

Optimal change point estimator for network data

Sharmodeep Bhattacharyya¹, Shirshendu Chatterjee², and Soumendu Sundar Mukherjee³

¹ Oregon State University, Corvallis OR 97331, USA,

Sharmodeep.Bhattacharyya@oregonstate.edu,

WWW home page: <https://stat.oregonstate.edu/people/bhattacharyya-sharmodeep>

² City University of New York, New York NY 10031, USA

shirshendu@ccny.cuny.edu,

WWW home page: shirshendu.ccny.cuny.edu

³ Indian Statistical Institute, Kolkata 700108, WB India,

soumendu041@gmail.com,

WWW home page: <https://soumendu041.gitlab.io/>

1 Introduction

Change-point detection is a classical problem in statistics which has gained significant importance and applicability in many fields including medical diagnostics, gene expression, spam email filtering, astronomy and finance. Such problems arise in the analysis of various data types including sequentially observed normally distributed data, time-series data and multivariate data for detecting change in different parameters of the data distribution such as mean, variance, correlation, density.

In this article, we tackle the problem of change-point detection in temporal network data. The observable is a sequence of networks indexed by time. The goal is to check if there is any time-point, which will be referred to as change-point, when there is a significant change in the structure of these networks and to estimate the location of such change-points. These problems arise in many applications including (i) brain image analysis, where one observes scanned images of brains collected over time and looks for abnormalities, (ii) ecological networks observed over time, where one checks whether there is any structural change. We stress here that we observe the whole time series ahead of our analysis, this is thus an *offline* or *a posteriori* change-point problem.

Recent work in this area include [6] (Bayesian procedure for hierarchical random graph model), [5, ?] (use of local graph statistics for anomaly detection in dynamic networks), [2] (eigenvalue based test to segregate graph models). Although much empirical work has been done, not much theory can be found (exception is [7]), and most theoretical results focus on particular structures or specialized models. Some recent works [8, ?] propose methods for change point detection in networks generated from block models and graphon models with some theoretical results on the consistency of the detection methods.

The classical CUSUM statistic [4] for univariate change point problems can be used in the network problem as well, and provides a unified way of constructing estimates of change points. A preliminary study of its theoretical properties was carried out in [3]. In this paper, we will further that investigation in a much more general setting under very mild assumptions.

2 Contribution of the paper

We consider the setup where one observes T networks with adjacency matrices $\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \dots, \mathbf{A}^{(T)}$ on the same set of nodes $\{v_1, v_2, \dots, v_n\}$, where the edges in $\mathbf{A}^{(1)}$ appear independently and $\mathbb{E}\mathbf{A}^{(t)} =$



The 8th International Conference on Complex Networks and Their Applications (COMPLEX NETWORKS 2020) will be held in Athens, Greece, from July 13 to 17, 2020.

P_t . The first problem that we address is testing the hypotheses

$$H_0 : P_i = P, 1 \leq i \leq T, \text{ versus } H_1 : \exists 1 \leq \tau \leq T-1 \text{ such that } P_i = \begin{cases} P & 1 \leq i \leq \tau \\ Q & \tau+1 \leq i \leq T. \end{cases}$$

and estimating τ when H_1 is true. Let $\tau \in (\kappa, T - \kappa)$ and Λ is the target precision for estimating τ . Also let $\bar{D}$ be the sample average degree of a node over all layers. Define

$$\begin{aligned} \text{For } \ell \in [T/\Lambda], \text{ let } T_\ell &:= (\ell-1)\Lambda \text{ and } \mathbb{J}_\ell := \cup_{\star \in \{+, -\}} \mathbb{I}_\star \left(T_\ell, T_{\ell+1}; \frac{1}{3}\Lambda \wedge \kappa \right), \\ \text{where } \mathbb{I}_-(a, b; c) &:= \{(a, t] : a+c \leq t \leq b\}, \mathbb{I}_+(a, b; c) := \{(t, b] : a \leq t \leq b-c\} \\ \Gamma &:= n \cdot \min \left\{ \frac{1}{2}, \left(\frac{\log(T/\Lambda)}{\bar{D}^3(\Lambda \wedge \kappa)} \right)^{1/2} \right\}, \\ D_{i,\ell} &:= \max_{J \in \mathbb{J}_\ell} \frac{1}{|J|} \sum_{j \in [n], s \in J} A_{ij}^{(s)} \text{ for } i \in [n] \text{ and } \ell \in [T/\Lambda], \end{aligned}$$

Algorithm 1: Change Point Detection

Input: Adjacency matrices $\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \dots, \mathbf{A}^{(T)}$; cushion κ , scanning window Λ .

Output: Change point estimate $\hat{\tau}$.

Obtain $\bar{D} = \frac{1}{nT} \sum_{i,j \in [n], s \in [T]} A_{ij}^{(s)}$.

Obtain Γ .

For $\ell = 1, 2, \dots, T/\Lambda$ do

 Obtain T_ℓ and $\mathbb{J}_\ell$.

 For $i = 1, 2, \dots, n$ do

 Obtain $D_{i,\ell}$.

 Order the values $D_{1,\ell}, \dots, D_{n,\ell}$ to get $D_{(1),\ell} \leq \dots \leq D_{(n),\ell}$.

 Obtain row indices $i_1, \dots, i_\Gamma$ such that $D_{i_k,\ell} \geq D_{(n+1-\Gamma),\ell}$.

 Obtain $\tilde{\mathbf{A}}^{(s)}$ from $\mathbf{A}^{(s)}$ for each $s \in (T_\ell, T_{\ell+1}]$ by removing rows and columns with indices $i_1, \dots, i_\Gamma$.

 For $t = \frac{1}{3}(\Lambda \wedge \kappa), \frac{1}{3}(\Lambda \wedge \kappa) + 1, \dots, \Lambda - \frac{1}{3}(\Lambda \wedge \kappa)$ do

 Obtain $\mathbf{G}^{(t)} := \frac{1}{t} \sum_{s \in (0,t]} \tilde{\mathbf{A}}^{(s+T_\ell)} - \frac{1}{\Lambda-t} \sum_{s \in (t,\Lambda]} \tilde{\mathbf{A}}^{(s+T_\ell)}$.

 Obtain $u = \arg \max_{t \in (T_\ell, T_{\ell+1}]} \|\mathbf{G}^{(t)}\|$

 If $\|\mathbf{G}^{(u)}\| > C\Psi \left[\frac{\bar{D}}{(\log n)\Lambda \wedge \kappa} \log(CT/\Lambda) \right]^{1/2}$, declare u as a change point.

A natural statistic to consider under our single change-point alternative is based on the cumulative averages of estimates of the P_t s. Such CUSUM statistics are very widely used in change-point problems. We use $\mathbf{A}^{(t)}$ as an estimate of P_t . Then we obtain submatrices $\tilde{\mathbf{A}}^{(t)}$ of $\mathbf{A}^{(t)}$ by removing some high degree vertices as described in Algorithm 1. Then we obtain



$$G_t := \frac{1}{t} \sum_{i=1}^t \mathbf{A}^{(i)} - \frac{1}{T-t} \sum_{i=t+1}^T \mathbf{A}^{(i)} \text{ for } \kappa \leq t \leq T-\kappa.$$

We accept H_0 (no change-point) if $\max_{\kappa \leq t \leq T-\kappa} \|G_t\| \leq C\sqrt{\bar{D}/T}$, where $\bar{D}$ is the estimated average degree, $\|\cdot\|$ denotes the spectral norm and C is a large constant. Otherwise, we accept H_1 (existence of a change-point) and obtain $\hat{\tau}$ (our estimate of the single change-point τ) by

$$\hat{\tau} := \arg \max_{\kappa \leq t \leq T-\kappa} \|G_t\|.$$

In many univariate settings, such CUSUM statistics are minimax optimal (see, e.g., [1]).

Theorem 1. *Let d be the average degree of a node among the networks $A_1, \dots, A_T$. Then $|\hat{\tau} - \tau| = o(T)$ when $\|P - Q\| \gg \sqrt{d/T}$. Also, if $\|P - Q\| \not\gg \sqrt{d/T}$, detection is not be possible.*

We also address the case of multiple change-points, where there are K (unknown) change-points $\tau_1 < \tau_2 < \dots < \tau_K$ and obtain consistent estimates $\hat{K}, \hat{\tau}_1, \dots, \tau_K$ under minimal assumptions on network parameters.

Comment: Note that, Theorem 2.1 states the optimal condition on operator norm of the difference between the connection probability matrices, $\|P - Q\|$ for recovery of change-point. The optimal condition depends on $\sqrt{d/T}$.

Summary. Based on a finite sequence of observed independent network data, we provide an algorithm to test if there is any change-point and estimate the location of the change-points (if any). The algorithm works effectively whenever the signal (norm of the difference of means of the networks before and after the change-point) is above the detectability threshold irrespective of whether the networks are sparse or dense.

References

1. E. Brodsky and B. S. Darkhovsky. *Nonparametric methods in change point problems*, volume 243. Springer Science & Business Media, 2013.
2. J. Cape, M. Tang, and C. E. Priebe. The kato–temple inequality and eigenvalue concentration with applications to graph inference. *Electronic Journal of Statistics*, 11(2):3954–3978, 2017.
3. Soumendu Sundar Mukherjee. *On Some Inference Problems for Networks*. PhD thesis, May 2018.
4. E. S. Page. On problems in which a change in a parameter occurs at an unknown point. *Biometrika*, 44(1/2):248–252, 1957.
5. Y. Park, C. E. Priebe, and A. Youssef. Anomaly detection in time series of graphs using fusion of graph invariants. *IEEE journal of selected topics in signal processing*, 7(1):67–75, 2013.
6. L. Peel and A. Clauset. Detecting change points in the large-scale structure of evolving networks. In *AAAI*, pages 2914–2920, 2015.
7. S. Roy, Y. Atchadé, and G. Michailidis. Change point estimation in high dimensional markov random-field models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(4):1187–1206, 2017.
8. Daren Wang, Yi Yu, and Alessandro Rinaldo. Optimal change point detection and localization in sparse dynamic networks. *arXiv preprint arXiv:1809.09602*, 2018.



The 8th International Conference on Complex Networks and Their Applications. 10 - 12 Dec., 2019, Lisbon, Portugal