Advances in Mathematics: Scientific Journal **9** (2020), no.9, 7397–7406
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.9.9.86

DESIGN OF THE SYNTHETIC MULTIVARIATE COEFFICIENT OF VARIATION CHART BASED ON THE MEDIAN RUN LENGTH

MING HA LEE[1], VICTOR Y. C. LIM[2], XIN YING CHEW[3], MAN F. LAU[4], SEBASTIAN YAKUB[5], AND PATRICK H. H. THEN[6]

ABSTRACT. **Objectives:** The aim of this study is to use median run length (MRL) as an alternative criterion to evaluate the performance of the synthetic multivariate coefficient of variation (Syn MCV) chart, where the MRL is the 50th percentile of the run length distribution. **Methods/Statistical analysis:** A Markov chain approach is used to compute the MRL and the probabilities of the run length distribution of the Syn MCV chart. The design procedure of the Syn MCV chart based on the MRL is presented to obtain the optimal design parameters by minimizing the out-of-control MRL at a specified shift. **Findings:** The plots of the probability of the run length distribution show that the skewness of the run length distribution for the Syn MCV chart decreases with the increase in the magnitude of shift. This indicates that when evaluating the performance of the Syn MCV chart, the MRL is a more meaningful criterion compared to the average run length (ARL). The comparison results reveal that the Syn MCV chart is superior against the standard MCV (Std MCV) chart in terms of the MRL. **Application/Improvements:** The synthetic scheme can be employed to improve the performance of the Std MCV chart based on the MRL. In addition, the numerical results show that the MRL is a better performance measure for the Syn MCV chart compared to the ARL.

[2]*corresponding author*

2010 *Mathematics Subject Classification.* 62H99, 60A99.

Key words and phrases. median run length, Markov chain, multivariate coefficient of variation chart, run length distribution, synthetic chart.

1. Introduction

Control chart is considered as one of the most useful tools in Statistical Process Control (SPC) to maintain and improve the quality of the goods and services. Control charts are commonly designed based on the average run length (ARL) and this criterion has been received much criticism. Since the run length distribution varies depending on the process shift, the median run length (MRL) can give a more reliable interpretation to measure the performance of a control chart, where the MRL is defined as the median number of samples that are plotted on a control chart before an out-of-control signal is triggered. Using the MRL as a performance measure has been advocated by many researchers such as [1–3].

The standard multivariate coefficient of variation (Std MCV) chart was first proposed by [4] for monitoring the processes when the mean and the variability of the process are not independent of each other, where the variance is a function of the mean in the multivariate data. Since then, the interest of applying the Std MCV chart in many areas increased tremendously and is grabbing more attention in SPC. Recently, the performance of the Std MCV chart has been further improved by many researchers such as [5–7]. The design of the synthetic MCV (Syn MCV) chart based on the ARL is proposed by [8]. The objective of this study is to extend the work by [8] to optimally design the Syn MCV chart based on the MRL, where it can be proven that the MRL is a more credible criterion to measure the performance of the Syn MCV chart.

2. The Synthetic Multivariate Coefficient of Variation Chart

The Syn MCV chart presented by [8] integrates the Std MCV sub-chart and the conforming run length (CRL) sub-chart. Note that this study presents only the upward Syn MCV chart for detecting the magnitude of shift $\tau > 1$ since [8] stated that detecting the increase in shift is more important than the decreasing shift since an increase in shift indicates the increase of relative variation to the mean. Let UCL denote the upper control limit of the Std MCV sub-chart and let L denote the lower control limit of the CRL sub-charts. The operational procedure of the Syn MCV chart is as follows:

Step 1: Take a random sample of size n, that is $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ from a p-variate normal distribution, where $\mathbf{X}_i = (X_{i1}, X_{i2}, \ldots, X_{ip})^T$ for $i = 1, 2, \ldots, n$.

Step 2: Compute the MCV sample as follows:

$$\hat{\gamma} = \left(\overline{\mathbf{X}}^T \mathbf{S}^{-1} \overline{\mathbf{X}}\right)^{-\frac{1}{2}}, \tag{2.1}$$

where $\overline{\mathbf{X}} = \sum_{i=1}^{n} \mathbf{X}_i / n$ is the sample mean vector and $\mathbf{S} = \sum_{i=1}^{n} (\mathbf{X}_i - \overline{\mathbf{X}})(\mathbf{X}_i - \overline{\mathbf{X}})^T / (n-1)$ is the sample covariance matrix.

Step 3: If $\hat{\gamma} < \mathrm{UCL}$, the MCV sample is considered as conforming and return to Step 1, where $\mathrm{UCL} > 0$. Otherwise, the MCV sample is considered as non-conforming and proceed to the next step.

Step 4: Count the CRL value, which is the number of samples between the current (include) and the last (exclude) nonconforming MCV samples.

Step 5: If $\mathrm{CRL} > L$, the process is considered as in-control state and return to Step 1, where L is a positive integer. Otherwise, proceed to the next step.

Step 6: An out-of-control signal is triggered. Correction action is taken to find and remove the assignable cause(s), then return to Step 1.

3. The Run Length Distribution of the Synthetic Multivariate Coefficient of Variation Chart

Let $F_{\hat{\gamma}}(\mathrm{UCL}\,|n, p, \delta) = 1 - F_F\left(\frac{n(n-p)}{(n-1)p\mathrm{UCL}^2}\,|p, n-p, \delta\right)$ be the cumulative distribution function (c.d.f.) of the MCV sample $\hat{\gamma}$. Here, $F_F(\cdot|p, n-p, \delta)$ is the non-central F distribution with n and $(n-p)$ degrees of freedom and non-centrality parameter δ, where n is the sample size, p is the number of quality characteristics and $\delta = n/(\tau\gamma_0)^2$, with τ is the magnitude of shift in the process MCV and γ_0 is the in-control MCV. Note that the process is considered as in-control when $\tau = 1$, while the process is considered as out-of-control when $\tau > 1$. The probability of the MCV sample to be plotted above the UCL of the Std MCV sub-chart is calculated as [8]

$$A = 1 - F_{\hat{\gamma}}\left(\mathrm{UCL}|\, n, p, \delta\right). \tag{3.1}$$

In general, the $(L+1)$ by $(L+1)$ transition probability matrix $\mathbf{Q}$ for the transient states of the Markov chain for a synthetic control chart can be obtained as follows [9]:

(1) In the first row, A is in the first column and B is in the second column.

(2) In the last row, A is in the first column.

(3) In all other rows, the entry above the diagonal is A.
(4) All other empty entries are filled with zeros.

Therefore, the matrix $\mathbf{Q}$ can be derived from the transition probabilities of the Markov chain. For example, if $L = 4$, the transition probability matrix corresponding to the transient states is [9]

$$\mathbf{Q} = \begin{bmatrix} A & B & 0 & 0 & 0 \\ 0 & 0 & A & 0 & 0 \\ 0 & 0 & 0 & A & 0 \\ 0 & 0 & 0 & 0 & A \\ A & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{3.2}$$

where $B = 1 - A$. The probability distribution function of the run length R is defined as [10]

$$\Pr(R = r) = \mathbf{s}^T \mathbf{Q}^{r-1} (\mathbf{I} - \mathbf{Q}) \mathbf{1}. \tag{3.3}$$

According to [10], the c.d.f. of the run length R is defined as

$$\Pr(R \leq r) = \mathbf{s}^T (\mathbf{I} - \mathbf{Q}^r) \mathbf{1}, \tag{3.4}$$

for $r = 1, 2, 3, \ldots$, where $\mathbf{I}$ is the $(L+1)$ by $(L+1)$ identity matrix, $\mathbf{1}$ is a vector with each of its $(L+1)$ elements equal to unity, and $\mathbf{s}$ is the initial probability column vector having $(L+1)$ elements. As stated by [9], since the initial state of the Markov chain for the synthetic control chart corresponds to the second row of $\mathbf{Q}$, then the second element of s is equal to 1 and 0 elsewhere.

The 100ρ percentile of the run length R can be determined as m_ρ value such that

$$\Pr(R \leq m_\rho - 1) \leq \rho \quad \text{and} \quad \Pr(R \leq m_\rho) > \rho \tag{3.5}$$

where $0 < \rho < 1$ [10]. If $\rho = 0$ in (3.5), the 50th percentile of the run length distribution, that is the MRL can be computed, where $m_{0.5} = \text{MRL}$. Note that $\text{MRL} = \text{MRL}_0$ when $\tau = 1$, whereas $\text{MRL} = \text{MRL}_1$ when $\tau > 1$, where MRL_0 and MRL_1 are the in-control and out-of-control MRLs, respectively.

Figures 1-4 present the graphs of the probability of the run length distribution for the Syn MCV chart, where the probability of the run length distribution is computed using (3.3) for the Syn MCV chart with $L = 7$ and $\text{UCL} = 0.922817$ for $\tau = 1.0, 1.2, 1.5$ and 2.0 when $p = 2$, $n = 5$, $\gamma_0 = 0.5$, and $\text{MRL}_0 = 200$.

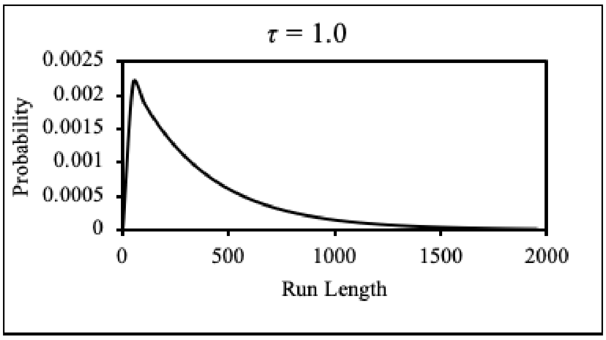

FIGURE 1. The run length distribution of the Syn MCV chart at $\tau = 1.0$ for $L = 7$ and $\mathrm{UCL} = 0.922817$ when $p = 2$, $n = 5$, $\gamma_0 = 0.5$ and $\mathrm{MRL}_0 = 200$

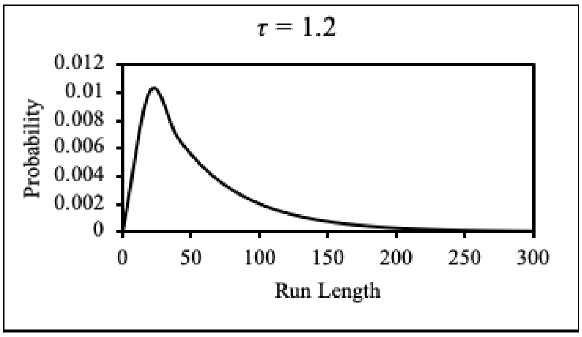

FIGURE 2. The run length distribution of the Syn MCV chart at $\tau = 1.2$ for $L = 7$ and $\mathrm{UCL} = 0.922817$ when $p = 2$, $n = 5$, $\gamma_0 = 0.5$ and $\mathrm{MRL}_0 = 200$

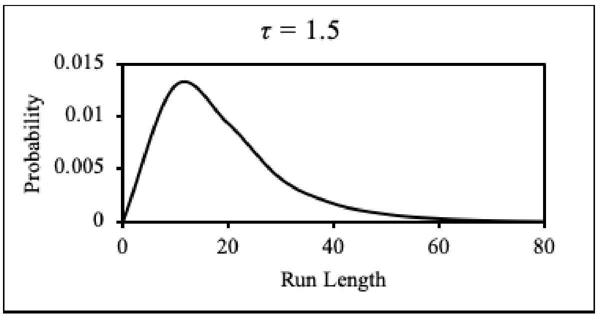

FIGURE 3. The run length distribution of the Syn MCV chart at $\tau = 1.5$ for $L = 7$ and $\mathrm{UCL} = 0.922817$ when $p = 2$, $n = 5$, $\gamma_0 = 0.5$ and $\mathrm{MRL}_0 = 200$

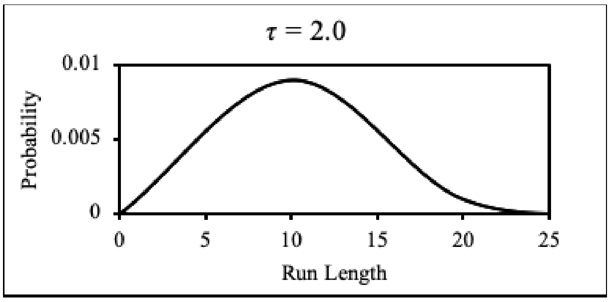

FIGURE 4. The run length distribution of the Syn MCV chart at $\tau = 2.0$ for $L = 7$ and $\text{UCL} = 0.922817$ when $p = 2$, $n = 5$, $\gamma_0 = 0.5$ and $\text{MRL}_0 = 200$

From Figures 1-4, it can be observed that the skewness of the run length distribution changes depending on the size of t, from highly skewed when the process is in-control (or at the small shift) to almost symmetric at the larger shift. This indicates that the skewness of the run length distribution decreases with the increase in the magnitude of t. Hence, solely depending on the ARL as the performance measure for the Syn MCV chart is not sufficient. For instance, using the MRL criterion can provide an alternative and more meaningful interpretation on the performance of the Syn MCV chart.

4. THE OPTIMAL DESIGN OF THE SYNTHETIC MULTIVARIATE COEFFICIENT OF VARIATION CHART BASED ON THE MEDIAN RUN LENGTH

This section presents the design procedure to obtain the optimal parameters L and UCL of the Syn MCV chart based on the MRL. The Syn MCV chart is optimized at a certain magnitude of shift τ such that the MRL_1 value is minimized to attain a specified MRL_0 value for the given values of the number of quality characteristics p, the number of sample size n and the in-control MCV γ_0. The optimal design procedure of the Syn MCV chart based on the MRL is given as follows:

Step 1: Specify the values of p, n, τ, γ_0, and MRL_0.
Step 2: Initialize $L = 1$.
Step 3: By setting $\rho = 0.5$, determine the UCL value using (3.5) such that $m_{0.5} = \text{MRL}_0$ when $\tau = 1$.

Step 4: By setting $\rho = 0.5$, calculate the MRL_1 value based on the current L and UCL values using (3.5) such that $m_{0.5} = \mathrm{MRL}_1$ when $\tau > 1$.

Step 5: When $L = 1$, increase L by one and return to Step 3.

When $L \geq 2$, if the MRL_1 value of the current L has been reduced compared to the MRL_1 value of the previous L (i.e. $L - 1$), increase L by one and return to Step 3; otherwise, proceed to the next step.

Step 6: Take the current L and UCL values as the optimal design parameters of the Syn MCV chart.

Note that the Syn MCV chart with the optimal design parameters in Step 6 gives the smallest MRL_1 value at the given shift.

5. THE PERFORMANCE OF THE SYNTHETIC MULTIVARIATE COEFFICIENT OF VARIATION CHART

In this study, the performance of the Syn MCV chart is evaluated based on the MRL. The MRL_0 value of the Syn MCV chart and the Std MCV chart is set as 200 to numerically evaluate the performance of the control charts at a desired shift. The control chart with the lower MRL_1 value is considered as the control chart with a better performance.

Table 1 shows the optimal parameter combination of L and UCL for the Syn MCV chart when $p \in \{2, 4\}$, $n \in \{5, 10, 15\}$, $\gamma_0 \in \{0.1, 0.3\}$ and $\mathrm{MRL}_0 \in \{200, 370, 500\}$, where the Syn MCV chart is optimized at $\tau \in \{1.2, 1.5, 2.0, 2.5, 3.0\}$. It can be observed that when sample size n increases, the L and UCL values decrease. For example, referring to Table 1, when $p = 2$, $\gamma_0 = 0.1$, $\mathrm{MRL}_0 = 200$ and $\tau = 1.2$, the L values for $n = 5, 10$ and 15 are 9, 5 and 3, respectively; while the UCL values for $n = 5, 10$ and 15 are 0.158305, 0.139692 and 0.130632, respectively.

From Table 1, it also can be noticed that when p increases, L increases while UCL decreases. For example, referring to Table 1, when $n = 5$, $\gamma_0 = 0.3$, $\mathrm{MRL}_0 = 200$ and $\tau = 1.2$, the L values for $p = 2$ and 4 are 10 and 18, respectively; while the UCL values for $p = 2$ and 4 are 0.506110 and 0.379253, respectively.

Table 2 compares the performance between the Syn MCV chart and the Std MCV chart based on the MRL when $p \in \{2, 4\}$, $n \in \{5, 10, 15\}$, $\gamma_0 \in \{0.1, 0.3\}$ and $\mathrm{MRL}_0 = 200$, where the Syn MCV chart is optimized at $\tau \in \{1.2, 1.5, 2.0, 2.5, 3.0\}$. The comparison results show that the MRL_1 values of the Syn MCV chart are significantly lower than that of the Std MCV chart for all the given shifts. Therefore, the Syn MCV chart outperforms the Std MCV chart by being able to detect the out-of-control signal faster than the Std MCV chart.

TABLE 1. The optimal design parameters L and UCL for the Syn MCV chart

			$p=2$						$p=4$					
			$n=5$		$n=10$		$n=15$		$n=5$		$n=10$		$n=15$	
MRL_0	γ_0	τ	L	UCL	L	UCL	L	UCL	L	UCL	L	UCL	L	UCL
200	0.1	1.2	9	0.158305	5	0.139692	3	0.130632	17	0.122965	6	0.127621	4	0.124031
		1.5	2	0.144274	1	0.129772	1	0.125217	5	0.111361	2	0.121001	1	0.117244
		2.0	1	0.137326	1	0.129772	1	0.125217	2	0.102133	1	0.116551	1	0.117244
		2.5	1	0.137326	1	0.129772	1	0.125217	2	0.102133	1	0.116551	1	0.117244
		3.0	1	0.137326	1	0.129772	1	0.125217	1	0.094773	1	0.116551	1	0.117244
	0.3	1.2	10	0.506110	5	0.435088	4	0.407777	18	0.379253	7	0.395777	4	0.380232
		1.5	3	0.465543	1	0.400263	1	0.384381	6	0.344746	2	0.370146	1	0.357271
		2.0	1	0.426985	1	0.400263	1	0.384381	3	0.32232	1	0.355199	1	0.357271
		2.5	1	0.426985	1	0.400263	1	0.384381	2	0.3089	1	0.355199	1	0.357271
		3.0	1	0.426985	1	0.400263	1	0.384381	2	0.3089	1	0.355199	1	0.357271
370	0.1	1.2	13	0.166806	6	0.144221	4	0.134775	24	0.131476	8	0.13272	5	0.127827
		1.5	3	0.153936	1	0.133715	1	0.128316	6	0.11898	2	0.124775	1	0.120355
		2.0	1	0.143532	1	0.133715	1	0.128316	3	0.112385	1	0.120526	1	0.120355
		2.5	1	0.143532	1	0.133715	1	0.128316	2	0.108401	1	0.120526	1	0.120355
		3.0	1	0.143532	1	0.133715	1	0.128316	2	0.108401	1	0.120526	1	0.120355
	0.3	1.2	14	0.537101	7	0.454414	5	0.421061	26	0.408855	9	0.412616	5	0.39326
		1.5	3	0.486465	2	0.428822	1	0.395049	7	0.368909	2	0.382959	1	0.367742
		2.0	1	0.448875	1	0.413985	1	0.395049	3	0.342271	1	0.368543	1	0.367742
		2.5	1	0.448875	1	0.413985	1	0.395049	2	0.329218	1	0.368543	1	0.367742
		3.0	1	0.448875	1	0.413985	1	0.395049	2	0.329218	1	0.368543	1	0.367742
500	0.1	1.2	15	0.170503	7	0.146709	4	0.136103	28	0.13537	9	0.134996	5	0.129141
		1.5	3	0.15667	2	0.139701	1	0.129778	7	0.123174	2	0.126564	1	0.121822
		2.0	1	0.146468	1	0.135576	1	0.129778	3	0.115288	1	0.122403	1	0.121822
		2.5	1	0.146468	1	0.135576	1	0.129778	2	0.111376	1	0.122403	1	0.121822
		3.0	1	0.146468	1	0.135576	1	0.129778	2	0.111376	1	0.122403	1	0.121822
	0.3	1.2	17	0.552961	8	0.462982	5	0.425695	31	0.42293	10	0.420305	6	0.40052
		1.5	4	0.506061	2	0.435121	1	0.400112	8	0.38227	2	0.389075	2	0.383783
		2.0	2	0.482982	1	0.420517	1	0.400112	3	0.351848	1	0.374891	1	0.372711
		2.5	1	0.459364	1	0.420517	1	0.400112	2	0.338956	1	0.374891	1	0.372711
		3.0	1	0.459364	1	0.420517	1	0.400112	2	0.338956	1	0.374891	1	0.372711

TABLE 2. The comparison of the MRL_1 value between the Syn MCV chart and the Std MCV chart when $\mathrm{MRL}_0 = 200$

		$p=2$						$p=4$					
		$n=5$		$n=10$		$n=15$		$n=5$		$n=10$		$n=15$	
γ_0	τ	Syn MCV	Std MCV	Syn MCV	Std MCV	Syn MCV	Std MCV	Syn MCV	Std MCV	Syn MCV	Std MCV	Syn MCV	Std MCV
0.1	1.2	9	30	5	16	3	11	17	47	6	20	4	13
	1.5	2	7	1	3	1	2	5	14	2	4	1	2
	2.0	1	2	1	1	1	1	2	5	1	1	1	1
	2.5	1	1	1	1	1	1	2	3	1	1	1	1
	3.0	1	1	1	1	1	1	1	2	1	1	1	1
0.3	1.2	10	33	5	19	4	13	18	51	7	23	4	15
	1.5	3	8	1	4	1	2	6	16	2	5	1	3
	2.0	1	3	1	1	1	1	3	6	1	2	1	1
	2.5	1	2	1	1	1	1	2	4	1	1	1	1
	3.0	1	1	1	1	1	1	2	3	1	1	1	1

6. CONCLUSIONS

Based on the numerical results, the in-control run length distribution of the Syn MCV chart is highly skewed to the right. In addition, this skewness changes from highly skewed when the process is in-control to almost symmetric at the larger shift. Hence, the MRL can be used as an alternative performance measure of the Syn MCV chart. The comparison results conclude that the Syn MCV chart is superior against the Std MCV chart in term of the MRL.

ACKNOWLEDGMENT

The authors would like to acknowledge the Sarawak Multimedia Authority (SMA) in financing this research through Sarawak Digital Economy Research Grant.

REFERENCES

[1] Y. QIAO, X. HU, J. SUN, Q. XU: *Optimal design of one-sided exponential EWMA charts with estimated parameters based on the median run length*, IEEE Access, **7** (2019), 76645–76658.

[2] M. H. LEE, M. B. C. KHOO: *Optimal design of synthetic np control chart based on median run length*, Commun. Stat. - Theory Methods, **46**(17) (2017), 8544–8556.

[3] W. L. TEOH, J. K. CHONG, M. B. C. KHOO, P. CASTAGLIOLA, W. C. YEONG: *Optimal designs of the variable sample size $\bar{X}$ chart based on median run length and expected median run length*, Qual. Reliab. Eng. Int., **33**(1) (2017), 121–134.

[4] W. C. YEONG, M. B. C. KHOO, W. L. TEOH, P. CASTAGLIOLA: *A control chart for the multivariate coefficient of variation*, Qual. Reliab. Eng. Int., **32**(3) (2016), 1213–1225.

[5] Q. T. NGUYEN, K. P. TRAN, H. L. HEUCHENNE, T. H. NGUYEN, H. D. NGUYEN: *Variable sampling interval Shewhart control charts for monitoring the multivariate coefficient of variation*, Appl. Stoch. Models Bus. and Ind., **35**(5) (2019), 1253–1268.

[6] A. HAQ, M. B. C. KHOO: *New adaptive EWMA control charts for monitoring univariate and multivariate coefficient of variation*, Comput. Ind. Eng., **131** (2019), 28–40.

[7] X. Y. CHEW, K. W. KHAW, W. C. YEONG: *The efficiency of run rules schemes for the multivariate coefficient of variation: a Markov chain approach*, J. Appl. Stat., **47**(3) (2020), 460–480.

[8] K. W. KHAW, X. Y. CHEW, W. C. YEONG, S. L. LIM: *Optimal design of the synthetic control chart for monitoring the multivariate coefficient of variation*, Chemometr. Intell. Lab. Syst., **186** (2019), 33–40.

[9] R. B. DAVIS, W. H. WOODALL: *Evaluating and improving the synthetic control chart*, J. Qual. Technol., **42**(2) (2002), 200–208.

[10] D. BROOK, D. A. EVANS: *An approach to the probability distribution of CUSUM run length*, Biometrika, **59**(3) (1972), 539–549.

[1,2,6]FACULTY OF ENGINEERING, COMPUTER AND SCIENCE, SWINBURNE UNIVERSITY OF TECHNOLOGY SARAWAK CAMPUS, JALAN SIMPANG TIGA, KUCHING, SARAWAK, MALAYSIA

Email address: [1]mhlee@swinburne.edu.my

Email address: [2]vlim@swinburne.edu.my

[3]SCHOOL OF COMPUTER SCIENCE, UNIVERSITY SAINS MALAYSIA, PENANG, MALAYSIA

Email address: xinying@usm.my

[4]SCHOOL OF SOFTWARE AND ELECTRICAL ENGINEERING, SWINBURNE UNIVERSITY OF TECHNOLOGY, MELBOURNE, AUSTRALIA

Email address: elau@swin.edu.au

[5]COLLEGE OF ENGINEERING, IT & ENVIRONMENT, CHARLES DARWIN UNIVERSITY, CASUARINA, AUSTRALIA

Email address: yakub.sebastian@cdu.edu.au

Email address: [6]pthen@swinburne.edu.my