Group2 (the young-turks)
  - Regulars (Group1 & 2 + indiv. not app in Group1 or 2)
  - Irregulars

The Data: Social Relationships on the Beach



Net Result

- 30 days of network dynamics and population dynamics
- Informal group interaction over a month
- 95 different individuals (ranging from 3 to 37)
- 4 Ethnographically defined classifications
  - Group1 (the old-timers)
  - Group2 (the young-turks)
  - Regulars (Group1 & 2 + indiv. not app in Group1 or 2)
  - Irregulars

Application of Dynamic Network logistic-Regression

- Note that this is the first time this data set has been analyzed in a dynamic way
- DNR with Vertex Dynamics allows one to ask interesting questions about group dynamics from observational data
- DNR with Vertex Dynamics allows one to examine the interplay between network dynamics and population dynamics

Mech. of Dynamic Interpersonal Communication

1) Population Process Model: What Influences Who Shows Up?
2) Network Model: What Influences Who Talks to Whom?

Population Process Model: What Influences Who Shows Up?

- Regularity effect (dummy for Regulars and Group1)
- Inertial effect (lag term)
- Triadic effect (3-cycle or triadic closure)
- Seasonality (dummy for each day of the week)

Mech. of Dynamic Interpersonal Communication

Population Process Model: What Influences Who Shows Up?

- Regularity effect (dummy for Regulars and Group1)
- Inertial effect (lag term)
  - e.g., Corten and Buskens (2010)
- Triadic effect (3-cycle or triadic closure)
- Seasonality (dummy for each day of the week)

Population Process Model: What Influences Who Shows Up?

- Regularity effect (dummy for Regulars and Group1)
- Inertial effect (lag term)
- Triadic effect (3-cycle or triadic closure)
  - e.g., Faust (2010); Wasserman and Faust (1994)
- Seasonality (dummy for each day of the week)

Mech. of Dynamic Interpersonal Communication

Population Process Model: What Influences Who Shows Up?

- Regularity effect (dummy for Regulars and Group1)
- Inertial effect (lag term)
- Triadic effect (3-cycle or triadic closure)
- Seasonality (dummy for each day of the week)
  - e.g., Butts and Cross (2009); Baker (1984)

Mech. of Dynamic Interpersonal Communication

Population Process Model: What Influences Who Shows Up?

- Regularity effect (dummy for Regulars and Group1)
- Inertial effect (lag term)
- Triadic effect (3-cycle or triadic closure)
- Seasonality (dummy for each day of the week)

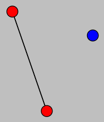
Mech. of Dynamic Interpersonal Communication

Network Model: What Influences Who Talks to Whom?

- Regularity of beach use (Mixing terms for regulars)
- Indiv. propensity effects for regularly occurring indiv. (dummy)
- Contagious participation ($\log(n_t)$)
- Inertial network effects (lag term)
- Embeddedness (log of 9-cycle effect)
- Seasonality (dummy for each day of the week)

Network Model: What Influences Who Talks to Whom?

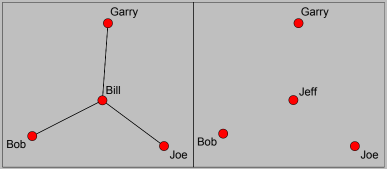
- Regularity of beach use (Mixing terms for regulars)
  - e.g., McPherson et al. (2001)



- Indiv. propensity effects for regularly occurring indiv. (dummy)
- Contagious participation ($\log(n_t)$)
- Inertial network effects (lag term)
- Embeddedness (log of 9-cycle effect)
- Seasonality (dummy for each day of the week)

Network Model: What Influences Who Talks to Whom?

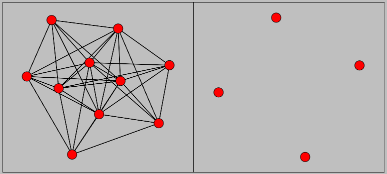
- Regularity of beach use (Mixing terms for regulars)
- Indiv. propensity effects for regularly occurring indiv. (dummy)



- Contagious participation ($\log(n_t)$)
- Inertial network effects (lag term)
- Embeddedness (log of 9-cycle effect)
- Seasonality (dummy for each day of the week)

Network Model: What Influences Who Talks to Whom?

- Regularity of beach use (Mixing terms for regulars)
- Indiv. propensity effects for regularly occurring indiv. (dummy)
- Contagious participation ($\log(n_t)$)

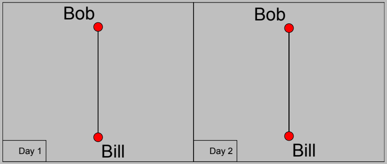


- Inertial network effects (lag term)
- Embeddedness (log of 9-cycle effect)
- Seasonality (dummy for each day of the week)

Mech. of Dynamic Interpersonal Communication

Network Model: What Influences Who Talks to Whom?

- Regularity of beach use (Mixing terms for regulars)
- Indiv. propensity effects for regularly occurring indiv. (dummy)
- Contagious participation ($\log(n_t)$)
- Inertial network effects (lag term)

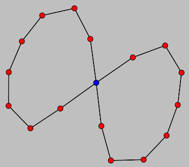


- Embeddedness (log of 9-cycle effect)
- Seasonality (dummy for each day of the week)

Mech. of Dynamic Interpersonal Communication

Network Model: What Influences Who Talks to Whom?

- Regularity of beach use (Mixing terms for regulars)
- Indiv. propensity effects for regularly occurring indiv. (dummy)
- Contagious participation ($\log(n_t)$)
- Inertial network effects (lag term)
- Embeddedness (log of 9-cycle effect)
  - e.g., Granovetter (1985)

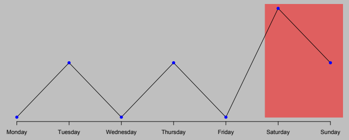


- Seasonality (dummy for each day of the week)

Mech. of Dynamic Interpersonal Communication

Network Model: What Influences Who Talks to Whom?

- Regularity of beach use (Mixing terms for regulars)
- Indiv. propensity effects for regularly occurring indiv. (dummy)
- Contagious participation ($\log(n_t)$)
- Inertial network effects (lag term)
- Embeddedness (log of 9-cycle effect)
- Seasonality (dummy for each day of the week)



Mech. of Dynamic Interpersonal Communication

Network Model: What Influences Who Talks to Whom?

- Regularity of beach use (Mixing terms for regulars)
- Indiv. propensity effects for regularly occurring indiv. (dummy)
- Contagious participation ($\log(n_t)$)
- Inertial network effects (lag term)
- Embeddedness (log of 9-cycle effect)
- Seasonality (dummy for each day of the week)

Model Selection and Assessment

Population Process Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	✓	✓		
$\mathbb{I}\{\text{Regular}\}$				✓
$\mathbb{I}\{\text{Group 1}\}$				✓
V_{t-1}		✓	✓	✓
3-Cycle$_{t-1}$			✓	✓
$\mathbb{I}\{\text{Monday}\}$			✓	✓
$\mathbb{I}\{\text{Tuesday}\}$			✓	✓
$\mathbb{I}\{\text{Wednesday}\}$			✓	✓
$\mathbb{I}\{\text{Thursday}\}$			✓	✓
$\mathbb{I}\{\text{Friday}\}$			✓	✓
$\mathbb{I}\{\text{Saturday}\}$			✓	✓
$\mathbb{I}\{\text{Sunday}\}$			✓	✓

Population Process Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	✓	✓		
$\mathbb{I}\{\text{Regular}\}$				✓
$\mathbb{I}\{\text{Group 1}\}$				✓
V_{t-1}		✓	✓	✓
3-Cycle$_{t-1}$			✓	✓
$\mathbb{I}\{\text{Monday}\}$			✓	✓
$\mathbb{I}\{\text{Tuesday}\}$			✓	✓
$\mathbb{I}\{\text{Wednesday}\}$			✓	✓
$\mathbb{I}\{\text{Thursday}\}$			✓	✓
$\mathbb{I}\{\text{Friday}\}$			✓	✓
$\mathbb{I}\{\text{Saturday}\}$			✓	✓
$\mathbb{I}\{\text{Sunday}\}$			✓	✓

Network Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	✓	✓		
Mixing{Regular (R)}			✓	✓
Mixing{¬ R ↔ R}			✓	✓
Mixing{¬ R}			–	–
$\mathbb{I}$\{Indiv 06\}				✓
⋮				⋮
$\mathbb{I}$\{Indiv 08\}				✓
$\log(n_t)$		✓	✓	✓
Y_{t-1}		✓	✓	✓
$\log(\text{9-Cycle}_{t-1} + 1)$			✓	✓
$\mathbb{I}$\{Monday\}			✓	✓
$\mathbb{I}$\{Tuesday\}			✓	✓
$\mathbb{I}$\{Wednesday\}			✓	✓
$\mathbb{I}$\{Thursday\}			✓	✓
$\mathbb{I}$\{Friday\}			✓	✓
$\mathbb{I}$\{Saturday\}			✓	✓
$\mathbb{I}$\{Sunday\}			✓	✓

Network Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	✓	✓		
Mixing{Regular (R)}			✓	✓
Mixing{¬ R ↔ R}			✓	✓
Mixing{¬ R}			–	–
$\mathbb{I}\{$Indiv 06$\}$				✓
⋮				⋮
$\mathbb{I}\{$Indiv 08$\}$				✓
$\log(n_t)$		✓	✓	✓
Y_{t-1}		✓	✓	✓
$\log(9\text{-Cycle}_{t-1} + 1)$			✓	✓
$\mathbb{I}\{$Monday$\}$			✓	✓
$\mathbb{I}\{$Tuesday$\}$			✓	✓
$\mathbb{I}\{$Wednesday$\}$			✓	✓
$\mathbb{I}\{$Thursday$\}$			✓	✓
$\mathbb{I}\{$Friday$\}$			✓	✓
$\mathbb{I}\{$Saturday$\}$			✓	✓
$\mathbb{I}\{$Sunday$\}$			✓	✓

Model Adequacy Checking

GLI One-Step Prediction Simulation Count ($\alpha \leq 0.95$)

GLI	Fraction Correct
Network Size	$26/28$
Density	$28/28$
Mean Degree	$28/28$
Degree Centralization	$20/28$
Krackhardt Connectedness	$28/28$
Triad Census: 0	$28/28$
Triad Census: 1	$27/28$
Triad Census: 2	$28/28$
Triad Census: 3	$28/28$

Table: Check of whether the $\alpha \leq 0.95$ simulation interval contains a given GLI. Total possible correct is 28.

Results

1) Population Process Model
2) Dynamic Network Model

Results: Population Process Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	-1.5893*	-1.9250*		
$\mathbb{I}\{\text{Regular}\}$				0.9319*
$\mathbb{I}\{\text{Group 1}\}$				0.7803*
V_{t-1}		1.5295*	1.1312*	0.7867*
3-Cycle$_{t-1}$			0.3780*	0.3520*
$\mathbb{I}\{\text{Monday}\}$			-2.7044*	-3.4522*
$\mathbb{I}\{\text{Tuesday}\}$			-2.5047*	-3.3754*
$\mathbb{I}\{\text{Wednesday}\}$			-2.1569*	-3.0218*
$\mathbb{I}\{\text{Thursday}\}$			-2.0497*	-2.8769*
$\mathbb{I}\{\text{Friday}\}$			-2.2750*	-3.0973*
$\mathbb{I}\{\text{Saturday}\}$			-1.1718*	-1.9258*
$\mathbb{I}\{\text{Sunday}\}$			-1.5207*	-2.2235*

Results: Population Process Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	-1.5893*	-1.9250*		
$\mathbb{I}${Regular}				0.9319*
$\mathbb{I}${Group 1}				0.7803*
				0.7867*
3-C				0.3520*
$\mathbb{I}${M				-3.4522*
$\mathbb{I}${Tu				-3.3754*
$\mathbb{I}${Wedn				-3.0218*
$\mathbb{I}${Thu				-2.8769*
$\mathbb{I}${Friday}			-2.2750*	-3.0973*
$\mathbb{I}${Saturday}			-1.1718*	-1.9258*
$\mathbb{I}${Sunday}			-1.5207*	-2.2235*

- Being in the "regular" category ↑ likelihood of appearance
- Being in Group 1 ↑ likelihood of appearing

Results: Population Process Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	-1.5893*	-1.9250*		
$\mathbb{I}\{\text{Regular}\}$				0.9319*
$\mathbb{I}\{\text{Group 1}\}$				0.7803*
V_{t-1}		1.5295*	1.1312*	0.7867*
3-Cycle$_{t-1}$			0.3780*	0.3520*
$\mathbb{I}\{$M				-3.4522*
$\mathbb{I}\{$Tu				-3.3754*
$\mathbb{I}\{$Wedn				-3.0218*
$\mathbb{I}\{$Thu				-2.8769*
$\mathbb{I}\{$				-3.0973*
$\mathbb{I}\{$Sat				-1.9258*
$\mathbb{I}\{$Sunday$\}$			[illegible]	-2.2235*

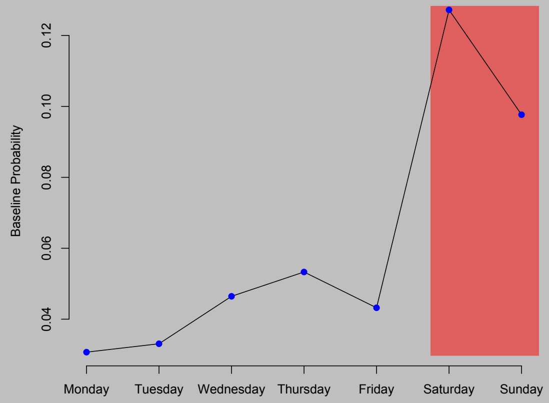
- Appearing today ↑ likelihood of appearing tomorrow!
- Each 3-clique ↑ her conditional odds of attendance by over 58%

Results: Population Process Model

	Model 1	Model 2	Model 3	Model 4
BI				66
In				
$\mathbb{I}\{$R				*
$\mathbb{I}\{$Gr				*
				*
3-Cy				*
$\mathbb{I}\{$Monday$\}$			-2.7044*	-3.4522*
$\mathbb{I}\{$Tuesday$\}$			-2.5047*	-3.3754*
$\mathbb{I}\{$Wednesday$\}$			-2.1569*	-3.0218*
$\mathbb{I}\{$Thursday$\}$			-2.0497*	-2.8769*
$\mathbb{I}\{$Friday$\}$			-2.2750*	-3.0973*
$\mathbb{I}\{$Saturday$\}$			-1.1718*	-1.9258*
$\mathbb{I}\{$Sunday$\}$			-1.5207*	-2.2235*

- These seasonal effects are comparable in magnitude to the effect of being a regular
- Exceed the effect of inertia (inertia + part. in 1-2 conv. simil. effect)

Results: Population Process Model



Results: Dynamic Network Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	-2.3162*	-4.3081*		
Mixing{Regular (R)}			1.1091*	1.1389*
Mixing{$\neg$ R $\leftrightarrow$ R}			0.5560*	0.5595*
Mixing{$\neg$ R}			–	–
$\mathbb{I}${Indiv 06}				-0.6563*
$\mathbb{I}${Indiv 17}				-0.9519*
$\mathbb{I}${Indiv 16}				-0.4602*
$\mathbb{I}${Indiv 05}				0.5852*
$\vdots$				$\vdots$
$\mathbb{I}${Indiv 42}				-1.1086*
$\mathbb{I}${Indiv 44}				1.4785*
$\mathbb{I}${Indiv 50}				-0.6624*
$\mathbb{I}${Indiv 08}				0.2206*
$\log(n_t)$		0.6394*	0.3884*	4.0946*
Y_{t-1}		0.8946*	0.3120*	0.2808*
$\log(\text{9-Cycle}_{t-1} + 1)$			0.0880*	0.1077*
$\mathbb{I}${Monday}			-5.9929*	-12.2986*
$\mathbb{I}${Tuesday}			-4.1357*	-9.5061*
$\mathbb{I}${Wednesday}			-3.7315*	-10.6229*
$\mathbb{I}${Thursday}			-4.0714*	-10.6006*
$\mathbb{I}${Friday}			-4.7021*	-11.6135*
$\mathbb{I}${Saturday}			-4.2353*	-11.4279*
$\mathbb{I}${Sunday}			-4.9328*	-12.5474*

Results: Dynamic Network Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	-2.3162*	-4.3081*		
Mixing{Regular (R)}			1.1091*	1.1389*
Mixing{$\neg$ R $\leftrightarrow$ R}			0.5560*	0.5595*
Mixing{$\neg$ R}			–	–
$\mathbb{I}${Indiv 06}				0.6563*
[illegible]				[illegible]9*
[illegible]				[illegible]2*
[illegible]				[illegible]2*
$\mathbb{I}${Indiv 42}				-1.1086*
$\mathbb{I}${Indiv 44}				1.4785*
$\mathbb{I}${Indiv 50}				-0.6624*
$\mathbb{I}${Indiv 08}				0.2206*
$\log(n_t)$		0.6394*	0.3884*	4.0946*
Y_{t-1}		0.8946*	0.3120*	0.2808*
$\log(\text{9-Cycle}_{t-1} + 1)$			0.0880*	0.1077*
$\mathbb{I}${Monday}			-5.9929*	-12.2986*
$\mathbb{I}${Tuesday}			-4.1357*	-9.5061*
$\mathbb{I}${Wednesday}			-3.7315*	-10.6229*
$\mathbb{I}${Thursday}			-4.0714*	-10.6006*
$\mathbb{I}${Friday}			-4.7021*	-11.6135*
$\mathbb{I}${Saturday}			-4.2353*	-11.4279*
$\mathbb{I}${Sunday}			-4.9328*	-12.5474*

- Regulars more likely to interact with regulars
- But irregulars no more likely to interact with regulars as irregulars

Results: Dynamic Network Model

- Individual propensity to interact is a large effect
- Range from (+)1.47 to (-)2.85

	Model 1	Model 2	Model 3	Model 4
$\mathbb{I}\{\text{Indiv } 06\}$				-0.6563*
$\mathbb{I}\{\text{Indiv } 17\}$				-0.9519*
$\mathbb{I}\{\text{Indiv } 16\}$				-0.4602*
$\mathbb{I}\{\text{Indiv } 05\}$				0.5852*
$\vdots$				$\vdots$
$\mathbb{I}\{\text{Indiv } 42\}$				-1.1086*
$\mathbb{I}\{\text{Indiv } 44\}$				1.4785*
$\mathbb{I}\{\text{Indiv } 50\}$				-0.6624*
$\mathbb{I}\{\text{Indiv } 08\}$				0.2206*
$\log(n_t)$		0.6394*	0.3884*	4.0946*
Y_{t-1}		0.8946*	0.3120*	0.2808*
$\log(\text{9-Cycle}_{t-1}+1)$			0.0880*	0.1077*
$\mathbb{I}\{\text{Monday}\}$			-5.9929*	-12.2986*
$\mathbb{I}\{\text{Tuesday}\}$			-4.1357*	-9.5061*
$\mathbb{I}\{\text{Wednesday}\}$			-3.7315*	-10.6229*
$\mathbb{I}\{\text{Thursday}\}$			-4.0714*	-10.6006*
$\mathbb{I}\{\text{Friday}\}$			-4.7021*	-11.6135*
$\mathbb{I}\{\text{Saturday}\}$			-4.2353*	-11.4279*
$\mathbb{I}\{\text{Sunday}\}$			-4.9328*	-12.5474*

Results: Dynamic Network Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	-2.3162*	-4.3081*		
Mixing{Regular (R)}			1.1091*	1.1389*
Mixing{¬ R ↔ R}			0.5560*	0.5595*
Mixing{¬ R}			–	–
				-0.6563*
				-0.9519*
				-0.4602*
				0.5852*
				⋮
				-1.1086*
				1.4785*
				-0.6624*
				0.2206*
$\log(n_t)$		0.6394*	0.3884*	4.0946*
Y_{t-1}		0.8946*	0.3120*	0.2808*
$\log(\text{9-Cycle}_{t-1} + 1)$			0.0880*	0.1077*
$\mathbb{I}$\{Monday\}			-5.9929*	-12.2986*
$\mathbb{I}$\{Tuesday\}			-4.1357*	-9.5061*
$\mathbb{I}$\{Wednesday\}			-3.7315*	-10.6229*
$\mathbb{I}$\{Thursday\}			-4.0714*	-10.6006*
$\mathbb{I}$\{Friday\}			-4.7021*	-11.6135*
$\mathbb{I}$\{Saturday\}			-4.2353*	-11.4279*
$\mathbb{I}$\{Sunday\}			-4.9328*	-12.5474*

- Contagious participation is strong positive effect!
- Max: 37 individuals appear $\log(37) \cdot 4.0946 = 6.42$
- Min: $\log(3) \cdot 4.0946 = 1.95$

Results: Dynamic Network Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	-2.3162*	-4.3081*		
Mixing{Regular (R)}			1.1091*	1.1389*
Mixing{$\neg$ R $\leftrightarrow$ R}			0.5560*	0.5595*
Mixing{$\neg$ R}			–	–
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]3*
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]9*
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]2*
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]2*
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]6*
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]5*
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]4*
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]6*
$\log(n_t)$		[illegible]	[illegible]	[illegible]46*
Y_{t-1}		0.8946*	0.3120*	0.2808*
$\log(9\text{-Cycle}_{t-1} + 1)$			0.0880*	0.1077*
$\mathbb{I}${Monday}			-5.9929*	-12.2986*
$\mathbb{I}${Tuesday}			-4.1357*	-9.5061*
$\mathbb{I}${Wednesday}			-3.7315*	-10.6229*
$\mathbb{I}${Thursday}			-4.0714*	-10.6006*
$\mathbb{I}${Friday}			-4.7021*	-11.6135*
$\mathbb{I}${Saturday}			-4.2353*	-11.4279*
$\mathbb{I}${Sunday}			-4.9328*	-12.5474*

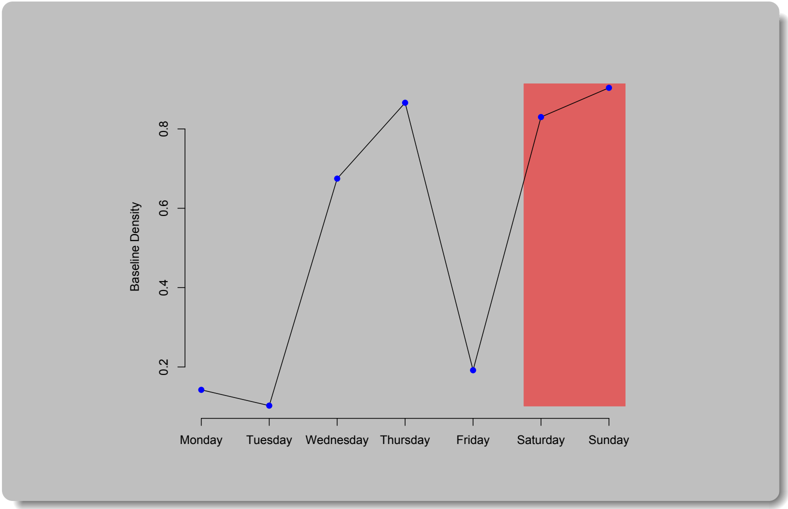
- Inertial effect
  - Past interaction increases likelihood of interaction in the future
- Embeddedness effect
  - Past embeddedness increases the likelihood of interaction in the future

Results: Dynamic Network Model

	Model 1	Model 2	Model 3	Model 4
BIC	8166.2395	7929.2997	7699.5743	7689.1466
Intercept	-2.3162*	-4.3081*		
Mixing{Regular (R)}			1.1091*	1.1389*
Mixing{$\neg$ R $\leftrightarrow$ R}			0.5560*	0.5595*
Mixing{$\neg$ R}			–	–
$\mathbb{I}${Indiv 06}				-0.6563*
$\mathbb{I}${Indiv 17}				-0.9519*
$\mathbb{I}${Indiv 16}				-0.4602*
$\mathbb{I}${Indiv 05}				0.5852*
[illegible]				6*
[illegible]				5*
[illegible]				4*
[illegible]				6*
[illegible]				6*
Y_{t-1}		0.8946*	0.3120*	0.2808*
$\log(\text{9-Cycle}_{t-1}+1)$			0.0880*	0.1077*
$\mathbb{I}${Monday}			-5.9929*	-12.2986*
$\mathbb{I}${Tuesday}			-4.1357*	-9.5061*
$\mathbb{I}${Wednesday}			-3.7315*	-10.6229*
$\mathbb{I}${Thursday}			-4.0714*	-10.6006*
$\mathbb{I}${Friday}			-4.7021*	-11.6135*
$\mathbb{I}${Saturday}			-4.2353*	-11.4279*
$\mathbb{I}${Sunday}			-4.9328*	-12.5474*

- Day of the week is a big effect!
- Max density effect (in comb. w/ cp): -3.8
- Min density effect (in comb. w/ cp): -6.4

Results: Dynamic Network Model



$\mathbb{I}\{\text{Friday}\}$	-4.7021*	-11.6135*
$\mathbb{I}\{\text{Saturday}\}$	-4.2353*	-11.4279*
$\mathbb{I}\{\text{Sunday}\}$	-4.9328*	-12.5474*

Broader Implications

A Few Empirical and Theoretical Insights

- Theoretical insight: Population and network processes interact in non-trivial ways
- Empirical insight: Ability to spot cohesive groups dynamically (e.g., the active group on Wednesday/Thursday)

Broader Implications

Methodological Take-away:

- Dynamic Network logistic-Regression can capture complex network dynamics
- It is scalable and interpretable
- Allows for population processes, which can have a large effect on the dynamic network processes

Some Basic Lessons:

- Seasonality and period effects are important!
- Processes can, themselves, evolve over time

Extensions and Expansions...

Core Question:

How can we integrate population process models into standard network practice?

Why is this hard?

- Small deviations in vertex (e.g., individuals, organizations) selection matter
- Getting the number of vertices right is not good enough
  - One needs to get the labels right!

One idea...

- Latent variable models (Valluvan et al., 2012)

Other extensions

- Open population model: In which those vertices that are present are drawn from some arbitrarily large set

Road Map

1) Overview of Population/Network Processes
2) Examples
3) A Model for Dynamic Networks with Population Processes
4) Empirical Example
5) **Further Reading**

Further Reading

- Zack W. Almquist and Carter T. Butts (2018). "Dynamic Network Analysis with Missing Data: Theory and Methods." Statistica Sinica 28(3) 1245-1264.
- Carter T. Butts and Zack W. Almquist. (2015). "A Flexible Parameterization for Baseline Mean Degree in Multiple-Network ERGMs." The Journal of Mathematical Sociology, 39(3), 163-167.
- Zack W. Almquist and Carter T. Butts. (2014). "Logistic Network Regression for Scalable Analysis of Networks with Joint Edge/Vertex Dynamics." Sociological Methodology, 44(1), 273-321.
- Zack W. Almquist and Carter T. Butts. (2013). "Dynamic Network Logistic Regression: A Logistic Choice Analysis of Inter- and Intra-group Blog Citation Dynamics in the 2004 US Presidential Election." Political Analysis, 21(4), 430-448.
- Zack W. Almquist, Emma S. Spiro, and Carter T. Butts (2016). "Shifting Attention: Modeling Follower Relationship Dynamics among US Emergency Management-related Organizations During a Colorado Wildfire." In: Social Network Analysis of Disaster Response, Recovery, and Adaptation. Ed. by A. Faas and E. Jones. Philadelphia, PA: Elsevier.
- Zack W. Almquist and Carter T. Butts. (2014). "Bayesian Analysis of Dynamic Network Regression with Joint Edge/Vertex Dynamics." In Bayesian Inference in the Social Sciences. Ed. by I. Jeliazkov and X.-S. Yang. Hoboken, New Jersey: John Wiley & Sons.


Thank You!

References I

Almquist, Zack W and Carter T Butts. 2012. "Logistic Network Regression for Scalable Analysis of Networks with Joint Edge/Vertex Dynamics." *Sociological Methodology* forthcoming.

Anderton, Douglas L., Joseph Conaty, and George A. Miller. 1983. "Structural Constraints on Organizational Change: a longitudinal analysis." *Journal of Business Research* 11:153–170.

Baker, Wayne E. 1984. "The Social Structure of a National Securities Market." *American Journal of Sociology* 89:775–811.

Bearman, Peter S., James Moody, and Katherine Stovel. 2004. "Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks." *American Journal of Sociology* 110:44–91.

Boyd, Monica. 1989. "Family and Personal Networks in International Migration: Recent Developments and New Agendas." *International Migration Review* 23:638–670.

References II

Buskens, Vincent and Arnout van de Rijt. 2008. "Dynamics of Networks if Everyone Strives for Structural Holes." *American Jounral of Sociology* 114:371–407.

Butts, Carter T. 2008. "A Relational Event Framework for Social Action." *Sciological Methodology* 38:155–200.

Butts, Carter T. 2009. "Revisiting the Foundations of Network Analysis." *Science* 325:414–416.

Butts, Carter T., Ryan M. Acton, and Christopher Steven Marcum. 2012. "Interorganizational Collaboration In the Hurricane Katrina Response." *Journal of Social Structure* 13.

Butts, Carter T. and B. Remy Cross. 2009. "Butts, Carter T., and B. Remy Cross. "Change and external events in computer-mediated citation networks: English language weblogs and the 2004 US electoral cycle." Journal of social structure 10.3 (2009)." *Journal of Social Structure* 10.

References III

Carley, Kathleen M. 1999. "On the evolution of social and organizational networks." *Research in the Sociology of Organizations* 16:3–30.

Chase, Ivan D. 1980. "Social Process and Hierarchy Formation in Small Groups: A comparative Perspective." *American Sociological Review* 45:905–924.

Coleman, James S. 1961. *The Adolescent Society*. Glencoe, IL: The Free Press.

Corten, Rense and Vincent Buskens. 2010. "Co-evolution of Conventions and Networks: An Experimental Study." *Social Networks* 32:4 – 15. Dynamics of Social Networks.

Doğan, Gönül, Marcel van Assen, Arnout van de Rijt, and Vincent Buskens. 2009. "The Stability of Exchange Networks." *Social Networks* 31:118–25.

Entwisle, Barbara, Katherine Faust, Ronald R. Rindfuss, and Toshiko Kaneda. 2007. "Networks and Contexts: Variation in the Structure of Social Ties." *American Journal of Sociology* 112:1495–1533.

References IV

Faust, Katherine. 2010. "A puzzle concerning triads in social networks: Graph constraints and the triad census." *Social Networks* 32:221 – 233.

Feldman, Kenneth A. 1972. "Some Theoretical Approaches to the Study of Change and Stability of College Students." *Review of Educational Research* 42:1–26.

Fischer, Claude S. 1982. *To Dwell Among Friends: Personal Networks in Town and City*. Chicago, IL: University of Chicago Press.

Freeman, Linton C. 1992. "The Sociological Concept of Group – An Emperical - Test of 2 Models." *American Journal of Sociology* 98:152–166.

Freeman, Linton C., Sue C. Freeman, and Alaina G. Michaelson. 1988. "On Human Social Intelligence." *Journal of Social Biological Structure* 11:415–425.

References V

Granovetter, Mark. 1985. "Economic Action and Social Structure: the Problem of Embeddedness." *American Journal of Sociology* 91:481–510.

Hannan, M.T. and J. Freeman. 1977. "The Population Ecology of Organizations." *American Journal of Sociology* 82:929–964.

Hannan, Michael T. and John Freeman. 1984. "Structural inertia and organizational change." *American Sociological Review* 49:149–164.

Hanneke, Steve, Wenjie Fu, and Eric P. Xing. 2010. "Discrete temporal models of social networks." *Electronic Journal of Statistics* 4:585–605.

Hanneke, Steve and Eric P. Xing. 2007. *Statistical Network Analysis: Models, Issues, and New Directions: ICML 2006 Workshop on Statistical Network Analysis, Pittsburgh, PA, USA, June 29, 2006, Revised Selected Papers*, volume 4503 of *Lecture Notes in Computer Science*, chapter Discrete Temporal Models of Social Networks, pp. 115–125. Springer-Verlag.

References VI

Hodgkinson, Harold L. 1985. *All One System: Demographics of Education, Kindergarten through Graduate School*. Institute for Educational Leadership, Washington, DC.

Hunter, David R., Steven M. Goodreau, and Mark S. Handcock. 2008. "Goodness of fit of social network models." *Journal of the American Statistical Association* 103:248–258.

Lee, Ronald. 1987. "'Population Dynamics of Humans and Other Animals', Presidential Address to the Population Association of America." *Demography* 24:443–466.

Lee, Ronald D. and Lawrence R. Carter. 1992. "Modeling and forecasting US mortality." *Journal of the American Statistical Association* 87:659–671.

Mayhew, Bruce H. and Roger L. Levinger. 1976. "Size and the Density of Interaction in Human Aggregates." *The American Journal of Sociology* 83.

References VII

McPherson, Miller, Lynn Smith-Lovin, and James M Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415–444.

Mercken, L., T.A.B. Snijders, C. Steglich, E. Vartiainen, and H. de Vries. 2010. "Dynamics of adolescent friendship networks and smoking behavior, Social Networks." *Social Networks* 32:72–81.

Mintz, Beth and Micheal Schwartz. 1981. "Interlocking Directorates and Interest Group Formation." *American Sociological Review* 46:851–869.

Morris, Martina and Mirjam Kretzschmar. 1997. "Concurrent partnerships and the spread of HIV." *AIDS* 11:641–648.

Newcomb, Theodore M. 1961. *The Acquaintance Process*. New York, NY: Holt, Reinhard & Winston.

Nordlie, P. 1958. *A Longitudinal Study of Interpersonal Attraction in a Natural Group Setting*. Ph.D. thesis, University of Michigan.

References VIII

Robins, Garry, Philipppa Pattison, and Stanley Wasserman. 1999. "Logit Models and Logistic Regressions for Social Networks: III Valued Relations." *Psychometrika* 64:371–394.

Romo, Frank P. and Michael Schwartz. 1995. "The Structural Embeddedness of Business Decisions: The Migration of Manufacturing Plants in New York State, 1960 to 1985." *American Sociological Review* 60:874–907.

Valluvan, Ragupathyraj, Zack W Almquist, Anima Anandkumar, and Carter T Butts. 2012. "Modeling Dynamic Social Networks with Vertex Evolution via Latent Graphical Models." Poster presentation at The 2012 NIPS Workshop on Social Network and Social Media Analysis, Methods, Models, and Applications, Lake Tahoe, NV.

van de Rijt, Arnout. 2011. "The Micro-Macro Link for the Theory of Structural Balance." *Journal of Mathematical Sociology* 35:94–113.

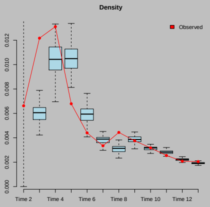
Wasserman, Stanley and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.

Why Has This Area Been Ignored in the Network Literature?

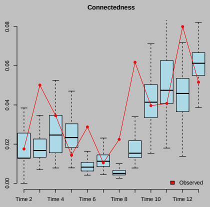
- Social networkers
  - Focused on small bounded groups
  - Focused on static case (and/or relatively stable relations)
  - Limited data
  - Computational and algorithmic complexity

What Happens When the Vertex Set is Wrong?

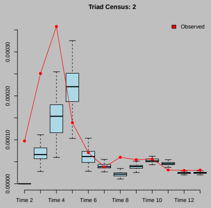
Correct Vertex Set



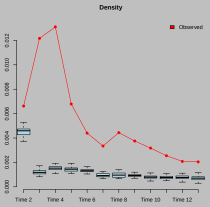
Correct Vertex Set



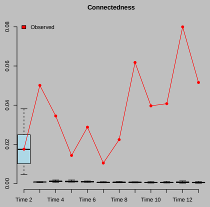
Correct Vertex Set



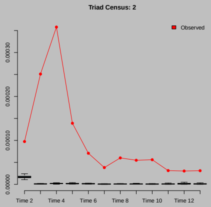
Naive Model



Naive Model



Naive Model



Getting to Scalability: Lagged Logistic Network Model

- Most serious obstacle to scalability: The ERG normalizing factor

$$Pr(G = g|s, \theta) = \frac{\exp\{\theta^T s(g)\}}{\sum_{g' \in \mathbb{G}} \exp(\theta^T s(g'))} \mathbb{I}_{\mathbb{G}}(g)$$

- Approximate workaround: Assume conditional independence of edges, given the past
  - Introduced by Robins et al. (1999), extended by Hanneke and Xing (2007), and others
  - We extend existing practice by relaxing Markov assumption, adding vertex set dynamics
- Consequence: lagged logistic network model w/ vertex dynamics (Almquist & Butts, Forthcoming)
  - Network at time t is modeled as logistic regression on prior states, along with covariates
  - Allows us to leverage extensive machine learning literature on fast logistic regression for sparse matrices

Adding in Vertex Dynamics

When the vertex set also evolves, we include it within the same framework

- Separability: $\Pr(\mathbb{I}\{v_t \in V_t\}) \times Pr(Y_t \mid V_t)$
- Support assumption: V_t is drawn from some maximal set V
- Dependence assumptions: Adjacency matrix Y_t depends on past $Z_{t-k} = (Y_{t-k}, V_{t-k})$ (generally up to some fixed $_k$)

$$
\begin{aligned}
&\Pr(G_t = g_t \mid G_{t-1}, \cdots, G_{t-k}, X) = \\
&\quad \Pr(V_t = v_t \mid Z_{t-k}, X) \times \Pr(Y_t = y_t \mid V_t, Z_{t-k}, X) = \\
&\quad \frac{\exp\left(\psi^T w(v_t, Z_{t-k}, X)\right)}{\sum_{v' \in \mathcal{V}} \exp\left(\psi^T w(v'_t, Z_{t-k}, X)\right)} \times \frac{\exp\left(\theta^T u(y_t, V_t, Z_{t-k}, X)\right)}{\sum_{y' \in \mathcal{Y}} \exp\left(\theta^T u(y'_t, V_t, Z_{t-k}, X)\right)}
\end{aligned}
$$

Adding in Population Dynamics

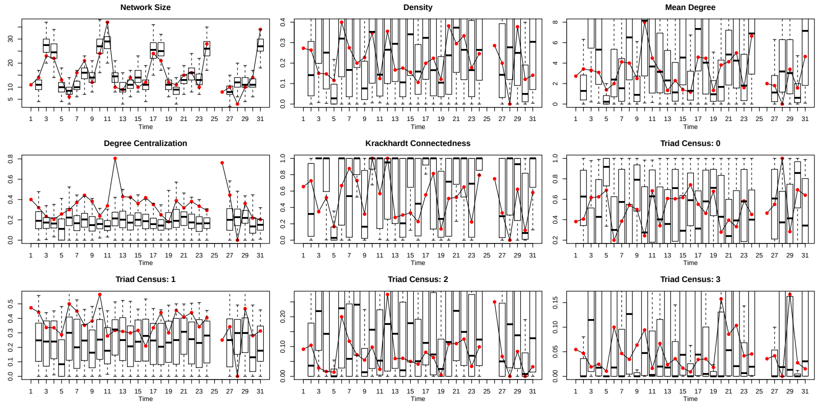
Conditional Independence and Vertex Dynamics

- Assuming that each $Y_{ij,t}, V_{i,t}$ are independent given the above leads to a joint logistic formulation

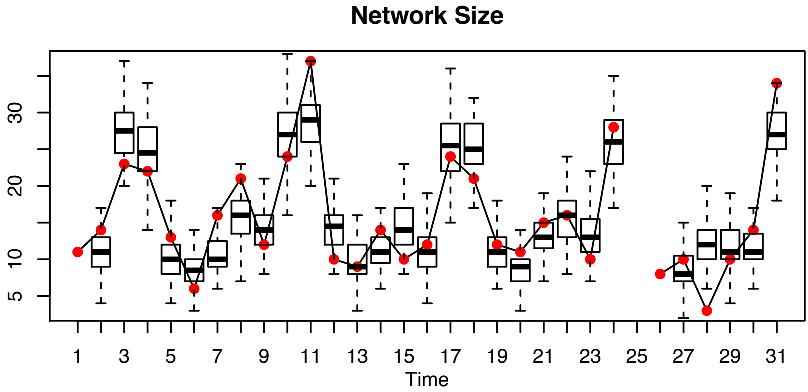
$$\begin{aligned}
&\Pr(V_t \mid Z_{t-1}, \ldots, Z_{t-k}, X) \\
&\qquad = \prod_{i=1}^{n} B\left(\mathbb{I}(v_i \in V_t) \,\middle|\, \mathrm{logit}^{-1}\left(\psi^T w(i, Z_{i-1}, \ldots, Z_{i-k}, X)\right)\right) \\
&\Pr(Y_t \mid V_t, Z_{t-1}, \ldots, Z_{t-k}, X) \\
&\qquad = \prod_{(i,j) \in V_t \times V_t}^{n} B\left(Y_{tij} \,\middle|\, \mathrm{logit}^{-1}\left(\theta^T u(i, j, V_t, Z_{i-1}, \ldots, Z_{i-k}, X)\right)\right)
\end{aligned}$$

Net result: flexible framework (TERGM subfamily) that readily scales to thousands of vertices/time points

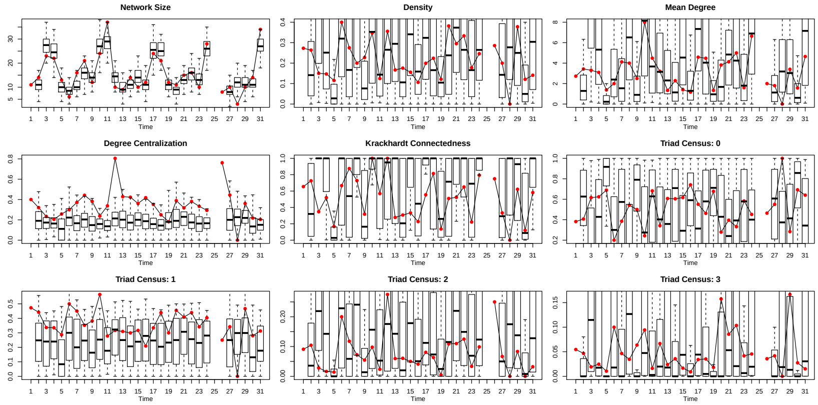
Model Assessment: Via Simulation



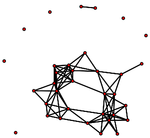
Model Assessment: Via Simulation



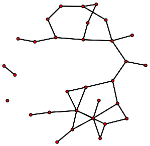
Model Assessment: Via Simulation



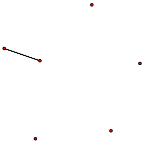
Model Assessment: n-step Projection



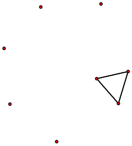
Day 31 (Saturday 9/27/86)



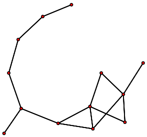
Day 32 (Sunday 9/28/86)



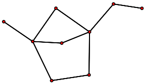
Day 33 (Monday 9/29/86)



Day 34 (Tuesday 9/30/86)



Day 35 (Wednesday 10/01/86)



Day 36 (Thursday 10/02/86)