Journal of Optimization in Industrial Engineering
Vol.13, Issue 1, Winter & Spring 2020, 131-143
DOI:10.22094/JOIE.2019.581174.1615

Research Article

Monitoring of Social Network and Change Detection by Applying Statistical Process: ERGM

Farshid Rajabi[a], Abbas Saghaei[a,*], Soheil Sadi-Nezhad[a]

[a] *Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran*

Received 15 December 2018; Revised 03 September 2019; Accepted 17 September 2019

Abstract

The statistical modeling of social network data needs much effort because of the complex dependence structure of the tie variables. In order to formulate such dependences, the statistical exponential families of distributions can provide a flexible structure. In this regard, the statistical characteristics of the network is provided to be encapsulated within an Exponential Random Graph Model (ERGM). Applying the ERGM, in this paper, we follow to design a statistical process control through network behavior. The results demonstrated the superiority of the designed chart over the existing change detection methods in controlling the states. Additionally, the detection process is formulated for the social networks and the results are statistically analyzed.

***Keywords*:** Statistical Process Control; ERGM, Social network; Change detection

1. Introduction

Social Network Analysis (SNA) has become an important analytical tool for the analysis of terrorist networks, dating, commerce, biological wars, disease spread, and so on. Tracking dynamic changes over time from the SNA perspective, one can detect signals of the change in the organization, and may even be able to predict important events or analyze behaviors. One of the most important issues that can be considered as a challenge or topic in social networks is to estimate the parameters and also monitor the changes. Since operations and interactions are so important in the networks and the outputs of these networks are used in important areas, such as manufacturing, distribution and supply chains, telecommunication, computer, biological, and social networks, finding a framework and structure for predicting the changes, which can be used in developing the action plan, is the key concern .

For a better description of the problem, assume that a network has a number of arcs and nodes and its parameters change over time. For example, the structure of the network changes over time or the relationships between the nodes experience a change. In this case, it is necessary to investigate such dynamic networks thoroughly, so that appropriate methods or even control charts can be proposed for their monitoring. In networks, there are four types of changes (as will be discussed in greater detail in the future): stability, evolution, shock, and jumps. One way of dealing with the subject is to define the problem variables in the form of criteria or measurements or ideal sizes in social networks including mean, maximum, minimum, standard deviation, closeness centrality, betweenness centrality, and eigenvector centrality, and monitor these criteria either individually (one variable) or together (multivariate), and provide each one with diagrams and instructions.

Many metrics exist to describe the structural features of an observed network such as the density, centrality, or assortativity. However, these metrics describe the observed network which is only one instance of a large number of possible alternative networks. This set of alternative networks may have similar or dissimilar structural features. To support statistical inference on the processes affecting the formation of network structure, a statistical model should consider the set of all possible alternative networks weighted on their similarity to an observed network. However, because network data is inherently relational, it violates the assumptions of independence and identical distribution of standard statistical models like linear regression. Alternative statistical models should reflect the uncertainty associated with a given observation, permit inference about the relative frequency about network substructures of theoretical interest, disambiguate the influence of confounding processes, efficiently represent complex structures, and link local-level processes to global-level properties. Degree-preserving randomization is, for example, a specific way in which an observed network could be considered in terms of multiple alternative networks.

There are two mechanisms for social network analysis: model-oriented approach (based on model) and model-free based approach (independent from the model) (Kevin Joe, 2012). Both are valuable and each one has its own applicability. When the proposed framework is based on collecting information from the nodes and the existing relationships, and then, designing the structure on the

*Corresponding author Email address: abassaghaie@gmail.com

logic defined by the researcher, the method is considered to be model-free based. If during the construction of the network the relevant mathematical models that exist in the literature are used, a model-oriented approach will be used. One of the most efficient mathematical models is ERGM, which is used in this paper. Quality and process control techniques are a statistical approach to detect abnormalities in the behavior of a random process over time. A control chart is a vital quality engineering tool that helps manufacturing companies to maintain profitability (Montgomery, 1991; Ryan, 2000). These approaches, as tools for quality control, are widely used in manufacturing topics. Manufacturing systems have similar experiences about high correlation issues which are common in network relational data. In this research, it is assumed that by using statistical process control, changes in networks with longitudinal data can be quickly tracked with graph-level network measures. In ERGM, it is possible to consider relationship dependencies and define relationships between nodes in case of having a shared node. The examples presented in this study show that identifying social network changes enables the analysts to identify significant changes in size over time. In addition, with a high probability, it is possible to see the time of the change event. This allows us to trace network behavior with the minimum resource allocation to appropriately respond to its consequences. Based on what was mentioned, it is crucial to monitor changes in the time-dependent network, and even to differentiate a change from accidental and random fluctuations. The strength of the proposed method is the statistical approach it uses so that it can be used in a wide range of networks, being capable of defining and identifying the change points in the behavior of the organization.

A prerequisite for statistical inference about social networks is to study the possible structure of the network connections. One can investigate what other features can be defined to be used in the monitoring of the network. Another topic is to discover social complications after identifying the change points to define recovery programs, for which a statistical based framework can be considered. Considering the importance of social network changes, . referred to by Woodall et al, the main challenge or question of this research is to provide a statistical technique that can be used in planning and controlling the future network behaviors. Generally, network monitoring is planning and controlling the future behavior of the network, where any failure in this regard makes change management impossible and can lead to an enormous cost. A good example of this is the terrorist attack to the Twin towers, where if an adequate monitoring of the changes had occured in the al-Qaeda terrorist network, it could have been prevented. A thorough review of the literature reveals the following shortcomings. First, most of the studies have concentrated on the monitoring of points, processes, and functions (profiles) rather than on a network, especially a dynamic network. Secondly, we know that the most accurate methods for monitoring a network are statistical ones that have been used rarely. Taking into account the shared nodes between the arcs and constructing new relationships using these nodes, ERGM-based models can be easily adapted to real networks. Thirdly, the form and the structure of some problems have good compatibility with networks, and due to the importance of network monitoring, a comprehensive research should be conducted in this regard. Finally, regarding the network diagnosis, after the detection of the change points and definition of the recovery plan, there has not been any specific work. This gap can be covered in this research.

As a result, two types of innovation can be imagined for this research: the innovation in the problem itself and the innovation in the problem solving methods. The problem itself is new and attractive regarding the introduction of statistical issues, and given the wide range and extensive use of social networks, there is ample opportunity for future activities in this area. Moreover, here, an ERGM based framework (Durin, 1997) is proposed for understanding longitudinal networks based on the activities in dynamic networks. A method for detecting changes in dynamic networks is introduced. An analytical tool is developed to determine the control limits of the control chart that can be used in monitoring changes in different domains. Exponential random graph models (ERGMs) are a family of statistical models for analyzing data about social and other networks. The Exponential family is a broad family of models for covering many types of data, not just networks. We applied ERGMs to explore the micro-level network structures of emergency management networks and their impact on performance. There is no research in the literature that addressed the statistical processes analysis by ERGMs approach. Therefore, this study aims to design statistical process control; by controlling the ERGM, we can understand the network and performance patterns. This research offers a triangulation method that utilizes both qualitative and quantitative approaches. The final output of this approach is a valid and reliable data collection method that facilitates the collection of both singular attribute and social network information. The data collection method is basically reasonable to apply, and it is time-efficient and simply replicable for further related studies.

2. Literature Review

There has been an increasing interest on the monitoring of social networks. An overview of methods was given in a recent review paper by Savage et al. (2014). They listed different applications of network monitoring including the detection of important and influential network participants, the detection of clandestine organizational structures, and the detection of fraudulent or predatory activity (Woodall et al., 2016). The basic idea in social network monitoring is to detect changes in the behavior of a subset of the individuals in the network. While global changes are typically the easiest one to be detected, significant increases in the communication levels of the entire network, smaller sub-networks, or individuals are often of primary interest in applications. In some cases, however, decreases are of interest (Woodall, 2016). Kendrick discussed several network metrics used for a change point detection problem and conducts an experimental, comparative analysis using the Enron and MIT networks. (LucyKendrick, 2018). Sheng et al. introduced a new community detection algorithm, called Commansor (Community Detection Based on Human Social Behavior), which automatically detects communities in the network by simulating human social behavior (Sheng et al., 2019). There are some monitoring methods including control chart and hypothesis testing methods, Bayesian methods, scan methods, time series models and other approaches. In this research, we will use statistical control chart. Statistical process control (SPC) is a method of quality control applied to monitor and control a process. Monitoring and controlling a process ensures that it operates at its full potential and makes as much conforming product as possible with the least possible waste (Barlow and Irony, 1992).

In recent years, various measures such as the density and centrality explain the structural features of an observed network. However, these metrics describe the observed network which is only one instance of a large number of possible alternative networks. This set of alternative networks may have similar or dissimilar structural features (Hubert L, Schultz, 1976; Robins et al., 2007). To support statistical inference based on the statistical processes affecting the network structure, a statistical model should consider the set of all possible alternative networks, weighted on their similarity to an observed network. However, because network data is inherently relational, it violates the assumptions of independence and identical distribution that underlie standard statistical models like linear regression. Alternative statistical models should reflect the uncertainty associated with a given observation, permit inference about the relative frequency of network substructures, disambiguate the influence of confounding processes, represent efficiently complex structures, and link local-level processes to global-level properties. Degree preserving randomization, for example, is a specific way in which an observed network could be considered in terms of multiple alternative networks (Everett and Borgatti, 2014). For the period 1990–2013, 202 countries were analyzed as nodes linked by in- or out-migration of substantial shares of the sending country's population. The resulting network shows regional, but also "cultural" clustering. Different of ERGMs are used to determine geographic, demographic, economic, religious, linguistic, and historical factors of migration (Windzio, 2018). Azarnoush et al. (2016) modeled a method for change detection of sub-network behavior. The authors developed a model based on the probabilities of connections between the nodes of the network using the logistic regression method, where the sub-network membership and covariant data on the individuals were used as explanatory variables. A likelihood ratio test was proposed to detect changes in that model, which is compatible with each new diagram. Three approaches of this model are static reference approach, in which each new sample graph is compared with the baseline diagram; the dynamic reference approach, in which each incoming diagram is compared to the previous ones; and the dynamic reference sliding windows approach, in which the newest diagram is compared to the most recent p diagram, where p stands for the size of moving windows.

Heard et al. (2010) developed a two-level Bayesian approach for anomaly detection. Its aim was to detect anomalous communication levels between the paired nodes. Once a specific pair is identified, it is used to develop a sub-network which can be analyzed then. Similar to the zero-inflated passion model, this method assumed either a Poisson conditional distribution or a hurdle Poisson conditional distribution for contact counting. In this model, it is used to control limits based on a Bayesian predictive distribution on the contacts between pairs to identify subset anomalous pairs. This method is used to analyze anomalous behavior of the system. Priebe et al. (2005) developed a new method, called scan-based network monitoring scheme, which is used to detect increment in the communication levels based on the size of k^{th} order neighborhood of each node, where k=0,1 or 2. The degree of a node refers to the size of 0^{th} order neighborhood. Statistical methods were conducted for each one of the three metrics of the nodes using a moving window of a particular length to set the baseline mean and standard deviation. A lower bound or a baseline is to set the lower limitation over the system as a measurement of more significant changes. Savage et al. (2014) expressed Pincombe (2005) as a support for a new network monitoring method based on the time series. This model can be conducted over the metrics of any system. Usually, large residuals indicate network variations. This method is widely used for process monitoring in public health, industries, and etc.

Many methods have been expressed in the network analysis literature which are used for detecting changes in network structure or trends over time. Examples include detecting fraudulent accounts, detecting unusual events affecting network trend, and detecting changes in community structure.

An overview of the works done in this field reveals that most of the studies have concentrated on the monitoring of points, processes, and functions (profiles) rather than on a network, especially a dynamic network. Statistical techniques as one of the most accurate methods in monitoring networks have been rarely used. The shape and structure of some problems have a lot of compatibility with the networks, which, due to the importance of their monitoring, requires a comprehensive research in this regard. Moreover, there has been no specific work in network diagnosing, change point detection, and definition of recovery plans. These are why this paper can be considered as a useful research. ERGMs are a family of statistical models for analyzing data on social and other networks. The Exponential family is a broad family of models covering many types of data, not just networks. An ERGM is a model from this family which describes networks. Formally a random graph Y consists of a set of n nodes and m edges $\{Y_{ij}$: i=1,…,n & j=1,…,n$\}$ where $Y_{ij}=1$ if the nodes (i,j) are connected and $Y_{ij}=0$ otherwise. The basic assumption of these models is that the structure in an observed graph y can be explained by any statistics s(y) depending on the observed network and nodal attributes. In this way, it is possible to describe any kind of dependence between the dyadic variables (Robins et al., 2007):

$$P(Y = y|\theta) = \frac{\exp(\theta^T s(y))}{c(\theta)} \tag{1}$$

where θ is a vector of model parameters associated with s(y) and c(θ) is a normalizing constant. These models represent a probability distribution on each possible network on n nodes. However, the size of the set of possible networks for an undirected network (simple graph) of size n is $2^{n(n-1)/2}$. Because the number of possible networks in the set vastly outnumbers the number of parameters which can constrain the model, the ideal probability distribution is the one which maximizes the Gibbs entropy (Robins et al., 2007). In another research, by proposing a temporal extension of ERGMs (TERGM), Block et al. (2017) concluded that the TERGM, in contrast to the ERGM, has no consistent interpretation on tie-level probabilities and also on processes of network change. Numerous social network analysis studies of disasters mostly demonstrate that informal individual and group relations, independent of government assistance and survivors' individual circumstances (such as income, education, the level of loss), play a major part in disaster relief efforts. Hossain and Kuti (2010) proposed a research model for exploring coordination preparedness based on network connectedness. They identified a positive relationship between social network connectedness and coordination within disasters. Schweinberger et al. (2011) have studied the inter-organizational network that emerged in response to the September 11, 2001 attacks by introducing ERG Models which can answer the substantive questions of interest while being as simple and parsimonious as possible. Hossain et al. (2015) proposed an approach to map the collaboration network among emergency management personnel. Then, we use ERGMs to explore the micro-level network structures of emergency management networks and their impact on performance. There is no research works with statistical processes analysis by ERGMs approach in the literature. Therefore, this study aims to design statistical process control in order to control the ERGM to understand network and performance patterns in statistical processes analysis. The proposed approach is based on statistical methods that have been followed. This method includes some stages and we need to understand the problem structure and every subject related to it. This method has three main part, simulation, estimation and detection capability that will be reviewed.

Another researcher found that dense substructures of collaboration increase the level of specificity in the plans in regards to explicating constraints on human activities, but it does not enhance the overall level of specificity (Bodin et al., 2016). To sample a graph *x*, we rely on the well-known procedure Markov Chain Monte Carlo (MCMC), which here consists of generating a sequence of M graphs successively updated through small changes. For large M, the last graph in the sequence is a draw from the target - in our case, the ERGM. In particular, we may use the Metropolis-Hastings algorithm (Chib & Greenberg, 1995; Hastings, 1970, Metropolis et al., 1953; Tierney, 1994) to sample from ERGM.

The updating rule in moving from the current (old) graph to a new graph in the simulation consists of choosing a pair of nodes at random and removing or adding a tie between them according to whether they are already tied. In another paper, the researchers first review the multilevel network data structure and multilevel ERGM specifications. Then they will apply these models to a dataset collected among 265 farmers and their communication network in a rural community in Ethiopia, and provide an interesting description of this farming community（Wang et al., 2015). There are other models for longitudinal network analysis, among them which TERGM and the stochastic actor-oriented model (SAOM, e.g., SIENA) are popular ones. To assess their relative pros and cons, we compare these models theoretically, via simulation, and through a real-data example in order. (Pilip Leifeld (a1) and Skyler J.Cranmer,2018).

The inferential goal is to center the distribution of statistics over the observed network and thus fit a model that we say gives maximal support to the data. We define the distribution as centered on the observed values when the values of the statistics from the distribution are the same as those observed on average. Formally, we want the expected value of the statistics $E_\theta(z(X))$ to be equal to the observed statistics (i.e., $E_\theta(z(X)) = z(x_{obs})$), where x_{obs} is the observed graph. Equivalently $E_\theta(z(X)) - z(x_{obs}) = 0$ is known as the moment equation. Solving the moment equation for θ gives the parameter values that provide maximal support to the data. The maximum likelihood estimator (MLE) of parameter for a given model and observed data is the value of θ that makes the observed data most likely - we want to find the vector θ that makes the probability $P_\theta(x_{obs})$ as large as possible (Crouch, Wasserman, and Trachtenberg (1998); Handcock (2003); snijder (2002)). Skyler J. Cranmer showed that all the three network models outperform standard logit estimates on multiple criteria. That article (Navigating the Range of Statistical Tools for Inferential Network Analysis) introduces political scientists to a class of network techniques beyond simple descriptive measures of network structure, and it helps researchers choose which model to use in their own research（Cranmer et al., 2016). Today, the discussion of networks and how they are monitored is one of the most important issues in controlling the status and one important decision making factor for managers. In customer's perspective, concepts such as quality, cost, and reliability are key and important factors that affect a business environment, which means they improve processes, reduce waste, and etc. Specifically, to maintain or enhance product quality, it is necessary to continuously monitor the system by using related techniques. The monitoring and controlling a product or process is not a new discussion, but it has a lot to do and is a good field to visualize and model it in the network, and especially in the dynamic networks. Different types of dynamic networks in engineering sciences such as computers, electronics, industrial, or medical sciences, such as protein chains or diseases, and even sociology, such as social networks, are becoming more and more up-to-date. In general, there are two types of approaches for monitoring the change in social networks, one based on the model and the other one, regardless of the model, where model-based approach and especially the statistical approach is more valuable. In order to analyze the social network, we develop a methodology using the statistical approach with ERGM, and the results are shown in step 4. The rest of this paper is organized as follows: in Section 2, the proposed methodology is illustrated in details. In Section 3, the results and numerical investigations are analyzed. Finally, in Section 4, the conclusions and further researches are given.

3. Proposed Methodology

The problem of this research is in the presence of ERGM concept. Therefore, social network monitoring by applying SPC charts is investigated by ERGM approach. In order to illustrate the methodology, we provide it in eight main steps as the following subsections.

Identify the network structure (1st Step)

In this step, we survey the observed network to see whether it is compatible with the ERGM network structure. In this way, the number of edges, 2-stars and transitive relationships between the nodes, can create new relationships that ERGM is able to count them. As already mentioned, We consider ERGM network with the following structure and parameters $\theta_{L_0}, \theta_{S_0}, \theta_{T_0}$:

$$P(X=x)=\exp(\theta_L L(x)+\theta_S S(x)+\theta_T T(x)) \tag{2}$$

where

x : a state of the sample network

θ_L = Edge parameter

θ_S = K-star parameter

θ_T = Triangle parameter

$x_{ij} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ if there is a relationship between i and j actors

$L(X)=\Sigma x_{ij}$; The number of arcs (edges)

$S(x)=\Sigma\begin{pmatrix} x_{i+} \\ k \end{pmatrix}, x_{i+}=\Sigma x_{ij}$; The number of k-stars

$T(x)=\Sigma_{i\prec j\prec k\prec l} x_{ij}x_{jk}x_{kl}$; The number of triangles

Control chart design (2nd Step)

Assuming the given parameters, it does not need phase I of control chart design.

Since, based on the assumptions, the network parameters are known, the variance-covariance matrix (Σ) is specified and the control limits can be calculated by the following equation:

$$UCL=\chi^2_{\alpha,P} \tag{3}$$

where P Shows the number of parameters (here is 3).
The above network with known parameters has the following upper limit (UCL):

$$UCL=\chi^2_{\alpha,P} \tag{4}$$

Note: If the parameters are not specified, Fisher distribution function is used. In addition, if parameters are not normal, some techniques such as Box-Cox or bootstrap function and the like can be used for the conversion.

Changing the parameters (3rd Step)

In the below relation, we can apply a change of size δ to the parameters:

we consider $\mu_1=\mu_0+\delta$ where

$$\delta=\sqrt{(\mu_1-\mu_0)\Sigma^{-1}(\mu_1-\mu_0)'} \tag{5}$$

and

$$\mu_0=(\theta_{L_0},\theta_{S_0},\theta_{T_0}),$$
$$\mu_1=(\theta_{L_1},\theta_{S_1},\theta_{T_1}) \tag{6}$$

Network simulation (4rd Step)

For different values δ (0.5, 1, 1.5, 2, and 2.5), we produce ERGM networks with the parameters $\alpha_1, \beta_1, \gamma_1$ obtained from the third step, with a sample size of 10:

$$P(X = x) = \exp(\theta_{L_1} L(x) + \theta_{S_1} S(x) + \theta_{T_1} T(x)) \quad (7)$$

Parameter estimation (5th Step)

For each sample, the statistic T^2 is calculated according to the following equation:

$$T^2 = n(\bar{X} - \mu)' \Sigma^{-1} (\bar{X} - \mu) \quad (8)$$

To be clear, $\bar{X}$ is the estimation vector of $\alpha_1, \beta_1, \gamma_1$, that are the counter of $L(x)$, $S(x)$, and $T(x)$, are calculated for each sample using the MLE, parameter estimation is performed.

where μ is

$$\mu_0 = (\theta_{L_0}, \theta_{S_0}, \theta_{T_0}) . \quad (9)$$

Calculate T^2 statistics (6th Step)

Repeat the steps four and five for 1,000 times.

Define the statistical power (7th Step)

For each iteration, T^2 is compared with the upper limit and the percentage of times that T^2 is more than UCL is calculated).

$$1 - \beta = P(T^2 \succ UCL) \quad (10)$$

Chart capability detection (8th Step)

At this stage, the performance of the graph in detecting the changes induced to the parameters is obtained:

$$ATTS = ARL * H$$
$$ARL = \frac{1}{1-\beta} \quad (11)$$

It should be noted here that because of the continuous intervals, the result ATTS (Average time to signal) is equivalent to ARL(Average run length) ; H shows sampling intervals.

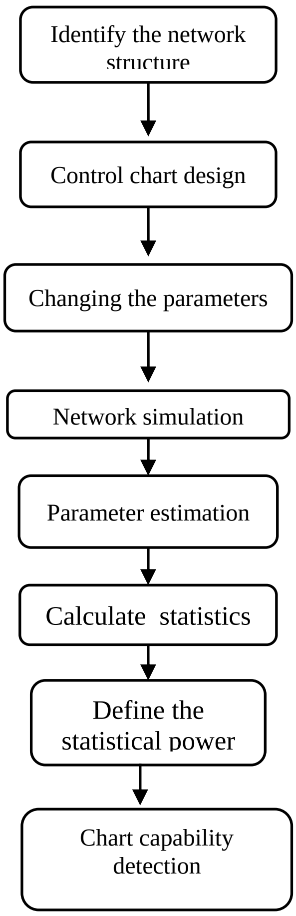

Fig. 1. Framework of the proposed methodology

4. Results

In order to evaluate the performance of the proposed methodology, we executed the methodology in eight steps. The MATLAB Software has been used for implementing the proposed methodology, and the program has been executed on a 2 GHz laptop with eight GB RAM. Tables 1 to 3 show different values of the Hotelling's T2 and EWMA statistics under different variations of the parameters. The EWMA chart has been used for validation testing. For each parameter, the parameters are changed individually or together, the length of the sequence (RL) is calculated, and this procedure is repeated ten times. Then, the average run length (ARL) of the sequence is calculated and their box plot is plotted.

The results show that as the size of change become greater, the power of the test increases, naturally. The proposed method provides a good diagnosis power so that when the size of change moves from 0.5 to 3 units, its detection power increases from 0.05 to about 1.

Based on Tables 1, for example, if the parameter α experiences 0.5 unit change and other parameters remain constant, the Hotelling's T2 chart shows it 95% of the time between samples 10 to 250. This value is better for the EWMA chart. In the case of small changes, the EWMA chart has better performance, but as the changes become larger, this trend turns to the T^2 chart.

Table 1
Designing charts for variable changes α

RL	Statistics											
	$\alpha_1=\alpha_0+0.5\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0$		$\alpha_1=\alpha_0+1\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0$		$\alpha_1=\alpha_0+1.5\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0$		$\alpha_1=\alpha_0+2\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0$		$\alpha_1=\alpha_0+2.5\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0$		$\alpha_1=\alpha_0+3\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0$	
	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA
RL_1	390	133	17	3	12	4	5	2	5	5	7	9
RL_2	41	14	46	8	68	19	23	7	5	5	3	5
RL_3	53	18	43	7	15	6	41	8	6	5	2	3
RL_4	20	7	59	10	4	2	3	9	6	5	2	3
RL_5	272	93	144	24	67	19	5	23	3	2	5	7
RL_6	167	57	105	17	70	21	41	20	11	9	2	3
RL_7	732	249	122	21	72	21	12	19	5	5	3	5
RL_8	66	23	4	1	24	7	4	6	6	5	4	6
RL_9	160	55	117	20	11	4	6	5	3	3	2	4
RL_{10}	40	14	2	1	14	5	19	4	15	12	4	7

Table 2
Designing charts for variable changes α,γ

RL	$\alpha_1=\alpha_0+0.5\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0+0.5\delta$		$\alpha_1=\alpha_0+1\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0+1.5\delta$		$\alpha_1=\alpha_0+1.5\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0+1.5\delta$		$\alpha_1=\alpha_0+2\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0+2\delta$		$\alpha_1=\alpha_0+2.5\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0+2.5\delta$		$\alpha_1=\alpha_0+3\delta$ $\beta_1=\beta_0$ $\gamma_1=\gamma_0+3\delta$	
	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA
RL_1	30	11	19	4	26	8	3	3	3	3	2	3
RL_2	72	26	31	6	7	2	2	1	3	3	2	3
RL_3	56	21	52	9	6	2	5	1	6	7	2	3
RL_4	398	137	7	2	2	1	8	1	2	2	2	3
RL_5	167	58	35	7	7	2	2	2	2	2	2	3
RL_6	17	6	30	6	10	3	5	1	2	2	2	3
RL_7	129	45	2	1	28	8	7	4	3	4	2	3
RL_8	187	65	16	4	43	12	4	3	2	2	2	3
RL_9	128	45	19	4	28	9	4	3	2	2	2	3
RL_{10}	193	67	9	2	16	5	6	3	2	2	2	3

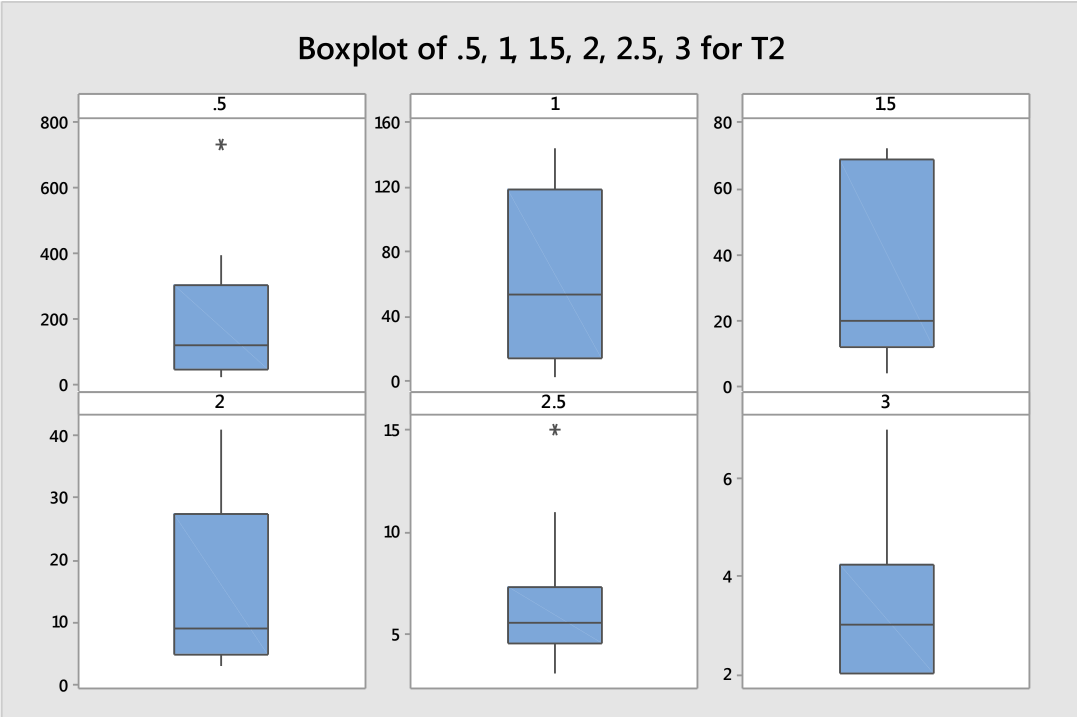

Fig. 1. Box plots for the Hotelling's T^2 control chart corresponding to Table 1

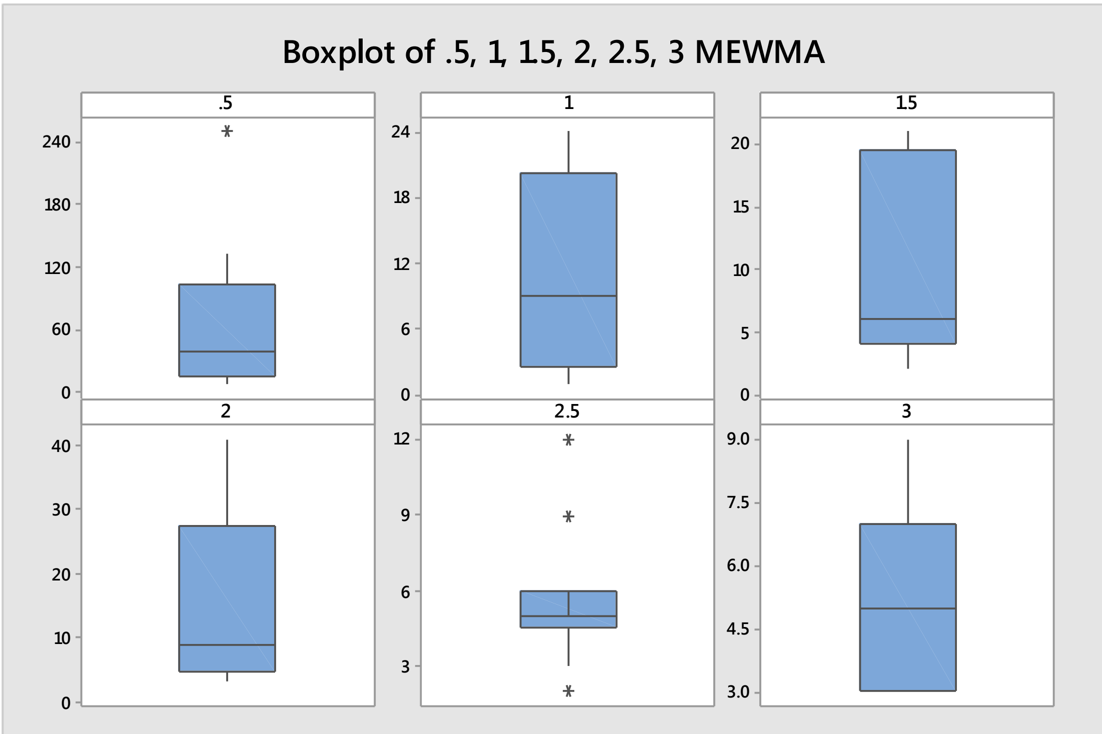

Fig. 2. Box plots for the EWMA control chart corresponding to Table 1

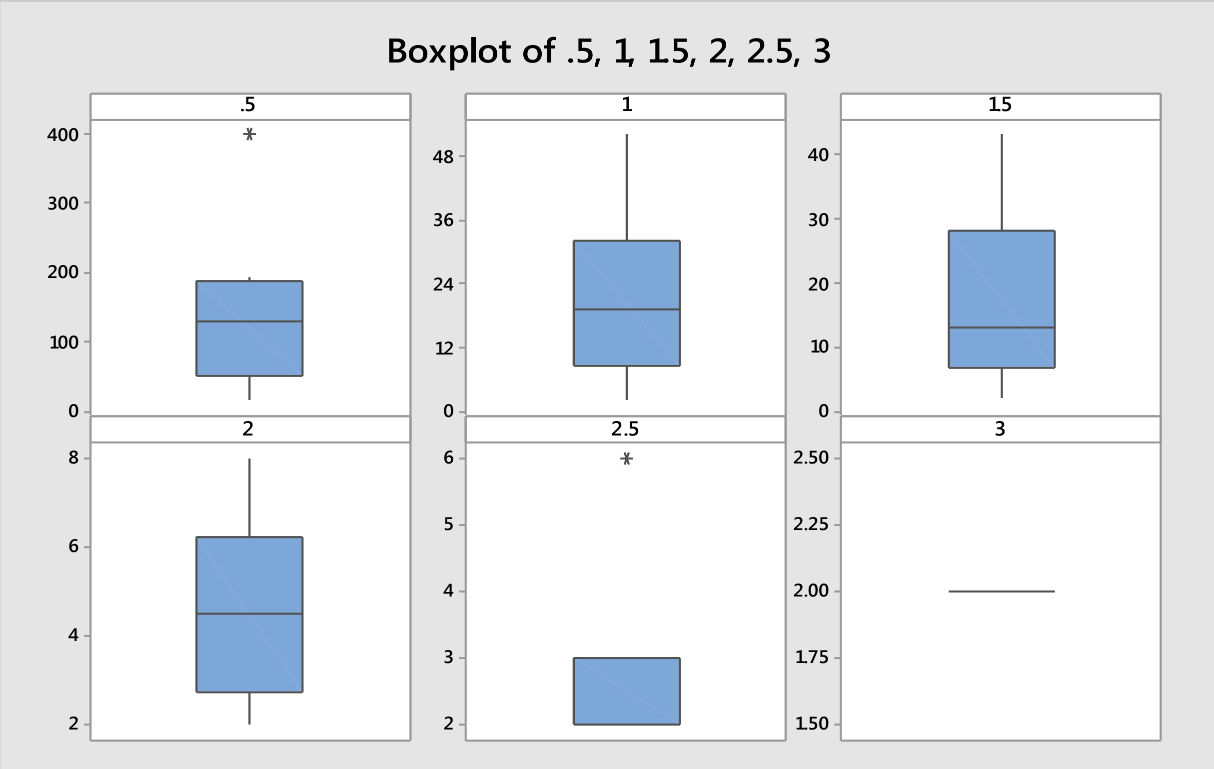

Fig. 3. Box plots for the Hotelling's T^2 control chart corresponding to Table 2

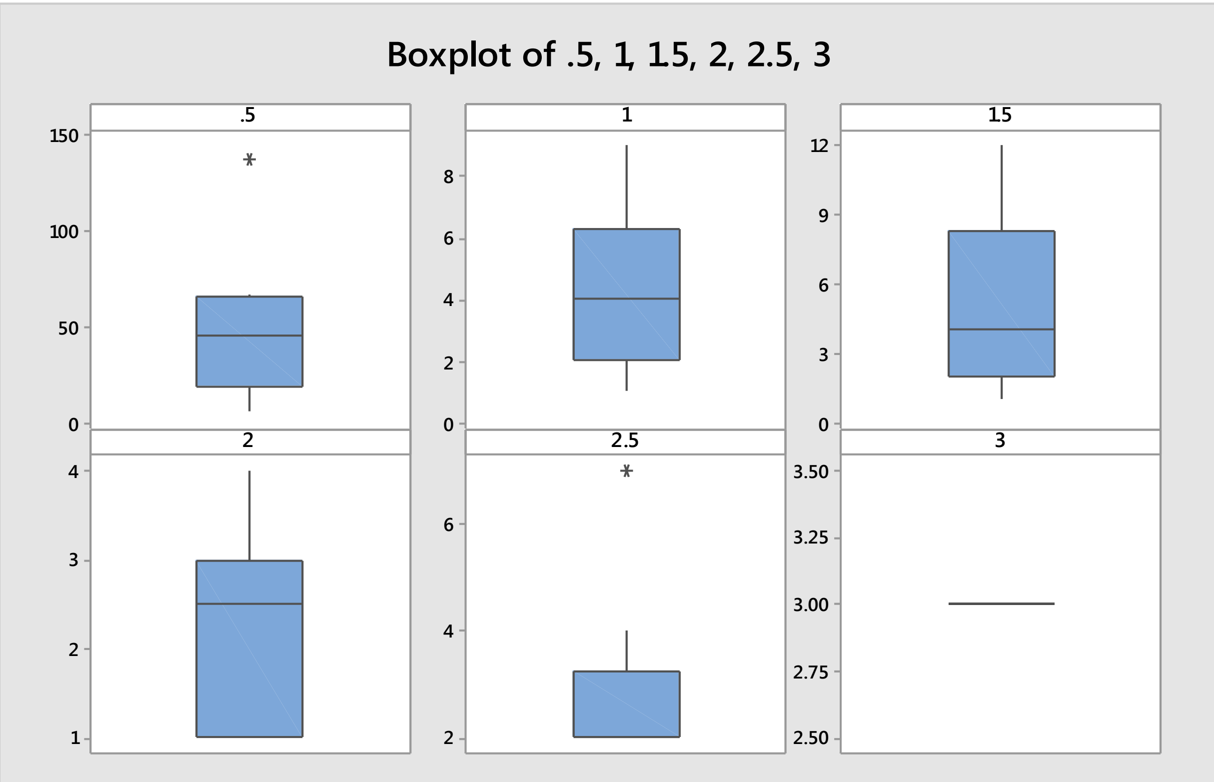

Fig. 4. Box plots for the EWMA control chart corresponding to Table 2

Table 3
Designing charts for variable changes

RL	$\alpha_1=\alpha_0+0.5\delta$, $\beta_1=\beta_0+0.5\delta$, $\gamma_1=\gamma_0+0.5\delta$		$\alpha_1=\alpha_0+1\delta$, $\beta_1=\beta_0+1\delta$, $\gamma_1=\gamma_0+1.5\delta$		$\alpha_1=\alpha_0+1.5\delta$, $\beta_1=\beta_0+1.5\delta$, $\gamma_1=\gamma_0+1.5\delta$		$\alpha_1=\alpha_0+2\delta$, $\beta_1=\beta_0+2\delta$, $\gamma_1=\gamma_0+2\delta$		$\alpha_1=\alpha_0+2.5\delta$, $\beta_1=\beta_0+2.5\delta$, $\gamma_1=\gamma_0+2.5\delta$		$\alpha_1=\alpha_0+3\delta$, $\beta_1=\beta_0+3\delta$, $\gamma_1=\gamma_0+3\delta$	
	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA	T^2	MEWMA
RL_1	4	2	15	3	5	2	2	1	2	2	2	3
RL_2	13	5	12	2	4	2	2	1	2	2	2	3
RL_3	36	13	6	1	6	3	3	2	2	2	2	3
RL_4	65	23	13	3	4	2	2	1	2	2	2	3
RL_5	9	4	34	7	4	2	4	2	2	2	2	3
RL_6	53	19	6	2	4	2	2	1	2	2	2	3
RL_7	20	7	16	3	6	3	2	1	2	2	2	3
RL_8	47	17	7	2	4	2	5	3	3	3	2	3
RL_9	31	12	5	1	2	1	2	1	2	2	2	3
RL_{10}	156	54	6	2	3	2	5	3	3	3	2	3

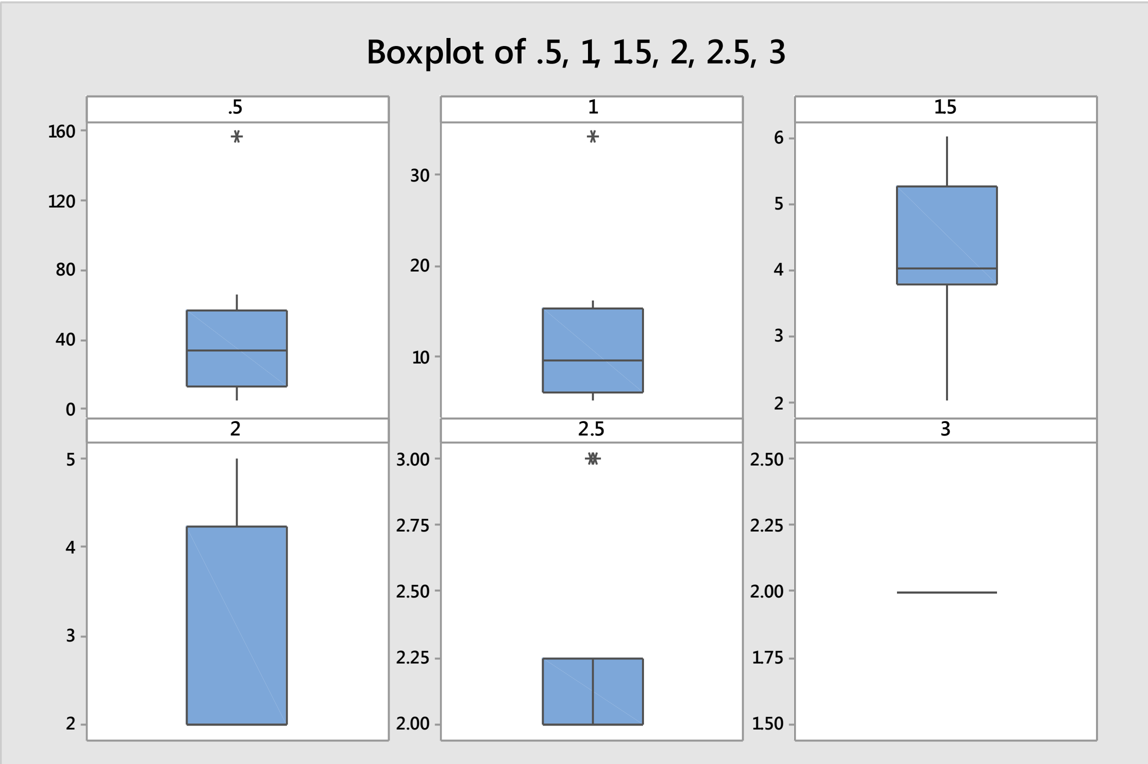

Fig. 5. Box plots for the Hotelling's T^2 control chart corresponding to Table 3

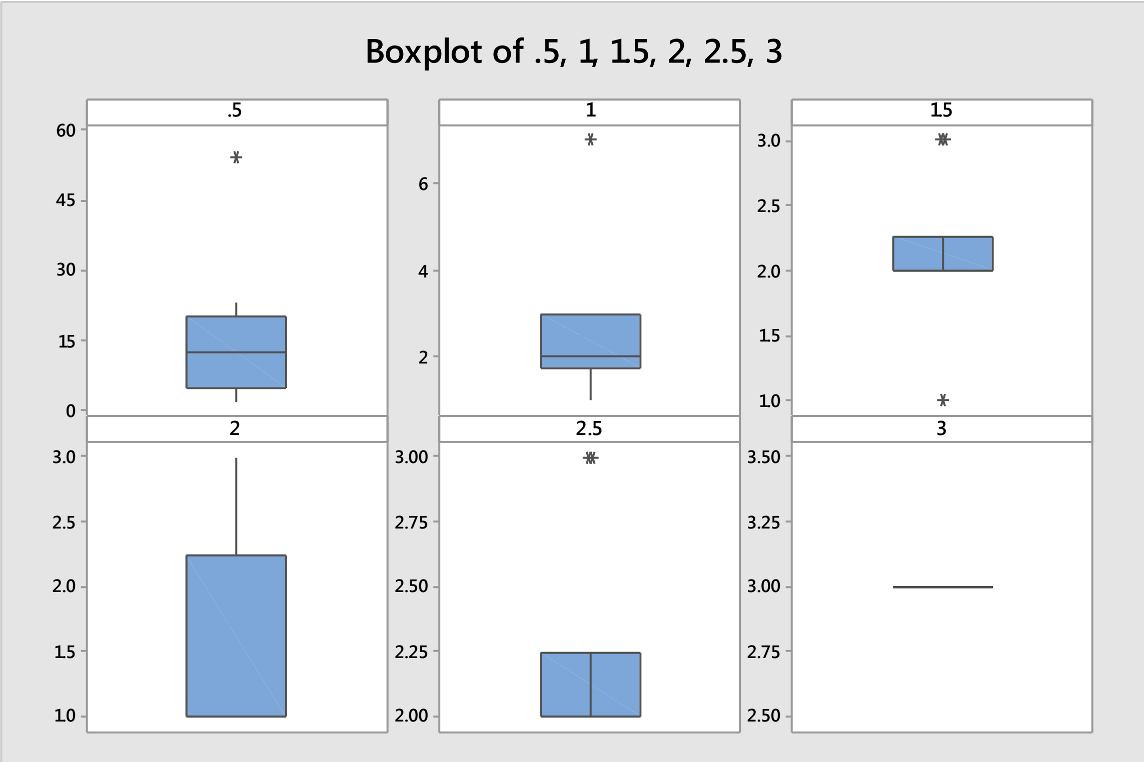

Fig. 6. Box plots for the EWMA control chart corresponding to Table 3

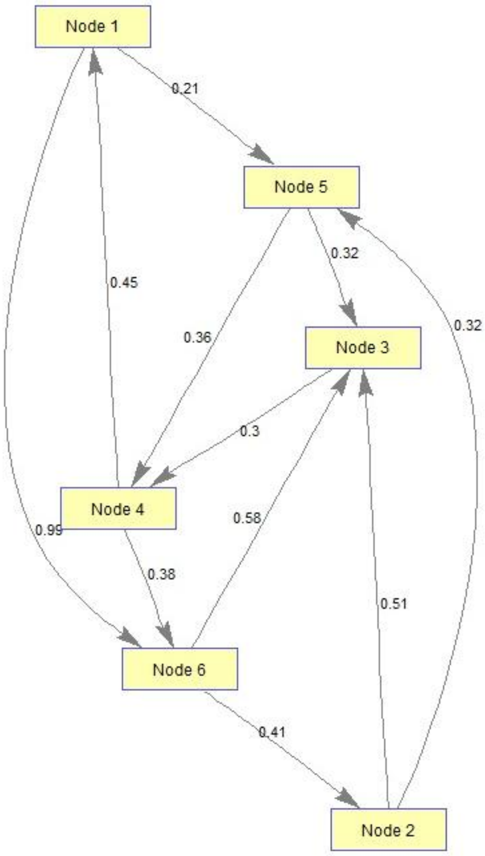

Fig. 7. The resulting graph corresponding to the changes in Table 1

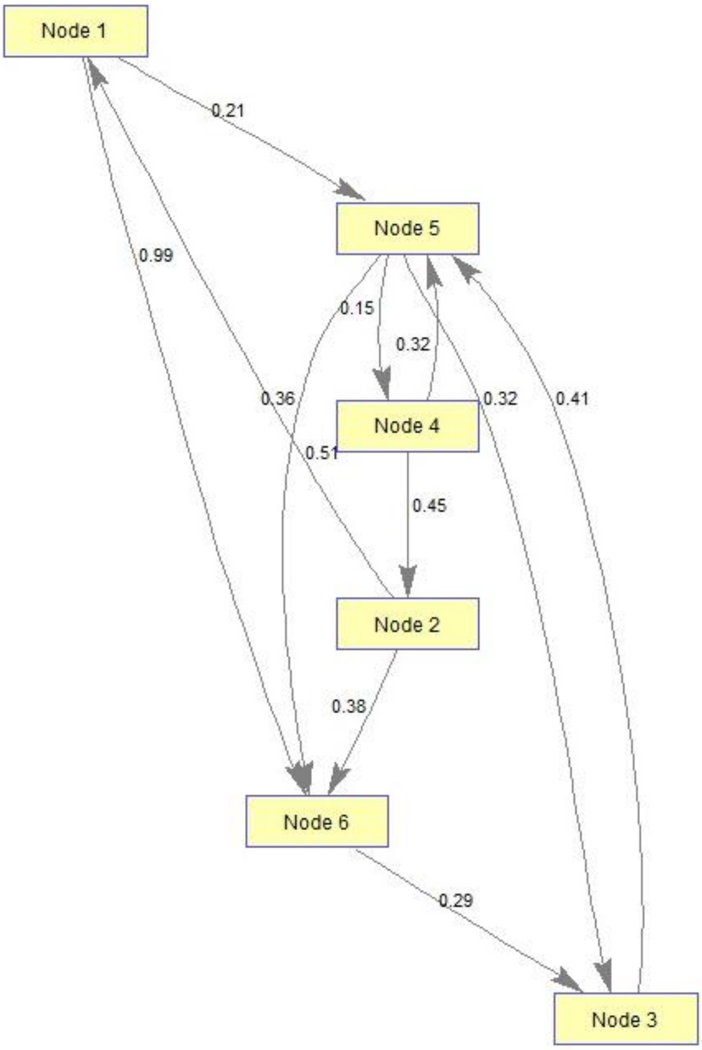

Fig. 8. The resulting graph corresponding to the changes in Table 2

The result showed that by increasing the amount of change of the parameters, the accuracy of the proposed method significantly increases. The issues have been made more complex due to their expansion and variation, and many of them have a network mode, so any analysis of knowledge and planning in dealing with them has a high scientific value. One of the factors that occurs dynamically over time is the changes that any possibility of managing and controlling them. Predicting the future behavior of the networks requires a reasonable and acceptable monitoring mechanism. The current research is moving in this direction.

5. Conclusion and Future Researches

The current research was aimed at providing a statistical method based on statistical knowledge of the network. For this, the network was depicted in the form of a mathematical model, and by applying the quality engineering issues and defining the appropriate statistics, an acceptable framework and methodology was developed to design the boundaries and control modules. Finally, the accuracy of the proposed method was examined and verified. In this research, a method for monitoring changes in social networks was provided. To this end, the ERGM model has been used and Hotelling's T2 and EWMA charts were developed to monitor the changes. In other words, assuming that the distribution function is based on the ERGM model and then using the EWMA statistics, the goal was to find a method for tracking and detecting deviations. In terms of the errors resulting from each change, along with the detection power of each method. As indicates before, the higher the rate of change, the more discriminating power is obtained, which is reasonable and the method has correctly identified it. As future researches, one can refer to:

- Monitoring the change in social networks with the combination of various features, including nodes and relationships
- Rooting the causes of deviations by finding the node or related relationship

- Monitoring the change in social networks using models other than ERGM
- Providing statistics or other methods for monitoring social networks
- Providing a model or framework for monitoring changes in the medical, biochemical, supply chain, and so on.

References

Barlow, R. E. & Irony, T. Z. (1992). Foundations of statistical quality control" in Ghosh, M. & Pathak, P.K. (eds.) Current Issues in Statistical Inference: Essays in Honor of D. Basu, Hayward, CA: Institute of Mathematical Statistics, 99-112.

Bersimis, S., Psarakis, S. & Panaretos, J. (November 2006). Multivariate Statistical Process Control Charts: An Overview Optimal Design of the Variable Sampling Size and Sampling Interval Variable Dimension T2 Control Chart for Monitoring the Mean Vector of a Multivariate Normal Process. Quality and Reliability Engineering International.

Block, P. & et al, (2017). Change we can believe in: Comparing longitudinal network models on consistency, interpretability and predictive power.

Bodin, Ö. & et al, (2016). Collaborative Networks for Effective Ecosystem- Based Management: A Set of Working Hypotheses, *Policy Studies Journal*, Volume 45, Issue2.

Cranmer, S. J. & et al, (2016), Navigating the Range of Statistical Tools for Inferential Network Analysis. *American Journal of Political science*.

Dekker D, Krackhardt D, Snijders TAB. (2007). Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. Psychometrika; 72(4):563–81.

Everett MG, Borgatti SP. (2014).Networks containing negative ties. *Social Networks*; 38:111–20.

Hossain L. Hamra J, Wigand R.T., Carlsson S., (2015). Exponential random graph modeling of emergency collaboration networks, Knowl. Based Syst., http://dx.doi.org/10.1016/j.knosys. 2014.12.029.

Hossain L., Kuti M., (2010). Disaster response preparedness coordination through social networks, Disasters 34 755–786.

Hubert L, Schultz J. (1976). Quadratic assignment as a general data analysis strategy. *Br J Math Stat Psychol*; 29(2):190–241.

Kendrick, L. (2018). Change point detection in social networks—Critical review with experiments, Computer Science Review 29:1-13.

Leifeld, P. & et al, (2018). Temporal exponential random graph models with btergm: estimation and bootstrap confidence intervals, *Journal of Statistical Software*, vol 83.

Leifeld, P. (a1) & Cranmer, S. J. (2018). A theoretical and empirical comparison of the temporal exponential random graph model and the stochastic actor-oriented model, *Network Science*.Volume 7, Issue 1 March 2019 , pp. 20-51

MATLAB Version 7.10.0.499 (2010). The MathWorks, Inc. Protected by U.S. and international patents.

Robins G, Pattison P, Kalish Y, Lusher D, (2007). An introduction to exponential random graph models for social networks, Social Networks; 29: 173–191.

Schweinberger M., Petrescu-Prahova M. & Vu D.Q., (2011). Disaster Response on September 11, 2001 Through the Lens of Statistical Network Analysis, Tech. Rep., Department of Statistics, Pennsylvania State University, *J Comput Graph Stat*. 2012 Dec 1; 21(4): 856–882.

Sheng, J. (2019). Community detection based on human social behavior, Physica A: Statistical Mechanics and its Applications 531:121765.

Wang, P. & et al, (2015). Multilevel Network Analysis Using ERGM and Its Extension.

William, H.,Woodall, Zhao, M. J., Paynabar, Sparks, K.R. & Wilson, J. D. (2016). An overview and perspective on social network monitoring.

Windzio, M. (2018). The network of global migration 1990–2013: Using ERGMs to test theories of migration between countries, Social Networks Volume 53, May 2018, Pages 20-29.

This article can be cited: Rajabi, F., Saghaei, A. & Sadi-Nezhad, S. (2020). Monitoring of Social Network and Change Detection by Applying Statistical Process: ERGM. *Journal of Optimization in Industrial Engineering*. 13 (1), 131-143.

http://www.qjie.ir/article_667925.html
DOI: 10.22094/JOIE.2019.581174.1615