Quickest Detection of the Change of Community via Stochastic Block Models

Fei Sha
Department of statistics
University of Nebraska-Lincoln
Lincoln, NE
fei.sha@huskers.unl.edu

Ruizhi Zhang
Department of statistics
University of Nebraska-Lincoln
Lincoln, NE
rzhang35@unl.edu

2022 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-2159-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIT50566.2022.9834562

***Abstract*—Community detection is a fundamental problem in network analysis and has important applications in sensor networks and social networks. In many cases, the community structure of the network may change at some unknown time and thus it is desirable to come up with efficient monitoring procedures that can detect the change as quickly as possible. In this work, we use the Erdős-Rényi model and the bisection stochastic block model (SBM) to model the pre-change and post-change distributions of the network, respectively. That is, initially, we assume there is no community in the network. However, at some unknown time, a change occurs, and two communities are formed in the network. We then propose an efficient monitoring procedure by using the number of k-cycles in the graph. The asymptotic detection properties of our proposed procedure are derived when all parameters are known. A generalized likelihood ratio (GLR) type detection procedure and an adaptive CUSUM type detection procedure are constructed to address the problem when parameters are unknown.**

***Index Terms*—Change detection, Community detection, Erdős-Rényi model, Bisection stochastic block model**

I. Introduction

Community detection is a fundamental problem in network analysis and has important applications in sensor networks [1] and social networks [2]. In this work, we investigate the quickest detection problem of the change of the community pattern in the network. Particularly, we assume initially, the network follows the Erdős-Rényi (ER) model [3], [4]. That is, it starts in the homogeneous state, all nodes are connected randomly with the same probability, and there is no community structure in the graph. Then, a change occurs at some unknown time, and two different communities emerge among individuals with similarity. This network with communities can be modeled by the bisection stochastic block model (SBM), which is widely studied in the literature of community detection [5]–[7]. Therefore, this quickest change detection problem of the community pattern can be formulated by the problem to detect the change between Erdős-Rényi model and the bisection SBM. This problem has many potential applications. For example, in an online retail network, identifying the community of customers with similar purchasing interests can be helpful to establish an efficient recommendation system [8]. However, it is non-trivial to develop efficient monitoring procedures due to the computational complexity of the likelihood function in bisection SBM.

We should mention that the Erdős-Rényi model and the stochastic block model are two fundamental models for network data in offline community detection research [5]–[7]. For a review of recent community detection work in the SBM, see [9]. In the Erdős-Rényi model, all nodes can be connected randomly with a constant probability. In the SBM, nodes are assigned into communities such that two nodes within the same community have a different probability of being connected than the nodes from different communities. For the problem of online community detection, [10] focuses on the Erdős-Rényi model and consider that change will affect the connectivity probability of edges in some unknown subgraph. However, no research has been done on the problem of quickest detection on the change from the Erdős-Rényi model to the SBM model.

In the literature of sequential change-point detection, when monitoring independent streaming data, many efficient likelihood-based procedures have been developed, such as Page's CUSUM procedure [11], Shiryaev-Roberts procedure [12]. These procedures enjoy nice optimality when the pre-change and post-change distribution functions are fully specified [12], [13]. When the post-change parameters are unknown, generalized likelihood ratio based procedures [14] and the adaptive estimation of the post-change parameters based procedures [15] are often employed to detect the possible change. However, all these methods are based on the likelihood functions, which are difficult to compute in the bisection SBM.

In this paper, we focus on sparse graphs with the constant average degree. To overcome the computational difficulty of likelihood function in SBM, we propose computationally efficient online monitoring procedures by the number of k-cycles in the graph. Our proposed methods are inspired by the fact that the number of cycles of length $\log^{1/4}(N)$ in graph is asymptotically Poisson distributed with different means in both Erdős-Rényi model and bisection SBM [16]. Thus, we propose to use the CUSUM type detection procedure based on the number of k-cycles instead of raw graphic data directly when parameters in the Erdős-Rényi model and the bisection SBM are both completely specified. Then, when parameters are unknown, we adopt the generalized likelihood ratio (GLR) test framework to form a GLR type detection procedure. To obtain a computationally efficient algorithm, we propose to construct an adaptive CUSUM approach borrowing the idea from [15]. Numerical results show that the adaptive CUSUM approach not only can be implemented in real-time due to its recursive form, but also is able to detect the unknown change of the community structure in the network fast.

The rest of the paper is organized as follows. We first provide the problem formulation in Section II. The detail of our proposed CUSUM type detection procedure and its theoretical property are presented in Section III. We develop the GLR type and adaptive CUSUM detection procedure in Section IV to address the detection problem with unknown parameters. The simulation results are presented in Section V. Finally Section VI contains the conclusion and further discussions.

II. Problem Formulation

For a given network with N nodes, the graphical structure can be observed as a sequence of undirected adjacency matrices, i.e., $G^{(1)}, G^{(2)}, \cdots$, where the adjacency matrix $G^{(t)} \in \{0,1\}^{N\times N}$ characterizes the interaction information between different nodes at each time t, i.e., $G_{ij}^{(t)} = 1$ if and only if there is an edge between node i and node j. Initially, we assume $G^{(t)}$ follows the classical Erdős-Rényi model, which is denoted by $\mathcal{G}(N, \frac{a+b}{2N})$, where a, b are positive fixed constants. In the Erdős-Rényi model, each edge between two nodes i, j is connected randomly and independently with fixed probability $\frac{a+b}{2N}$, i.e., $\mathbf{P}(G_{ij}^{(t)} = 1) = \frac{a+b}{2N}$. Then, at an unknown time τ, the distribution of $G^{(t)}$ changes and follows the bisection stochastic block model (SBM) with the same average edge degree. We denote the bisection SBM as $\mathcal{G}_2(N, \frac{a}{N}, \frac{b}{N})$. In the bisection SBM, each node $i \in \{1, \cdots, N\}$ is assigned with a label $\sigma_i \in \{\pm\}$ with probability $\frac{1}{2}$. Then each edge between two nodes i, j is connected with probability $\frac{a}{N}$ if $\sigma_i = \sigma_j$ and probability $\frac{b}{N}$ if $\sigma_i \neq \sigma_j$, i.e., $\mathbf{P}(G_{ij}^{(t)} = 1|\sigma_i = \sigma_j) = \frac{a}{N}$ and $\mathbf{P}(G_{ij}^{(t)} = 1|\sigma_i \neq \sigma_j) = \frac{b}{N}$. Therefore, after the change, the network has two different communities. In our study, we focused on the assortative network, which implies that two nodes from the same community are more likely to be connected, i.e., $a > b$.

In this case, the problem of change-point detection between the Erdős-Rényi model and bisection SBM can be formulated by the following hypothesis testing problem:

$$\begin{aligned} H_0 :& G^{(t)} \sim \mathcal{G}(N, \frac{a+b}{2N}); \quad t = 1, 2, \cdots, \\ H_1 :& G^{(t)} \sim \mathcal{G}(N, \frac{a+b}{2N}); \quad t = 1, 2, \cdots, \tau - 1, \\ & G^{(t)} \sim \mathcal{G}_2(N, \frac{a}{N}, \frac{b}{N}); \quad t = \tau, \cdots, \end{aligned} \tag{1}$$

where τ is the unknown change point. The problem (1) is illustrated in Fig 1. The objective is to raise an alarm as quickly as possible after the change occurs.

An online monitoring scheme for detecting the change-point can be defined as a stopping time T, which can be viewed as the time when we raise an alarm to declare that a change has occurred. Here T is an integer-valued random variable, the decision $\{T = t\}$ is based only on the observations in the first t time steps. To evaluate the performance of T, we denote the probability measure and expectation as $\mathbf{P}_{(\infty)}$ and $\mathbf{E}_{(\infty)}$ when the data $G^{(t)}$'s are i.i.d. with distribution $\mathcal{G}(N, \frac{a+b}{2N})$, and use $\mathbf{P}_{(\tau)}$ and $\mathbf{E}_{(\tau)}$ to denote the same when the change occurs at time τ and $G^{(t)}$'s have the post-change distribution $\mathcal{G}_2(N, \frac{a}{N}, \frac{b}{N})$. Under the standard minimax formulation for online change-point detection [13], the performance of a stopping time T is evaluated by the average run length (ARL) to the false alarm, $\mathbf{E}_{(\infty)}(T)$ and the worst-case detection delay

$$\mathbf{D}(T) = \sup_{\tau \geq 1} \operatorname{ess\,sup} \mathbf{E}_\tau[(T-\tau)^+ | G^{(1)}, \cdots, G^{(\tau-1)}]. \tag{2}$$

An efficient online monitoring scheme T should have a small detection delay $\mathbf{D}(T)$ subject to the false alarm constraint

$$\mathbf{E}_{(\infty)}(T) \geq \gamma \tag{3}$$

for some pre-specified large constant $\gamma > 0$.

We should emphasize that the reason why we consider the change-point detection problem where the pre-change distribution is $\mathcal{G}(N, \frac{a+b}{2N})$ and the post-change distribution is $\mathcal{G}_2(N, \frac{a}{N}, \frac{b}{N})$ is that both models have the same average number of edges as $(a+b)(N-1)/4$. Thus, it is non-trivial to detect such change by simply monitoring the number of edges in the graphic network.

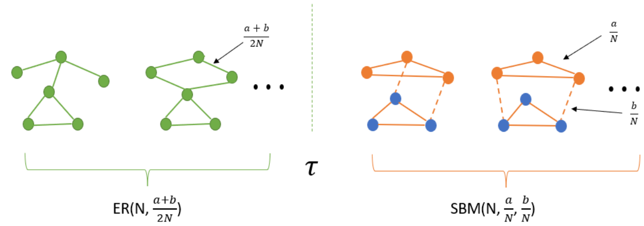

Fig. 1. The pre-change graphs follow the classical Erdős-Rényi model $\mathcal{G}(N, \frac{a+b}{2N})$. After time τ, the structure of graphs changed. The post-change graphs follow the bisection SBM $\mathcal{G}_2(N, \frac{a}{N}, \frac{b}{N})$. There are two communities in the bisection SBM, the nodes in the same community have the connectivity probability $\frac{a}{N}$(solid line), while the nodes belong to the different communities have the connectivity probability $\frac{b}{N}$(dashed line).

III. Our Proposed Detection Method with Known Parameters

In this section, we proposed our detection procedure to solve the change detection problem in (1). We start with the cumulative sum (or CUSUM) procedure, which has been widely used when the pre-change and post-change distributions are completely specified.

Specifically, for the detection problem (1), under the null hypothesis, i.e., $G^{(t)} \sim \mathcal{G}(N, \frac{a+b}{2N})$, the likelihood function of $G^{(t)}$ for can be written as

$$L_0(G^{(t)}) = \prod_{1 \leq i < j \leq N} (\frac{a+b}{2N})^{G_{ij}^{(t)}} (1 - \frac{a+b}{2N})^{1 - G_{ij}^{(t)}}. \tag{4}$$

Under the alternative hypothesis, after the change point τ, the graph $G^{(t)}$ follows $\mathcal{G}_2(N, \frac{a}{N}, \frac{b}{N})$, and the likelihood function of $G^{(t)}$ can be expressed as

$$L_1(G^{(t)}) = \frac{\sum_{\sigma\in\{\pm\}^N} \prod_{1\le i<j\le N} (p_{ij}(\sigma))^{G_{ij}^{(t)}} (1-p_{ij}(\sigma))^{1-G_{ij}^{(t)}}}{2^N}, \quad (5)$$

where

$$p_{ij}(\sigma) = \begin{cases} \frac{a}{N}, \sigma_i = \sigma_j, \\ \frac{b}{N}, \sigma_i \neq \sigma_j. \end{cases}$$

Then, the CUSUM statistic can be written as a recursive form:

$$W_{\text{CUSUM}}^{(t)} = \max(W_{\text{CUSUM}}^{(t-1)} + \log(\frac{L_1(G^{(t)})}{L_0(G^{(t)})}), 0). \quad (6)$$

The CUSUM procedure is then defined as the first time when the CUSUM statistic exceeds some pre-defined threshold c. That is, the CUSUM procedure is given by

$$T_{\text{CUSUM}}(c) = \inf\left\{t : W_{\text{CUSUM}}^{(t)} \ge c\right\}, \quad (7)$$

where the threshold c is a pre-set constant to control the false alarm rates. The CUSUM procedure enjoys nice optimal properties when the pre- and post-change distributions are known [13].

Although the CUSUM statistic has a recursive form, it is computationally expensive to compute the updated CUSUM statistic in (6) sequentially due to the 2^N terms in the summation when calculating $L_1(G^{(t)})$ in (5), especially when the number of nodes N is large. To overcome the computational difficulty, we proposed a detection procedure based on the number of k-cycles of the network. In graph theory, k-cycles refers to a circuit with k vertices in which only the first vertex (which is also the last) appears twice. Our proposed procedure is inspired by the fact that the number of k-cycles are approximately Poisson distributed in both the Erdős-Rényi model and the bisection SBM but with different means, see [16]. This result is summarized in the following lemma.

Lemma 1. *For a graph $G^{(t)}$ with N nodes, let's denote $X_k^{(t)}$ as the number of k-cycles of graph $G^{(t)}$. If $k = O(\log^{1/4}(N))$, we have*

1) *If $G^{(t)} \sim \mathcal{G}(N, \frac{a+b}{2N})$, then $X_k^{(t)} \xrightarrow{d}$ Poi(λ_k);*
2) *If $G^{(t)} \sim \mathcal{G}_2(N, \frac{a}{N}, \frac{b}{N})$, then $X_k^{(t)} \xrightarrow{d}$ Poi$(\lambda_k(1+\delta_k))$,*

where

$$\lambda_k = \frac{1}{2k}(\frac{a+b}{2})^k, \delta_k = (\frac{a-b}{a+b})^k.$$

Therefore, instead of constructing the CUSUM procedure in (7) by using the pre-change and post-change distributions of the graph $G^{(t)}$ directly, we can use the pre-change and post-change distributions of the number of the k-cycles $X_k^{(t)}$. Specifically, we propose to consider the following change detection problem of $X_k^{(t)}$:

$$\begin{aligned} H_0 :& X_k^{(t)} \sim \text{Poi}(\lambda_k); \quad t = 1, 2, \cdots \\ H_1 :& X_k^{(t)} \sim \text{Poi}(\lambda_k); \quad t = 1, \cdots, \tau - 1 \\ & X_k^{(t)} \sim \text{Poi}(\lambda_k(1+\delta_k)); \quad t = \tau, \cdots, \end{aligned} \quad (8)$$

where

$$\lambda_k = \frac{1}{2k}(\frac{a+b}{2})^k, \delta_k = (\frac{a-b}{a+b})^k. \quad (9)$$

Then, when the parameters a, b are known, we can derive the CUSUM statistic of $X_k^{(t)}$ by

$$W_{\text{C}}^{(t)} = \max(W_{\text{C}}^{(t-1)} + X_k^{(t)} \log(1+\delta_k) - \lambda_k\delta_k, 0), \quad (10)$$

and the corresponding monitoring procedure can be written by

$$T_{\text{C}}(c) = \inf\{t : W_{\text{C}}^{(t)} > c\}. \quad (11)$$

Note that the monitoring procedure proposed in (11) is computationally simple because at each time, it only need to count the number of k-cycles in each graph, which can be obtained by efficient algorithm such as depth first search (DFS) [17].

Moreover, by the properties of classical CUSUM [13], we can derive theoretical properties of our proposed method.

Theorem 1. *If $k = O(\log^{1/4}(N))$, we have $\mathbf{E}_\infty(T_C(c)) \ge e^c$. Moreover, when $\gamma \to \infty$, we have*

$$\mathbf{D}(T_C(c)) \sim \frac{c}{I}, \quad (12)$$

where

$$I = \lambda_k(1+\delta_k)\log(1+\delta_k) - \lambda_k\delta_k,$$

is the Kullback–Leibler (KL) divergence between the distribution of Poi(λ_k) and Poi$(\lambda_k(1+\delta_k))$.

By Theorem 1, we can see by setting $c = c_\gamma = \log\gamma$, our proposed method $T_C(c_\gamma)$ satisfies the false alarm constraint in (3). In the meantime, the corresponding detection delay satisfies $\mathbf{D}(T_{\text{C}}(c_\gamma)) \sim \log(\gamma)/I$. Let $m = \frac{(a-b)^2}{2(a+b)}$. Then we should notice the detection delay $\mathbf{D}(T_{\text{C}}(c_\gamma)) \le 6k\log(\gamma)/(m^k)$. This result implies the detection performance of our proposed method depends on whether $m > 1$ or not. If $m > 1$ and $k \gg \log\log\gamma$, the upper bound of the detection delay goes to 0, which implies that our proposed procedure (10) can detect the change very quickly. However, if $m < 1$, the upper bound will go to infinity as $k \to \infty$. This result implies the detection delay of our proposed procedure will be large when $m < 1$. This phase transition phenomena is consistent to the well studied phenomena in the offline community detection literature [16]: if $m < 1$, no consistent tests exist to distinguish the ER model $\mathcal{G}(N, \frac{a+b}{2N})$ and the bisection SBM $\mathcal{G}_2(N, \frac{a}{N}, \frac{b}{N})$. If $m > 1$, consistent tests are available to distinguish these two models.

IV. Detection procedures with unknown parameters

Although our proposed detection procedure $T_{\text{C}}(c)$ in (11) is computationally simple and also enjoys good theoretical properties, its implementation requires the full information of parameters a, b, which may be unknown in practice. In this section, we proposed a GLR type detection procedure and an adaptive detection procedure to address the problem when the parameters a, b are unknown. On the high level, we still monitor the process by counting the number of k-cycles $X_k^{(t)}$. Based on Lemma 1, when $k = O(\log^{1/4}(N))$, the pre-change and post-change distributions of $X_k^{(t)}$ follow Poisson distribution with different means depending on the unknown parameters λ_k, δ_k. Thus, it is suffice to construct detection procedures that can estimate λ_k, δ_k sequentially.

For the estimation of $\lambda_k = \frac{1}{2k}(\frac{a+b}{2})^k$, note either in pre-change or post-change distribution, the expectation of number of edges is $(a+b)(N-1)/4$. Therefore, we can use the number of edges in each graph to estimate λ_k. Specifically, for the observed graph $G^{(t)}$, denote the corresponding number of edges by $|E^{(t)}|$. Then, we propose to estimate λ_k by

$$\hat{\lambda}_{k,t} = \frac{1}{2k}(\frac{2|E^{(t)}|}{N})^k, \tag{13}$$

which is consistent in both pre-change and post-change distributions. Then, we adopt the generalized likelihood ratio (GLR) framework to solve the change-point detection problem with unknown δ_k. The key idea of GLR is that, when the post-change model involves unknown parameters, GLR statistic finds the maximum likelihood estimate (MLE) of the post-change parameter and plugs it back into the likelihood ratio to form the detection statistic. Without knowing where change occurs beforehand when calculating the GLR statistic, we need to search over all possible change locations.

Specifically, the log-likelihood ratio statistic of problem (8) when $t \geq \tau$ is

$$R(\tau, \lambda_k, \delta_k) = \sum_{i=\tau}^{t}(X_k^{(i)} \log(1+\delta_k) - \lambda_k\delta_k). \tag{14}$$

Using the observed post-change samples of the number of k-cycles $X_k^{(\tau)}, X_k^{(\tau+1)}, \cdots, X_k^{(t)}$, we can get the MLE of the unknown parameter δ_k by

$$\hat{\delta}_{k,t} = \frac{M_t - M_{\tau-1}}{\hat{\lambda}_{k,t}(t-\tau+1)} - 1,$$

where $M_t = \sum_{i=1}^{t} X_k^{(i)}$ and $\hat{\lambda}_{k,t}$ defined in (13).

Then, we can get

$$\begin{aligned} R(\tau, \hat{\lambda}_{k,t}, \hat{\delta}_{k,t}) &= (M_t - M_{\tau-1})[\log\frac{(M_t - M_{\tau-1})}{(t-\tau+1)\hat{\lambda}_{k,t}} - 1] \\ &+ (t-\tau+1)\hat{\lambda}_{k,t}. \end{aligned}$$

The GLR statistic at time t is

$$W_{\mathrm{G}}^{(t)} = \max_{1\leq\tau\leq t} R(\tau, \hat{\lambda}_{k,t}, \hat{\delta}_{k,t}), \tag{15}$$

and the corresponding detection procedure is defined by

$$T_{\mathrm{G}}(c) = \inf\{t : W_{\mathrm{G}}^{(t)} > c\}, \tag{16}$$

where the threshold c is a pre-set constant to control the false alarm rates.

However, it is not computationally efficient to apply this procedure because we need to recalculate MLE and update the GLR statistic in (15) when a new data is coming. To save the computation cost, we adopt the adaptive CUSUM approach proposed by Lorden and Pollak [15] and propose an adaptive detection procedure for the online community detection problem in (1). The key idea is to replace the unknown parameters λ_k, δ_k in the CUSUM statistic in (10) by their adaptive estimations. In that way, we can get the following adaptive CUSUM statistic:

$$W_A^{(t)} = \max(W_A^{(t-1)} + X_k^{(t)} \log(1+\hat{\delta}_{k,t}) - \hat{\lambda}_{k,t}\hat{\delta}_{k,t}, 0). \tag{17}$$

Here, we propose to use the same $\hat{\lambda}_{k,t}$ as in (13). For $\hat{\delta}_{k,t}$, we propose to adopt the post-change parameter estimation approach in [15], which has a nice recursive form. That is, at each time t, the CUSUM-type detection statistics can produce a candidate post-change time $\hat{v} \in \{0, 1, \cdots, t-1\}$ and thus the observations $X_k^{(\hat{v})}, X_k^{(\hat{v}+1)}, \cdots, X_k^{(t-1)}$ can be used to estimated the unknown parameter δ_k.

Specifically, at time t, denote $\hat{v}$ as the largest $i \in [0, 1, \cdots, t-1]$ such that $W_A^{(i)} = 0$, and denote T_t and S_t as the total number and the summation of observations $X_k^{(i)}$'s between the candidate post-change time $\hat{v}$ and time step $t-1$, i.e.,

$$T_t = t - \hat{v}, S_t = \sum_{i=\hat{v}}^{t-1} X_k^{(i)}.$$

For the Poisson distribution, the MLE of the post-change mean $\lambda_k(1+\delta_k)$ is S_t/T_t. Assume ρ is the smallest shift of δ_k that is meaningful in practice, we can then estimate δ_k at time t by

$$\hat{\delta}_{k,t} = \max(\rho, \frac{S_t}{T_t\hat{\lambda}_{k,t}} - 1). \tag{18}$$

Then, the adaptive CUSUM detection procedure is given by

$$T_{\mathrm{A}}(c) = \inf\{t : W_{\mathrm{A}}^{(t)} > c\}. \tag{19}$$

From the algorithm viewpoint, our proposed algorithm can be recursively implemented as follows. Let $S_0 = T_0 = W_{\mathrm{A}}^{(0)} = X_k^{(0)} = 0$, and $\hat{\delta}_{k,0} = 0$. For time $t \geq 1$, $W_{\mathrm{A}}^{(t)}$ is updated by (17), where $\hat{\lambda}_{k,t}$ is defined in (13), $\hat{\delta}_{k,t}$ is defined in (18) and

$$\begin{pmatrix} S_t \\ T_t \end{pmatrix} = \begin{cases} \begin{pmatrix} S_{t-1}+X_k^{(t-1)} \\ T_{t-1}+1 \end{pmatrix}, & \text{if } W_{\mathrm{A}}^{(t-1)} > 0 \\ \begin{pmatrix} X_k^{(t-1)} \\ 1 \end{pmatrix}, & \text{if } W_{\mathrm{A}}^{(t-1)} = 0 \end{cases} \tag{20}$$

From above equation (17) and (20), we see the detection statistics $W_{\mathrm{A}}^{(t)}$'s can be computed recursively as the part of three-dimensional vectors $(S_t, T_t, W_{\mathrm{A}}^{(t)})$.

Therefore, our proposed detection procedure $T_{\mathrm{A}}(c)$ in (19) can be implemented fast.

V. Simulation Study

In this section, we report the simulation results for our proposed CUSUM type detection procedure T_{C} in (11), the GLR type detection procedure T_{G} in (16), and the adaptive CUSUM detection procedure T_{A} in (19). We assume that before the change, the graph $G^{(t)}$ follows the Erdős-Rényi model $\mathcal{G}(N, \frac{a+b}{2N})$. After the change, the graph $G^{(t)}$ follows the bisection SBM $\mathcal{G}_2(N, \frac{a}{N}, \frac{b}{N})$. We set two different sizes of graph N, the total number of nodes, as 100 and 500. For connection probability, we consider two settings $(a, b) = (5, 2), (8, 2)$, where the corresponding threshold $m = \frac{(a-b)^2}{2(a+b)}$ is $m = 0.64$ and $m = 1.8$, respectively. Moreover, we set $k = 3$ for all of our methods, so that the number of cycles with 3 edges for each graph is used to detect the change of community pattern.

For all detection procedures, we conduct 100 Monte Carlo simulations to search the thresholds c to satisfy the false alarm constraint with $\gamma = 1000$. Using the obtained threshold for each method, we then simulate the detection delay when the change occurs at time $\tau = 1$ based on 100 Monte Carlo simulations. The average and standard deviation of detection delays of three procedures under different parameter settings for $N = 100$ and $N = 500$ are summarized in Table I and Table II, respectively.

	$a=5, b=2 (m=0.64)$
$T_{\mathrm{C}}(c=2.8)$	79.4(2.62)
$T_{\mathrm{G}}(c=9.3)$	80.2(3.10)
$T_{\mathrm{A}}(c=5.3)$	81.6(3.04)
	$a=8, b=2 (m=1.8)$
$T_{\mathrm{C}}(c=6.6)$	16.9(1.18)
$T_{\mathrm{G}}(c=11.9)$	12.4(1.10)
$T_{\mathrm{A}}(c=5.3)$	14.8(1.14)

TABLE I
COMPARISON OF DETECTION DELAY FOR THREE DETECTION PROCEDURES WHEN $N = 100$ AND $\gamma = 1000$

	$a=5, b=2 (m=0.64)$
$T_{\mathrm{C}}(c=3.0)$	77.5(2.54)
$T_{\mathrm{G}}(c=7.3)$	79.9(2.92)
$T_{\mathrm{A}}(c=5.5)$	79.8(3.13)
	$a=8, b=2 (m=1.8)$
$T_{\mathrm{C}}(c=5.3)$	11.6(0.78)
$T_{\mathrm{G}}(c=7.8)$	13.9(0.69)
$T_{\mathrm{A}}(c=5.9)$	10.9(0.69)

TABLE II
COMPARISON OF DETECTION DELAY FOR THREE DETECTION PROCEDURES WHEN $N = 500$ AND $\gamma = 1000$

From Table I, by comparing the detection delays of the same procedure under different settings when $N = 100$, we see when $m > 1$, all three procedures have much smaller detection delays compared with the situation when $m < 1$. This result matches our theoretical results in Theorem 1 that when the graph size N is large, it is easier to detect the change of the community pattern when $m > 1$ compared to the case when $m < 1$. From Table II, we get similar results when $N = 500$. That is, when $m > 1$, our proposed procedures can detect the change quickly.

Moreover, when $a = 8, b = 2$, we can see an interesting and counterintuitive result. In this case, we find the detection delays of the GLR type and adaptive CUSUM procedure have smaller detection delays compared with the CUSUM type procedure, which is constructed using the true parameters. One possible explanation for this phenomena maybe because all of these methods are trying to detect the change of community patterns by counting the number of k-cycles in each graph. The Poisson approximation in Lemma 1 is used to construct these methods. However, for some finite values of k and N, the estimation error may be large so that we may lose some information of original graphs. Therefore, CUSUM type detection procedure T_{C} with correct parameters may not has the smallest detection delay.

All simulations were conducted on a Windows 10 Laptop with Intel i7-8750H CPU 2.20GHz using Python 3.8. In the procedure of conducting 100 Monte Carlo simulations for T_{G} and T_{A}, computation time for the GLR type procedure T_{G} and adaptive procedure T_{A} differ greatly. When we determine the thresholds for different procedures, we can simulate the detection delays of T_{G} and T_{A}. For example, when $a = 8, b = 2$, it takes about 93.8 minutes for GLR type procedure T_{G} to simulate the detection delays in that 100 Monte Carlo simulations. For adaptive procedure T_{A}, it takes about 1.0 minute to simulate the detection delays for 100 Monte Carlo simulations. We see although the GLR type procedure T_{G} has smaller detection delay compared to the adaptive CUSUM procedure T_{A}, it takes much longer time to compute and implement that method. The computational advantage of T_{A} is evident.

VI. Conclusion

We study the problem of quickest detection of the change of the community structure from the ER model to the bisection SBM model. Computationally and statistically efficient detection procedures are constructed by using the number of k-cycles in the graph. An interesting future direction would be to explore the optimality for such change detection problem. Moreover, the proposed framework may be extended to solve the quickest detection problem of the general stochastic block models.

Acknowledgements

R. Zhang's research was supported in part by National Science Foundation (NSF) grant ECCS-2112740.

References

[1] B. Pásztor, L. Mottola, C. Mascolo, G. P. Picco, S. Ellwood, and D. Macdonald, "Selective reprogramming of mobile sensor networks through social community detection," in *European Conference on Wireless Sensor Networks*. Springer, 2010, pp. 178–193.

[2] P. Basuchowdhuri, S. Sikdar, V. Nagarajan, K. Mishra, S. Gupta, and S. Majumder, "Fast detection of community structures using graph traversal in social networks," *Knowledge and Information Systems*, vol. 59, no. 1, pp. 1–31, 2019.

[3] P. Erdös and A. Rényi, "On random graphs i," *Publicationes Mathematicae Debrecen*, vol. 6, p. 290, 1959.

[4] P. Erdos, A. Rényi *et al.*, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci*, vol. 5, no. 1, pp. 17–60, 1960.
[5] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, "Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications," *Physical Review E*, vol. 84, no. 6, p. 066106, 2011.
[6] S. Heimlicher, M. Lelarge, and L. Massoulié, "Community detection in the labelled stochastic block model," *arXiv preprint arXiv:1209.2910*, 2012.
[7] R. R. Nadakuditi and M. E. Newman, "Graph spectra and the detectability of community structure in networks," *Physical review letters*, vol. 108, no. 18, p. 188701, 2012.
[8] P. K. Reddy, M. Kitsuregawa, P. Sreekanth, and S. S. Rao, "A graph based approach to extract a neighborhood customer community for collaborative filtering," in *International Workshop on Databases in Networked Information Systems*. Springer, 2002, pp. 188–200.
[9] E. Abbe, "Community detection and stochastic block models: recent developments," *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 6446–6531, 2017.
[10] L. Xie and Y. Xie, "Optimality of graph scanning statistic for online community detection," *arXiv preprint arXiv:2102.05821*, 2021.
[11] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–115, 1954.
[12] A. N. Shiryaev, "On optimum methods in quickest detection problems," *Theory of Probability & Its Applications*, vol. 8, no. 1, pp. 22–46, 1963.
[13] G. Lorden, "Procedures for reacting to a change in distribution," *The Annals of Mathematical Statistics*, pp. 1897–1908, 1971.
[14] T. L. Lai, "Information bounds and quick detection of parameter changes in stochastic systems," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2917–2929, 1998.
[15] G. Lorden and M. Pollak, "Sequential change-point detection procedures that are nearly optimal and computationally simple," *Sequential Analysis*, vol. 27, no. 4, pp. 476–512, 2008.
[16] E. Mossel, J. Neeman, and A. Sly, "Reconstruction and estimation in the planted partition model," *Probability Theory and Related Fields*, vol. 162, no. 3, pp. 431–461, 2015.
[17] R. Tarjan, "Depth-first search and linear graph algorithms," *SIAM journal on computing*, vol. 1, no. 2, pp. 146–160, 1972.