Change Point Detection in Weighted and Directed Random Dot Product Graphs

Federico Larroca*, Paola Bermolen*, Marcelo Fiori* and Gonzalo Mateos†

*Facultad de Ingeniería, Universidad de la República, Uruguay
†Dept. of Electrical and Computer Engineering, University of Rochester, Rochester, NY, USA
Introduction

Detecting changes on a sequence of random graphs has many applications:

- Social networks
- Neuronal activity
- Wireless network monitoring
Introduction

Detecting changes on a sequence of random graphs has many applications:
- Social networks
- Neuronal activity
- Wireless network monitoring

Two possible approaches:
1. Embedding-based: limited interpretability and lack of theoretical guarantees
2. Probabilistic generative models: theoretically sound results but lack of generality of classic models (e.g. ER or SBM)
Introduction

- Detecting changes on a sequence of random graphs has many applications:
  - Social networks
  - Neuronal activity
  - Wireless network monitoring

- Two possible approaches:
  1. Embedding-based: limited interpretability and lack of theoretical guarantees
  2. Probabilistic generative models: theoretically sound results but lack of generality of classic models (e.g. ER or SBM)

- We resort to the very versatile Random Dot-Product Graph (RDPG) model:
  - Each node $i$ has an associated vector $\mathbf{x}_i \in \mathbb{R}^d$
  - A link exists between nodes $i$ and $j$ with probability $\mathbf{x}_i^T \mathbf{x}_j$
  - Attention: rotation ambiguity
RDPG: intuition

- Vectors may be estimated by spectral decomposition of the adjacency matrix
- Example: the Zachary’s Karate Club graph

**Figure:** Zachary’s Karate Club graph and its resulting $\mathbf{x}_i$ for $d = 2$.

- Intuition: larger vectors tend to be more connected and angle between vectors indicate affinity
Weighted graphs?

- We propose to extend the RDPG model to the weighted case.
- Assume weights $\omega_{i,j}$ are independently drawn from distributions with moment-generating function $M_{\omega_{i,j}}(t) = \mathbb{E}[e^{t\omega_{i,j}}] = \sum_{m=0}^{\infty} \frac{t^m\mathbb{E}[\omega_{i,j}^m]}{m!}$.

Weighted RDPG: Each node has a sequence of vectors $x_i[m] \in \mathbb{R}^{dm}$ where $\mathbb{E}[\omega_{i,j}^m] = x_i[m]^T x_j[m]$.

- Vanilla RDPG is recovered by setting $x_i[m] = x_i \forall m$. 

Example: weighted SBM with $p = 0.5$ and weights $N(\mu = 5, \sigma = 0.1)$ except between a group of nodes where the distribution is Poisson ($\lambda = 5$).
Weighted graphs?

- We propose to extend the RDPG model to the weighted case
- Assume weights $\omega_{i,j}$ are independently drawn from distributions with moment-generating function $M_{\omega_{i,j}}(t) = \mathbb{E}[e^{t\omega_{i,j}}] = \sum_{m=0}^{\infty} \frac{t^m\mathbb{E}[\omega_{i,j}^m]}{m!}$

**Weighted RDPG:** Each node has a sequence of vectors $x_i[m] \in \mathbb{R}^{dm}$ where $\mathbb{E}[\omega_{i,j}^m] = x_i[m]^Tx_j[m]$

- Vanilla RDPG is recovered by setting $x_i[m] = x_i \forall m$
- Vectors $x_i[m]$ may be estimated as in the RDPG case by considering $A^{(m)} = [\omega_{i,j}^m]$
- Example: weighted SBM with $p = 0.5$ and weights $N(\mu = 5, \sigma = 0.1)$ except between a group of nodes where the distribution is Poisson($\lambda = 5$)

(a) $m = 1$  
(b) $m = 2$  
(c) $m = 3$
Application: Change-Point Detection

Due to the rotation ambiguity, vectors $\mathbf{x}_i[m]$ cannot be used directly to detect changes:

- Use the entries of $\hat{Y}[m] = \hat{\mathbf{X}}[m]\hat{\mathbf{X}}[m]^T$ instead (in particular, the entries that do not share nodes such as $(i,j) \in \mathcal{O} = \{(i, i + n/2) \forall i = 1, \ldots, n/2\}$)

---

Example: Wi-Fi network with $n = 6$ APs, with measurements collected hourly during almost four weeks, where an AP was moved at $t \approx 3100$. The evolution of $\hat{Y}_t[1]$ for entries $(i,j) \in \mathcal{O}$ in the RSSI graph. The background color indicates the change-point estimated through our method and the vertical lines by applying different thresholds $\theta$ to the graph.
Application: Change-Point Detection

Due to the rotation ambiguity, vectors $\mathbf{x}_i[m]$ cannot be used directly to detect changes:

- Use the entries of $\hat{\mathbf{Y}}[m] = \hat{\mathbf{X}}[m] \hat{\mathbf{X}}[m]^T$ instead (in particular, the entries that do not share nodes such as $(i,j) \in \mathcal{O} = \{(i, i + n/2) \forall i = 1, \ldots, n/2\}$)

Example: Wi-Fi network with $n = 6$ APs, with measurements collected hourly during almost four weeks, where an AP was moved at $t \approx 310$

Figure: The evolution of $\hat{\mathbf{Y}}_t[1]$ for entries $(i,j) \in \mathcal{O}$ in the RSSI graph. The background color indicates the change-point estimated through our method and the vertical lines by applying different thresholds $th$ to the graph.