On the Statistical Monitoring of Interpersonal Organizational Networks

Marcus B. Perry

Department of Information Systems, Statistics, and Management Science
The University of Alabama
Tuscaloosa, AL

mperry@cba.ua.edu

October 6, 2017

Overview

Introduction

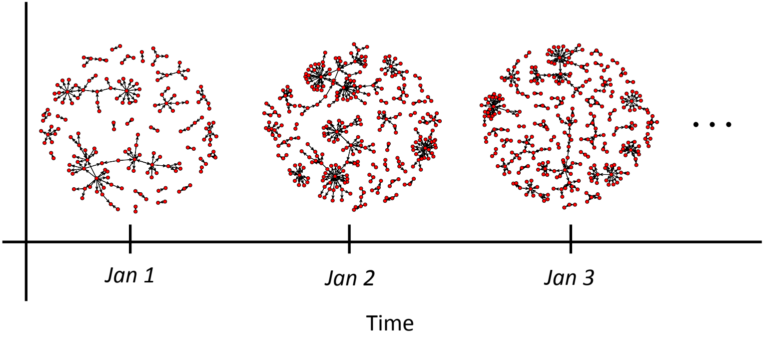
- Statistical process control (SPC) methods have traditionally been applied to industrial processes as a means to monitor the quality of manufacturing processes.
- More recently, however, the application of SPC methods to non-manufacturing processes has become more prevalent.
- For example:
    - Service processes.
    - Healthcare processes.
    - Environmental processes.
    - **Social processes**.

Introduction

- While discussing the “big-data” challenges for SPC practitioners and researchers, Lin (2014) categorizes the types of data to be monitored using SPC methods as follows:
  - Time series 101 - each observation is a scalar
  - Time series 201 - each observation is a vector
  - Time series 301 - each observation is a profile
  - Time series 401 - each observation is a network
  - Time series 501 - each observation is a graphic

Introduction

- The first three data types in the list above are fairly well studied; however, the development and evaluation of methods for monitoring networks and/or graphs is lacking.



What's been done?

- Although the development and evaluation of SPC methods for network-based systems is in its early stages, the literature does contain some relevant work.
- Woodall et al. (2017) recently provided a review paper outlining the current state of the literature on SPC methods developed for **undirected** social network data.
  - Many of the methods cited considered monitoring for changes in the communication *levels* of a network or subnetwork.
  - Others apply traditional SPC methods to monitor for changes in the average closeness or betweenness of the network.
- Additional works have emphasized detecting changes in the *community structure* of undirected networks over time, e.g., see Wilson et al. (2016).

What about organizational networks?

- Nohria (1992) argued that organizations can be commonly viewed as communication (or social) networks, and these need to be analyzed and addressed as such.
- Organizational networks typically model how information flows amongst the organization's members.
- Typically these are *directed* networks, or digraphs.
- An important question about these networks concerns whether or not they are hierarchically structured.
    - e.g., is there a point in time where interpersonal communications within an organization shifted from decentralized to centralized?
    - Such a shift could indicate the onset of a hierarchical decision-making structure, where all decisions and processes are handled at the top of the organization.

What about organizational networks?

- An increase in the "hierarchical tendency" of an organization's communication structure could lead to:
  1. Slower decision making $\Rightarrow$ response to changing market is slow.
  2. Increased global efficiency $\Rightarrow$ decreased local effectiveness.
  3. Less innovation $\Rightarrow$ poor creativity.
- Further, the root cause of such an increase could be associated with **conflict** or **crisis** within an organization.
  - e.g., see Danowski and Edison-Swift (1985), Davis et al. (2007), and Diesner et al. (2005)
- One means to assessing the hierarchical tendency of a digraph is to assess the relationship between its **reciprocal** and **transitive** flows.

Reciprocity and Transitivity

Let $c(i,j) = 1$ if information flows from node i to node j, and zero otherwise.

- A **reciprocal** flow between nodes i and j exists if $c(i,j) = c(j,i) = 1$.
- Let i, j, and k denote any three nodes in the network, then there are 6 possible ways to define a **transitive** flow (shown below).

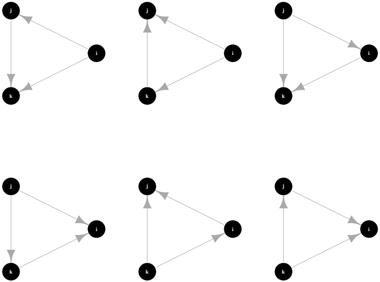


Figure: Six possible transitive flows for the triad (i, j, k).

Reciprocity and Transitivity

- Reciprocity and transitivity of digraphs are well studied in the social sciences literature:
  - Reciprocity is linked theoretically to exchange and resource dependence theories, e.g., see Cropanzano and Mitchell (2005) and Hillman et al. (2009),
  - Transitivity is linked theoretically to balance theory, originally developed as a theory of cognitive consistency for understanding interpersonal structures in a network, e.g., see Heider (1958).
- While theories of exchange and resource-dependency posit structural tendencies at the dyadic level, theories of cognitive consistency posit structural tendencies at the triadic level.

Interpretation of Reciprocity and Transitivity in Organizational Networks

- Holland and Leinhardt (1971) posited that transitivity, as a generalization of *structural balance theory*, can lead to the development of hierarchies and cliques.
- Contractor et al. (2006) explained that transitive triads within organizational networks point to the presence of hierarchy or centralization.
- Kilduff and Tsai (2003) suggested that if an organizational network tends towards low reciprocity and high transitivity, this results in a hierarchical flow of communication.

  $\rightarrow$ Easy to see if one considers that, in the limit of low reciprocity and high transitivity, one obtains a **fully transitive tournament**.

Fully Transitive Tournament

- To construct this type of graph, consider ordering all vertices on the real line, and then placing (i, j) edge in all i, j pairs such that $i > j$.
- The Figure below shows a fully transitive tournament with 4 vertices constructed this way.
- Notice the hierarchical nature of the flows for this type of graph.
- In practice, one does not see anything quite as extreme, but certainly some networks can tend in such a direction.

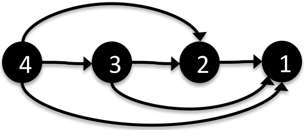


Figure: Fully transitive tournament with 4 vertices.

Hypothetical Example

- Consider the digraph shown in the Figure below.
- Suppose this network represents the communication flows expected under normal conditions between faculty members from two different departments within a college.

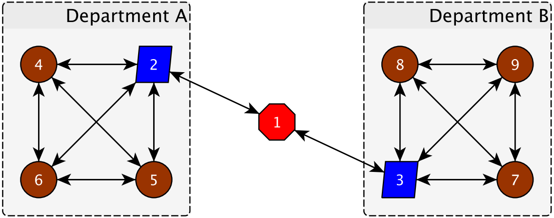


Figure: Directed graph for hypothetical example showing expected flows under normal conditions.

Hypothetical Example

- Expected communication flow between dean and department heads is mutual, and the expected communication flow between faculty members within each department is mutual.
- Note that the count of reciprocal dyads for this digraph is 14, while the count of transitive triads is 48.

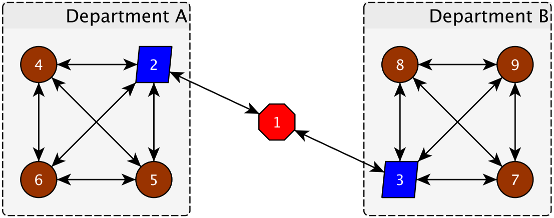


Figure: Directed graph for hypothetical example showing expected flows under normal conditions.

Hypothetical Example

- Suppose that budget constraints have resulted in the two departments being considered for merger into a single department.
- Then, one might expect the dean to communicate information relevant to the merger with all faculty directly, i.e., not just through the department heads.

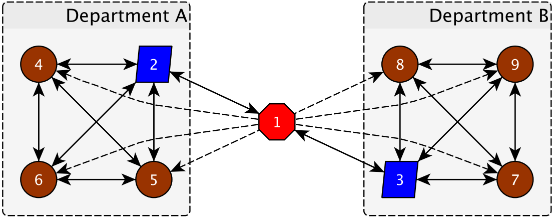


Figure: Directed graph for hypothetical example showing expected flows under normal conditions.

Hypothetical Example

- Notice that the number of reciprocal dyads remains at 14, while the number of transitive triads increases to 84.
- As a consequence of the threat of a merger, an increase in the *hierarchical tendency* of the communication flows is observed.

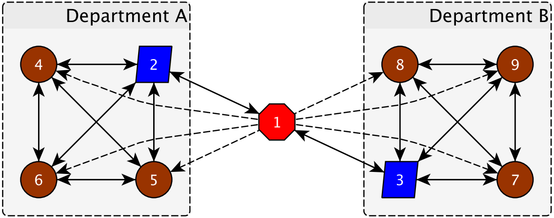


Figure: Directed graph for hypothetical example showing expected flows under normal conditions.

Enron Email Corpus

- To motivate the methodological developments that follow, consider the open-source Enron email corpus.
- Dataset was made public during the federal investigation of the Enron scandal.
- We consider the daily Enron email networks observed between January 1, 2001 through December 2, 2001, resulting in a time series of 240 consecutive days of *directed* email communications.
- However, we removed the outlier digraphs corresponding to Independence Day (July 4), Labor Day (Sept 3), and the day of the Sept 11 terrorist attacks, resulting in a total of 237 days.

Enron Email Corpus

- Suppose the first 100 days are representative of the in-control communication flows at Enron.
  - i.e., communications that are consistent with normal operations.
- Figure below shows the time series of daily *proportions* of mutual (or reciprocal) dyads and transitive triads over the first 100 days.
- Note because the sizes of the email networks is not constant over time, we use proportions of mutual dyads and transitive triads instead of counts.

Enron Email Corpus

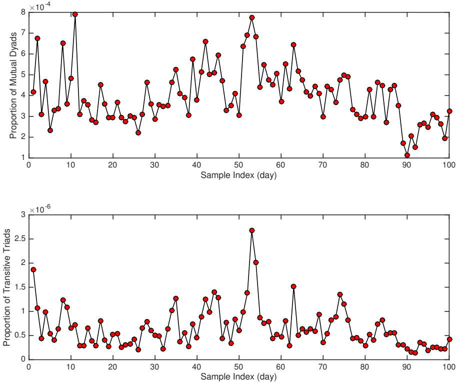


Figure: Time series of proportions of mutual dyads and transitive triads for the daily Enron email networks over the first 100 days of 2001 calendar year.

Enron Email Corpus

- It is important to note that our goal is NOT to detect sustained shifts in the mean of the joint process.
- Instead, recall that in the limit of low reciprocity and high transitivity one obtains the fully transitive tournament.
- Our interest then lies in detecting increases in the proportion of digraphs having a reciprocal tendency lower than expected *and* a transitive tendency higher than expected.

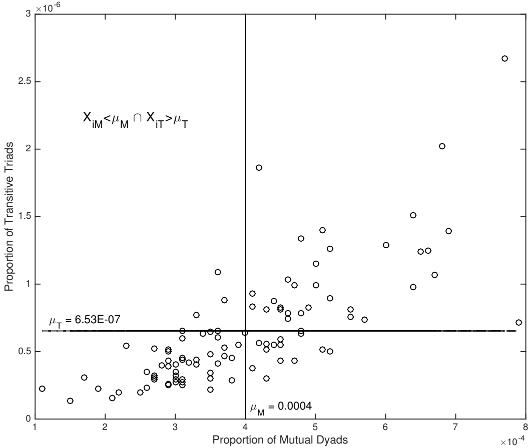


Figure: Scatterplot of proportion of mutual dyads versus proportion of transitive triads over first 100 days of the daily Enron email networks.

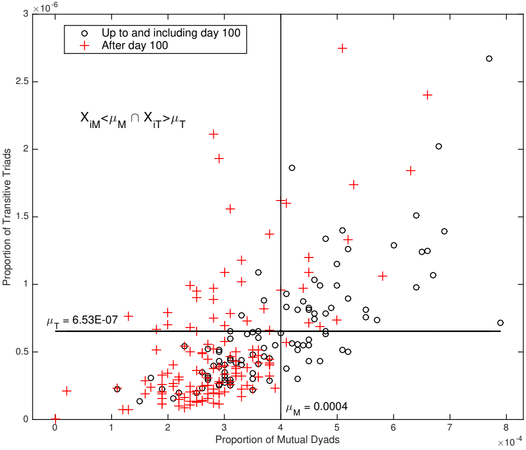


Figure: Scatterplot of proportion of mutual dyads versus proportion of transitive triads for all 237 daily Enron email networks. Note that the first 100 digraphs are assumed to be from the in-control process.

Enron Email Corpus

- Clearly, the structure of Enron's email communication flows shifted to a more hierarchical or centralized tendency at some point following day 100.
- 260% relative increase from the in-control proportion of hierarchical digraphs.
- Such a large observed increase could in fact be due to the now well known crisis that was going on within Enron during that time.
- In retrospect, if Enron stakeholders had implemented a monitoring scheme to detect increases in the hierarchical tendencies of the communication flows, this could have potentially provided some early indication of the serious internal problems Enron was facing.

Statistical Process Control for Detecting Shifts in Hierarchical Tendency

- For the Enron example, daily digraphs were classified as either hierarchical with probability p or not hierarchical with probability $1 - p$, so that Y was a Bernoulli random variable.
  - $\Rightarrow$ Enron network monitoring problem reduced to monitoring a Bernoulli process.
- One can generally consider classifying the observed digraph at time i into one of $k \geq 2$ classes.
- Remainder of this talk conerned with detecting change points in a k-class multinomial process with $n = 1$.

Control Charts for Monitoring Multinomial Processes

- Extensive review of control charts for multinomial and multiattribute data is given in Topalidou and Psarakis (2009);
- Developments that are most relevant to our general problem:
  - Shewhart p-chart is the traditional approach to monitoring binomial processes, but is not useful when $n = 1$.
  - Cumulative sum (CUSUM) control charts, e.g., see Reynolds and Stoumbos (1999) for $k = 2$ and Ryan et al. (2011) for $k \geq 2$.
  - Marcucci (1985) proposed a generalized p-chart to account for $k > 2$ using Pearson's chi-square statistic.
  - Weiβ (2012) proposed two control charting strategies using "comparative statistics" and a w-size moving average estimator of the k-class probabilities.
  - Yashchin (2012) discussed multinomial control charting strategies using generalized likelihood ratio tests.

The EWMA Control Chart

- EWMA control chart is one of the most popular tools in today's SPC practice, it is surprising to find the literature lacking in EWMA control charting strategies for multinomial processes.
- One exception is Gan (1990), who proposed a modified EWMA control chart for use with binomial data.
- Steiner (1998) proposed a grouped-data EWMA control chart for continuous data distributions – could be applicable to our problem if
  1. we were interested in monitoring for shifts in the mean of one of the marginal variables, i.e., reciprocity or transitivity, and
  2. we could only observe enough information on the variables to classify observations into one of the k groups.

The EWMA Control Chart

- Historically, the EWMA control charting strategy has been an attractive tool for the SPC practitioner.
- It is simple to design and implement, and is known to have detection properties similar to that of the CUSUM control chart.
- Further, upon signaling, the EWMA control chart offers reliable estimates of the process change point and magnitude of change.
- In what follows we propose a new EWMA control charting strategy for the general multinomial process.

Multinomial EWMA Control Chart

- Let $\mathbf{Y} \sim \text{multinomial}(n, \mathbf{p})$.
  - n denotes the number of independent trials.
  - $\mathbf{Y}$ is a $k \times 1$ random vector of class counts with $\sum_{j=1}^{k} Y_j = n$
  - $\mathbf{p}$ is a $k \times 1$ parameter vector with $\sum_{j=1}^{k} p_j = 1$
- We consider the linear combination of the class counts Z given by

$$Z = \sum_{j=1}^{k} w_j Y_j$$

where w_j is a weight applied to the j^{th} category.

Multinomial EWMA Control Chart

- The mean and variance of Z are easily shown to be

$$\mu_Z = n \sum_{j=1}^{k} w_j p_j$$

and

$$\sigma_Z^2 = n \left(\sum_{j=1}^{k} w_j^2 p_j (1 - p_j) - 2 \sum_{j=1}^{k-1} \sum_{j'=j+1}^{k} w_j w_{j'} p_j p_{j'} \right)$$

Multinomial EWMA Control Chart

- Denote the in-control vector of multinomial class probabilities by $\mathbf{p}_0$, and the out-of-control vector by $\mathbf{p}_1$, where $\mathbf{p}_1 \neq \mathbf{p}_0$.
- In this effort, we define the weight for the j^{th} class by $w_j = (np_{0j})^{-1}$, or the inverse of class j's in-control expected count.
- If the process is in-control

$$\mu_{Z|\mathbf{p}=\mathbf{p}_0} = k$$

and

$$\sigma^2_{Z|\mathbf{p}=\mathbf{p}_0} = \frac{1}{n}\left(\sum_{j=1}^{k} \frac{1-p_{0j}}{p_{0j}} - k(k-1)\right)$$

Multinomial EWMA Control Chart

- At this point our goal is to set up an EWMA control chart on the random variable Z.
- Let $\mathbf{Y}_i$ denote the vector of multinomial cell counts observed at time i so that $Z_i = \sum_{j=1}^{k} w_j Y_{ij}$ denotes the linear combination of the cell counts at time i.
- Define the standardized linear combination at time i by

$$U_i = \frac{Z_i - \mu_{Z|\mathbf{p}=\mathbf{p}_0}}{\sigma_{Z|\mathbf{p}=\mathbf{p}_0}}$$

- Note that $E(U_i) = 0$ and $Var(U_i) = 1$ if the process is in-control.

Multinomial EWMA Control Chart

- Let $0 < r < 1$ and $G_0 = 0$, then one can chart

$$G_i = rU_i + (1-r)G_{i-1}, \quad i = 1, 2, ..., N$$

- Compare G_i at time i to the control limits

$$LCL_i/UCL_i = \pm L\sqrt{\frac{r}{2-r}[1-(1-r)^{2i}]}$$

- Steady-state control limits given by

$$LCL/UCL = \pm L\sqrt{\frac{r}{2-r}}$$

EWMA Control Chart Design

- To design the control chart, user must specify r and L.
- In practice, r is typically chosen between 0.05 and 0.20, e.g., see Hawkins and Wu (2014).
- In general, one can use Monte Carlo simulation to design the control chart.
- For $n = 1$, a slightly modified version of the Markov chain model discussed in Borror et al. (1998) can be used to select a value for L, given an acceptable *in-control* ARL, or ARL_0.
    - Provides analytical approximations to the chart's ARLs
    - Can be used to approximate both initial state and steady-state ARLs.

Markov Chain Approximations to In-Control ARLs

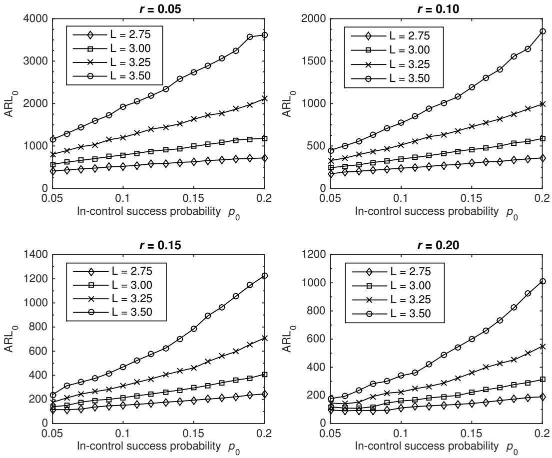


Figure: Initial-state in-control ARL approximations for the proposed EWMA control chart for a Bernoulli process with $p_0 \in [0.05, 0.20]$.

Average Run Length Performance

- ARL performance of our proposed EWMA control chart is studied.
- For the $k = 2$ case, we compare performance to those obtained from the Bernoulli CUSUM discussed in Reynolds and Stoumbos (1999).
- For the $k > 2$ case, we compare our results to those obtained from the multinomial CUSUM control chart published in Ryan et al. (2011).
- Only the $n = 1$ case is considered.
- Initial-state control limits are considered in our study.
- Because most processes are likely to operate in-control for some period of time, and then shift out-of-control, only the steady-state ARLs are considered for the out-of-control cases.

Average Run Length Performance

- To efficiently summarize ARL performance across the settings of the out-of-control parameter values considered, we compute the relative mean index, or *RMI*, given by

$$RMI(\nu) = \frac{1}{m}\sum_{\ell=1}^{m}(ARL_{\mathbf{p}_{1\ell}}(\nu) - ARL^*_{\mathbf{p}_{1\ell}})(ARL^*_{\mathbf{p}_{1\ell}})^{-1}$$

  - $ARL_{\mathbf{p}_{1\ell}}(\nu)$ is the ARL of control chart ν at the ℓ^{th} setting of $\mathbf{p}_1$.
  - m is the total number of out-of-control settings of $\mathbf{p}_1$ being studied.
  - $ARL^*_{\mathbf{p}_{1\ell}}$ is the smallest ARL across all control charts being compared for the ℓ^{th} setting of $\mathbf{p}_1$.
- Control charting strategy with best relative detection performance across range of out-of-control parameter settings considered will have smallest *RMI*.

ARL Performance for $k = 2$ (Bernoulli Process)

Table: In-control success probability $p_0 = 0.05$. Control charts calibrated to have in-control initial-state ARL of approximately 500.

p_1	CUSUM $p^* = 0.10$	CUSUM $p^* = 0.30$	EWMA $r = 0.05$	EWMA $r = 0.10$
0.10	66.68	88.04	76.99	85.57
0.15	30.86	36.38	33.36	35.73
0.20	19.78	21.17	20.56	21.05
0.25	14.58	14.53	14.76	14.59
0.30	11.56	10.98	11.48	11.07
0.35	9.60	8.81	9.40	8.92
0.40	8.23	7.39	7.96	7.45
0.45	7.20	6.37	6.90	6.41
0.50	6.42	5.61	6.10	5.62
0.60	5.28	4.56	4.94	4.51
0.70	4.48	3.86	4.16	3.78
0.80	3.90	3.35	3.59	3.27
0.90	3.45	2.96	3.15	2.87
$RMI =$	0.0988	0.0506	0.0801	0.0420

ARL Performance for $k = 2$ (Bernoulli Process)

Table: In-control success probability $p_0 = 0.10$. Control charts calibrated to have in-control initial-state ARL of approximately 500.

p_1	CUSUM $p^* = 0.15$	CUSUM $p^* = 0.35$	EWMA $r = 0.05$	EWMA $r = 0.10$
0.15	88.24	119.45	109.45	119.12
0.20	41.21	50.73	46.86	50.71
0.25	26.33	28.76	27.99	29.05
0.30	19.33	19.26	19.64	19.62
0.35	15.28	14.32	15.01	14.56
0.40	12.63	11.33	12.15	11.51
0.45	10.78	9.37	10.19	9.48
0.50	9.40	8.00	8.77	8.05
0.60	7.48	6.20	6.86	6.19
0.70	6.22	5.10	5.63	5.02
0.80	5.82	4.36	4.79	4.22
0.90	4.59	3.82	4.16	3.65
0.95	4.28	3.60	3.91	3.42
$RMI =$	0.1420	0.0636	0.1086	0.0580

ARL Performance for $k = 3$

Table: ARL comparisons between proposed EWMA and the multinomial CUSUM published in Ryan et al. (2011) for Case 1. First row corresponds to the in-control case. Parameter values for which the CUSUM was designed are shown in bold.

p_{11}	p_{12}	p_{13}	Multinomial CUSUM $h = 3.47$	EWMA $r = 0.05$ $L = 2.702$	EWMA $r = 0.10$ 3.081
0.65	*0.25*	*0.10*	508.31	508.31	508.43
0.625	0.255	0.12	236.73	247.72	250.18
0.6	0.26	0.14	132.43	132.36	141.94
0.55	0.27	0.18	57.28	57.26	62.32
0.5	0.28	0.22	32.68	33.72	35.52
0.4517	**0.2999**	**0.2484**	23.44	24.96	25.64
0.35	0.35	0.30	14.69	16.45	16.10
0.25	0.4	0.35	10.65	12.25	11.60
0.15	0.45	0.40	8.34	9.77	9.01
0.05	0.5	0.45	6.85	8.13	7.33
		$RMI=$	~ 0.0000	0.0857	0.0815

ARL Performance for $k = 3$

Table: ARL comparisons between proposed EWMA and the multinomial CUSUM published in Ryan et al. (2011) for misspecified Case 1. First row corresponds to the in-control case.

p_{11}	p_{12}	p_{13}	Multinomial CUSUM $h = 3.47$	EWMA $r = 0.05$ $L = 2.702$	EWMA $r = 0.10$ $L = 3.081$
0.65	*0.25*	*0.10*	508.31	508.31	508.43
0.66	0.22	0.12	315.95	316.06	299.56
0.67	0.20	0.13	271.98	255.29	243.73
0.68	0.18	0.14	237.18	211.15	202.97
0.69	0.15	0.16	174.90	144.73	141.03
0.72	0.10	0.18	151.97	116.59	112.56
0.73	0.07	0.20	120.17	89.07	86.45
0.74	0.05	0.21	111.28	80.63	78.13
0.75	0.02	0.23	91.77	65.24	63.03
		$RMI=$	0.2750	0.0378	0.0000

ARL Performance for $k = 4$

Table: ARL comparisons between proposed EWMA and multinomial CUSUM published in Ryan et al. (2011) for Case 4. First row corresponds to the in-control case. Parameter values for which the CUSUM was designed are shown in bold.

p_{11}	p_{12}	p_{13}	p_{14}	Multinomial CUSUM $h = 3.7992$	EWMA $r = 0.05$ $L = 2.721$	EWMA $r = 0.10$ $L = 3.216$
0.65	*0.20*	*0.10*	*0.05*	499.62	500.20	499.47
0.60	0.22	0.12	0.06	179.87	207.16	239.25
0.55	0.24	0.14	0.07	83.72	104.05	128.34
0.50	0.26	0.16	0.08	46.93	62.78	78.30
0.46	0.28	0.17	0.09	32.70	45.86	55.94
0.396	**0.3283**	**0.1734**	**0.1023**	21.31	33.73	39.94
0.35	0.34	0.19	0.12	16.59	25.01	28.31
0.30	0.35	0.21	0.14	13.27	19.27	20.86
0.25	0.36	0.23	0.16	11.04	15.61	16.36
0.20	0.38	0.24	0.18	9.46	13.30	13.57
0.15	0.40	0.25	0.20	8.27	11.61	11.59
0.10	0.42	0.26	0.22	7.34	10.28	10.10
0.05	0.43	0.28	0.24	6.59	9.13	8.87
			$RMI=$	0.0000	0.3905	0.5361

ARL Performance for $k = 4$

Table: ARL comparisons between proposed EWMA and multinomial CUSUM published in Ryan et al. (2011) for misspecified Case 4. First row corresponds to the in-control case.

p_{11}	p_{12}	p_{13}	p_{14}	Multinomial CUSUM $h = 3.7992$	EWMA $r = 0.05$ $L = 2.721$	EWMA $r = 0.10$ $L = 3.216$
0.65	*0.20*	*0.10*	*0.05*	499.62	500.20	499.47
0.625	0.15	0.15	0.075	238.35	113.77	131.91
0.60	0.10	0.20	0.10	133.49	49.36	56.70
0.55	0.05	0.25	0.15	58.70	23.02	24.34
0.50	0.025	0.30	0.175	35.00	16.71	17.05
0.45	0.02	0.35	0.18	24.49	14.23	14.40
0.40	0.015	0.40	0.185	18.43	12.38	12.37
0.35	0.01	0.45	0.19	14.66	10.93	10.81
0.30	0.005	0.50	0.195	12.10	9.77	9.57
0.25	0.0025	0.55	0.1975	10.29	8.90	8.64
0.20	0.001	0.60	0.199	8.95	8.18	7.89
			$RMI=$	0.7601	0.0101	0.0397

Performance Results - Conclusions

- General results suggest the proposed EWMA control chart is a viable alternative to the CUSUM chart when:
  1. for $k = 2$, the magnitude of the actual shift is different than what the CUSUM chart was designed to detect
  2. for $k > 2$ the direction of the actual shift is different than what the CUSUM chart was designed to detect.

$\Rightarrow$ When the practitioner cannot specify the magnitude and/or direction of the shift in **p** *a priori*, we recommend use of the proposed EWMA control chart.

Enron Email Corpus - Revisited

- Let's revisit the Enron email corpus discussed earlier and apply proposed EWMA chart in efforts to detect if and when an increase in the hierarchical tendency of the organization's communication structure occurred.
- Let $y_i = 1$ if digraph i was classified as hierarchical and zero otherwise.
- Recall for this data set $k = 2$ and $\mathbf{p}_0 = [0.05, 0.95]$, so that the linear combination of the class counts at time i is given by

$$z_i = 0.05^{-1} y_i + 0.95^{-1}(1 - y_i)$$

Enron Email Corpus - Revisited

- Note that $\mu_Z = 2$ and $\sigma_Z \approx 4.13$, so that the standardized linear combination at time i is
$$u_i = \frac{z_i - 2}{4.13}$$
- Using $r = 0.10$, the EWMA statistic $G_i = 0.10u_i + 0.90G_{i-1}$ for $i = 1, 2, ..., 237$ was plotted and compared to the upper control limit
$$UCL_i = 3.64\sqrt{\frac{r}{2-r}[1-(1-r)^{2i}]}$$
- Control chart produces the first signal at digraph 119, corresponding to June 16, 2001.

Enron Email Corpus - Revisited

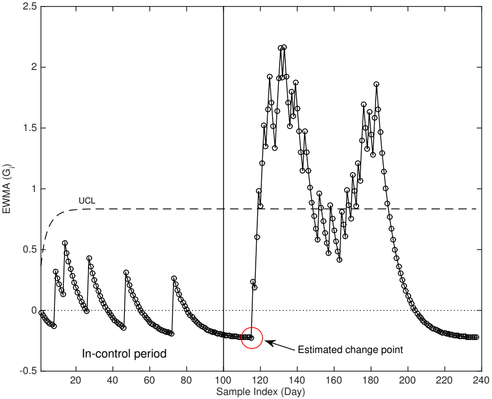


Figure: Proposed EWMA control chart applied to Enron email networks with $p_0 = 0.05$, $r = 0.10$ and $L = 3.64$. In-control ARL is approximately 500. Control chart signaled on day 119, or June 14, 2001. Large shift observed from mid-June to mid-September.

Enron Email Corpus - Revisited

- If we take a retrospective look at Enron's process, we see a very large increase in the hierarchical tendency of the communications from approximately mid-June to mid-September.
- Note that the proposed control chart detected the onset of the shift very quickly.
- Using the EWMA change point estimator, process change point τ is then estimated at $\hat{\tau} = 115$, or June 11, 2001.
- Also we compute $\hat{p}_1 = 1/2$, suggesting an estimated relative percent increase of approximately 900% from the in-control value.

Enron Email Corpus - Revisited

- Well known that Enron's organizational structure was very decentralized.
- Such a large increase in the proportion of hierarchical digraphs is alarming, and may suggest mid-June was the beginning of the end of Enron.
- Interesting to note that around the time of the estimated change point, two significant events did occur:
  1. Enron closed down the controversial Dabhol Power Plant (DPP) in Maharashtra, India, due to a payment and contract dispute between the local state government and the plant owners $\rightarrow$ \$ 1 Billion loss.
  2. Federal Energy Regulatory Commission (FERC) instituted price caps across the western United States, putting an end to the California energy crisis $\rightarrow$ End to Enron's cash flows.

Concluding Remarks

- Increases in the hierarchical tendency of an organization's communications network might suggest a shift to a more hierarchical decision-making structure.
- Unfortunately, such shift in structure can often lead to slower decision making, decreased operational efficiency, and less innovation.
- Further, the literature on organizational research suggests the root cause of such an increase could be associated with conflict or crisis within the organization.

Concluding Remarks

- In this talk we considered the important problem of detecting shifts in the hierarchical tendency of interpersonal networks over time.
- The early detection of increases can potentially reveal the onset of conflict or crisis within the organization.
- Consequently, organizational managers can implement a crisis or conflict management plan sooner so that the effects of the crisis on the health of the organization are minimized.

Concluding Remarks

- To this end, it was demonstrated that the network monitoring problem considered here can be reduced to that of monitoring a Bernoulli process.
- Justified by the fact that in the limit of low reciprocity and high transitivity one obtains the fully transitive tournament.
- In this talk, an EWMA control charting strategy for the general multinomial process was proposed, where the Bernoulli process is an important special case.
- ARL performance of proposed chart was compared to optimal CUSUM strategies $\rightarrow$ results suggest using EWMA when:
  - For $k = 2$, if magnitude of shift is unknown *a priori*.
  - For $k > 2$, if direction of shift is unknown *a priori*.

Concluding Remarks

- To demonstrate how the proposed control chart is applied in practice, we considered the daily Enron email networks observed during the year 2001, or Enron's crisis year.
- Through our analysis it was estimated that a significant increase in the hierarchical tendency of the email communications network at Enron occurred at or around June 11, 2001, approximately 6 months before the organization filed for bankruptcy.
- Assuming the observed increase was a result of the ongoing crisis within the Enron organization during that time, then the proposed control chart was able to detect the onset of this crisis rather quickly.

Concluding Remarks

- Finally, in addition to the network monitoring application discussed in this paper, the proposed EWMA control chart is applicable to a wide range of other applications.
- For example, the proposed strategy is generally applicable to multinomial processes, which are prevalent in the manufacturing, service and healthcare industries.

Thank You!

Questions or Comments?
Email: mperry@cba.ua.edu

References I

Borror, C. M., C. W. Champ, and S. E. Rigdon (1998). Poisson EWMA control charts. *Journal of Quality Technology 30*(4), 352–361.

Contractor, N. S., S. Wasserman, and K. Faust (2006). Testing multitheoretical, multilevel hypotheses about organizational networks: An analytical framework and empirical example. *Academy of Management Review 31*(3), 681–703.

Cropanzano, R. and M. S. Mitchell (2005). Social exchange theory: An interdisciplinary review. *Journal of Management 31*(6), 874–900.

Gan, F. F. (1990). Monitoring observations generated from a binomial distribution using modified exponentially weighted moving average control chart. *Journal of Statistical Computation and Simulation 37*, 45–60.

Hawkins, D. M. and Q. Wu (2014). The CUSUM and the EWMA head-to-head. *Quality Engineering 26*(2), 215–222.

References II

Heider, F. (1958). *The Psychology of Interpersonal Relations*. John Wiley.

Hillman, A. J., M. C. Withers, and B. J. Collins (2009). Resource dependence theory: A review. *Journal of Management 35*(6), 1404–1427.

Holland, P. and S. Leinhardt (1971). Transitivity in structural models of small groups. *Small Group Research 2*(2), 107–124.

Kilduff, M. and W. Tsai (2003). *Social Networks and Organizations*. Sage.

Lin, D. K. J. (2014). Comments: Some challenges for multivariate statistical process control. *Quality Engineering 98*, 26–96.

Marcucci, M. (1985). Monitoring multinomial processes. *Journal of Quality Technology 17*(2), 86–91.

Nohria, N. (1992). *Is a network perspective a useful way of studying organizations*. Network and Organizations: Structure, Form, and Action. Harvard Business School Press.

References III

Reynolds, M. and Z. Stoumbos (1999). A CUSUM chart for monitoring a proportion when inspecting continuously. *Journal of Quality Technology 31*, 87–108.

Ryan, A. G., L. J. Wells, and W. H. Woodall (2011). Methods for monitoring multiple proportions when inspecting continuously. *Journal of Quality Technology 43*(3), 237–248.

Steiner, S. H. (1998). Grouped data exponentially weighted moving average control charts. *Journal of the Royal Statistical Society, Series C 47*(2), 203–216.

Topalidou, E. and S. Psarakis (2009). Review of mulrinomial and multiattribute quality control charts. *Quality and Reliability Engineering International 25*(7), 773–804.

Weiβ, C. H. (2012). Continuously monitoring categorical processes. *Quality Technology and Quantitative Management 9*(2), 171–188.

References IV

Wilson, J. D., N. Stevens, and W. H. Woodall (2016). Modeling and detecting change in temporal networks via a dynamic degree corrected stochastic blockmodel. *http://arxiv.org/abs/1605.04049*.

Woodall, W. H., M. J. Zhao, K. Paynabar, R. Sparks, and J. D. Wilson (2017). An overview and perspective on social network monitoring. *IISE Transactions 49*, 354–365.

Yashchin, E. (2012). On detection of changes in categorical processes. *Quality Technology and Quantitative Management 9*(1), 79–96.