JOURNAL OF APPLIED STATISTICS
2024, VOL. 51, NO. 9, 1621–1641
https://doi.org/10.1080/02664763.2023.2230530

Modelling and monitoring social network change based on exponential random graph models

Yantao Cai[a], Liu Liu[b] and Zhonghua Li [a]

[a]School of Statistics and Data Science, LPMC, LEBPS and KLMDASR, Nankai University, Tianjin, People's Republic of China; [b]College of Mathematics and Physics, Chengdu University of Technology, Chengdu, People's Republic of China

ABSTRACT

This paper aims to detect anomalous changes in social network structure in real time and to offer early warnings by phase II monitoring social networks. First, the exponential random graph model is used to model social networks. Then, a test and online monitoring technique of the exponential random graph model is developed based on the split likelihood-ratio test after determining the model and its parameters for a specific data set. This proposed approach uses pseudo-maximum likelihood estimation and likelihood ratio to construct the test statistics, avoiding the several steps of discovering Monte Carlo Markov Chain maximum likelihood estimation through an iterative method. A bisection algorithm for the control limit is given. Simulations on three data sets Flobusiness, Kapferer and Faux.mesa.high are presented to study the performance of the procedure. Different change points and shift sizes are compared to see how they affect the average run length. A real application example on the MIT reality mining social proximity network is used to illustrate the proposed modelling and online monitoring methods.

ARTICLE HISTORY
Received 2 November 2022
Accepted 13 June 2023

KEYWORDS
Exponentially random graph model; online monitoring; social network; split likelihood ratio test; statistical process control

1. Introduction

Social network has infiltrated practically every part of our lives. Many things, such as transportation, telephone or Internet communication, disease transmission and criminal activity, can be defined as social networks [4,8]. A social network may be described as a graph with nodes (people) and edges (social ties), which can allow users to investigate its structural aspects using mathematical and statistical methods. The social network in a group may change dynamically over time. When a major event occurs, the network structure frequently shifts dramatically. Expecting that by monitoring the social network, we will be able to detect structural changes in the social network in time and offer an early warning when anomalous occurrences occur, such as criminal activities, public opinion storms and so on. Social network analysis and statistical process control (SPC) are the aims of this paper.

CONTACT Zhonghua Li zli@nankai.edu.cn School of Statistics and Data Science, LPMC, LEBPS and KLMDASR, Nankai University, Tianjin 300071, People's Republic of China

SPC is a technique widely used to monitor industrial processes [13,27,32,38,41]. Applications of SPC are divided into two stages: Phase I and Phase II. In Phase I, a set of process data is collected and analysed all at once to determine whether the process has been in-control (IC). After Phase I, we have a set of data that is representative of IC process. In Phase II, control charts are used to monitor the process by comparing the sample statistic to the control limits. This paper focuses on the Phase II monitoring methods for social networks.

When it comes to combining social network analysis and SPC methods, the overview papers [30,40] gave comprehensive overviews of some statistical methods for the monitoring of social networks before 2018. Besides the references therein, [5] suggested a change point detection approach for correlation networks and [20] discussed several possible network metrics to be used for a change point detection problem. More recently, [35] fitted the network data with an exponential random graph model (ERGM) to obtain maximum likelihood estimation (MLE) of the model parameters, and applied Hotelling T^2 and multivariate exponentially weighted moving average (MEWMA) control charts to monitor the parameters, and [26] used ERGM and Hotelling's T^2 and likelihood ratio test control charts in Phase I. From the perspective of ERGMs, the global structure is made by a combination of different substructures, which is similar to the multiple regression analysis. ERGMs use some summary statistics as covariates, rather than the entire information of network structure such as adjacency matrix, which leads to some information loss of the network structure. However, unlike models that require an independent and specific distribution of the observation data, ERGMs may describe related network data with ease, and the node attributes can be simply added to the model. As a result, despite the limitation of ERGMs, they can be used to explain the network data with complicated associations in a variety of ways, which hence motivates us to use ERGMs to model social networks. In addition to ERGMs, recently, [1] fitted the network data with degree corrected stochastic block model (DCSBM) and used three multivariate process monitoring techniques Hotelling T^2, MEWMA, and MCUSUM to monitor the model parameters simultaneously. Then [2] proposed a labelled DCSBM (LDCSBM) and used Shewhart control chart to monitor the LDCSBM. [25] considered monitoring attributed social networks based on count data and random effects. Considering the interaction between nodes as a function of nodal similarities, [29] used zero-inflated Poisson regression to model sparse and attributed network. Furthermore, they proposed a likelihood ratio test and an EWMA chart to monitor social networks. [9] integrated the Hurdle model with a state-space model to capture temporal dynamics of the edge formation process and used EWMA control charts to monitor the residuals and detect change points. What is more, [28] combined zero-inflated generalized linear mixed models (ZI-GLMMs) with the state-space model to fit the sequences of sparse, attributed, and weighted multilevel networks and monitor changes in them.

In aspect of ERGM, much effort is devoted to the estimation of the parameters of ERGMs [10,14,16,18,37]. Assuming a finite dimension of the parameter space, [34] showed that the MLE is not consistent in the ERGMs when the sufficient statistics involve k-stars, triangles and motifs of k-nodes ($k \geq 2$). When dealing with a higher-order complex network sequence, the computational burden will be very high, as the MLE of ERGM is derived using the Markov Chain Monte Carlo (MCMC) approach [36]. [39] proposed the sample split likelihood ratio test (Split LRT). This is a fairly basic test procedure that does not require any regular conditions. Therefore it is ideal for statistical inferences in complicated scenarios. The modelling and on-line monitoring method of social networks based on ERGM is provided in this paper, which is combined with the sample Split LRT.

The rest of this paper is organized as follows. The ERGMs and Split LRT of ERGMs are briefly introduced in Section 2. The proposed online monitoring methods of ERGMs and a searching algorithm to find the control limit are presented in Section 3. Several computer simulations are carried out in Section 4, in which ERGMs are used to fit three datasets: Flobusiness, Kapferer and Faux.mesa.high, the power of the test is displayed, and the performance comparisons of the proposed control chart with MEWMA and T^2 control charts are given. Section 5 studies a real application example on the MIT reality mining social proximity network to illustrate the proposed modelling and on-line monitoring methods. Section 6 outlines the merits of the proposed methods, as well as future study possibilities.

2. Background

2.1. Brief introduction to ERGMs

For a graph with adjacency matrix y, ERGMs have the following form [10,14]:

$$P_\theta(Y = y) = \exp\{\theta^\top Z(y) - \phi(\theta)\}, \tag{1}$$

where $Z(y)$ is the chosen statistics, $\theta \in R^q$ is the matching parameter, and $\phi(\theta) = \log \sum_y \exp\{\theta^\top Z(y)\}$ is a normalized parameter that assures Equation (1) is a suitable probability distribution.

The items in $Z(y)$ denote the number of edges and substructures in the graph (k-star, k-triangles, k-twopath, etc.). The selection of statistics usually depends on the presupposition of the correlation between elements y_{ij} in the adjacency matrix [10]. When considering the simplest case, in which the graph is undirected and $Z(y)$ is the number of edges, the model degenerates into the Erds–Rényi model, in which each element $y_{ij} (i < j)$ of the adjacency matrix y is independent and follows the Bernoulli distribution. ERGMs takes the elements y_{ij} as samples, and $P_\theta(Y = y)$ can be seen as the joint probability density of these samples.

Some widely selected statistics are as follows [16,37]:

Edges:

$$E(y) = \sum_{i,j} y_{ij};$$

geometrically weighted degree statistic (gwdegree(γ_s)) :

$$gwdegree(y, \gamma_s) = \mathrm{e}^{\gamma_s} \sum_{k=1}^{n-1} \{1 - (1 - e^{-\gamma_s})^k\} D_k(y);$$

geometrically weighted edgewise shared partner statistic (gwesp(γ_t)):

$$gwesp(y, \gamma_t) = \mathrm{e}^{\gamma_t} \sum_{k=1}^{n-2} \{1 - (1 - \mathrm{e}^{-\gamma_t})^k\} EP_k(y);$$

geometrically weighted dyad-wise shared partner statistic (gwdsp(γ_p)):

$$gwdsp(y, \gamma_p) = e^{\gamma_p} \sum_{k=1}^{n-2} \{1 - (1 - e^{-\gamma_p})^k\} DP_k(y);$$

where $D_k(y)$ is defined to be the number of nodes in network y whose degree equals k, $EP_k(y)$ is defined as the number of unordered pairs $\{i, j\}$ in network y such that $y_{ij} = 1$ and nodes i and j have exactly k common neighbours (edgewise shared partner), and $DP_k(y)$ is the number of unordered pairs $\{i, j\}$ such that i and j have exactly k common neighbours whatever the value of y_{ij} (dyadic shared partner).

By these definitions, it is always true that $DP_k(y) \geq EP_k(y)$, and in fact $DP_k(y) - EP_k(y)$ equals the number of unordered pairs $\{i, j\}$ for which $y_{ij} = 0$ and i and j share exactly k common neighbours. The gwdegree, gwesp and gwdsp can be rewritten from alternating k-star statistic, alternating k-triangle statistic and alternating k-twopath statistic respectively, where

$$\text{alternating k-star}(y, \lambda_s) = S_2(y) - \frac{S_3(y)}{\lambda_s} + \cdots + (-1)^{n-3} \frac{S_{n-1}(y)}{\lambda_s^{n-3}};$$

$$\text{alternating k-triangle}(y, \lambda_t) = 3T_1(y) - \frac{T_2(y)}{\lambda_t} + \cdots + (-1)^{n-3} \frac{T_{n-2}(y)}{\lambda_t^{n-3}};$$

$$\text{alternating k-twopath}(y, \lambda_p) = P_1(y) - \frac{2P_2(y)}{\lambda_p} + \cdots + (-1)^{n-3} \frac{P_{n-2}(y)}{\lambda_p^{n-3}},$$

and $S_k(y)$, $T_k(y)$, $P_k(y)$ are k-star, k-triangle and the k-twopath respectively. What is more, degree statistics $D_i(y)$ and the edge wise and dyadic shared partner statistics $EP_i(y)$, $DP_i(y)$ are related to k-star statistics $S_k(y)$, k-triangle $T_k(y)$ and k-twopath statistics $P_k(y)$, respectively, by the following equations:

$$S_k(y) = \begin{cases} \frac{1}{2} \sum_{i=1}^{n-1} i D_i(y), & k = 1, \\ \sum_{i=k}^{n-1} \binom{i}{k} D_i(y), & 2 \leq k \leq n-2; \end{cases}$$

$$T_k(y) = \begin{cases} \frac{1}{3} \sum_{i=0}^{n-2} i EP_i(y), & k = 1, \\ \sum_{i=k}^{n-2} \binom{i}{k} EP_i(y), & 2 \leq k \leq n-2; \end{cases}$$

$$P_k(y) = \begin{cases} \frac{1}{2} \sum_{i=2}^{n-2} \binom{i}{2} DP_i(y), & k = 2, \\ \sum_{i=k}^{n-2} \binom{i}{k} DP_i(y), & 1 \leq k \leq n-2,\ k \neq 2. \end{cases}$$

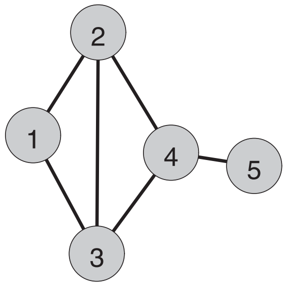

Figure 1. A five-node network example.

To make these concepts concrete, consider the simple undirected graph in Figure 1 presented by [16]. For this five-node network, the edgewise and dyadic shared partner distributions are $(EP_0, EP_1, EP_2, EP_3) = (1, 4, 1, 0)$ and $(DP_0, DP_1, DP_2, DP_3) = (2, 6, 2, 0)$, respectively. The k-triangle and k-twopath distributions are $(T_1, T_2, T_3) = (2, 1, 0)$ and $(P_1, P_2, P_3) = (10, 1, 0)$, respectively.

To rewrite the alternating k-star, alternating k-triangle and alternating k-twopath statistics in terms of the $D_k(y)$, $EP_k(y)$ and $DP_k(y)$, begin by substituting $S_k(y)$, $T_k(y)$ and $P_k(y)$ into the equations above, then introduce the parameters γ_s=log λ_s, γ_t=log λ_t and γ_p=log λ_p. Finally, after simplification, the statistics gwdegree, gwesp and gwdsp can be calculated directly by $D_k(y), EP_k(y), DP_k(y)$. These statistics may be expected to avoid the large degeneracy problems associated with the traditional specification such as k-star statistic and k-triangle statistic, and lead to much better results. [16,37] provided more details about these statistics. In this article, we will consider them as candidate variables for the model. For the sake of simplicity, we will assume that $\gamma_s, \gamma_t, \gamma_p$ are fixed at 0.7 in the following of this paper.

2.2. Split LRT of ERGMs

Classical LRT requires that the statistical model satisfy certain regularity conditions such that the log-likelihood-ratio has asymptotic chi-square distribution. But if the regularity conditions do not hold, like ERGMs, the limiting distribution of likelihood-ratio will be intractable. Therefore Split LRT is a simple method without any regularity conditions, which can be used for any parametric model. Although [3,6,11,12,33] have given approaches to calculate the likelihood ratio of ERGMs, the asymptotic distribution of the estimated likelihood ratio remains unknown due to the limitations of the model and the complex correlation between edges (sample) y_{ij}, $1 \le i < j \le n$ [34]. In such irregular statistical models, the sample Split LRT [39] can be employed to test composite null hypotheses. The main idea is to split the data into two sets. Under the null hypothesis, the two sets of data are independent and identically distributed (i.i.d.). One set of data is used to estimate the parameters, while the other set is utilized to generate the likelihood ratio, which is then used as a test statistic. The advantage of this method is that it does not require any extra conditions, but only i.i.d. of the two sets of samples.

[39] considered a possibly composite null set Θ_0 and the testing

$$H_0 : \theta \in \Theta_0 \quad \text{versus } H_1 : \theta \notin \Theta_0,$$

and $\arg\max_{\theta\in\Theta_0} L_0(\theta)$ is used in likelihood ratio. To apply Split LRT to the test of ERGMs, in our setting, we assume that the graphs Y_0 and Y_1 are identically distributed $P_\theta(Y)$, and we consider the simple null hypothesis

$$H_0 : \theta = \theta^\star \quad \text{versus } H_1 : \theta \neq \theta^\star. \tag{2}$$

Here we suppose $\theta^\star$ is known. Let $L_0(\theta) = P_\theta(Y_0), L_1(\theta) = P_\theta(Y_1)$, and $\hat{\theta}_0, \hat{\theta}_1$ be the estimation of θ based on sample Y_0, Y_1, respectively. Let $U = L_0(\hat{\theta}_1)/L_0(\theta^\star)$, and $U^{swap} = L_1(\hat{\theta}_0)/L_1(\theta^\star)$. Since

$$\begin{aligned} E_{\theta^\star}\left[\frac{L_0(\hat{\theta}_1)}{L_0(\theta^\star)}\right] &= \sum_{Y_0} \frac{L_0(\hat{\theta}_1)}{L_0(\theta^\star)} P_{\theta^\star}(Y = Y_0) \\ &= \sum_{Y_0} L_0(\hat{\theta}_1) = 1, \end{aligned}$$

in the case of the null hypothesis, and by Markov inequality, we have

$$P_{\theta^\star}\left(U > \frac{1}{\alpha}\right) \leq \alpha E_{\theta^\star}\left[\frac{L_0(\hat{\theta}_1)}{L_0(\theta^\star)}\right] = \alpha.$$

Denote the Split LRT statistics recommended in [39] as

$$SL = \frac{U + U^{swap}}{2}. \tag{3}$$

It can be similarly seen that

$$P_{\theta^\star}\left(SL > \frac{1}{\alpha}\right) \leq \frac{\alpha}{2} E_{\theta^\star}\left[\frac{L_0(\hat{\theta}_1)}{L_0(\theta^\star)} + \frac{L_1(\hat{\theta}_0)}{L_1(\theta^\star)}\right] = \alpha. \tag{4}$$

When the significance level α is set, the null hypothesis H_0 will be rejected if $SL > 1/\alpha$, hence the probability of the Type I error can be controlled less than α.

It can be seen from the construction process that we do not need the particular form of the statistic $\hat{\theta}_1, \hat{\theta}_0$, and all we need is a solid estimate of θ, which can be MLE, Bayesian estimation, or any other estimation [3,33,39].

3. Online monitoring of ERGMs

3.1. Online monitoring of ERGMs based on split LRT

In this section, we construct a split LRT-based online monitoring method of ERGMs. Assuming that $Y_1, Y_2, \ldots, Y_t, \ldots$ are an online sequence of independent un-directed graphs observed during Phase II process monitoring. It is necessary to monitor whether the distribution of the new sample Y_t changes in comparison to the previous observations, that is, whether the parameter θ in the ERGM shifts. Assuming that the IC distribution is

$P_{\theta^\star}(Y)$. We consider the hypothesis:

$$\begin{gathered} H_0 : Y_1, Y_2, \ldots, Y_t, \ldots \sim P_{\theta^\star}(Y) \\ \text{versus} \\ H_1 : \text{Exist} t_0 \geq 1, Y_1, Y_2, \ldots, Y_{t_0-1} \sim P_{\theta^\star}(Y) \text{and } Y_{t_0},\ Y_{t_0+1} \cdots \sim P_\theta(Y),\ \theta \neq \theta^\star. \end{gathered} \tag{5}$$

In statistical process monitoring, the two phases, Phase I and Phase II, are different in their goals. In Phase I, it aims to get a stream of IC observations and estimate the IC parameters. For example, the IC parameters can be estimated by the mean of MLEs of the IC observations. In Phase II, it attempts to detect anomalies on the foundation of known parameter estimation obtained in Phase I. In this paper, we focus on the Phase II monitoring method, so we assume that $\theta^\star$ is known or can be estimated through a set m of Phase I observations, in which case it is denoted as $\hat{\theta}_m^\star$.

In the case of offline, we might use $SL = (U + U^{swap})/2$ as test statistic. But in the case of online, we have only one new sample at a time. It is not straightforward to use the cross-fit statistic as SL recommended in [39], so we consider the sequential version of the Split LRT statistic. The performance of SL is shown based on simulation study, and we found that the application of Split LRT to test ERGMs is effective. And then we can expect the on-line version to perform well in monitoring. To build an on-line monitoring method, let $\hat{\theta}_{1,t} := \sum_{i=1}^{t} \hat{\theta}_i / t$ denote the mean value of the previous t estimates. The disadvantage of the average is that as the running time t increases, the impact of the new estimations on the average $\hat{\theta}_{1,t}$ gets smaller. Therefore, the exponential weighted average approach is utilized, that is

$$\hat{\theta}_{1,t} = \begin{cases} \hat{\theta}_1, & t = 1, \\ (1-w)\hat{\theta}_{1,t-1} + w\hat{\theta}_t, & t \geq 2, \end{cases}$$

where $\hat{\theta}_t$ is the pseudo MLE of $P_\theta(Y = Y_i)$ [6], and the weight $0 < w < 1$ is defined as smoothing constant, which ensures that each new estimate $\hat{\theta}_t$ has the same level of influence on $\hat{\theta}_{1,t}$. It can be set to 0.2 for instance. It can be seen that if the first $t-1$ samples come from the same distribution $P_{\theta^\star}(Y)$, $\hat{\theta}_{1,t-1}$ is still a solid estimate of $\theta^\star$. In the previous section, we know that the test does not require $\hat{\theta}_i$ to be a specific estimation. As the calculation speed of pseudo MLE is substantially faster than that of MCMC MLE [33], pseudo MLE is chosen as the estimation of θ in our detection method.

Specifically, when $\theta^\star$ is known, we consider the test statistics

$$M_t = \begin{cases} 1 & t = 1, \\ \dfrac{P_{\hat{\theta}_{1,t-1}}(Y_t)}{P_{\theta^\star}(Y_t)}, & t \geq 2. \end{cases}$$

And when $\theta^\star$ is unknown, let $\hat{\theta}_m^\star$ be the estimate obtained in Phase I. We consider the test statistics

$$M_t = \begin{cases} 1, & t = 1, \\ \dfrac{P_{\hat{\theta}_{1,t-1}}(Y_t)}{P_{\hat{\theta}_m^\star}(Y_t)}, & t \geq 2. \end{cases}$$

EWMA is a famous method in the SPC theory [22,23,42], which is more effective for small changes of process. The following EWMA procedure is suggested at time t,

$$S_t = \begin{cases} 1 & t = 1, \\ (1-\lambda)S_{t-1} + \lambda M_t, & t \geq 2, \end{cases} \tag{6}$$

where $0 < \lambda \leq 1$ is defined as smoothing constant. For a given IC average run length ARL_0, the control limit h is determined. The process is out-of-control (OC) when $S_t > h$. The control limit h of the EWMA control chart can be derived from the formula

$$h = \mu_0 + L\sigma\sqrt{\frac{\lambda}{2-\lambda}},$$

where L is a constant, $\mu_0 = E_{\theta^\star}[M_t] = 1$, $\sigma = \sqrt{Var_{\theta^\star}(M_t)}$. Assuming that $M_t, t = 1, 2 \cdots$ is i.i.d., σ can be replaced by the sample standard deviation.

For the proposed EWMA statistic S_t in (6), we have the following property.

Proposition 3.1: *Define the stopping time* $\tau = \inf\{t | S_t > 1/\alpha, t \geq 1\}$. *Then*

$$P_{\theta^\star}(\tau < \infty) < \alpha.$$

Proof 3.1: *Define the natural filtration* $\mathcal{F}_t = \sigma(Y_1, Y_2, \ldots, Y_t)$. *Noting that*

$$E_{\theta^\star}[M_t|\mathcal{F}_{t-1}] = E_{\theta^\star}\left[\frac{P_{\hat{\theta}_{1,t-1}}(Y_t)}{P_{\hat{\theta}^\star_m}(Y_t)}|\mathcal{F}_{t-1}\right] \leq E_{\theta^\star}\left[\frac{P_{\hat{\theta}_{1,t-1}}(Y_t)}{P_{\theta^\star}(Y_t)}|\mathcal{F}_{t-1}\right] = 1, \tag{7}$$

$$\begin{aligned} E_{\theta^\star}[S_\tau|\mathcal{F}_{\tau-1}] &= E_{\theta^\star}[S_\tau|\mathcal{F}_{\tau-1}, \tau < \infty]P_{\theta^\star}(\tau < \infty) \\ &\quad + E_{\theta^\star}[S_\tau|\mathcal{F}_{\tau-1}, \tau = \infty]P_{\theta^\star}(\tau = \infty), \end{aligned} \tag{8}$$

$$\begin{aligned} E_{\theta^\star}[S_\tau|\mathcal{F}_{\tau-1}] &= E_{\theta^\star}[(1-\lambda)S_{\tau-1}|\mathcal{F}_{\tau-1}, \tau < \infty]P_{\theta^\star}(\tau < \infty) \\ &\quad + E_{\theta^\star}[(1-\lambda)S_{\tau-1}|\mathcal{F}_{\tau-1}, \tau = \infty]P_{\theta^\star}(\tau = \infty) \\ &\quad + \lambda E_{\theta^\star}[M_\tau|\mathcal{F}_{\tau-1}]. \end{aligned} \tag{9}$$

Combining (7)–(9), then

$$E_{\theta^\star}[(1-\lambda)S_{\tau-1}|\mathcal{F}_{\tau-1}, \tau < \infty]P_{\theta^\star}(\tau < \infty) + \lambda \geq E_{\theta^\star}[S_\tau|\mathcal{F}_{\tau-1}, \tau < \infty]P_{\theta^\star}(\tau < \infty).$$

Knowing that $S_{\tau-1} \leq 1/\alpha$ *and* $S_\tau > 1/\alpha$, *then*

$$\frac{1}{\alpha}(1-\lambda)P_{\theta^\star}(\tau < \infty) + \lambda > \frac{1}{\alpha}P_{\theta^\star}(\tau < \infty).$$

So we have $P_{\theta^\star}(\tau < \infty) < \alpha$. ■

The property above is similar to [15,39]. This conclusion demonstrates that the probability of the type I error can be controlled below α for online split LRT tests with rejection region $\{S_t > 1/\alpha, t \geq 2\}$. Note $\hat{\theta}_{1,t-1}$ is still an estimate of $\theta^\star$ when the tth sample Y_t is the first OC sample, so the value of M_t should be smaller than 1. When the OC process proceeds, the new sample Y_{t+1} is OC and i.i.d. with Y_t. No matter the shift of parameter

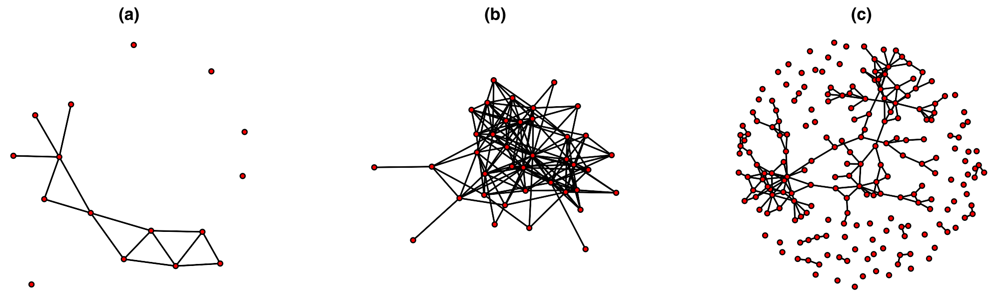

Figure 2. Plots of the networks for Flobusiness (a), Kapferer (b) and Faux.mesa.high (c).

$\theta^{\star}$ is decreasing or increasing, since $\hat{\theta}_{1,t}$ is more close to the MLE of Y_t and Y_{t+1}, the LRT $P_{\hat{\theta}_{1,t}}(Y_{t+1})/P_{\theta^{\star}}(Y_{t+1})$ will increase and so as the values of M_{t+1} and S_{t+1}. Therefore, we only need to define an upper control limit for the EWMA statistic S_t. It will be alerted if the control limit h is exceeded. The performance of the EWMA statistic S_t in simulation for both decreasing and increasing parameter shifts is shown in Section 4. However, this rejection region is too conservative, and the performance is not as good when the new sample changes slightly. In the following section, we will find a more appropriate control limit by simulation [24].

3.2. Searching algorithm for control limit h

The control limit h of EWMA with smoothing constant λ is adjusted to have an IC values of ARL_0. Following [23,24], MC simulation and the bisection method are used to calculate the values of the control limit h under various parameters. The detailed procedure is in Algorithm 1, where ARL_0 is the average run length under IC process, λ is the smoothing constant in EWMA, h_{min} and h_{max} are the initial values of the bisection method.

4. Simulation study

All the results of the simulation experiments are implemented in R 4.1.0. The package mainly used is 'ergm' [19] with version of 4.0 [21].

4.1. Datasets and model selection

The social networks of the three data sets Flobusiness, Kapferer and Faux.mesa.high contained in the 'ergm' package are shown in Figure 2.

Taking Flobusiness as an example, the data come from the business contacts between Renaissance Florentine families. Each vertex (family) in Figure 2(a) has three attributes: (1) wealth: family wealth level; (2) priorates: seats in the civic committee; (3) totalities: the total number of business contacts and marriages.

With the data set Flobusiness, the ERGM of Equation (1) is fitted. Edges of the network are always a term of the model when choosing statistics. In addition, the three statistics gwesp(0.7), gwdsp(0.7) and gwdegree(0.7) are chosen as the statistics, as well as the three

Algorithm 1 Search for control limit h.

Require: ARL_0, λ, h_{min} *and* h_{max}
Ensure: h
1: Set B, $\epsilon_1,\ \epsilon_2$;
2: Let $h = (h_{min} + h_{max})/2$;
3: **for** each $b \in [1, B]$ **do**
4: Set $M_1 = 1, S_1 = 1$, generate sample $Y_1 \sim P_{\theta^\star}(Y)$;
5: $h = (h_{min} + h_{max})/2, t = 1$;
6: **while** $S_t \leq h$ **do**
7: $t \leftarrow t + 1$;
8: Generate sample $Y_t \sim P_{\theta^\star}(Y)$;
9: Calculate statistics M_t, S_t;
10: Update parameters $\hat{\theta}_{1,t}$;
11: **end while**
12: $RL_b \leftarrow t$
13: **end for**
14: $ARL \leftarrow \frac{1}{B}\sum_{b=1}^{B} RL_b$
15: **if** $ARL < ARL_0 - \epsilon_1$ and $h_{max} - h_{min} > \epsilon_2$ **then**
16: Update $h_{min} \leftarrow h$;
17: Update $h = (h_{min} + h_{max})/2$;
18: Turn to line 3;
19: **end if**
20: **if** $ARL > ARL_0 + \epsilon_1$ and $h_{max} - h_{min} > \epsilon_2$ **then**
21: Update $h_{max} \leftarrow h$;
22: Update $h = (h_{min} + h_{max})/2$;
23: Turn to line 3;
24: **end if**
25: The algorithm will stop at $|ARL - ARL_0|$ or $h_{max} - h_{min}$ being small enough.
26: Output h.

vertex attributes wealth, priorates and totalities. Since there are at least one term, combinations of edges and the additional six statistics yield a total of $2^7 - 1 = 127$ models. For each model, the AIC value [17] is obtained using MCMC MLE [33]. Table 1 lists the selected statistics and estimations with the first five of the smallest AIC values.

According to the results in Table 1, edges and gwesp(0.7) are selected as the statistics of the ERGM. The results of MCMC MLE are displayed in Table 2. Then the final model is

$$P(Y) = \exp\{-2.9304 \times edges + 0.8600 \times gwesp(0.7) - \phi(-2.9304, 0.8600)\}. \qquad (10)$$

A large number of network data can be simulated using the given model (10) and Metropolis–Hastings algorithm, which is used to sample exponential random graphs [36]. Figure 3 shows the goodness-of-fit diagnostics, which depicts a comparison between the sample distribution of many typical network data statistics and the real observation [17]. The following statistics were chosen: degree distribution, edge-wise shared partners, minimum geometric distance, and the two statistics used in the model: edges and gwesp(0.7). The

Table 1. Model selection of ERGM for data set Flobusiness.

Models	Statistic	Estimation	Significance	AIC
1	*edges*	−2.9304	***	84.59
	gwesp(0.7)	0.8600	***	
2	*edges*	−2.9178	***	86.26
	gwesp(0.7)	0.9025	***	
	priorates	−0.7179		
3	*edges*	−3.0250	***	86.36
	gwesp(0.7)	0.8884	***	
	totalties	0.6209		
4	*edges*	−3.1164	*	86.61
	gwesp(0.7)	0.9276	*	
	gwdegree(0.7)	0.1391		
5	*edges*	−2.8582	***	86.89
	gwesp(0.7)	0.8096	*	
	gwdsp(0.7)	0.0042		

Table 2. MCMC results of the best model of ERGM for data set Flobusiness.

Statistic	Estimate	Std. Error	*z* value	Pr(> |*z*|)
edges	−2.9304	0.4231	−6.926	< 1e-04
gwesp(0.7)	0.8600	0.2521	3.412	0.000645

Table 3. Information and percentiles of $-2\log(SL)$ of the three datasets.

	Flobusiness	Kapferer	Faux.mesa.high
vertex	16	39	205
edges	15	158	203
5% percentile	−1.9177	−1.9077	−40.2981
25% percentile	0.0253	0.4776	0.5431
50% percentile	1.3432	2.1708	3.0410
75% percentile	3.2073	5.4429	56.7055
95% percentile	8.6733	17.4505	400.0304

sample distributions of the statistics acquired from the generated data are shown in the box plots in Figure 3, and the real line in the center represents the true statistics of the network for Flobusiness. The real statistics are essentially close to the mean value of the box plots, showing that the model is well-fitted to the data.

4.2. Sample split LRT performance

Flobusiness and the other two datasets, Kapferer and Faux.mesa.high, were used to fit the model

$$P(Y) = \exp\{\theta_1 \times edges + \theta_2 \times gwesp(0.7) - \phi(\theta_1, \theta_2)\}$$

and 3000 samples of the statistics SL in Equation (3) are generated separately. Table 3 displays the number of vertices and edges, and some percentiles for the three datasets.

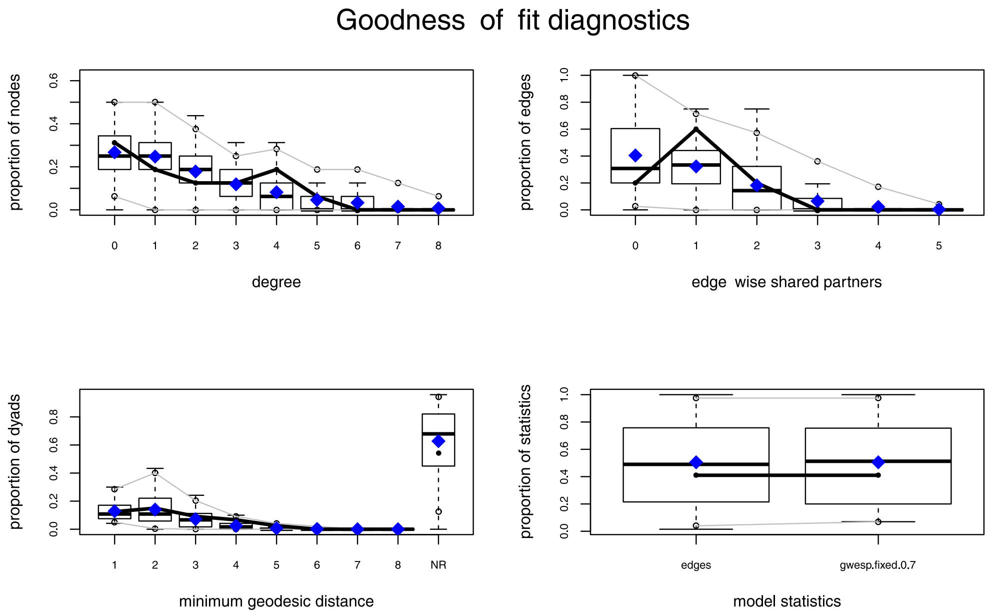

Figure 3. Goodness-of-fit diagnostics of Model (10).

The sample distribution is depicted in Figure 4, with chisq(1) serving as the benchmark, showing a chi-square distribution with 1 degree of freedom. It can be seen from Figure 4 that there are some discrepancies between the distributions.

The sample distribution of $-2\log(SL)$ is presented in Figure 5, with reference to the form of likelihood ratio statistics. It is clear from Figure 5 that they do not follow the chi-square distribution, which is different from the conventional likelihood ratio statistics.

The percentiles of the sample distribution in Figure 5 are also shown in Table 3. It can be seen that as the number of network edges and vertices increases, the distribution of the statistics becomes more concentrated to zero and the tail becomes more heavier. There are also some large numbers of SL for the dataset Faux.mesa.high, indicating that statistical properties do not perform well when the network order is high. The explanation for this might be that when the order is huge, fitting the model becomes more difficult, and the parameter estimation becomes worse. When the order is high, however, even if the network topology or parameters change slightly, the model's likelihood changes dramatically, and the statistic SL is more susceptible to extreme instances.

To produce simulation data, we use the data set Flobusiness as an example. Set the parameter shift to δ times the standard deviation (see Table 2) of estimation, and use the model (10). That is,

$$P(Y) = \frac{\exp\{(-2.9304 + 0.4231\delta) \times edges + (0.8600 + 0.2521\delta) \times gwesp(0.7)\}}{c(-2.9304 + 0.4231\delta, 0.8600 + 0.2521\delta)}. \quad (11)$$

For the significant level α, the rejection region is $\{SL|SL > 1/\alpha\}$. The simulated power of the test in Equation (4) is shown in Figure 8. In this simulation, δ varies from 0.05 to 0.7 in steps of 0.005. At each of these 131 δs, 150 simulations were run. Two networks are

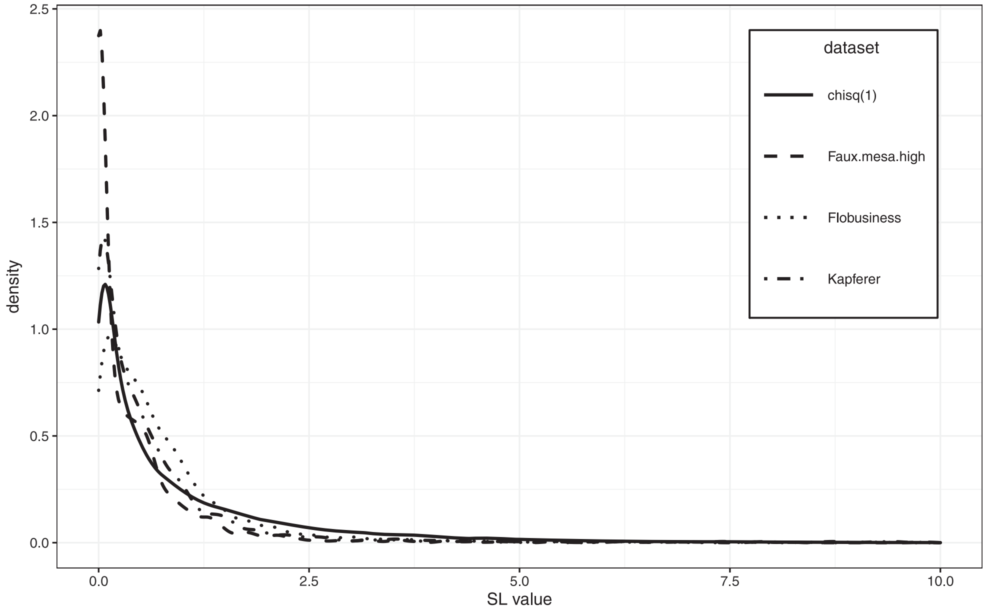

Figure 4. The sample distributions of the statistics *SL* for the three datasets.

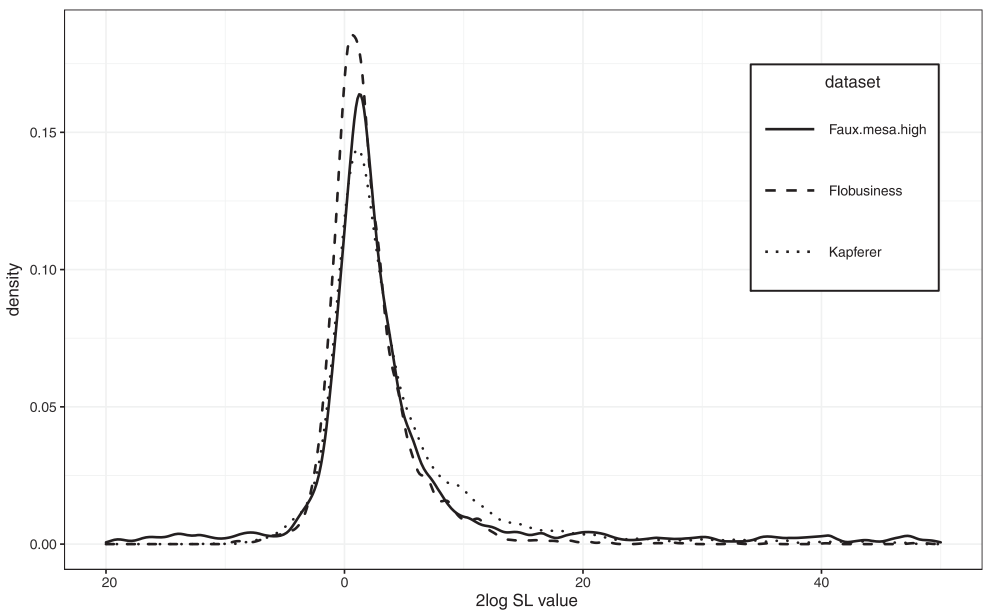

Figure 5. The sample distributions of the statistics $-2\log(SL)$ for the three datasets.

generated following the model (11) and the results of the test are record in every simulation. The solid lines in Figure 8 are the smoothing lines of the corresponding points.

Based on the same simulation data, the bootstrap method is used to calculate the 25%, 50%, 75%, 95% percentiles for the 150 *SL* samples at each δ, as illustrated in Figure 7.

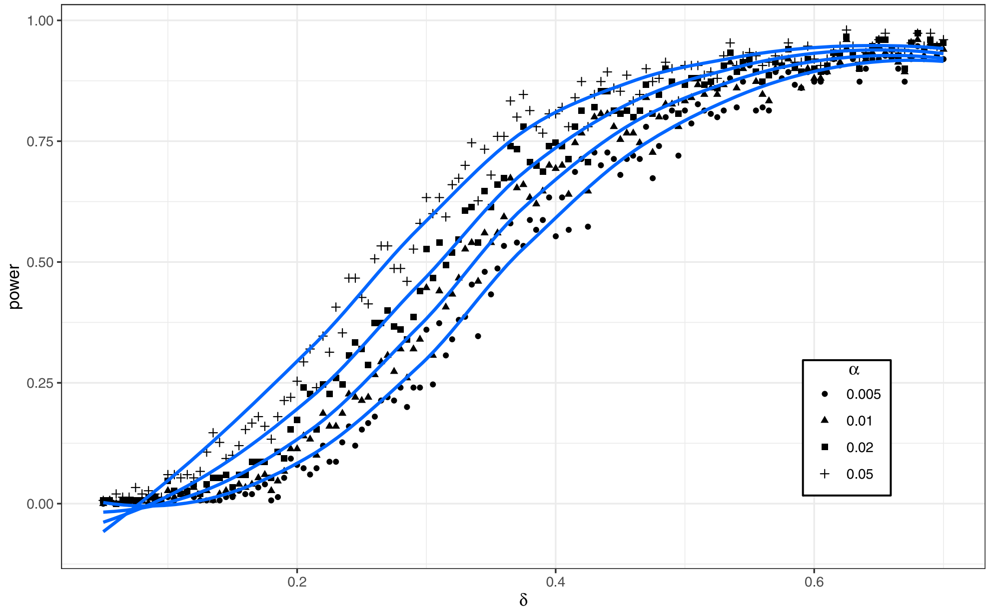

Figure 6. Percentiles of $-2\log(SL)$ changes with δ.

Table 4. The control limits h for various combinations of λ and ARL_0 of Split LRT based EWMA control chart.

	λ			
ARL_0	1	0.2	0.1	0.05
100	5.2140	1.6875	1.3010	1.0640
200	7.9347	2.3924	1.6084	1.2595

The trend of the sample percentiles of $-2\log(SL)$ with δ is shown in Figure 6 and the solid lines are smoothing lines of the corresponding points. As shown in Figures 8–6, the statistic SL grows exponentially with δ. When the shift is greater than 0.5 times the standard deviation, the power is nearly 1, indicating that the shift can be detected even if the control limit is set to $1/\alpha$. As a result, we will discuss the case when the shift is less than 0.5 times the standard deviation in the following section.

4.3. Control limit h and ARL performance

In this section, we only consider the situation when $\theta^\star$ is known. When $\theta^\star$ is unknown, we can replace $\theta^\star$ with the estimated $\hat{\theta}^\star$ obtained in Phase I. Set model in Equation (10) as the null hypothesis, with $\theta^\star = (-2.9304, 0.8600)$. We consider the performance of EWMA statistic in Equation (6).

At first, we set the smoothing parameters $\lambda = 1, 0.2, 0.1, 0.05$ and $ARL_0 = 100, 200$, and Algorithm 1 is used to get the control limit h. The results are shown in Table 4.

The first experiment is to show the performance of ARL with increasing shift. For the parameters $\lambda = 1, 0.2, 0.1, 0.05$, $ARL_0 = 100, 200$ and the change point $\tau = 0, 20, 50$, the corresponding control limits h in Table 4 are set. The ARL results are displayed in Table 5.

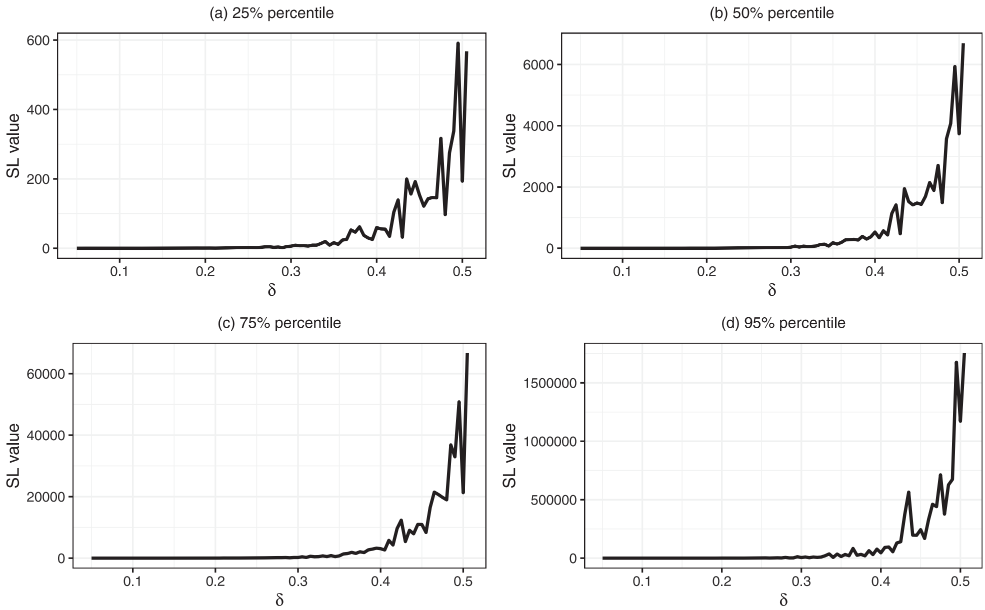

Figure 7. 25%, 50%, 75% and 95% percentiles of *SL* changes with δ.

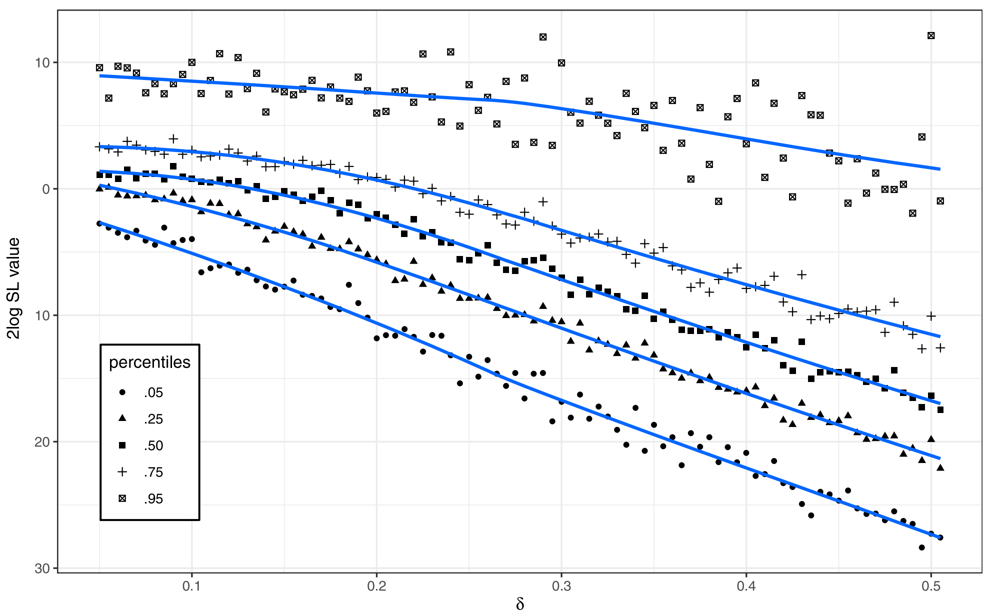

Figure 8. The power of the test in Equation (4) varies with δ.

The table shows that as δ increases, the ARL at each λ value drops. The ARL values fluctuate with δ under these four λ, as shown in Figure 9. When $\lambda = 1$, the ARL values of $\tau = 20, 50$ are nearly identical. And when $\lambda = 0.2, 0.1, 0.05$, the bigger the τ, the smaller the ARL, showing that the monitoring method is more effective when the length of the IC process is longer.

The next experiment is to show the performance of ARL under decreasing shift. We consider negative δ, i.e. $\delta = -0.2, -0.3, -0.5$. We fix $\tau = 0$, as the effect of τ are the same as before. The results are displayed in Table 6. We can conclude that the proposed monitoring method still work well when decreasing shift occurs.

To compare the performance of the proposed control chart with the methods discussed in [35], we introduce Hotelling's T^2 and MEWMA control charts here. Hotelling's T^2 is also known as multivariate Shewhart control chart. Let $\hat{\theta}_i$ be a new sample, and Σ be the covariance matrix of $\hat{\theta}_i$. The Hotelling T^2 statistic is

$$T_i^2 = (\hat{\theta}_i - \theta^\star)^\top \Sigma^{-1} (\hat{\theta}_i - \theta^\star).$$

And the MEWMA statistic is

$$T_i^2 = (z_i - \theta^\star)^\top \Sigma_{z_i}^{-1} (z_i - \theta^\star),$$

where λ is the smoothing constant, $z_i = \lambda \hat{\theta}_i + (1 - \lambda) z_{i-1}$, $z_0 = \theta^\star$ and $\Sigma_{z_i} \approx \frac{\lambda}{2-\lambda} \Sigma$. We can replace Σ by sample covariance matrix in simulation. The upper control limit can also be computed by Algorithm 1. The results are shown in Table 7 for $ARL_0 = 100, 200$, respectively.

The ARL performances are shown in Table 8. Compared with Table 5, it is obvious that the performances of proposed method are uniformly better than MEWMA and T^2 control charts.

5. Real application example

The real dataset to illustrate the proposed modelling and on-line monitoring methods is the MIT reality mining social proximity network. The dataset contains call records and other information of students and staff at a major university during the months between July 2004 and June 2005. The website of the dataset is http://realitycommons.media.mit.edu/realitymining.html. And more detailed information on the data can be found in [8].

After cleaning and preprocessing, the dataset contains 80 subjects (nodes) during the months from 2004-07-20 to 2005-06-14. To be comparable with other studies [20,31], the datasets were split up into weekly time windows. Finally, according to the time sequence, we get 48 undirected networks (observations), and each network has 80 nodes. Each edge in the graph indicates that two nodes are connected at least once in a week. We assume that networks at different weeks are independent.

For the selection of models, we still use pseudo MLE parameters [6] and AIC minimum criterion [17] to determine the model. Therefore, gwdsp(0.7) and gwdegree(0.7) are selected as the sufficient statistics of the ERGM. The AIC of this model is approximately equal to 77.74. For the monitoring process, we choose the first two observations as IC samples and the mean of their pseudo MLE is set as IC parameters, $\hat{\theta}_m^\star = (-0.2392, -3.7467)$. The parameter λ is set to 1. When the monitoring statistic SL is greater than 20 (α=5%), the control charts will send OC signal. Figure 10 shows the OC signals against the selected metrics of networks. These three metrics are centralized and standardized.

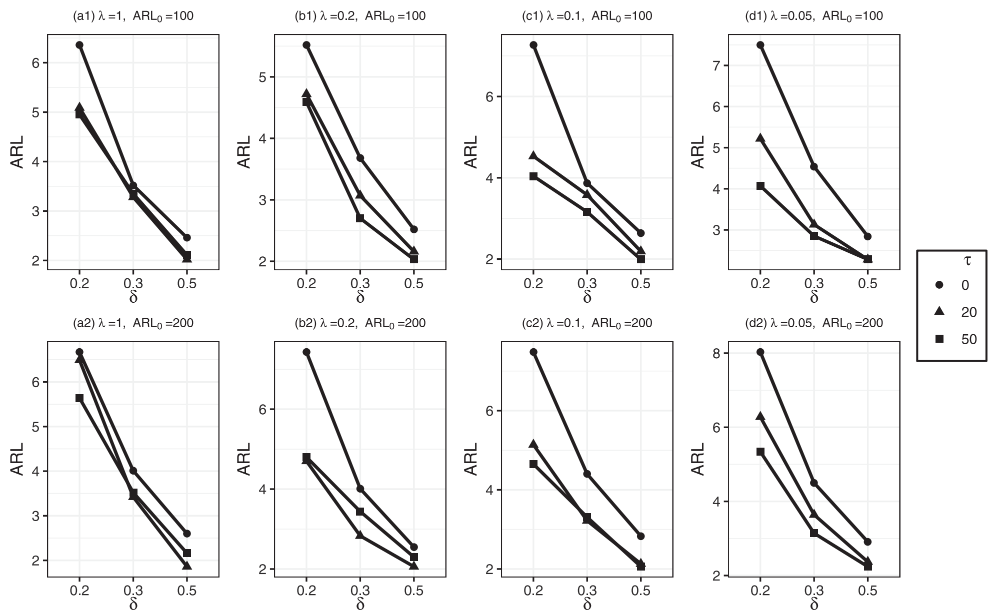

Figure 9. ARL varies with δ and τ when $\lambda = 1, 0.2, 0.1, 0.05$ and $ARL_0 = 100, 200$.

Table 5. ARL values for different λ, δ, τ and ARL_0 of Split LRT-based EWMA control chart.

		$ARL_0 = 100$				$ARL_0 = 200$			
		λ							
τ	δ	1	0.2	0.1	0.05	1	0.2	0.1	0.05
0	0	102.50	99.11	101.10	99.25	200.85	202.34	199.51	198.64
	0.2	6.36	5.52	7.27	7.50	6.67	7.43	7.48	8.03
	0.3	3.52	3.68	3.87	4.54	4.01	4.01	4.40	4.50
	0.5	2.46	2.52	2.64	2.84	2.60	2.55	2.83	2.91
20	0	102.50	99.11	101.10	99.25	200.85	202.34	199.51	198.64
	0.2	5.09	4.72	4.53	5.22	6.49	4.70	5.14	6.28
	0.3	3.28	3.07	3.58	3.13	3.42	2.83	3.22	3.64
	0.5	2.02	2.16	2.19	2.28	1.86	2.06	2.13	2.37
50	0	102.50	99.11	101.10	99.25	200.85	202.34	199.51	198.64
	0.2	4.95	4.59	4.04	4.07	5.64	4.80	4.65	5.34
	0.3	3.35	2.70	3.16	2.85	3.52	3.44	3.31	3.14
	0.5	2.11	2.03	1.99	2.28	2.16	2.30	2.06	2.24

From Figure 10, in this application example, the charts give OC signals at the 11th, 20th, 24th, 28th, 43rd observations, respectively. The dates corresponding to the observations are 2004/09/28-2004/10/04, 2004/11/30-2004/12/06, 2004/12/28-2005/01/03, 2005/01/25-2005/01/31 and 2005/05/10-2005/05/16, respectively. Consistent with [31], 2004/11/30-2004/12/06 is the last week of classes and the independent activities were held from 2004/12/28 to 2005/01/31, so the OC signals are given at the 20th, 24th, 28th observation. Moreover, we found that 2004/09/06 was the start of the semester, which may lead to the OC signal at 2004/09/28-2004/10/04. And the OC signal at 2005/05/10-2005/05/16 means the end of this semester.

Table 6. ARL values for decreasing δ of Split LRT-based EWMA control chart.

		$ARL_0 = 100$				$ARL_0 = 200$			
		λ							
τ	δ	1	0.2	0.1	0.05	1	0.2	0.1	0.05
0	0	102.50	99.11	101.10	99.25	200.85	202.34	199.51	198.64
	−0.2	3.51	4.91	7.13	9.24	9.60	7.40	7.98	11.58
	−0.3	2.77	4.03	5.35	7.70	9.90	5.62	6.52	9.61
	−0.5	2.03	2.99	4.33	5.95	6.30	4.22	4.73	7.23

Table 7. The control limits h for various combinations of λ and ARL_0 of MEWMA control chart.

	λ			
ARL_0	1	0.2	0.1	0.05
100	14.25	17.45	17.96	18.18
200	19.40	21.85	23.53	24.87

Table 8. ARL values for different λ, δ and ARL_0 of MEWMA control chart.

		$ARL_0 = 100$				$ARL_0 = 200$			
		λ							
τ	δ	1	0.2	0.1	0.05	1	0.2	0.1	0.05
0	0	99.77	97.19	101.34	97.49	200.65	202.25	202.77	200.48
	0.2	27.35	58.99	91.58	82.53	43.34	108.24	175.66	174.11
	0.3	16.06	29.80	41.37	69.21	22.49	50.89	69.18	104.94
	0.5	6.98	12.72	24.25	32.72	8.81	19.19	34.87	44.59

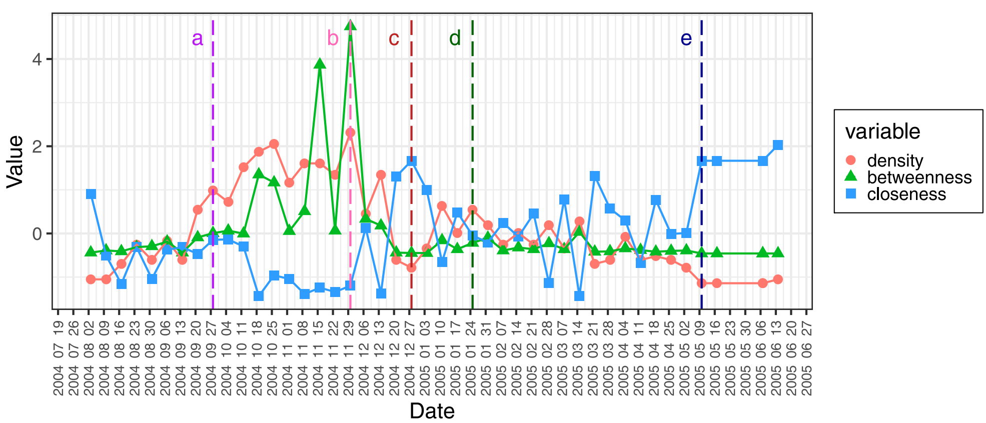

Figure 10. Density, average betweenness and average closeness of the networks for the MIT data. Vertical dotted lines are indicated the OC signals, including (a) the start of the semester, (b) the last week of classes, (c) the beginning of independent activities, (d) the ending of independent activities, and (e) the end of the semester.

6. Concluding remarks

This paper proposes methods to detect anomalous changes in social network structure in real time. First, we model social networks by ERGMs. Then by merging sample Split LRT with SPC, this paper proposes a testing and online monitoring technique. The statistic M_t utilized in the approach is sensitive to parameter changes. According to the simulation studies, the performance is quite good when the parameter shift is not so small (0.5 times standard deviation or above). For tiny shifts, EWMA control charts and simulation search are proposed to find the control limit. Three data sets Flobusiness, Kapferer and Faux.mesa.high are used in the numerical simulations, and one data on the MIT reality mining social proximity network is used to illustrate the proposed modelling and online monitoring methods.

However, a limitation is that ERGMs use summary statistics rather than the entire information of network structure such as adjacency matrix. Future research may find a way to model the network data and explain the complicated associations behind it, reducing information loss of network structure.

Although we focus on Phase II monitoring, the Phase I sample size m may cause estimation error for parameter estimates. It is worth exploring how to get a stream of IC networks for Phase I, and the relationship between the size m and the performance of our monitoring method in Phase II. And, in monitoring process, we assume the independence of social networks at different time. To reflect the evolution of social networks more accurate, we need to develop some new methods to model and monitor the temporal dynamic structures of the networks, which is an interesting area of future research.

The characteristics of the statistic SL require more investigation to reach more concentrated results than $P_{\theta^\star}(SL > 1/\alpha) \leq \alpha$. [7] proved and characterized the degeneracy observed in the ERGM with the counts of edges and triangles as the exclusively sufficient statistics. When the network order is high, however, the statistic SL is more sensitive to network structure changes and is more prone to extreme big values. As a result, the approach needs to be enhanced for high-order network data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The authors are grateful to the editor, the associate editor and three anonymous referees for their comments that have greatly improved this paper. This research was supported by National Key R &D Program of China [grant 2022ZD0114801], National Natural Science Foundation of China [grants 12071233, 12075162 and 11971247], Natural Science Foundation of Fujian Province, China [grant 2022J02050] VC&VR Key Lab of Sichuan Province.

ORCID

Zhonghua Li http://orcid.org/0000-0002-0927-226X

References

[1] S. Abossedgh, A. Saghaei, and A. Amiri, *Enhancing the performance of monitoring the DCSBM using multivariate control charts with estimated parameters*, Adv. Ind. Eng. 56 (2022), pp. 73–86.

[2] S. Abossedgh, A. Saghaei, and A. Amiri, *Monitoring a labeled degree corrected stochastic block model*, Qual. Rel. Eng. Int. 39 (2023), pp. 99–112.

[3] C.J. Anderson, S. Wasserman, and B. Crouch, *A p* primer: logit models for social networks*, Soc. Netw. 21 (1999), pp. 37–66.

[4] M.P. Arrieta-Paredes and B. Cronin, *Exponential random graph models for management research: A case study of executive recruitment*, Eur. Manag. J. 35 (2017), pp. 373–382.
[5] I. Barnett and J.P. Onnela, *Change point detection in correlation networks*, Sci. Rep. 6 (2016), pp. 1–11.
[6] J. Besag, *Statistical analysis of non-lattice data*, J. R. Stat. Soc. Ser. D. Stat. 24 (1975), pp. 179–195.
[7] S. Chatterjee and P. Diaconis, *Estimating and understanding exponential random graph models*, Ann. Statist. 41 (2013), pp. 2428–2461 MR3127871.
[8] N. Eagle, A. Pentland, and D. Lazer, *Inferring friendship network structure by using mobile phone data*, Proc. Natl. Acad. Sci. USA 106 (2009), pp. 15274–15278.
[9] S. Ebrahimi, M. Reisi-Gahrooei, K. Paynabar, and S. Mankad, *Monitoring sparse and attributed networks with online hurdle models*, IISE Trans. 54 (2021), pp. 91–104.
[10] O. Frank and D. Strauss, *Markov graphs*, J. Amer. Statist. Assoc. 81 (1986), pp. 832–842 MR0860518.
[11] C.J. Geyer and E.A. Thompson, *Constrained monte carlo maximum likelihood for dependent data*, J. R. Stat. Soc. Ser. B. Stat. Methodol. 54 (1992), pp. 657–683 MR1185217.
[12] A. Gelman and X.-L. Meng, *Simulating normalizing constants: from importance sampling to bridge sampling to path sampling*, Statist. Sci. 13 (1998), pp. 163–185 MR1647507.
[13] P.J. Harrison, *Statistical process control and model monitoring*, J. Appl. Stat. 26 (1999), pp. 273–292.
[14] P.W. Holland and S. Leinhardt, *An exponential family of probability distributions for directed graphs*, J. Amer. Statist. Assoc. 76 (1981), pp. 33–50 MR0608176.
[15] S.R. Howard, A. Ramdas, J. McAuliffe, and J. Sekhon, *Time-uniform Chernoff bounds via nonnegative supermartingales*, Probab. Surv. 17 (2020), pp. 257–317 MR4100718.
[16] D.R. Hunter, *Curved exponential family models for social networks*, Soc. Netw. 29 (2007), pp. 216–230.
[17] D.R. Hunter, S.M. Goodreau, and M.S. Handcock, *Goodness of fit of social network models*, J. Amer. Statist. Assoc. 103 (2008), pp. 248–258 MR2394635.
[18] D.R. Hunter and M.S. Handcock, *Inference in curved exponential family models for networks*, J. Comput. Graph. Statist. 15 (2006), pp. 565–583 MR2291264.
[19] D.R. Hunter, M.S. Handcock, C.T. Butts, S.M. Goodreau, and M. Morris, *ergm: A package to fit, simulate and diagnose exponential-family models for networks*, J. Stat. Softw. 24 (2008), pp. 1–29.
[20] L. Kendrick, K. Musial, and B. Gabrys, *Change point detection in social networks critical review with experiments*, Comput. Sci. Rev. 29 (2018), pp. 1–13 MR3847498.
[21] P.N. Krivitsky, D.R. Hunter, M. Morris, and C. Klumb, *ergm 4.0: New features and improvements*, preprint (2021). Available at https://arxiv.org/pdf/2106.04997.pdf.
[22] S.Y. Li, L.C. Tang, and S.H. Ng, *Nonparametric cusum and ewma control charts for detecting mean shifts*, J. Qual. Technol. 42 (2010), pp. 209–226.
[23] Z. Li, M. Xie, and M. Zhou, *Rank-based EWMA procedure for sequentially detecting changes of process location and variability*, Qual. Technol. Quant. M. 15 (2018), pp. 354–373.
[24] Z. Li, C. Zou, Z. Gong, and Z. Wang, *The computation of average run length and average time to signal: an overview*, J. Stat. Comput. Simul. 84 (2014), pp. 1779–1802 MR3215723.
[25] H. Mogouie, G.A.R. Ardali, A. Amiri, and E.B. Samani, *Monitoring attributed social networks based on count data and random effects*, Sci. Iran. 29 (2022), pp. 1581–1591.
[26] M. Mohebbi, A. Amiri, and A.R. Taheriyoun, *Monitoring the structure of social networks based on exponential random graph model*, Comm. Statist. Theory Methods (2023). doi: 10.1080/03610926.2022.2163366.
[27] D.C. Montgomery, *Introduction to Statistical Quality Control*, John Wiley & Sons, 2020.
[28] M. Mostafapour, F. Movahedi Sobhani, and A. Saghaei, *Monitoring sparse and attributed network streams with multilevel and dynamic structures*, Mathematics 10 (2022), pp. 4483.
[29] N. Motalebi, M.S. Owlia, A. Amiri, and M.S. Fallahnezhad, *Monitoring social networks based on zero-inflated poisson regression model*, Comm. Statist. Theory Methods 52 (2023), pp. 2099–2115.

[30] R. Noorossana, S.S. Hosseini, and A. Heydarzade, *An overview of dynamic anomaly detection in social networks via control charts*, Qual. Reliab. Eng. Int. 34 (2018), pp. 641–648.
[31] L. Peel and A. Clauset, *Detecting change points in the large-scale structure of evolving networks*, in: Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015.
[32] D. Qi, Z. Wang, X. Zi, and Z. Li, *Phase ii monitoring of generalized linear profiles using weighted likelihood ratio charts*, Comput. Ind. Eng. 94 (2016), pp. 178–187.
[33] G. Robins, P. Pattison, Y. Kalish, and D. Lusher, *An introduction to exponential random graph (p*) models for social networks*, Soc. Netw. 29 (2007), pp. 173–191.
[34] C.R. Shalizi and A. Rinaldo, *Consistency under sampling of exponential random graph models*, Ann. Statist. 41 (2013), pp. 508–535 MR3099112.
[35] S.G. Sharifnia and A. Saghaei, *A statistical approach for social network change detection: an ergm based framework*, Comm. Stat. Theory Methods 51 (2022), pp. 2259–2280 MR4391565.
[36] T.A. Snijders, *Markov chain monte carlo estimation of exponential random graph models*, J. Soc. Struct. 3 (2002), pp. 1–40.
[37] T.A. Snijders, P.E. Pattison, G.L. Robins, and M.S. Handcock, *New specifications for exponential random graph models*, Sociol. Methodol. 36 (2006), pp. 99–153.
[38] A.R. Torres, A.D.P. de Oliveira, S. Grangeiro, and W.D. Fragoso, *Multivariate statistical process control in annual pharmaceutical product review*, J. Process Control 69 (2018), pp. 97–102.
[39] L. Wasserman, A. Ramdas, and S. Balakrishnan, *Universal inference*, Proc. Natl. Acad. Sci. USA 117 (2020), pp. 16880–16890 MR4242731.
[40] W.H. Woodall, M.J. Zhao, K. Paynabar, R. Sparks, and J.D. Wilson, *An overview and perspective on social network monitoring*, IISE Trans. 49 (2017), pp. 354–365.
[41] J. Yue and L. Liu, *Multivariate nonparametric control chart with variable sampling interval*, Appl. Math. Model. 52 (2017), pp. 603–612 MR3716312.
[42] C. Zou and F. Tsung, *Likelihood ratio-based distribution-free ewma control charts*, J. Qual. Technol. 42 (2010), pp. 174–196.