Lecture 3 – Part 2
Research on Network Data
(Conference Slides 2014)

(January 23, 2015)

Mu Zhu
University of Waterloo

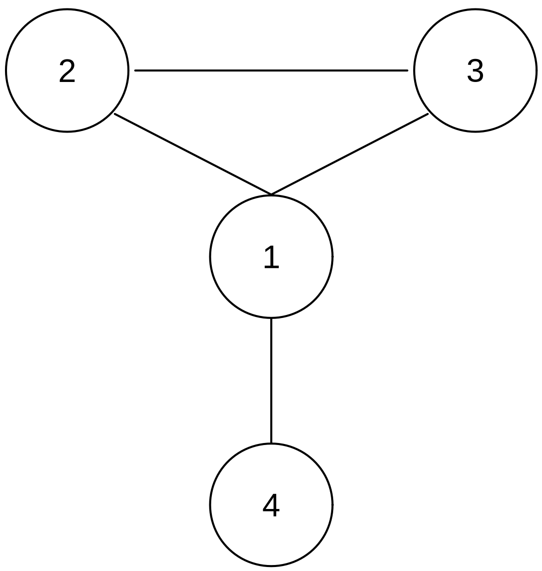
Continuous-time Stochastic Block Models for Transactional Networks

Lu Xin
University of Waterloo

Mu Zhu
University of Waterloo

Hugh A. Chipman
Acadia University

Networks



$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Stochastic Block Models (SBMs)

- Notion of "community", "group", "block".
- Let $e_i, e_j \in \{1, 2, ..., K\}$ be the group memberships of i and j.
- Given $\mathbf{E} = \{e_i\}$, SBM assumes

$$\mathbb{P}(\mathbf{A}|\mathbf{E}) = \prod_{i,j} \left(P_{e_i e_j}\right)^{A_{ij}} \left(1 - P_{e_i e_j}\right)^{1-A_{ij}}.$$

- But e_i, e_j are unknown. Given observed connections A_{ij}, would like to cluster the nodes (i.e., estimate e_i, e_j).
- Snijders & Nowicki (1997), Zhao, Levina & Zhu (2011), others.

Dynamic Networks

- Discrete time:

$$\mathbf{A}^{(0)} \longrightarrow \mathbf{A}^{(1)} \longrightarrow \ldots \longrightarrow \mathbf{A}^{(t)} \longrightarrow \ldots$$

- Continuous time:

i	j	time of transaction (t)
1	4	06/29/2014, 08:05
1	7	06/29/2014, 08:07
5	1	07/02/2014, 17:21
⋮	⋮	⋮

Non-homogeneous Poisson Process

Suppose we observe m events at $0 < t_1 < t_2 < ... < t_m < \tau$, then the joint density of arrival times $T_1, T_2, ..., T_m$ is

$$f(t_1, t_2, ..., t_m) = \underbrace{\left(e^{-\int_0^\tau \rho(u)du}\right)}_{J(\rho)} \times \prod_{h=1}^{m} \rho(t_h)$$

where $\rho(t)$ is the rate function of the underlying process.

SBMs for Transactional Networks

- Model events (transactions) between each (i, j) with a non-homogeneous Poisson process.

- Otherwise, inherit key features of (regular) SBMs, i.e.,
  - rate function governed by group membership, $\rho_{e_i e_j}(t)$;
  - transactions conditionally independent.

Likelihood

Given group memberships $e_i, e_j \in \{1, 2, ..., K\}$,

$$\text{likelihood} = \prod_{i,j} \left\{ J(\rho_{e_i e_j}) \times \prod_{h=1}^{m_{ij}} \rho_{e_i e_j}(t_{ijh}) \right\}$$

where

- t_{ijh} = time of h-th transaction between i and j,
- m_{ij} = total number of transactions between i and j.

But, again, the group memberships e_i, e_j are unknown.

Mixture Model

- Introduce latent variables

$$z_{ik} = \begin{cases} 1, & \text{if } i \text{ belongs to group } k; \\ 0, & \text{otherwise.} \end{cases}$$

- Complete likelihood is

$$\mathcal{L}(\Theta; \mathbf{T}, \mathbf{Z}) = \prod_{i,j} \left\{ \prod_{k,l} \left[J(\rho_{kl}) \times \prod_{h=1}^{m_{ij}} \rho_{kl}(t_{ijh}) \right]^{z_{ik} z_{jl}} \right\} \times \prod_{i} \prod_{k} \pi_k^{z_{ik}},$$

where $\pi_k \equiv \mathbb{P}(z_{ik} = 1)$.

Parameters and Data

- Parameters:

$$\Theta = \{\rho_{kl}(t), \pi_k : \quad 1 \leq k, l \leq K\}.$$

- Observed Data:

$$\mathbf{T} = \{t_{ijh} : \quad 1 \leq i, j \leq n; \quad h = 1, 2, ..., m_{ij}\}.$$

- Latent "Data":

$$\mathbf{Z} = \{z_{ik} : \qquad 1 \leq i \leq n; \quad 1 \leq k \leq K\}.$$

EM Algorithm

Log-likelihood

$$\ell(\Theta; \mathbf{T}, \mathbf{Z}) = \sum_{i,j} \left\{ \sum_{k,l} z_{ik} z_{jl} \left[\log J(\rho_{kl}) + \sum_{h=1}^{m_{ij}} \log \rho_{kl}(t_{ijh}) \right] \right\} + \sum_{i} \sum_{k} z_{ik} \log \pi_k .$$

E-Step Compute $\mathbb{E}(z_{ik} z_{jl} | \Theta, \mathbf{T})$, $\mathbb{E}(z_{ik} | \Theta, \mathbf{T})$.

M-Step Update $\rho_{kl}(\cdot)$ and π_k by maximizing $\ell(\Theta; \mathbf{T}, \mathbf{Z})$, replacing latent quantities with their respective expectations.

E-Step

- Exact $\mathbb{E}(\cdot|\Theta, \mathbf{T})$ hard to compute.
- Use Monte Carlo estimate.
- By drawing $\mathbf{z}_i$ from the (conditional) distribution of

$$\mathbf{z}_i \big| \mathbf{z}_1, ..., \mathbf{z}_{i-1}, \mathbf{z}_{i+1}, ..., \mathbf{z}_n; \Theta, \mathbf{T}$$

  for $i = 1, 2, ..., n$, using the Gibbs Sampler.

Gibbs Sampler

Recall likelihood

$$\mathcal{L}(\Theta; \mathbf{T}, \mathbf{Z}) = \prod_{i,j} \left\{ \prod_{k,l} \left[J(\rho_{kl}) \times \prod_{h=1}^{m_{ij}} \rho_{kl}(t_{ijh}) \right]^{z_{ik} z_{jl}} \right\} \times \prod_{i} \prod_{k} \pi_k^{z_{ik}}.$$

Let $\alpha_{ij,kl}$ denote "stuff in square bracket". Can see that

$$\mathbb{P}(z_{ik} = 1 | \mathbf{z}_{j \neq i}; \Theta, \mathbf{T}) = \frac{\left[\pi_k \prod_{j \neq i} \left(\prod_{l=1}^{K} {\alpha_{ij,kl}}^{z_{jl}} \right)\right]}{\sum_{k'=1}^{K} \left[\pi_{k'} \prod_{j \neq i} \left(\prod_{l=1}^{K} {\alpha_{ij,k'l}}^{z_{jl}} \right)\right]}$$

for $k = 1, 2, ..., K \Rightarrow$ drawing $\mathbf{z}_i \sim$ tossing a (loaded) K-die.

M-Step

Updating $\rho_{kl}(\cdot)$ Model

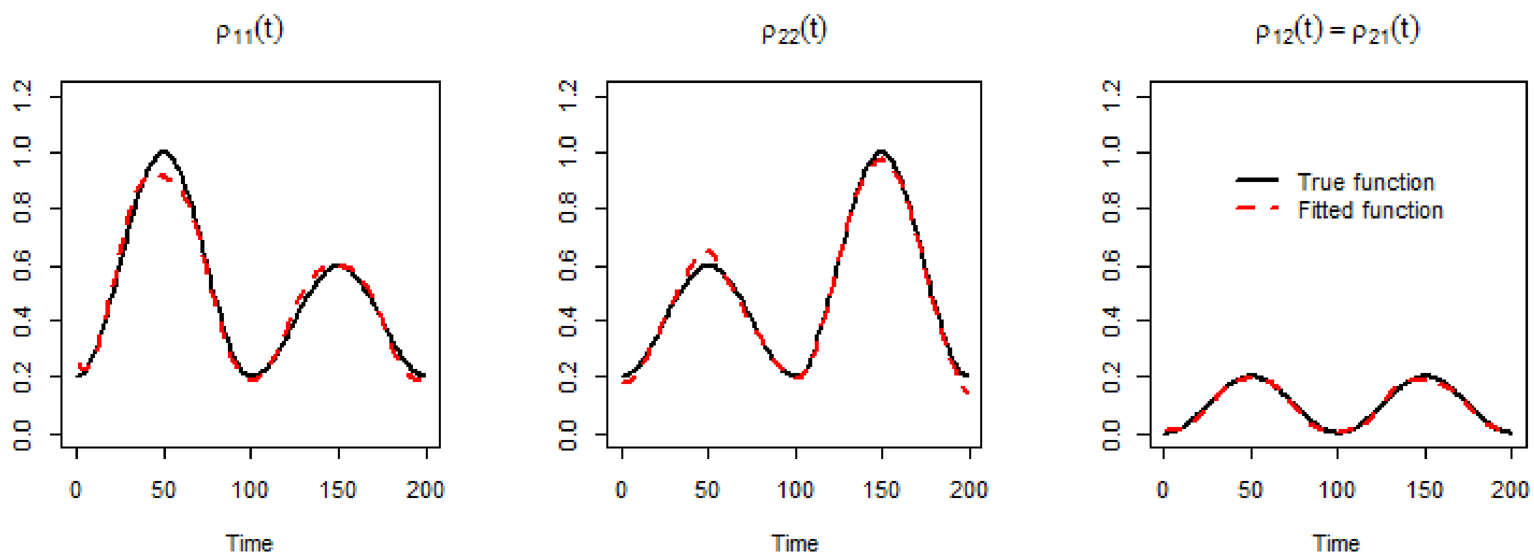
$$\rho_{kl}(t) = \sum_{p=1}^{P} e^{\beta_{klp}} B_p(t),$$

where $B_1(t), B_2(t), ..., B_P(t)$ are B-spline basis functions, and maximize over β_{klp} with quasi-Newton. (Parameterization $e^{\beta_{klp}}$ used to ensure $\rho_{kl}(t) \geq 0$.)

Updating π_k

$$\pi_k \longleftarrow \frac{\sum_{i=1}^{n} \mathbb{E}(z_{ik}|\Theta, \mathbf{T})}{\sum_{k=1}^{K} \sum_{i=1}^{n} \mathbb{E}(z_{ik}|\Theta, \mathbf{T})} = \frac{\sum_{i=1}^{n} \mathbb{E}(z_{ik}|\Theta, \mathbf{T})}{n}.$$

Simple Simulation



Two groups with ten nodes each. True rate functions are piecewise cosines. Estimates based on fitting B-splines with 15 knots to $11,605$ simulated "events", and using the BIC to choose K.

Application: NBA Games

initial state	i	j	absorbing state	t (sec)
inbound	-	2	-	00
-	2	5	-	02
-	5	3	-	05
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
-	4	1	-	22
-	1	-	score	23

Model Adaptation (I)

- Recall

$$\mathcal{L}(\Theta; \mathbf{T}, \mathbf{Z}) = \prod_{i,j} \left\{ \prod_{k,l} \left[J(\rho_{kl}) \times \prod_{h=1}^{m_{ij}} \rho_{kl}(t_{ijh}) \right]^{z_{ik} z_{jl}} \right\} \times \cdots .$$

- Modification (Ia)

$$\rho_{kl}(\cdot) \quad \Longrightarrow \quad \rho_{kl}(\cdot) \times \frac{1}{|G_l^{-i}(\cdot)|}, \quad |G_l^{-i}(\cdot)| = \sum_{r \neq i} z_{rl} \times \mathbf{I}_r(\cdot)$$

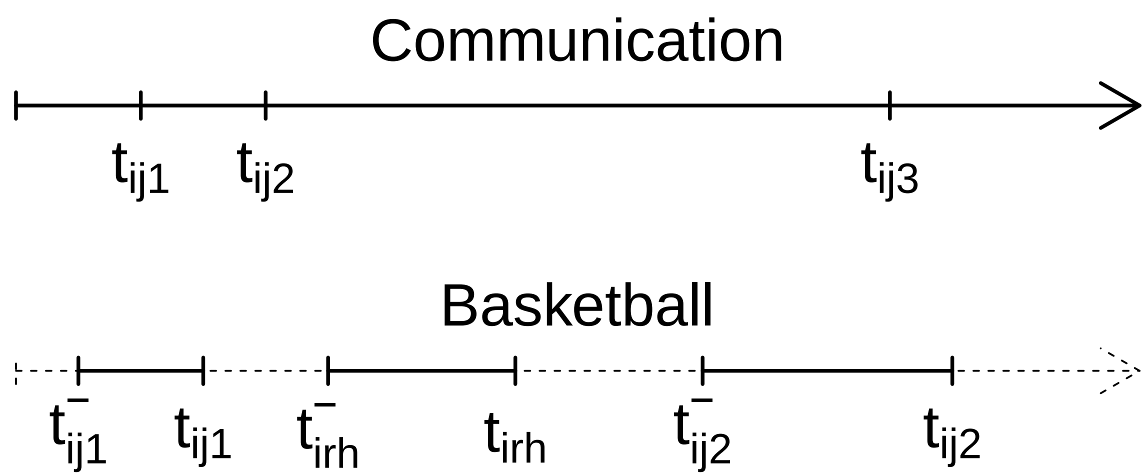
necessary because $\exists$ only ONE ball, so belonging to a large group reduces the chance of j getting the ball.

- Modification (Ib)

$$J(\rho_{kl}) \quad \Longrightarrow \quad J_i(\rho_{kl})$$

also due to fact $\exists$ only ONE ball; see next slide.

From $J(\rho_{kl})$ to $J_i(\rho_{kl})$



$$J(\rho_{kl}) \equiv \exp\left[-\int_0^{\tau} \rho_{kl}(u)du\right] \quad \Longrightarrow$$

$$J_i(\rho_{kl}) \equiv \exp\left[-\sum_{r\neq i}\sum_{h=1}^{m_{ir}}\int_{t^-_{irh}}^{t_{irh}} \rho_{kl}(u) \times \frac{1}{|G_l^{-i}(u)|}du\right]$$

Model Adaptation (II)

- Considered re-parameterization of $\rho_{kl}(t)$ as

$$\rho_{kl}(t) = \lambda_k(t) \times P_{kl},$$

  where

$$\begin{aligned} \lambda_k(t) &= \text{rate function for group } k \text{ \underline{to pass}, and} \\ P_{kl} &= \mathbb{P}(\text{group } k \text{ passes to group } l). \end{aligned}$$

- Now, K rate functions rather than K^2 — expected to be especially useful for larger K.

Model Adaptation (III)

- Considered two initial states (start a possession from inbound or rebound), and augmented likelihood function by

$$\left\{ \prod_{i=1}^{n} \prod_{s=1}^{2} \left[\prod_{k=1}^{K} \left(Q_{sk} \cdot \frac{1}{|G_k|} \right)^{z_{ik}} \right]^{m_{si}} \right\} \times (\text{“old” likelihood}),$$

  where

$$\begin{aligned} Q_{sk} &= \mathbb{P}(\text{group } k \text{ ... from initial state } s); \\ m_{si} &= \#\text{ of times player } i \text{ ... from initial state } s. \end{aligned}$$

- Considered transitions to six absorbing states (scored or missed a 2- or 3-pointer; fouled; turnover) $\Longrightarrow$ in addition to P_{kl}, also $P_{ka} = \mathbb{P}(\text{group } k \text{ ends play in state } a)$, $a = K+1, ..., K+6$.

Example

- Manually annotated two games (Game 1 and Game 5) from the 2012 Eastern Conference finals between the Miami Heat and the Boston Celtics.

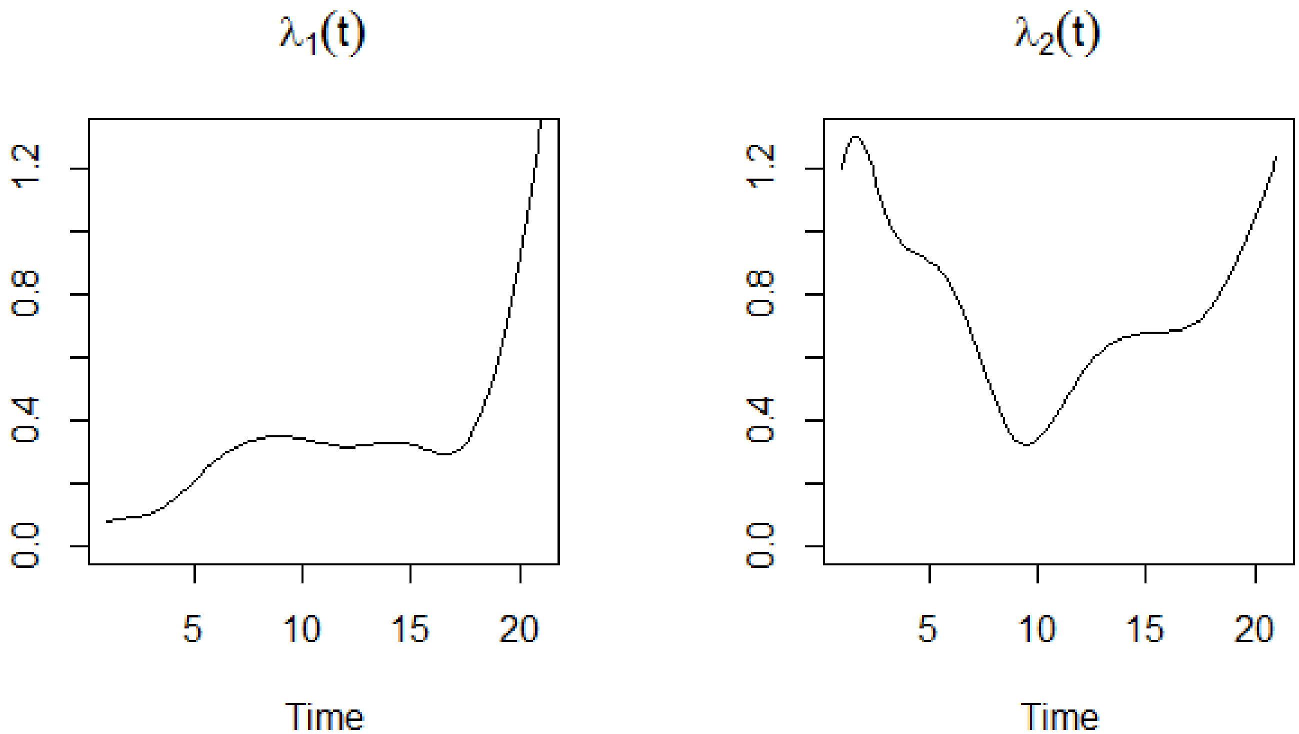
- For each game, only used data from the first 3 quarters.

- For Miami Heat, a total of 541 "transactions" in our data set (85 initial, 334 player-to-player, 122 absorbing).

Result for Miami Heat with $K = 2$

Player	Position	$\mathbb{E}(z_{i1}\|\Theta, \mathbf{T})$	$\mathbb{E}(z_{i2}\|\Theta, \mathbf{T})$
Mario Chalmers	Guard	1.00	0.00
Norris Cole	Guard	1.00	0.00
Dwayne Wade	Guard	1.00	0.00
James Jones	Guard	0.87	0.13
Joel Anthony	Center	0.98	0.02
Lebron James	Forward	1.00	0.00
Shane Battier	Forward	0.00	1.00
Mike Miller	Forward	0.00	1.00
Ronny Turiaf	Center	0.00	1.00

Result for $\lambda_k(t)$



$$
\begin{aligned}
\text{group 1} &= \{\text{mostly guards, Lebron James}\} \\
\text{group 2} &= \{\text{mostly forwards}\}
\end{aligned}
$$

Result for P_{kl}, P_{ka}, Q_{sk}

Q_{sk}	$k = 1$	$k = 2$
inbound	**0.92**	**0.08**
rebound	**0.69**	**0.31**

P_{kl}	$l = 1$	$l = 2$	$l =$ 'a'
$k = 1$	**0.52**	**0.20**	0.28
$k = 2$	**0.71**	**0.09**	0.20

P_{ka}	scored 2pt	scored 3pt	missed 2pt	missed 3pt	fouled	turnover
$k = 1$	0.09	0.01	0.07	0.04	0.05	0.02
$k = 2$	0.03	0.05	0.02	0.09	0.00	0.01

again, group 1 = {mostly guards, Lebron James}

group 2 = {mostly forwards}

Summary

- Modeling:
  - non-homogeneous Poisson process; stochastic block model.
- Inference:
  - latent variables; mixtures; EM algorithm.
- Computation:
  - Gibbs sampler; B-splines; quasi-Newton.
- Application:
  - clustering of NBA players by their playing style.