IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received June 7, 2020, accepted June 30, 2020, date of publication July 13, 2020, date of current version July 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3008708

Local Statistic With Dynamic Vertex Selection for Change-Point Detection in Stochastic Block Networks

YANPING ZHAO, BO WANG, MINGAN LUAN, (Graduate Student Member, IEEE), AND FENGYE HU, (Member, IEEE)
College of Communication Engineering, Jilin University, Changchun 130025, China
Corresponding author: Bo Wang (roywang@jlu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771217, in part by the Science and Technology Project of Jilin Province Department of Education under Grant JJKH20190164KJ, and in part by the Natural Science Foundation of Jilin Province under Grant 20180101051JC.

ABSTRACT Change-point detection within random networks is essential for many applications. Generally, the typical methods focus on the Erdős-Rényi random networks, or assume that the anomalous subnetworks only have high link probability with the fixed membership. In this paper, we consider the stochastic block model of random graphs, and study the change-point detection regarding to the scenario that after a change-point, the connectivity of subnetworks becomes denser or sparser while the membership of nodes also changes. Based on local graph features, we explore a local statistic with dynamic vertex selection for detecting the emergence of an abrupt change-point. In addition, we derive an analytic expression with respect to average run length to set detection threshold in a theoretical fashion, and achieve the probability bounds related to the dynamic vertex selection to characterize the performance of the presented algorithm. As a result, the proposed scheme can provide performance improvement as well as reduce the computational complexity. The proposed algorithm can address a more general problem than the typical methods. Numerical experiments are provided to show the effectiveness of our method.

INDEX TERMS Anomaly detection, change-point detection, stochastic block model, random networks, local statistic.

I. INTRODUCTION

The study of anomaly detection or change-point detection within networks is a fundamental but important topic in signal processing, and has a wide range of applications ranging from sensor networks, smart grid, vehicle networks to internet of things [1]–[8]. In one type of scenarios, the network data of interest are composed of nodes and the relationships between nodes, where the relationships can be the communication links [11] or correlation structures [9]. Conventionally, the network data are named as the relational data, and can be naturally modeled using random graph form. In this paper, the problem under consideration is to detect the emergence of a change-point that affects the statistical behaviors of a small fraction of the network data. In other words, given time series of random graphs over time $G_1, G_2, \cdots$, in a streaming fashion, we are interested in obtaining a change-point as soon as possible after it occurs subject to the constraint of the false alarm probability or average run length (ARL).

Statistical models, involving a form of graphical representation, is of significant importance for both analysis of the network data as well as change-point detection in random networks. Accordingly, many graph models have been developed over the past several decades [10]. For example, given a vertex set, the Erdős-Rényi random graphs can represent the random networks where the edges between any two vertices are the Bernoulli random variables with a same probability parameter. The preferential attachment models, however, are a popular class of evolving networks, and have many variants. Moreover, the stochastic block models [12]–[14], which divide the network into subnetworks, are the typical graph models for representing random networks. Additionally, the dynamic version of the stochastic block models have also been developed to explain the structures of the network

The associate editor coordinating the review of this manuscript and approving it for publication was Chengpeng Hao.

data [15]. In particular, the reference [16] has investigated spectral features of the stochastic block models, which can allow one to obtain insights for analyzing the statistical behaviors of random graph models.

Since the spectral characteristics or graph features can capture the changes of the network data, many detection statistics based on the above information have been proposed for detecting the anomalous behaviors of random networks. Specifically, in [17], a graph Fourier scan statistic has been presented to detect anomalous activity over graphs. Similarly, the spectral methods [18], [19], based on the spectral features of matrix representations of random graphs, is another popular class of anomaly detection methods. Recently, some statistical properties of these algorithms [20] have been revealed. In addition, the optimal statistics, i.e., likelihood ratio statistics [21], [22], and belief propagation [23] have been applied to address anomalous detection relevant to Erdős-Rényi random graph model. Although all the above methods can detect the anomalous patterns of the network data, most of these references only consider the single snapshot or an offline case, and the simply extension of the methods to the sequential observations for change-point detection may not be suitable.

With respect to the streaming data scenario, there has been some investigation of change-point detection in the context of time series of random graphs, including sequential likelihood ratio statistics [24], and locality statistics [25]. Amongst them, based on sequential likelihood ratios, Marangoni-Simonsen *et al.* developed three algorithms for online community detection [24]. These methods have optimal statistical properties, and assume that the model of networks is the Erdős-Rényi random graphs. Focusing on stochastic block models, Wang *et al.* studied two classes of locality statistics, and derived the corresponding asymptotical distribution [25].

In order to detect a possible change-point that minimizes the detection delay under the constraint of ARL, some methods mentioned above employ graph features to construct the corresponding statistics over the whole vertex set of random graphs, which can result in high computational complexity. Especially, with the growth of the size of networks, it becomes evidently difficult to detect the emergence in real time. Moreover, some of these methods, such as [24] and [25], exploit both current and past data to extract the anomalous patterns of the streaming data for improving detection performance.

In this paper, we design a low-complexity algorithm of change-point detection, which also provide detection performance improvement. Here, we restrict our attention to the stochastic block model. In particular, we assume that after the change, the connectivity of the anomalous subnetworks can be sparser or denser than before, while the memberships of nodes also change. In practice, the assumption is more general than that of the typical algorithm, where the subnetworks only become denser after a change-point. In order to overcome the problem of change-point detection, this paper develops a dynamic vertex selection scheme to choose a few vertices for building the detection statistic. Therefore, the presented method has low computational burden. Furthermore, the designed scheme can efficiently capture the statistical features that leads to detect the emergence of an abrupt change with low detection delay.

The main contributions of the paper include:

- We design a dynamic vertex selection scheme for choosing a few vertices associated with change of networks, and then develop a local statistic to address the change-point detection problem. The dynamic vertex selection method can discard most of vertices of random graphs during the construction of the local statistic. Consequently, the proposed algorithm of change-point detection has low complexity of computation. Moreover, the chosen vertices can characterize the change of random graphs that leads to high detection performance. Simulations show that the dynamic vertex selection can capture the anomalous patterns of the network data, and guarantee the high detection performance of the presented algorithm.
- We derive the probability distribution of the proposed statistic under the null scenario that there is no change in the network data, and then can theoretically set the detection threshold associated with the ARL, which is corresponding to false alarm level and a key challenge of change-point detection. More specifically, the theoretic result can avoid the large amount of computation of setting detection threshold involved in the Monte Carlo simulations.
- We analyze the properties of the dynamic vertex selection scheme and derive the related probability bounds with the aid of concentration measure theory [26], [27]. The analytic expressions illustrate that the dynamic vertex selection method can characterize the anomalous patterns of the sequential observations, and ensure the performance of change-point detection algorithm.

The rest of this paper is outlined as follows. In section II, we provide a brief description of stochastic block model and formulate the problem of change-point detection. Section III introduces the dynamic vertex selection scheme and the proposed detection statistic to solve the change-point detection problem in detail. In section IV, the theoretical analysis of the presented algorithms is derived to validate the local statistic. Simulation examples are provided in section V. Finally, the main conclusions of this paper are presented in section VI.

II. BACKGROUND AND PROBLEM STATEMENT

A. NOTATION AND STOCHASTIC BLOCK MODEL

Consider a random network with N nodes, and represent the network by time series of random graphs, with $G_t = (V, E_t)$ denoting the observed graph at each time t, where V and E_t stand for vertex (node) set and edge set of random graphs, respectively. In this paper, we assume that the random graphs are undirected and unweighted graphs without self-loops. Suppose that the corresponding vertex set V is fixed across time. Let A_t denote the adjacency matrix of the observed graph G_t. The adjacency matrix A_t is a $N \times N$ symmetric matrix. The (i, j) *th* element $[A_t]_{ij}$ of A_t is a Bernoulli random variable, which denotes the connectivity between vertices v_i and v_j.

As mentioned above, there are some random graphs to model random networks. Among these random graph models, the stochastic block model is widely used in the random networks [28], [29]. In this paper, we concentrate on the change-point detection of the stochastic block model. For a graph G_t in stochastic block model, the vertices in V are partitioned into M blocks corresponding to the subsets $V_{t,1}, V_{t,2}, \cdots, V_{t,M}$, and the numbers of vertices and the adjacency submatrices of $V_{t,1}, V_{t,2}, \cdots, V_{t,M}$ are corresponding to $N_{t,1}, N_{t,2}, \cdots, N_{t,M}$ and $A_{t,1}, A_{t,2}, \cdots, A_{t,M}$, respectively. Without loss of generality, we have $V_{t,1} = \{v_1, \cdots, v_{N_{t,1}}\}$, $V_{t,2} = \{v_{N_{t,1}+1}, \cdots, v_{N_{t,1}+N_{t,2}}\}$, $\cdots$, $V_{t,M} = \{v_{N-N_{t,M}+1}, \cdots, v_N\}$, and the vertices in each subset share the same membership. Additionally, we refer P_t to the $M \times M$ connectivity probability matrix of G_t. Any one of the diagonal elements of P_t is the connectivity probability between the two verteices which belong to the same block (or membership), while the off-diagonal elements of P_t represent the connectivity probabilities between the two blocks. For example, let $[P_t]_{rs}$ denote the (r, s) *th* element of P_t, assume that the *ith* and *jth* vertices v_i, v_j belong to the subsets $V_{t,r}$ and $V_{t,s}$, respectively, then we have

$$[A_t]_{ij} = \begin{cases} 1 & \textit{with probability } [P_t]_{rs} \\ 0 & \textit{otherwise}. \end{cases} \tag{1}$$

It is worth noting that in general the diagonal elements of P_t are bigger than the off-diagonal elements P_t, which implies that the connectivity of the vertex pairs within each block is denser than that of the vertex pairs between blocks, and displays the community structures of stochastic block model.

B. CHANGE-POINT DETECTION PROBLEM

The observed data is a sequence of random graphs with N vertices drawn from stochastic block model associated with a $M_1 \times M_1$ connectivity probability matrix $P_t = P_0$. Let $[P_0]_{rs}$ represent the (r, s) *th* element of P_0. The corresponding adjacency matrices $A_1, A_2, \cdots$ with $A_t \in R^{N \times N}$ of the observed sequence of random graphs are mutually independent. At an unknown but non-random time point τ, the observed sequence is still independent and follows from stochastic block model but with a new $M_2 \times M_2$ connectivity probability matrix $P_t = P_1$.

Generally, the change-point detection can be formulated as a sequential binary hypothesises. The null hypothesis with $t < \tau$ is that the observed graph G_t related to a network is a realization of a stochastic block model, and we have

$$H_0 : [A_t]_{ij} = \begin{cases} 1 & \textit{with probability } [P_0]_{rs} \\ 0 & \textit{otherwise}, \end{cases} \tag{2}$$

where $[A_t]_{ij}$ is the (i, j) *th* element of the adjacency matrix A_t, which denotes the connectivity between the vertex v_i in block r and the vertex v_j in block s. The alternative hypothesis with $t \geq \tau$ is that the statistical behavior of the observed graph G_t depends on the connectivity probability matrix P_1 with size $M_2 \times M_2$, i.e.,

$$H_1 : [A_t]_{ij} = \begin{cases} 1 & \textit{with probability } [P_1]_{rs} \\ 0 & \textit{otherwise}. \end{cases} \tag{3}$$

Under the alterative hypothesis, the statistical behavior of random graph G_t changes from P_0 to P_1. Generally speaking, because of malicious traffic, surreptitious behavior or suspicious activity [18], [24], [30], a small fraction of the network data exhibits anomalous behavior. More specifically, in the cyber security application [30], since suspicious activity occurs, the rates of communication between a small subset of nodes are notable higher than that of the normal scenario. Additionally, as shown in [24], after a change-point time, nodes inside a subset have much higher frequencies of interaction. In practice, these examples can be reviewed as the scenario that the connectivity of subnetworks becomes denser while the membership of nodes also changes. For instance, we assume that before a change-point τ, the M_1 vertex blocks of G_t are given by $V_{0,1}, V_{0,2}, V_{0,3}, \cdots, V_{0,M_1}$, the vertex numbers of these blocks are corresponding to $N_{0,1}, N_{0,2}, \cdots, N_{0,M_1}$, respectively, and the connectivity probability matrix P_0 is

$$P_0 = \begin{bmatrix} [P_0]_{11} & [P_0]_{12} & [P_0]_{13} & \cdots & [P_0]_{1M_1} \\ [P_0]_{21} & [P_0]_{22} & [P_0]_{23} & \cdots & [P_0]_{2M_1} \\ & \vdots & \vdots & \vdots & \vdots \\ [P_0]_{M_1 1} & [P_0]_{M_1 3} & [P_0]_{M_1 3} & \cdots & [P_0]_{M_1 M_1} \end{bmatrix}. \tag{4}$$

After the change-point time τ, we suppose without loss of generality that only the vertices in blocks $V_{0,1} = \{v_1, \cdots, v_{N_{0,1}}\}$ and $V_{0,2} = \{v_{N_{0,1}+1}, \cdots, v_{N_{0,1}+N_{0,2}}\}$ exhibit abnormal behavior related to memberships and the connectivity probabilities, i.e, $V_{0,1}$ and $V_{0,2}$ are divided into three blocks $V_{1,1}$, $V_{1,2}$, and $V_{1,3}$. Please note that the vertices in the same block have the same membership. The new $(M_1 + 1)$ vertex blocks are $V_{1,1}, V_{1,2}, V_{1,3}, V_{0,3}, \cdots, V_{0,M_1}$ and the vertex set $\{V_{1,1}, V_{1,2}, V_{1,3}\}$ is same as $\{V_{0,1}, V_{0,2}\}$. The connectivity probability matrix P_1 is given by

$$P_1 = \begin{bmatrix} P_a & P_b \\ P_c & P_d \end{bmatrix}, \tag{5}$$

where

$$P_a = \begin{bmatrix} [P_1]_{11} & [P_1]_{12} & [P_1]_{13} \\ [P_1]_{21} & [P_1]_{22} & [P_1]_{23} \\ [P_1]_{31} & [P_1]_{32} & [P_1]_{33} \end{bmatrix}. \tag{6}$$

where P_a is a connectivity probability submatrix corresponding to the vertex pairs in $\{V_{1,1}, V_{1,2}, V_{1,3}\}$ after the change-point τ, and the submatrices P_b, P_c and P_d can be given by the connectivity probability matrix P_0. It means that the connectivity probability between the vertex in $\{V_{1,1}, V_{1,2}, V_{1,3}\} = \{V_{0,1}, V_{0,2}\}$ (or $\{V_{0,3}, \cdots, V_{0,M_1}\}$) and the vertex in $\{V_{0,3}, \cdots, V_{0,M_1}\}$ remains unchanged and is given by the corresponding element of P_0. For example, the parameter of the Bernoulli random variable $[A_t]_{ij}$ associated with the vertices $v_i,\ v_j \in \{V_{1,1}, V_{1,2}, V_{1,3}\}$ (or $v_i,\ v_j \in \{V_{0,1}, V_{0,2}\}$) is given by P_a, while $v_i,\ v_j \in \{V_{0,3}, \cdots, V_{0,M_1}\}$ or $v_i \in \{V_{0,1}, V_{0,2}\},\ v_j \in \{V_{0,3}, \cdots, V_{0,M_1}\}$, the success probability of $[A_t]_{ij}$ remains unchanged, which is given by P_0. Note that G_t is undirected graph, we have $[P_0]_{ij} = [P_0]_{ji}$ ($[P_1]_{ij} = [P_1]_{ji}$). Here, in this example, we assume that the blocks regrading to $V_{0,1}$ and $V_{0,2}$ spilt into three blocks associated with $V_{1,1}, V_{1,2}, V_{1,3}$. It is observed from P_0 and P_1 that the detection problem of significant interest involves in the settings that the dimension N of network is typically larger than $(N_{0,1} + N_{0,2})$ related to abnormal vertices. Moreover, in this paper, the self-connectivity probabilities of the vertex pairs of the anomalous blocks in post-change can be higher or lower than that in pre-change. For example, the diagonal elements of P_a may be bigger or smaller than the diagonal elements $[P_0]_{11}$ (or $[P_0]_{22}$). However, the typical detection algorithms [25] assume that the self-connectivity probability of the vertex pairs with altered statistic behaviour in one block is just higher than before. Therefore, from this point of view, the assumption of this paper is more general than that of the typical algorithms. Additionally, we consider that the memberships $V_{0,1}, V_{0,2}, V_{0,3}, \cdots, V_{0,M_1}$ and the connectivity probability matrix P_0 are known, which can be obtained from historic data [24]. We also assume that the change parameters, i.e., P_1, τ, and the memberships related to P_1, are unknown.

The goal of change-point detection problem is to design a detection statistic to decide a stopping time T for declaring the emergence of the change. In order to quantify the performance of the algorithm, we define average run length (ARL) as the expectation $\mathbb{E}(T)$ of the stopping time T under the null hypothesis, which is another form of the probability of false alarm in the off-line detection problem. The performance metric of detection is given by the expected detection delay $\mathbb{E}(T - \tau | T > \tau)$, which represents the expected delay to fire an alarm after a change-point occurs. Therefore, the change-point detection aims to minimize the expected detection delay under the constraint of a given ARL. In the following section, based on the sequential observations, we will present a stopping rule to detect the occurrence of a change-point.

III. PROPOSED ALGORITHM

The previous works consider the fusion of the multiple features of the observed graphs, or exploit both the current data and the past data for change-point detection. The detection statistics are built on the whole vertex set, which may result in huge computational burden, especially in the large scale networks. In addition, the typical algorithms always assume that the connectivity probabilities of the vertex pairs corresponding to the anomalous subnetworks (anomalous blocks) just become higher, and the memberships are fixed across time. However, these algorithms may not be suitable for the case that after a change, the connectivity of anomalous subnetworks can be sparser or denser than before, while the memberships of vertices also change. In practice, the latter assumption is more general than the former.

Aiming to solve the problem, we propose a local detection statistic based on a dynamic vertex selection strategy. The objective of the developed vertex selection scheme is to reduce the computational complexity of the change-point detection algorithm and improve its detection performance. First of all, the low computational complexity of the algorithm is especially in favor of real-time requirement of change-point detection on the sequential observations. In order to reduce the computational complexity, the dynamic vertex selection scheme just chooses a few vertices to construct the detection statistic. On the other hand, to guarantee the detection performance of the proposed algorithm, the graph features related to the selected vertex pairs shall efficiently capture the anomalous behavior. The dynamic vertex selection scheme will be discussed in detail below.

A. DYNAMIC VERTEX SELECTION

In order to capture the complex structure of the random graphs, we first choose one referenced vertex in each block for capturing the statistical pattern of subnetwork. We denote by v_m^R the referenced vertex in mth block with $m \in \{1, 2, \cdots, M_1\}$, and refer to $A_{t,m}$ as the $N_{0,m} \times N_{0,m}$ submatrix of the adjacency matrix A_t at time t, which corresponds to the mth block. Note that since the subgraph corresponding to the mth block is an Erdős-Rényi random graph before a change-point, which has homogeneous connectivity within the block, we can choose any one of vertex from the vertex set $V_{0,m}$ as the referenced vertex v_m^R associated with the mth block. For convenience, we denote the index of v_m^R in $V_{0,m}$ by 1. Then, we define the degree of the vertex v_m^R in $V_{0,m}$ as

$$F_{0,m} = \deg_{V_{0,m}}\left(v_m^R\right) = \sum_{i=1}^{N_{0,m}} \left[A_{t,m}\right]_{1i}, \tag{7}$$

where $\deg_{V_{0,m}}\left(v_m^R\right)$ denotes the vertex degree of the referenced vertex v_m^R in the mth block with the vertex set $V_{0,m}$, and can be used for expressing the graph feature of the mth block. The value of $\deg_{V_{0,M}}\left(v_m^R\right)$ is an evidence of connectivity characteristic within the block. We further define the following localized graph feature

$$F_{1,m} = N_{0,m} - 2 - \deg_{V_{0,m}}\left(v_m^R\right). \tag{8}$$

Recall that the connectivity probability within a block is evidently bigger than that between blocks. Consider the scenario that after a change-point, the mth block may be partitioned into multiple blocks. Therefore, the change that the localized feature $F_{1,m}$ suddenly increase, is a clear evidence of anomalous pattern of random graphs and implies that a change-point appears. However, besides the change related to the memberships of some of vertices in V, we also consider a more general assumption that the connectivity probability associated with the anomalous subgraph can be sparser or denser, or even remain unchange. Namely, the growth of $F_{1,m}$ caused by the membership change may be taken out by the change that the block contained the vertex v_m^R become excessive interconnectivity. Consequently, the graph feature $F_{1,m}$ that just relies on the vertex v_m^R can not robustly indicate the presence of a change-point under the general assumption.

Without loss of generality, assume that the mth block is partitioned into two blocks after a change-point, and let the vertices of $V_{0,m}$ be divided into two vertex set $V_{1,m}^*$ and $V_{2,m}^*$, which correspond to the two new blocks. Let the vertex v_m^R belong to $V_{1,m}^*$. Naturally, the graph features corresponding to the vertices in $V_{1,m}^*$ have similar properties because the connectivity characteristic within the subgraph related to $V_{1,m}^*$ is homogeneous. In other words, the graph pattern related to the vertices in $V_{2,m}^*$ is different from that of the referenced vertex v_m^R since the related vertices belong to two different blocks. Consequently, except the referenced vertex v_m^R, we also need to choose one vertex in $V_{2,m}^*$ for adequately capturing the anomalous pattern of the mth block. However, both the change-point and the change of memberships are unknown for detector. In this paper, we develop a dynamic vertex selection scheme to decide one vertex v_m^*, which belongs to $V_{2,m}^*$ with high probability. Therefore, we can combine the graph features related to both vertices v_m^R and v_m^* for uncovering the hidden structure of the mth block. It is worth emphasizing that although the discussion above supposes that the mth block is only partitioned into two blocks, the graph features related to both vertices v_m^R and v_m^* are still applicable for the more complex change of membership when the selected vertices (v_m^R and v_m^*) belong to the different memberships.

Let $V_{0,m}^n$ denote the neighborhood vertex set in $V_{0,m}$ regarding to the vertex v_m^R. Consider a pair of vertices (v_1, v_2) with $\{v_1, v_2\} \subseteq V_{0,m}$, and let $d_m(v_1, v_2)$ represent the shortest path distance between v_1 and v_2 in the mth block. $path_{l,m}^k(v_1, v_2)$ stands for one path of the vertex pair (v_1, v_2) with $d_m(v_1, v_2) = k$ and index l in the mth block. By counting the number of paths with path distance $d_m\left(v_m^R, v_i\right) = 2$ between v_m^R and $v_i \in V_{0,m} \backslash V_{0,m}^n$, we can obtain v_m^* through the following formula

$$v_m^* = \mathop{arg\,\min}_{v_i \in V_{0,m} \backslash V_{0,m}^n} \sum_l path_l^2\left(v_i, v_m^R\right). \tag{9}$$

We further write (9) into the form of the adjacency matrix $A_{t,m}$, given as

$$v_m^* = \mathop{arg\,\min}_{v_i \in \left\{V_{0,m} \backslash V_{0,m}^n \cup v_m^R\right\}, v_j \in V_{0,m}^n} \sum_j \left[A_{t,m}\right]_{ij}, \tag{10}$$

where the notation $\cup$ denotes the union operator. The above formula depends on the fact that if the block associated with $V_{0,m}$ disjoins into several blocks, the shortest path distance between $v_m^R \in V_{1,m}^*$ and $v_i \in V_{1,m}^*$ is generally smaller than that between $v_m^R \in V_{1,m}^*$ and $v_i \in V_{2,m}^*$. The expression implies that when a sudden change related to memberships occurs, the referenced vertex v_m^R and the vertex v_m^* selected from (9) or (10) belong to two different blocks with high probability. Similar to (8), we define the graph feature related to v_m^* as

$$F_{2,m} = \deg_{V_{0,m}}\left(v_m^*\right), \tag{11}$$

and

$$F_{3,m} = N_{0,m} - 2 - \deg_{V_{0,m}}\left(v_m^*\right). \tag{12}$$

Obviously, when the membership of the mth block changes, the graph feature $F_{1,m}$ or $F_{3,m}$ just can capture one part of structure of the mth block. Therefore, we combine $F_{1,m}$ and $F_{3,m}$ into the following form

$$\begin{aligned} F_{4,m} &= \max\left(F_{1,m}, F_{3,m}\right) \\ &= N_{0,m} - 2 - \min\left(\deg_{V_{0,m}}\left(v_m^R\right), \deg_{V_{0,m}}\left(v_m^*\right)\right). \end{aligned} \tag{13}$$

This means that the graph feature $F_{4,m}$ related to the vertex pair $\left(v_m^R, v_m^*\right)$ can adequately express the statistical behavior of the block, and also give an evidence of an abrupt change related to membership and connectivity probability. Note that since the elements of $A_{t,m}$ in (10) is Bernoulli random variables, the index of v_m^* also changes over time. Specifically, before a change-point, the above procedure can select any one of vertices in $V_{0,m}^n \backslash v_m^R$ as v_m^* with the same probability because of homogeneous connectivity within the block. However, when the change about membership emerges in the mth block, one vertex in $V_{2,m}^*$ can be randomly chosen as v_m^* with high probability. Because of the dynamic property about the index of v_m^*, we name the procedure of selecting the vertex pair $\left(v_m^R, v_m^*\right)$ as dynamic vertex selection scheme in this paper. Additionally, with respect to the statistical behavior of the selection procedure, we will present theoretical analysis in the next section.

B. LOCAL STATISTIC FOR CHANGE-POINT DETECTION

Consider the alternative hypothesis that at a unknown time τ, a small subset of V changes the corresponding probabilistic behavior, while the majority of vertices in V still remain their normal pattern. Meanwhile, the subset associated with anomalous vertices is unknown to detector. Moreover, the vertex pair $\left(v_m^R, v_m^*\right)$ can just be utilized to capture the local behavior within the mth block. Therefore, we search M_1 pairs of vertices $\left(v_m^R, v_m^*\right)$, $m \in \{1, 2, \cdots, M_1\}$ with respect to M_1 blocks for expressing the graph features via recursive implementation of the dynamic vertex selection scheme mentioned above, so that the hidden change can be detected efficiently. Let the vertex set V^* with dimension $2M_1$ denote the M_1 vertex pairs, that is

$$V^* = \left\{v_1^R, v_1^*, v_2^R, v_2^*, \cdots, v_{M_1}^R, v_{M_1}^*\right\}. \tag{14}$$

It is observed from (13) that the graph features $F_{4,m}$, $m \in \{1, 2, \cdots, M_1\}$ can respectively characterize the probabilistic behavior within the corresponding blocks. Nonetheless, the characteristic associated with $F_{4,m}$ does not contain the connectivity feature between-blocks. In practice, the connectivity between-blocks may also include the anomalous pattern generated by the change, and it is necessary to construct a graph feature that can demonstrate the probabilistic behaviors of both within-block and between-blocks. Here, we denote the indices of the vertices in $V \backslash \left\{V_{0,m} \cup V^*\right\}$ by $\left\{1, 2, \cdots, N-2\left(M_1-1\right)-N_{0,m}\right\}$. Then, let l represent the index of v_m^*, define the degree of v_m^* related to the outside of the mth block, given by

$$F_{5,m} = \deg_{V \backslash \{V_{0,m} \cup V^*\}}\left(v_m^*\right) = \sum_{i=1}^{N-2(M_1-1)-N_{0,m}} [A_t]_{li}. \tag{15}$$

where $\deg_{V \backslash \{V_{0,m} \cup V^*\}}\left(v_m^*\right)$ denotes the vertex degree of the vertex v_m^* regarding to the vertex set $V \backslash \left\{V_{0,m} \cup V^*\right\}$. Consequently, a graph feature $F_{6,m}$ associated with the vertex pair $\left(v_m^R, v_m^*\right)$ is defined as

$$\begin{aligned} F_{6,m} &= \max\left(F_{1,m}, F_{3,m}\right) + \deg_{V \backslash \{V_{0,m} \cup V^*\}}\left(v_m^*\right) \\ &= N_{0,m} - 2 - \min\left(\deg_{V_{0,m}}\left(v_m^R\right), \deg_{V_{0,m}}\left(v_m^*\right)\right) \\ &\quad + \sum_{i=1}^{N-2(M_1-1)-N_{0,m}} [A_t]_{li}. \end{aligned} \tag{16}$$

Obviously, the graph feature $F_{6,m}$ can measure the probability behavior of local part of the observed graph G_t, which corresponds to the connectivity within the mth block and between-blocks specified by v_m^* and $V \backslash \left\{V_{0,m} \cup V^*\right\}$. More specifically, $F_{6,m}$ can be used to assess the probabilistic behavior change of the corresponding subgraph of G_t. It is worth pointing out that the stochastic block graphs have inhomogeneous property, in other words, the probability parameters of within blocks and between blocks, including the connectivity probabilities and block dimensions, can be different. Therefore, one or several graph features in $F_{6,m}$, $m \in \{1, 2, \cdots, M_1\}$, can dominate these features under the null hypothesis, which is disadvantageous for detecting the change. In order to prevent this, we must next perform vertex pair normalization for standardizing the scales of the graph feature $F_{6,m}$, $m \in \{1, 2, \cdots, M_1\}$, namely,

$$\bar{F}_{6,m} = \frac{F_{6,m} - \mu_{F_{6,m}}}{\sigma_{F_{6,m}}}. \tag{17}$$

where $\mu_{F_{6,m}}$ and $\sigma_{F_{6,m}}$ represent the mean value and the standard deviation of the graph feature $F_{6,m}$, respectively. For overcoming the inhomogeneous property, some typical algorithms, such as [25], also define the similar normalized operators. However, the corresponding means and standard deviations of the statistics are estimated using the recent past observations, which may be improper for the change-point detection on the streaming data scenario because of the memory and real-time restrictions. In order to provide a simpler way for normalizing graph features, we derive the probability distribution of $F_{6,m}$ in this paper, and then $\mu_{F_{6,m}}$ and $\sigma_{F_{6,m}}$ can be pre-computed. The analytical expressions for $\mu_{F_{6,m}}$ and $\sigma_{F_{6,m}}$ will be presented in the next section.

Based on the standardized graph features, we can construct the detection statistic F_t at each time t for change-point detection, given by

$$F_t = \max_{m \in \{1,2,\cdots,M_1\}} \bar{F}_{6,m}. \tag{18}$$

where the maximum is over all graph features related to the vertex pairs $\left(v_m^R, v_m^*\right)$, $m \in \{1, 2, \cdots, M_1\}$. In this paper, we call the constructed statistic F_t as local statistic because F_t is one kind of local graph features, which can be used to capture the probabilistic change arising from a small amount of vertex set of V. Note that the observation G_t arrives in a streaming style, we shall use the local statistic constantly test whether G_t belongs to the null hypothesis H_0 or the alternative hypothesis H_1, and determine the stopping time T, which declares an alarm that a change-point comes. Namely, given a pre-specified threshold λ, if

$$F_t \geq \lambda, \tag{19}$$

we will have a stopping time

$$T = t, \tag{20}$$

and fire an alarm, where the parameter λ is specified to meet the ARL requirement, which is the expected value of T under the null hypothesis. Otherwise, we need to continually test the coming observations until find a change-point. Based on the analysis mentioned above, the procedure of change-point detection is summarized in Algorithm 1.

Algorithm 1 Change-Point Detection

1: Input: Observed graph G_t on a streaming fashion.
2: Output: Stopping time (change-point) T.
3: Initialize: Detection threshold λ with a given ARL; $\mu_{F_{6,m}}$ and $\sigma_{F_{6,m}}$ for $m = 1, \cdots, M_1$.
4: **for** each observed graph G_t, $t = 1, \cdots,$ **do**
5: **for** $m = 1, \cdots, M_1$, **do**
6: Compute local feature $F_{1,m}$ by using (8);
7: Compute local feature $F_{3,m}$ by using (12);
8: Compute local feature $F_{4,m}$ within the mth block by using (13);
9: Compute graph feature $F_{6,m}$ related to the vertex pair $\left(v_m^R, v_m^*\right)$ by using (16);
10: Normalize the scale of $F_{6,m}$ by using (17) and obtain the standardized graph feature $\bar{F}_{6,m}$;
11: **end for**
12: Construct local statistic F_t with (18);
13: **if** $F_t \geq \lambda$, then
14: Load stopping time T with time t;
15: Break;
16: **end if**
17: **end for**

Remark 1: Similar to some typical algorithms of change-point detection, the proposed algorithm also utilize the corresponding graph features to construct detection statistic. In order to reduce the computational complexity and improve the detection performance, the considerations of developing the presented local statistic, which are the main difference of the traditional methods from our proposed detection statistic, are multifold: first, through the implementation of dynamic vertex selection scheme, the proposed method just chooses M_1 pairs of vertices, where $2M_1 \ll N$, to develop the detection statistic. Hence, compared with the classic work, the presented method has lower computational burden, which is especially in favor of real-time requirement of change-point detection on the sequential observations. Furthermore, each pair of vertices chosen by the dynamic vertex selection scheme is used to build the corresponding graph feature, which can capture the local behavior of the observed data efficiently. Additionally, because both the subset of vertices related to the anomalous pattern and the change time are unknown, we takes the maximum of these graph features to form the detection statistic for characterizing the time series of random graphs adequately. Therefore, the proposed method can minimize the detection delay with a given ARL.

Remark 2: In order to control false alarm rate, we need to quantify the ARL and set the threshold λ, which is a challenge of change-point detection. Although the direct numerical simulations can be utilized to obtain the detection threshold related to the corresponding ARL, it requires the huge computational cost, especially in the large networks. Consequently, an analytical expression of ARL is quite useful to change-point detection. In the next section, we will present the probabilistic distribution of the local statistic so that we can compute the theoretical threshold without resorting to the Monte Carlo method. In addition, the parameters $\mu_{F_{6,m}}$ and $\sigma_{F_{6,m}}$ for standardizing the graph features $F_{6,m}, \quad m \in \{1, 2, \cdots, M_1\}$, can also be obtained by the theoretical expression.

C. COMPUTATIONAL COMPLEXITY

The computational complexity of the presented algorithm mainly depends on computation associated with the dynamic vertex selection scheme for selecting M_1 pairs of vertices, and computation of local graph features related to the M_1 vertex pairs. Specifically, Consider a time t before the change-point, the computational cost related to the vertex selection of v_m^* is $O\left(\left(N_{0,m}-1\right)^2 (1-[P_0]_{mm})[P_0]_{mm}\right)$. In terms of the computation of the graph feature $F_{1,m}$, the computational complexity is $O\left(\left(N_{0,m}-1\right)[P_0]_{mm}\right)$. Similarly, the computational burden regarding to $F_{3,m}$ is also $O\left(\left(N_{0,m}-1\right)[P_0]_{mm}\right)$. Additionally, the graph feature which characterizes the local pattern outside of the mth block related to the vertex pair $\left(v_m^R, v_m^*\right)$ costs $O\left(\sum_{i=1,i\neq m}^{M_1}[P_0]_{mi}\left(N_{0,m}-2\right)\right)$. As a result, the total complexity of the proposed algorithm is $O\left(M_1\left(2\left(N_{0,m}-1\right)[P_0]_{mm} + \left(N_{0,m}(1-[P_0]_{mm})-1\right)N_{0,m}[P_0]_{mm} + \sum_{i=1,i\neq m}^{M_1}[P_0]_{mi}\left(N_{0,m}-2\right)\right)\right)$. For simplification, we assume that each block has the same dimension $N_{0,1} = \cdots = N_{0,M_1} = N_0$, and the same connectivity probabilities $[P_0]_{mm} = p_0, \quad m \in \{1, 2, \cdots, M_1\}$, $[P_0]_{mi} = p_1, \quad m \neq i, \quad m, i \in \{1, 2, \cdots, M_1\}$. So, the total complexity at time t can be rewritten as $O\left(M_1^2\left(N_0^2+1\right)\left(p_0-p_0^2\right)+2M_1N_0p_0^2+M_1(M_1-1)(N_0-2)p_1\right)$. It is worth noting that no matter that the time t belongs to pre-change or post-change, the computational cost of the proposed algorithm almost remains unchanged because only a small fraction of vertices changes their probability behavior.

IV. THEORETICAL ANALYSIS

Through the dynamic vertex scheme, the proposed algorithm first obtains M_1 pairs of vertices $\left(v_m^R, v_m^*\right), \quad m \in \{1, 2, \cdots, M_1\}$, and then compute the corresponding local features $F_{6,m}, \quad m \in \{1, 2, \cdots, M_1\}$. Finally, the detection statistic can be achieved by taking the maximum of the local features. Note that at the null hypothesis, the local behavior associated with each vertex pair, including block size, the connectivity probabilities within the block and between-blocks, may be different. Naturally, the local graph features $F_{6,m}, \ m \in \{1, 2, \cdots, M_1\}$, are always unequal, and we need to standardize these local features. Otherwise, at the alternative hypothesis, the maximum of $F_{6,m}, \ m \in \{1, 2, \cdots, M_1\}$, may not uncover the anomalous pattern caused by the probabilistic behavior change of a small subset of vertices because one certain $F_{6,m}$ no involved in anomalous behavior probably dominates all graph features.

In this section, we derive the probability distributions of the developed graph features $F_{4,m}$ and $F_{6,m}$ so that we can compute the means and standard deviations efficiently. Furthermore, in order to decide the detection threshold λ, we also achieve the ARL expression with respect to the corresponding λ, which is also a key challenge in change-point detection problem. Therefore, we can set the detection threshold theoretically rather than through the simulation fashion, which requires expensive computation. In addition, we investigate the probabilistic behavior of the dynamic vertex scheme, and then demonstrate the detection performance of the proposed local statistic.

In order to obtain $\mu_{F_{6,m}}$ and $\sigma_{F_{6,m}}$ for standardizing the graph feature $F_{6,m}$, we first achieve the the cumulative distribution function (cdf) $G_{4,m}(K)$ of the graph feature $F_{4,m}$ demonstrated in the following lemma, which is the basis of the mean and variance expressions of $F_{6,m}$.

Lemma 1: Let $G_t = (V, E_t)$ denote the random graph drawn from stochastic block model concerned with a $M_1 \times M_1$ connectivity probability matrix P_0 and N vertices. Assume that the mth block has $N_{0,m}$ vertices, and the connectivity probability in the mth block is $p = [P_0]_{mm}$, then the cdf $G_{4,m}(K)$ of graph feature $F_{4,m}$ can be given by

$$\begin{aligned} G_{4,m}(K) &= pr\left(F_{4,m} < K\right) \\ &= 1 - \bar{G}_{4,m}\left(N_{0,m} - 1 - K\right), \end{aligned} \tag{21}$$

where

$$\bar{G}_{4,m}(K) = \sum_{n=0}^{N_{0,m}-1} C_{N_{0,m}-1}^{n} p^n (1-p)^L \bar{G}_{4,m,n}(K), \tag{22}$$

$$L = N_{0,m} - 1 - n, \tag{23}$$

$$\bar{G}_{4,m,n}(K) = \begin{cases} G_{2,m,n}(K), & K \le n \\ 1\,, & K > n, \end{cases} \tag{24}$$

$$G_{2,m,n}(K) = \begin{cases} 1 - \left(1 - \bar{G}_{2,m,n}(K)\right)^L, & 1 \le n \le N_{0,m}-2 \\ 1\,, \quad n = 0 \text{ or } n = N_{0,m} - 1\,, \end{cases} \tag{25}$$

$$\bar{G}_{2,m,n}(K) = \sum_{k=0}^{K} C_n^k p^k (1-p)^{n-k}, \tag{26}$$

and $C_{N_{0,m}-1}^{n}$ and C_n^k denote combination operator.

Proof: See Appendix A. □

In order to normalize the graph feature $F_{6,m}$ related to the vertex pair $\left(v_m^R, v_m^*\right)$, we shall further obtain the cdf $G_{5,m}(K)$ of the graph feature $F_{5,m}$ characterizing the connectivity feature between blocks. Note that $F_{5,m}$ is the sum of the independent Bernoulli random variables, which is a Poisson Binomial distribution, and its cdf $G_{5,m}(K)$ can be given by [31]

$$G_{5,m}(K) = pr\left(F_{5,m} < K\right) = \sum_{k=0}^{K-1} \left\{ \sum_{A \in \Gamma_k} \prod_{i \in A} p_i \prod_{i \in A^c} (1-p_i) \right\}, \tag{27}$$

where $K \in \left\{0, 1, \cdots, \bar{N}_m\right\}$, $\bar{N}_m = N - 2(M_1 - 1) - N_{0,m}$, and $p_i \in \left\{[P_0]_{m1}, [P_0]_{m2}, \cdots, [P_0]_{mM_1}\right\}$ represents the connectivity probability between the mth block and the other block. Γ_k denotes the set of all subsets of k integers chosen from $\left\{0, 1, \cdots, \bar{N}_m\right\}$, and A^c represents the complementary set of A. $pr\,(.)$ represents the probability notation. It is worth pointing out that the formula (27) is just a cdf expression of Poisson binomial distribution, and a more efficiently computational method can be found in [31].

Based on the Lemma 1 and the cdf $G_{5,m}(K)$, we obtain the analytical expressions $\mu_{F_{6,m}}$ and $\sigma^2_{F_{6,m}}$, shown in Theorem 1.

Theorem 1: Consider a stochastic block random graph $G_t = (V, E_t)$ with N vertices and a $M_1 \times M_1$ connectivity probability matrix P_0. Let $\tilde{N}_m = N_{0,m} - 2$, and $N_{0,m}$ denotes the vertex number of the mth block, then the mean $\mu_{F_{6,m}}$ and the variance $\sigma_{F_{6,m}}$ of the graph feature $F_{6,m}$ are given by

$$\mu_{F_{6,m}} = \sum_{K=0}^{\bar{N}_m+\tilde{N}_m} K pr\left(F_{6,m} = K\right), \tag{28}$$

$$\sigma^2_{F_{6,m}} = \sum_{K=0}^{\bar{N}_m+\tilde{N}_m} \left(K - \mu_{F_{6,m}}\right)^2 pr\left(F_{6,m} = K\right), \tag{29}$$

where

$$pr\left(F_{6,m} = K\right) = \sum_{1 \le i \le \tilde{N}_m, 1 \le j \le \bar{N}_m, i+j = K+1} \bar{p}_i \tilde{p}_j, \tag{30}$$

$$\bar{p}_i = G_{4,m}(i+1) - G_{4,m}(i), \tag{31}$$

$$\tilde{p}_j = G_{5,m}(j+1) - G_{5,m}(j), \tag{32}$$

and $K \in \left\{0, 1, \cdots, \bar{N}_m + \tilde{N}_m\right\}$, $i \in \left\{0, 1, \cdots, \tilde{N}_m\right\}$, $j \in \left\{0, 1, \cdots, \bar{N}_m\right\}$.

Proof: See Appendix B. □

As mentioned in section IV, $F_{6,m}$ must be normalized as $\bar{F}_{6,m} = \frac{F_{6,m} - \mu_{F_{6,m}}}{\sigma_{F_{6,m}}}$ to avoid that a ceratin graph feature in $F_{6,m}$, $m \in \{0, 1, \cdots, M_1\}$ dominates the behavior of the random network. The probability mass function (pmf) of $\bar{F}_{6,m}$ is given by

$$pr\left(\bar{F}_{6,m} = \frac{K - \mu_{F_{6,m}}}{\sigma_{F_{6,m}}}\right) = pr\left(\bar{F}_{6,m} = x\right) = pr\left(F_{6,m} = K\right), \tag{33}$$

where

$$x \in \left\{ \frac{0 - \mu_{F_{6,m}}}{\sigma_{F_{6,m}}}, \frac{1 - \mu_{F_{6,m}}}{\sigma_{F_{6,m}}}, \cdots, \frac{\bar{N}_m + \tilde{N}_m - \mu_{F_{6,m}}}{\sigma_{F_{6,m}}} \right\}. \tag{34}$$

Therefore, the cdf $\bar{G}_{6,m}(x)$ of $\bar{F}_{6,m}$ can be obtained, given by

$$\begin{aligned} \bar{G}_{6,m}(x) &= pr\left(\bar{F}_{6,m} < x\right) \\ &= pr\left(F_{6,m} < x\sigma_{F_{6,m}} + \mu_{F_{6,m}}\right) \\ &= pr\left(F_{6,m} < K\right) \\ &= \sum_{k=0}^{K-1} pr\left(F_{6,m} = K\right), \end{aligned} \tag{35}$$

where $K = x\sigma_{F_{6,m}} + \mu_{F_{6,m}}$.

To derive detection threshold of local statistic F_t under the constraint of ARL, we need to obtain the cdf of F_t. Based on the definition of F_t, which is the biggest value of $\bar{F}_{6,m}$, $m \in \{0, 1, \cdots, M_1\}$, the cdf $G_t(x)$ of F_t is given by

$$G_t(x) = pr\left(F_t < x\right) = \prod_{m=1}^{M_1} \bar{G}_{6,m}(x), \tag{36}$$

where x is discrete and takes value from $x = \frac{K - \mu_{F_{6,m}}}{\sigma_{F_{6,m}}}$, $K \in \left\{0, 1, \cdots, \bar{N}_m + \tilde{N}_m\right\}$, $m \in \{0, 1, \cdots, M_1\}$, and the second equality follows from the fact that the random variables $\bar{F}_{6,m}$, $m \in \{0, 1, \cdots, M_1\}$ are mutual independent, which can be inferred from the formulas (16) and (17).

Based on the description of Algorithm 1, the detection statistic F_t of change-point detection declares that a change point of time series of random graphs occurs when $F_t \ge \lambda$, and the stopping time T is t. Note that judging $F_t \ge \lambda$ or $F_t < \lambda$ at time t is a binary detection problem, and can be viewed as a Bernoulli random variable. Thus, the success probability p_t, which announces $F_t \ge \lambda$, can be given by

$$p_t = 1 - G_t(x) = 1 - \prod_{m=1}^{M_1} \bar{G}_{6,m}(x). \tag{37}$$

The probability distribution of the stopping time T belongs to geometric distribution with parameter p_t. Consequently, the ARL of the proposed algorithm related to the detection threshold λ can be given by

$$ARL = \mathbb{E}(T) = \frac{1}{p_F}, \tag{38}$$

where $p_F = p_t$.

Remark 3: The analytical results presented in Theorem 1 and the related analysis mentioned above can provide several merits for change-point detection. On the one hand, the analytical expressions in (28) and (29) can be used to standardize the graph features $F_{6,m}$, $m \in \{0, 1, \cdots, M_1\}$, without resorting to obtaining sample estimation of $\mu_{F_{6,m}}$ and $\sigma^2_{F_{6,m}}$. As a result, it is helpful to save storing space and operating time, compared with the traditional way requiring sample estimation of mean and variance of graph features. On the other hand, given ARL, we have to set the corresponding detection threshold to decide whether anomalous behaviors in random networks happen or not. In this paper, we obtain the ARL formula (38) and the cdf $G_t(x)$ (36) of detection statistic so that the detection threshold of the proposed algorithm can be given in a theoretical fashion rather than Monte Carlo simulations, which suffer from costly computational complexity.

As described in section III.A, after the membership of the mth block changes, the vertices $\{v_m^R, v_m^*\}$ selected by the proposed scheme can belong to two different memberships with high probability so that the local statistic can detect change-point efficiently. In this section, we derive the theoretical results to discuss the dynamic vertex selection scheme. We have the following theorem.

Theorem 2: Assume that the 1st block with $V_{0,1}$ and $N_{0,1}$ of G_t breaks into two subblocks with $V_{1,1}$, $V_{1,2}$, which are corresponding to the vertex numbers $N_{1,1}$. $N_{1,2}$, respectively. Let X_{ij} represent $[A_t]_{ij}$. The probability parameters p_1, p_2, p_3 of X_{ij} are corresponding to $v_i, v_j \in V_{1,1}$, $v_i, v_j \in V_{1,2}$, $v_i \in V_{1,1}$, $v_j \in V_{1,2}$ (or $v_i \in V_{1,2}$, $v_j \in V_{1,1}$), respectively. Let $v_i \in V_{0,1}$. When $v_k \in V_{1,1}$, define $Z_{ik} = X_{1i}X_{ki}$ and $Z_k = \sum_{i=2}^{N_{0,1}} Z_{ik}$. Then, for $\varepsilon > 0$, we have the probability tail bound of Z_k, given by

$$\begin{aligned} &pr\left(|Z_k - \mathbb{E}(Z_k)| \geq \varepsilon, X_{1k} = 0\right) \\ &\leq 2(1-p_1)\exp\left(-\min\left\{\frac{5\varepsilon^2}{12\sigma^2_{Z_k}}, 3\varepsilon - \frac{9}{5}\sigma^2_{Z_k}\right\}\right), \end{aligned} \tag{39}$$

where $\sigma^2_{Z_k} = \sum_{i=2}^{N_{0,1}} \sigma^2_{Z_{ik}}$, $\sigma^2_{Z_{ik}} = p_1^2\left(1-p_1^2\right)$ when $v_i \in V_{1,1}$, or $\sigma^2_{Z_{ik}} = p_3^2\left(1-p_3^2\right)$ when $v_i \in V_{1,2}$, and $\mathbb{E}(Z_k) = \left(N_{1,1}-2\right)p_1^2 + N_{1,2}p_3^2$.

Proof: See Appendix C. □

$pr\left(|Z_k - \mathbb{E}(Z_k)| \geq \varepsilon, X_{1k} = 0\right)$ denotes the probability behavior of the shortest path number between v_1 and v_k ($v_1, v_k \in V_{1,1}$) with $d(v_1, v_k) = 2$. Namely, the bound gives the probability that the random variables Z_k under $X_{1k} = 0$ and $v_k \in V_{1,1}$ are far from their mean values $\mathbb{E}(Z_k)$. Similarly, let $v_k \in V_{1,2}$, $\bar{Z}_{ik} = X_{1i}X_{ki}$ and $\bar{Z}_k = \sum_{i=2}^{N_{0,1}} \bar{Z}_{ik}$. Then, for $\varepsilon > 0$, we also have

$$\begin{aligned} &pr\left(\left|\bar{Z}_k - \mathbb{E}\left(\bar{Z}_k\right)\right| \geq \varepsilon, X_{1k} = 0\right) \\ &\leq 2(1-p_3)\exp\left(-\min\left\{\frac{5\varepsilon^2}{12\sigma^2_{\bar{Z}_k}}, 3\varepsilon - \frac{9}{5}\sigma^2_{\bar{Z}_k}\right\}\right), \end{aligned} \tag{40}$$

where $\sigma^2_{\bar{Z}_k} = \sum_{i=2}^{N_{0,1}} \sigma^2_{\bar{Z}_{ik}}$, $\sigma^2_{\bar{Z}_{ik}} = p_2p_3(1-p_2p_3)$ when $v_i \in V_{1,2}$, $\sigma^2_{\bar{Z}_{ik}} = p_1p_3(1-p_1p_3)$ when $v_i \in V_{1,1}$, and $\mathbb{E}\left(\bar{Z}_k\right) = \left(N_{1,1}-1\right)p_1p_3 + \left(N_{1,2}-1\right)p_2p_3$.

Remark 4: Note that $pr\left(|Z_k - \mathbb{E}(Z_k)| \geq \varepsilon, X_{1k} = 0\right)$ represents the probability behavior of the shortest path number between v_1 and v_k with $d(v_1, v_k) = 2$, where v_1 and v_k belong to the same membership. Additionally, $pr\left(\left|\bar{Z}_k - \mathbb{E}\left(\bar{Z}_k\right)\right| \geq \varepsilon, X_{1k} = 0\right)$ is corresponding to the scenario that v_1 and v_k belong to the two different subblocks. The probability bounds (39) and (40) show that Z_k and $\bar{Z}_k$ concentrate theirs mean values because of exponential decay of tail bounds about ε. Moreover, when $N_{1,1}$ and $N_{1,2}$ are similar or $N_{1,1} > N_{1,2}$, $\mathbb{E}(Z_k)$ is distinctly larger than $\mathbb{E}\left(\bar{Z}_k\right)$ since the connectivity probabilities p_1, p_2 within-blocks are larger than the probability p_3 between blocks. Namely, based on the probability bounds, (9) and (10), it can be seen that the vertex pair $\left(v_m^R, v_m^*\right)$ selected by the dynamic vertex selection scheme can belong to two different memberships with high probability so that the detection statistic can capture the anomalous features efficiently. Consequently, the graph feature $F_{6,m}$ based on the selected vertex pair can capture graph behaviors of both memberships and connectivity probability, which guarantees that F_t can detect change-point timely.

V. SIMULATION RESULTS

In this section, we provide several simulation examples to illustrate the performance of the proposed algorithm, compared with the typical change-point detection methods in [25]. For convenience, the two detection statistics$S_{\tau,l,k}(t;\Phi)$ and $S_{\tau,l,k}(t;\Psi)$ expressed in [25] are named as compared method 1 and compared method 2, respectively. The time series of random graph are generated by stochastic block model. With respect to anomalous pattern, the anomalous subnetworks have sparser or denser connectivity than before, while the memberships also change.

A. SIMULATION EXAMPLE 1 FOR LOCAL STATISTIC UNDER H_0

In order to verify the theoretic ARL expression, the comparisons between simulations and the derived expression in (38) are presented. We consider both the stochastic block mode with same block size (case 1) and the stochastic model with different block dimension (case 2). Specifically, in the first case, the total nodes of random networks with $M_1 = 6$ is $N = 240$, the vertex number of the M_1 block are given by $N_{0,1} = N_{0,2} = \cdots = N_{0,6} = 40$. The diagonal elements

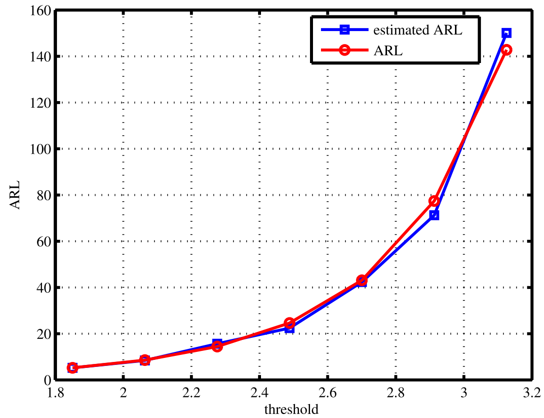


FIGURE 1. **ARL versus detection threshold λ with respect to case 1 for simulations and theoretical expression given in (38).**

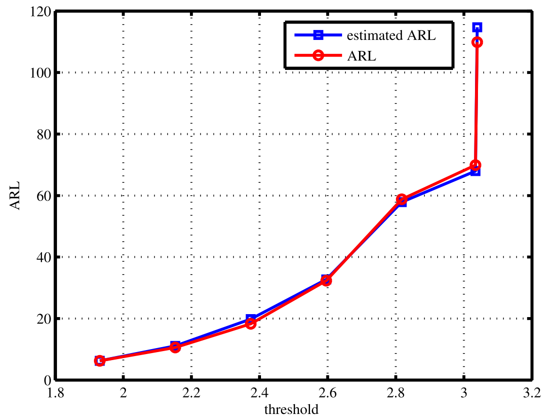


FIGURE 2. **ARL versus detection threshold λ with respect to case 2 for simulations and theoretical expression given in (38).**

of the connectivity probability matrix P_0 are 0.5, and the off-diagonal elements of P_0 are 0.1. For the second case, the stochastic block model has same block number M_1 and probability matrix P_0 as the first case but with the different block sizes, i.e., $N_{0,1} = N_{0,4} = N_{0,5} = N_{0,6} = 40$, and $N_{0,2} = N_{0,3} = 30$. The Monte-Carlo times are 200, and the simulation results are provided in Figs. 1 and 2. It can be seen from the results that the estimated ARL from simulations fits very well to the derived theoretic expression of ARL. As a result, the proposed algorithm can use the theoretic way to set detection threshold, without resorting to the Monte Carlo simulations.

B. SIMULATION EXAMPLE 2 FOR DYNAMIC VERTEX SELECTION

In this subsection, two simulations are provided to illustrate the performance of dynamic vertex selection for choosing the vertex pair to capture the anomalous change within block. In the first simulation, we consider two scenarios, i.e., after a change point, one block with vertex number $N_{0,1} = 50$ breaks into two blocks with $N_{1,1} = N_{1,2} = 25$ in the scenario 1, while the block is divided into two blocks with $N_{1,1} = 20$ and $N_{1,2} = 30$ in the second scenario 2. The

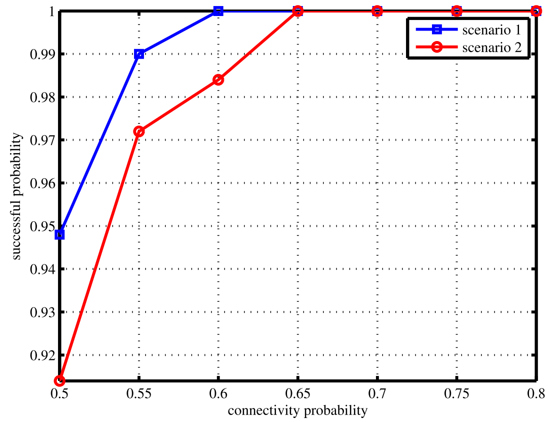


FIGURE 3. **Successful probability of dynamic vertex selection versus the connectivity probability of the first block after change point for two scenarios.**

two scenarios have the same connectivity probabilities where $[P_1]_{22} = 0.5$, $[P_1]_{12} = [P_1]_{12} = 0.1$, and the connectivity probability $[P_1]_{11}$ in the first block varies from 0.5 to 0.8. Additionally, assume that the mth block with vertex set $V_{0,m}$ is partitioned into $V^*_{1,m}$ and $V^*_{2,m}$, we define the successful probability to characterize the vertex pair $\left(v^R_m, v^*_m\right)$ chosen by dynamic vertex selection scheme,

$$\bar{p}_m = \sum_{i=1}^{MC} I\left\{v^R_m \in V^*_{1,m}, v^*_m \in V^*_{2,m}\right\}/MC, \tag{41}$$

where MC denotes the Monte Carlo times, and $I\{.\}$ represents indicator function where $I\{.\} = 1$ when the condition holds, otherwise $I\{.\} = 0$. Let $MC = 500$, the simulation for both scenarios is shown in Fig. 3. The results indicates that the vertex pair selected by dynamic vertex selection blongs to two different blocks with high probability.

In the second simulation, two blocks $V_{0,1}$ and $V_{0,2}$ corresponding to $N_{0,1} = N_{0,2} = 45$ are divided into three blocks $V_{1,1}$, $V_{1,2}$ and $V_{1,3}$ with $N_{1,1} = 30$, $N_{1,2} = 35$, and $N_{1,3} = 25$ after a change point. Note that $V_{1,2}$ includes 15 vertices in $V_{0,1}$ and 20 vertices in $V_{0,2}$. Let $MC = 500$, $[P_1]_{11} = [P_1]_{33} = 0.4$, and off-diagonal elements of $[P_1]$ is 0.1. $[P_1]_{22}$ varies from 0.4 to 0.7. The successful probabilities versus $[P_1]_{22}$ are presented in Fig. 4, where scenario 1 and 2 are corresponding to the blocks $V_{0,1}$ and $V_{0,2}$, respectively. The successful probability in scenario 3 is given by

$$\tilde{p}_m = \frac{\sum_{i=1}^{MC} I\left\{v^R_1 \in V^*_{1,1}, v^*_1 \in V^*_{2,1} \; or \; v^R_2 \in V^*_{1,2}, v^*_2 \in V^*_{2,2}\right\}}{MC}. \tag{42}$$

The simulation results show that when $[P_1]_{22}$ reaches to 0.5, scenario 2 have high probability to guarantee that the selected vertices from $V_{0,1}$ belongs to $V_{1,2}$ and $V_{1,3}$, respectively. The successful probabilities of scenario 1 slightly decay along with the growth of $[P_1]_{22}$. The main reason is that in scenario 1 the anomalous pattern related to membership is compensated by the increase of the connectivity probability $[P_1]_{22}$,

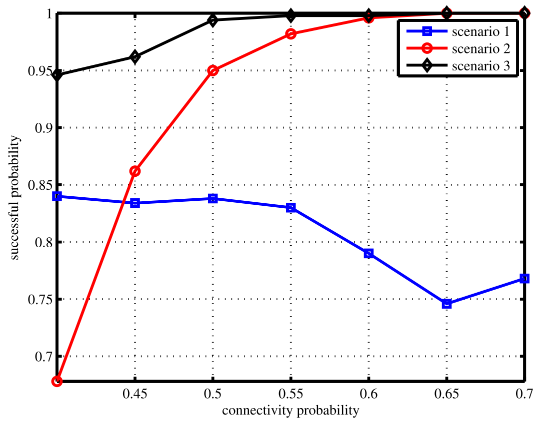


FIGURE 4. Successful probability of dynamic vertex selection versus the connectivity probability of the second block after change point for three scenarios.

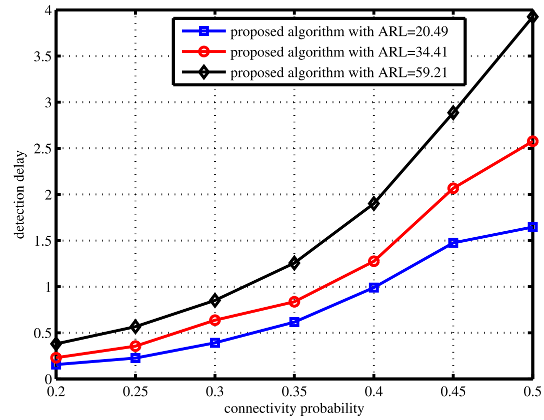


FIGURE 5. Detection delay of the proposed algorithm versus the connectivity probability of the second block after change point under 240 vertices.

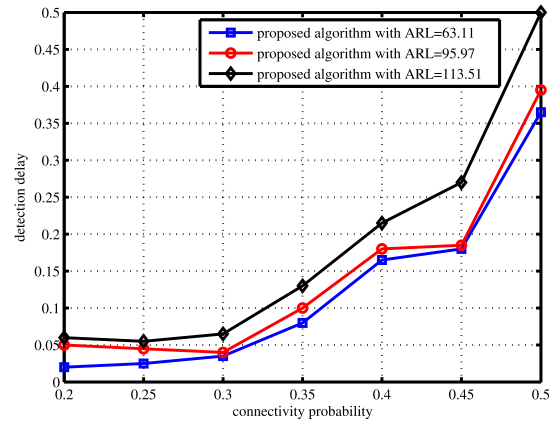


FIGURE 6. Detection delay of the proposed algorithm versus the connectivity probability of the second block after change point under 220 vertices.

which results in performance attenuation. However, it can be seen from scenario 3 that there are at least one vertex pair with high probability, which can characterize the anomalous of random networks. In other words, the developed vertex selection scheme of this paper can efficiently capture the anomalous features of random networks and guarantee the detection performance of the change-point detection algorithm. Furthermore, since the proposed algorithm does not need to construct the detection statistic over the whole vertex set of random graphs, it avoids the high computational complexity that uses all vertices for constructing detection statistic.

C. SIMULATION EXAMPLE 3 FOR ALGORITHM 1

In this subsection, we verify the detection performance of the proposed algorithm with several simulations. In the first simulation, the stochastic block networks with $M_1 = 4$, $N_{0,1} = \cdots = N_{0,4} = 60$ are divided into the random graphs with $M_2 = 5$, $N_{1,1} = N_{1,2} = 30$, , $N_{1,3} = \cdots = N_{1,5} = 60$ after a change point. Note that in this simulation the block $V_{0,1}$ beaks into two blocks $V_{1,1}$ and $V_{1,2}$ with the same size. The off-diagonal elements of P_0 and P_1 are 0.1, and except that $[P_1]_{22}$ varies from 0.2 to 0.5, the diagonal elements of P_0 and P_1 are 0.5. The Fig. 5 provides the results. In the second simulation, we consider that the stochastic block networks with $M_1 = 4$, $N_{0,1} = N_{0,4} = 60$, $N_{0,2} = N_{0,3} = 50$ are partitioned into the random graphs with $M_2 = 5$, $N_{1,1} = N_{1,3} = 35$, $N_{1,2} = 40$, $N_{1,4} = 50$, $N_{1,5} = 60$, and the vertex set $V_{1,2}$ contains 25 vertices in $V_{0,1}$ and 15 vertices in $V_{0,2}$. The settings of probability matrix are same as the first simulation, and the simulation results are given by Fig. 6. Figs. 5 and 6 show that since the anomalous feature of connectivity weakens the anomalous pattern related to memberships, the low detection delay of the proposed algorithm slightly increase along with the connectivity probability $[P_1]_{22}$. Even so, the proposed algorithm still achieves the low detection delay.

In the following simulations, we compare the detection performance of the proposed algorithm with the two detection methods in [25], where the parameters of vertex-dependent normalization, temporal normalization, and distance are given by 1, 0, 1, respectively. We keep the same settings of memberships and connectivity probability as the second simulation, except $[P_1]_{22}$. In the scenario 1, let $[P_1]_{22} = 0.4$ and the experimental results are provided in Fig. 7. In the scenario 2, $[P_1]_{22}$ is 0.8, and Fig. 8 gives the detection

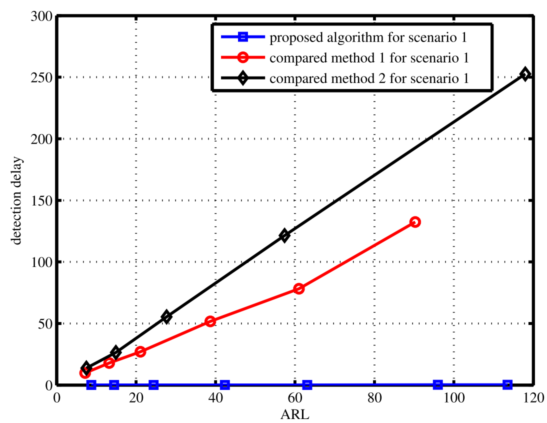


FIGURE 7. Detection delay versus ARL with $[P_1]_{22} = 0.4$.

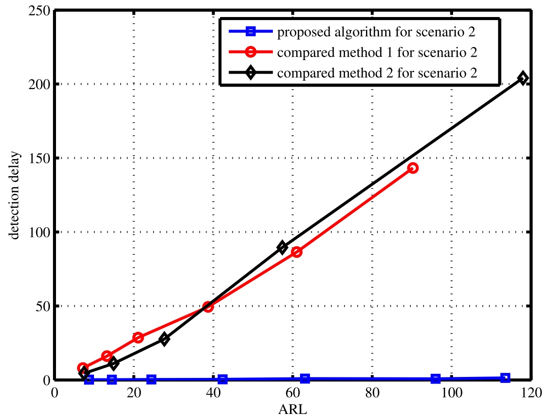


FIGURE 8. Detection delay versus ARL with $[P_1]_{22} = 0.8$.

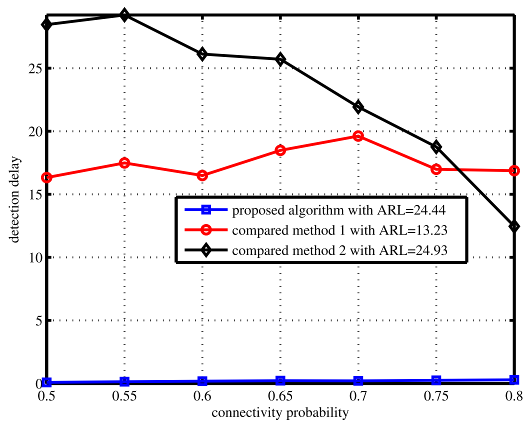


FIGURE 9. Detection delay versus the connectivity probability $[P_1]_{22}$ with a fixed ARL.

performance of three compared algorithms. In the last simulation, we assume that $[P_1]_{22}$ varies from 0.5 to 0.8, and the corresponding results are presented in Fig. 9. It is worth pointing out that the average run length (ARL) is a performance metric, which is corresponding to false alarm level. Therefore, the curves regarding to detection delay versus ARL in Figs. 7, 8 and 9 can measure the detection performance of the change-point detection algorithm. In addition, since the detection statistics of the proposed algorithm and the compared methods belong to the discrete distributions and have different statistic behaviors, we cannot achieve the same ARL for these algorithms. As a result, the number of labelled points shown in Figs. 7 and 8 is different. Even so, the curves in Figs. 7 and 8 can provide performance comparison for these algorithms. It can be seen from Figs. 7, 8 and 9 that the proposed algorithm can capture the anomalous pattern of membership and connectivity probability, and provide better change-point detection performance than the compared methods.

VI. CONCLUSION

In this paper, we consider the change-point detection of stochastic block networks. A dynamic vertex selection scheme is presented to efficiently capture the anomalous features of networks related to memberships and connectivity of subnetworks. Based on the selected vertex pairs, the constructed detection statistic achieves lower detection delay than the typical detection algorithms. Moreover, the theoretical expression of ARL of the local statistic is derived so that the proposed algorithm can set detection threshold in theoretical way without resorting to Monte Carlo simulations. In addition, the probability bounds with respect to dynamic vertex selection are also obtained to illustrate that the proposed scheme can capture the anomalous change with high probability.

APPENDIX A
PROOF OF LEMMA 1

To prove Lemma 1, we first obtain the cdf $pr\left(F_{0,m}=n\right)$ of the graph feature $F_{0,m}$, which is a binomial distribution with parameter $p=[P_0]_{mm}$. The cdf $pr\left(F_{0,m}=n\right)$ can be given by

$$pr\left(F_{0,m,n}\right) = pr\left(F_{0,m}=n\right) = \sum_{n=0}^{N_{0,m}-1} C_{N_{0,m}-1}^{n} p^n (1-p)^L, \tag{43}$$

where $L = N_{0,m}-1-n$, $n \in \left\{0, 1, \cdots, N_{0,m}-1\right\}$. Let $F_{0,m}=n$, then $F_{2,m}=F_{2,m,n}$. Assume that v_i is any one vertex in $V_{0,m}\backslash V_{0,m}^n$, then let $\bar{F}_{2,m,n}$ denote the degree of the vertex v_i under $F_{0,m}=n$. When $1 \le n \le N_{0,m}-2$, we have

$$\bar{G}_{2,m,n}(K) = pr\left(\bar{F}_{2,m,n} < K\right) = \sum_{k=0}^{K-1} C_n^k p^k (1-p)^{n-k}, \tag{44}$$

where $K \in \left\{0, 1, \cdots, N_{0,m}-1\right\}$. When $1 \le n \le N_{0,m}-2$, we have

$$\bar{G}_{2,m,n}(K) = 1, \ \ K \in \left\{0, 1, \cdots, N_{0,m}-1\right\}. \tag{45}$$

Let $G_{2,m,n}(K) = pr\left(F_{2,m,n} < K\right)$ represent the cdf of the graph feature $F_{2,m,n}$. Note that $F_{2,m,n}$ is the smallest value in regard to $\bar{F}_{2,m,n}$ related to the vertices in $V_{0,m}\backslash V_{0,m}^n$, which are independent random variables. Then, $\bar{G}_{2,m,n}(K)$ can be given by

$$\bar{G}_{2,m,n}(K) = \begin{cases} 1-\left(1-\bar{G}_{2,m,n}(K)\right)^L, & 1 \le n \le N_{0,m}-2 \\ 1, \quad n=0 \ or \ n=N_{0,m}-1, \end{cases} \tag{46}$$

where $L = N_{0,m}-1-n$. Let $\bar{F}_{4,m,n} = \min\left(F_{0,m,n}, F_{2,m,n}\right)$, then the cdf $\bar{G}_{4,m,n}(K)$ of $\bar{F}_{4,m,n}$ can be given by

$$\bar{G}_{4,m,n}(K) = pr\left(\bar{F}_{4,m,n} < K\right) = \begin{cases} G_{2,m,n}(K), & K \le n \\ 1, & K > n. \end{cases} \tag{47}$$

Let $\bar{F}_{4,m} = \min\left(F_{0,m}, F_{2,m}\right)$, and the cdf $\bar{G}_{4,m}(K)$ of $\bar{F}_{4,m}$ belongs to a compound extreme value distribution associated with the binomial distribution $pr\left(F_{0,m}=n\right)$ of graph feature $F_{0,m}$ and the distribution $\bar{G}_{4,m,n}(K)$. Then, we have

$$\bar{G}_{4,m}(K) = pr\left(\bar{F}_{4,m} < K\right)$$

$$= \sum_{n=0}^{N_{0,m}-1} pr\left(F_{0,m}=n\right)\bar{G}_{4,m,n}(K)$$
$$= \sum_{n=0}^{N_{0,m}-1} C_{N_{0,m}-1}^{n} p^{n}(1-p)^{N_{0,m}-n-1}\bar{G}_{4,m,n}(K). \quad (48)$$

Since $F_{4,m}=N_{0,m}-2-\min\left(F_{0,m},F_{2,m}\right)$, the cdf $G_{4,m}(K)$ of the graph feature $F_{4,m}$ can be given by

$$\begin{aligned} G_{4,m}(K) &= pr\left(F_{4,m}<K\right) \\ &= pr\left(\bar{F}_{4,m}>N_{0,m}-2-K\right) \\ &= 1-\bar{G}_{4,m}\left(N_{0,m}-2-K\right), \end{aligned} \quad (49)$$

where $K\in\left\{0,1,\cdots,N_{0,m}-2\right\}$. As a result, we achieve the desired result.

APPENDIX B
PROOF OF THEOREM 1

Since $F_{4,m}$ and $F_{5,m}$ are both discrete random variables, and take values in $\left\{0,1,\cdots,\bar{N}_m\right\}$ and $\left\{0,1,\cdots,\tilde{N}_m\right\}$, respectively, we can construct an $\tilde{N}_m\times\bar{N}_m$ element matrix B related to the graph feature $F_{6,m}=F_{4,m}+F_{5,m}$. The (i,j) th element $[B]_{ij}$ of B can be given by

$$[B]_{ij}=(i-1)+(j-1). \quad (50)$$

The random variable $F_{6,m}$ takes values in the element matrix B, which has the same values in the anti-diagonal lines. In order to achieve the probability mass function (pmf) of $F_{6,m}$, we further define a probability matrix P, where $[P]_{ij}$ denotes the probability that $F_{6,m}$ takes value $[B]_{ij}$. Because the random variables $F_{4,m}$ and $F_{5,m}$ are mutual independent, and $F_{6,m}$ meet with Poisson binomial distribution $G_{5,m}(K)$, the cdf of $F_{4,m}$ is given by $G_{4,m}(K)$, we have

$$[P]_{ij}=\bar{p}_i\tilde{p}_j, \quad (51)$$

where

$$\bar{p}_i = G_{4,m}(i+1)-G_{4,m}(i), \quad (52)$$
$$\tilde{p}_j = G_{5,m}(j+1)-G_{5,m}(j). \quad (53)$$

Note that the elements in the anti-diagonal lines are same, then the pmf $pr\left(F_{6,m}=K\right)$ of $F_{6,m}$ is obtained by

$$pr\left(F_{6,m}=K\right)=\sum_{1\le i\le\bar{N}_m,1\le j\le\tilde{N}_m,i+j=K+1}[P]_{ij}, \quad (54)$$

where $K\in\left\{0,1,\cdots,\bar{N}_m+\tilde{N}_m\right\}$. Thus, we achieve the desired the mean $\mu_{F_{6,m}}$ and the variance $\sigma_{F_{6,m}}$ associated with $F_{6,m}$.

APPENDIX C
PROOF OF THEOREM 2

Based on the definition of $z_{ik}=x_{1i}x_{ki}$, z_{ik}, $i\in\left\{2,3,\cdots,N_{0,1}\right\}$ are independent Bernoulli random variables with parameter p_1^2 when $v_i\in V_{1,1}$ (or p_3^2 when $v_i\in V_{1,2}$). Then, we have the variance and expectation of Z_{ik}

$$\sigma_{Z_{ik}}^2=\begin{cases} p_1^2\left(1-p_1^2\right), & v_i\in V_{1,1} \\ p_3^2\left(1-p_3^2\right), & v_i\in V_{1,2}, \end{cases} \quad (55)$$

$$\sigma_{Z_k}^2=\sum_{i=2}^{N_{0,1}}\sigma_{Z_{ik}}^2, \quad (56)$$

$$\mathbb{E}\left(Z_k\right)=\left(N_{1,1}-2\right)p_1^2+N_{1,2}p_3^2. \quad (57)$$

For $\rho>0$, we have the inequality related to moment generating function $\mathbb{E}\left(\exp\left(\rho\left(Z_{ik}-\mathbb{E}\left(Z_{ik}\right)\right)\right)\right)$, shown as

$$\begin{aligned} &\mathbb{E}\left(\exp\left(\rho\left(Z_{ik}-\mathbb{E}\left(Z_{ik}\right)\right)\right)\right) \\ &= \mathbb{E}\Big(1+\rho\left(Z_{ik}-\mathbb{E}\left(Z_{ik}\right)\right) \\ &\quad +\rho^2/2\mathbb{E}(Z_{ik}-\mathbb{E}\left(Z_{ik}\right))^2+\sum\nolimits_{l=3}^{\infty}1/l!\mathbb{E}(Z_{ik}-\mathbb{E}\left(Z_{ik}\right))^l\Big) \\ &= 1+\rho^2\sigma_{Z_{ik}}^2/2+\sum\nolimits_{l=3}^{\infty}1/l!\mathbb{E}(Z_{ik}-\mathbb{E}\left(Z_{ik}\right))^l \\ &\le 1+\rho^2\sigma_{Z_{ik}}^2/2+\sum\nolimits_{l=3}^{\infty}1/l!\mathbb{E}(Z_{ik}-\mathbb{E}\left(Z_{ik}\right))^2 \\ &\le 1+\rho^2\sigma_{Z_{ik}}^2/2+\rho^2\sigma_{Z_{ik}}^2\sum\nolimits_{l=1}^{\infty}(\rho/3)^l/2 \\ &= 1+3\rho^2\sigma_{Z_{ik}}^2/5 \\ &\le \exp\left(3\rho^2\sigma_{Z_{ik}}^2/5\right), \end{aligned} \quad (58)$$

where the first equality follows from the tailor expansion, and the first inequality holds for $|Z_{ik}-\mathbb{E}\left(Z_{ik}\right)|\le 1$. The third equality holds as $\rho<3$. As a result, we have

$$\begin{aligned} \mathbb{E}\left(\exp\left(\rho\left(Z_k-\mathbb{E}\left(Z_k\right)\right)\right)\right) &\le \exp\left(3\rho^2\sum\nolimits_{i=2}^{N_{0,1}}\sigma_{Z_{ik}}^2/5\right) \\ &= \exp\left(3\rho^2\sigma_{Z_k}^2/5\right) \end{aligned} \quad (59)$$

The upper tail bound can be given by

$$\begin{aligned} &pr\left(Z_k-\left(Z_k\right)\ge\varepsilon\right)\le\left(\exp\left(\rho\left(Z_k-\mathbb{E}\left(Z_k\right)\right)\right)\right)/\left(\exp\left(\rho\varepsilon\right)\right) \\ &\le \exp\left(3\rho^2\sigma_{Z_k}^2/5-\rho\varepsilon\right) \\ &= \exp\left(-\min\left\{\frac{5\varepsilon^2}{12\sigma_{Z_k}^2},3\varepsilon-\frac{9}{5}\sigma_{Z_k}^2\right\}\right), \end{aligned} \quad (60)$$

where the first inequality holds for Markov's inequality [26]. The equality holds by taking $\rho=\min\left\{5\varepsilon/\left(6\sigma_{Z_k}^2\right),3\right\}$ to minimize the term in the second inequality. ε is a constant. Then, we obtain

$$\begin{aligned} &pr\left(Z_k-\mathbb{E}\left(Z_k\right)\ge\varepsilon,X_{1k}=0\right) \\ &= pr\left(Z_k-\mathbb{E}\left(Z_k\right)\ge\varepsilon|X_{1k}=0\right)pr\left(X_{1k}=0\right) \\ &= \left(1-p_1\right)pr\left(Z_k-\mathbb{E}\left(Z_k\right)\ge\varepsilon|X_{1k}=0\right) \\ &= \left(1-p_1\right)pr\left(Z_k-\mathbb{E}\left(Z_k\right)\ge\varepsilon\right) \\ &\le \left(1-p_1\right)\exp\left(-\min\left\{\frac{5\varepsilon^2}{12\sigma_{Z_k}^2},3\varepsilon-\frac{9}{5}\sigma_{Z_k}^2\right\}\right), \end{aligned} \quad (61)$$

where the last equality holds for Z_k and X_{1k} are independent. Based on the above proof, the lower tail bound of $pr\left(Z_k-\mathbb{E}\left(Z_k\right)\le\varepsilon,X_{1k}=0\right)$ has same form as (61). Consequently, the desired bound is given.

REFERENCES

[1] R. Couillet and W. Hachem, "Fluctuations of spiked random matrix models and failure diagnosis in sensor networks," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 509–525, Jan. 2013.

[2] S. Abbas, M. Faisal, H. Ur Rahman, M. Z. Khan, M. Merabti, and A. U. R. Khan, "Masquerading attacks detection in mobile ad hoc networks," *IEEE Access*, vol. 6, pp. 55013–55025, 2018.

[3] M. He and J. Zhang, "A dependency graph approach for fault detection and localization towards secure smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 342–351, Jun. 2011.

[4] H. Olufowobi, C. Young, J. Zambreno, and G. Bloom, "SAIDuCANT: specification-based automotive intrusion detection using controller area network (CAN) timing," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1484–1494, Feb. 2020.

[5] H. Sedjelmaci, S. M. Senouci, and T. Taleb, "An accurate security game for low-resource IoT devices," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9381–9393, Oct. 2017.

[6] J. Cadena, F. Chen, and A. Vullikanti, "Graph anomaly detection based on steiner connectivity and density," *Proc. IEEE*, vol. 106, no. 5, pp. 829–845, May 2018.

[7] B. Wang, Y. Zhao, F. Hu, and Y.-C. Liang, "Anomaly detection with subgraph search and vertex classification preprocessing in chung-lu random networks," *IEEE Trans. Signal Process.*, vol. 66, no. 20, pp. 5255–5268, Oct. 2018.

[8] F. Zheng and Q. Liu, "Anomalous telecom customer behavior detection and clustering analysis based on ISP's operating data," *IEEE Access*, vol. 8, pp. 42734–42748, 2020.

[9] J. Heydari and A. Tajer, "Quickest search for local structures in random graphs," *IEEE Trans. Signal Inf. Process. over Netw.*, vol. 3, no. 3, pp. 526–538, Sep. 2017.

[10] A. Goldenberg, "A survey of statistical network models," *Found. Trends Mach. Learn.*, vol. 2, no. 2, pp. 129–233, 2009.

[11] S. Kar and J. M. F. Moura, "Sensor networks with random links: Topology design for distributed consensus," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3315–3326, Jul. 2008.

[12] P. W. Holland, K. B. Laskey, and S. Leinhardt, "Stochastic blockmodels: First steps," *Social Netw.*, vol. 5, no. 2, pp. 109–137, Jun. 1983.

[13] E. M. Airoldi, D. M. Blei, S. E. Fienberg, and E. P. Xing, "Mixed membership stochastic blockmodels," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 33–40.

[14] K. Nowicki and T. A. B. Snijders, "Estimation and prediction for stochastic blockstructures," *J. Amer. Stat. Assoc.*, vol. 96, no. 455, pp. 1077–1087, Sep. 2001.

[15] K. S. Xu and A. O. Hero, "Dynamic stochastic blockmodels for time-evolving social networks," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 4, pp. 552–562, Aug. 2014.

[16] A. Kadavankandy, L. Cottatellucci, and K. Avrachenkov, "Characterization of random matrix eigenvectors for stochastic block model," in *Proc. 49th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2015, pp. 861–865.

[17] J. Sharpnack, A. Rinaldo, and A. Singh, "Detecting anomalous activity on networks with the graph Fourier scan statistic," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 364–379, Jan. 2016.

[18] B. A. Miller, M. S. Beard, P. J. Wolfe, and N. T. Bliss, "A spectral framework for anomalous subgraph detection," *IEEE Trans. Signal Process.*, vol. 63, no. 16, pp. 4191–4206, Aug. 2015.

[19] B. A. Miller, N. T. Bliss, and P. J. Wolfe, "Subgraph detection using eigenvector L1 norms," in *Proc. Adv. Neural Inf. Process Syst.*, 2010, pp. 1633–1641.

[20] T. Komolafe, A. Valeria Quevedo, S. Sengupta, and W. H. Woodall, "Statistical evaluation of spectral methods for anomaly detection in networks," 2017, *arXiv:1711.01378*. [Online]. Available: http://arxiv.org/abs/1711.01378

[21] T. Mifflin, "Detection theory on random graphs," in *Proc. 12th Int. Conf. Inf. Fusion*, Seattle, WA, USA, Jul. 2009, pp. 954–959.

[22] T. L. Mifflin, C. Boner, G. A. Godfrey, and J. Skokan, "A random graph model for terrorist transactions," in *Proc. IEEE Aerosp. Conf.*, Mar. 2004, pp. 3258–3264.

[23] A. Kadavankandy, K. Avrachenkov, L. Cottatellucci, and R. Sundaresan, "The power of side-information in subgraph detection," *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1905–1919, Apr. 2018.

[24] D. Marangoni-Simonsen and Y. Xie, "Sequential changepoint approach for online community detection," *IEEE Signal Process. Lett.*, vol. 22, no. 8, pp. 1035–1039, Aug. 2015.

[25] H. Wang, M. Tang, Y. Park, and C. E. Priebe, "Locality statistics for anomaly detection in time series of graphs," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 703–717, Feb. 2014.

[26] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford, U.K.: Oxford Univ. Press, 2013.

[27] B. Wang, R. C. Qiu, and Y. Zhao, "Distributed source detection with dimension reduction in multiple-antenna wireless networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 2966–2980, Apr. 2017.

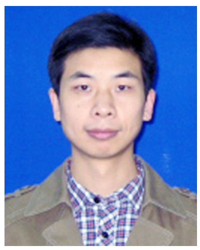
[28] H. Saad and A. Nosratinia, "Community detection with side information: Exact recovery under the stochastic block model," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 5, pp. 944–958, Oct. 2018.

[29] P. K. Gopalan and D. M. Blei, "Efficient discovery of overlapping communities in massive networks," *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 36, pp. 14534–14539, Sep. 2013.

[30] B. A. Miller, N. T. Bliss, P. J. Wolfe, and M. S. Beard, "Detection theory for graphs," *Lincoln Lab. J.*, vol. 20, no. 1, pp. 10–30, 2013.

[31] Y. Hong, "On computing the distribution function for the Poisson binomial distribution," *Comput. Statist. Data Anal.*, vol. 59, pp. 41–51, Mar. 2013.

YANPING ZHAO received the Ph.D. degree in communication and information systems from Jilin University, China, in 2014. She is currently a Lecturer with the College of Communication Engineering, Jilin University. Her research interests include speech signal processing and sparse signal processing.



BO WANG received the Ph.D. degree in communication and information systems from Jilin University, Changchun, China, in 2006. From July 2014 to July 2015, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, Tennessee Technological University (TTU), Cookeville, TN, USA. He is currently a Professor with the College of Communication Engineering, Jilin University. His research interests include graph signal processing, array signal processing, signal detection, and source localization.

MINGAN LUAN (Graduate Student Member, IEEE) received the B.S. degree in electronics and information engineering from the Changchun Institute of Technology, China, in 2018. He is currently pursuing the M.S. degree with the Information and Signal Processing Laboratory, College of Communication Engineering, Jilin University. His research interests include signal detection, machine learning, and graph signal processing.

FENGYE HU (Member, IEEE) received the B.S. degree from the Department of Precision Instrument, Xi'an University of Technology, China, in 1996, and the M.S. and Ph.D. degrees in communication and information systems from Jilin University, China, in 2000 and 2007, respectively. He was a Visiting Scholar in electrical and electronic engineering from Nanyang Technological University (NTU), Singapore, in 2011. He is currently a Professor with the College of Communication Engineering, Jilin University. His research interests include wireless body area networks, wireless energy and information transfer, energy harvesting, cognitive radio, and space-time communication.