A GENERALIZED APPROACH IN MULTIPROCESSOR ENVIRONMENT USING REGRESSION TYPE ESTIMATOR AND COST ANALYSIS

Sarla More, Diwakar Shukla
•
Department of Computer Science and Applications
Dr. Harisingh Gour University Sagar (MP) India
sarlamore@gmail.com, diwakarshukla@rediffmail.com

Abstract

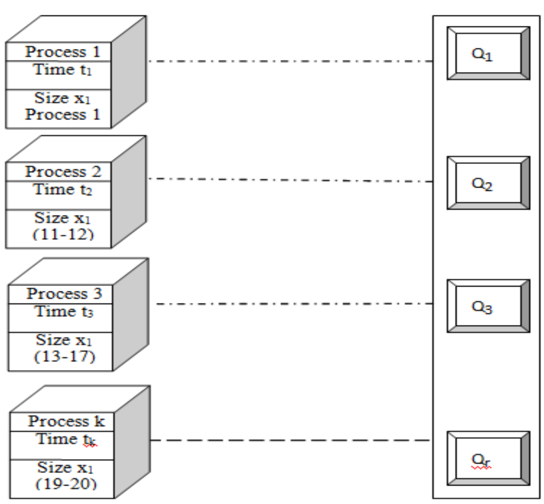
Consider a multi-processors computer system consisting of a ready queue of different jobs to be executed/processed. Lottery scheduling is fair enough to schedule the resources for each and every job. The research idea assumes condition where one can observe some processes to be fully executed; some partially executed few blocked/suspended/ terminated, after sudden system breakdown. An estimation strategy has been designed for the estimation of the total time required to process all these types of processes (processed, partially processed and blocked processes). How much time is required to process the remaining in any hazardous situation? A regression type estimator of sampling theory is used to perform this task. This remaining time estimation technique deals with the backup cost and recovery management as well. Sampling techniques are used in proposed approach for the testing purpose and a simulation has been performed. Another tool adopted is the confidence intervals which are calculated and gives proper précised values in comparison to the true mean for the total remaining time. The linear, square root and square cost function model are adopted for the calculation of backup cost and recovery management. In addition some auxiliary information is also incorporated in the form of size measure of the processes which is an effective approach to calculate the complete remaining time of the processes in multiprocessor environment. The purpose of the proposed research has been served effectively as one can observe the results of disaster and recovery management of the computer system.

Keywords: Ready Queue, Lottery scheduling, Multiprocessors, Simulation, Random Sampling, Estimation, Confidence Interval, Jobs(Processes), Size measure, Estimator

I. Introduction

In the scenario of cloud computing, ready queue is a setup among many servers and processors. For optimal resource allocation there exists several priority scheduling methodologies in the literature of scheduling schemes. In same way lottery scheduling scheme works on randomness of selection of process and distribution of resources providing fair chances. A random number is generated by processors in multiprocessor environment and some token numbers are assigned to each of the process. The execution of process depends upon the condition when the token number of a process is matches with the token number of the processor. The process which has the highest number of tokens has the chance to be allocated the resource for execution of the task. The jobs waiting in the queue always have the chance to be allocated the resource. lottery scheduling maintains the fairness between processes and gives equal chance to each and every process to be allocated the resource. Due to this reason Lottery scheduling is also known as starvation free scheme. In multiprocessor cloud based environment working of Lottery scheduling scheme is similar to draw a random sample through the sampling technique. The remaining time parameter estimation of the ready queue can be executed using the sampling techniques. A job in the ready queue has its process ID, the CPU time(in terms of bytes) as well as the process size (in terms of bytes). With the use of information of process size, it is expected to estimate better the unknown parameter. This paper exploit the approach of use of size measure information for efficient prediction.

Let (t_1, x_1), (t_2, x_2), (t_3, x_3)…….....(t_i, x_i)........(t_k, x_k) be the time consumed by i^{th} process in the waiting queue having size measure x_i. Further let Q_1, Q_2, Q_3, ...Q_r be the r processors ($r < k$) in a computer system who generate random numbers to select processes for resource allocation. Figure 1 describes the general setup of multiprocessors and ready queue. The Figure 2 and 3 are showing the same but in the classified and categorized manner.



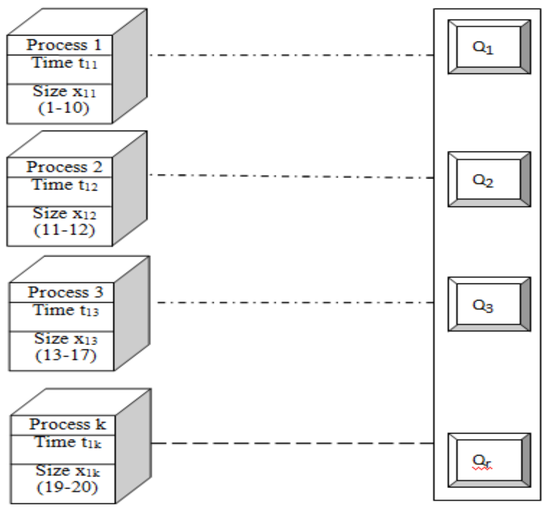


Figure 1: *Ready queue with waiting Processes and Multiprocessor*, **Figure 2:** *Small size processes and Multiprocessors*

This paper takes into account the approach of [4] but adds additional feature of partially processed, blocked processes and size measure of processes for time estimation. All these features are under assumption that the multiprocessor computer system fails at an instant due to unavoidable reasons and backup/recovery management is required. How much the backup cost is needed while sudden breakdown is a question of interest and can be predicted by using the suggested methodology of this paper.

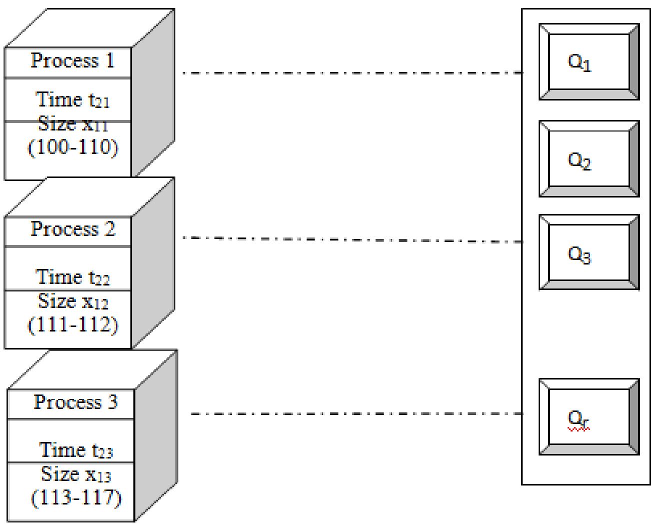


Figure 3: *Big size processes and Multiprocessors*

II. A Review

The priority scheduling is used when any of the jobs is to prefer over others in the waiting queue. Lottery scheduling is one such similar [8] where the job having highest number of tickets has the high chance of being allocated the desired resource. In Linux kernel setup, the lottery scheduling is useful [18] and it could be utilized as a framework [5, 7] for applying the sampling techniques. The similar job group formation scheme for mean time estimation of a ready queue [6] came into picture using lottery scheduling. A review on ready queue mean estimation [3] has opened up avenues for developing new methods in this area. The lottery scheduling types and model based utilization [16, 17] exists in literature as hybrid multilevel structure using Markov chain model along with analysis and chance based prediction. A sample can be used as a suitable input source for mean value prediction [9, 11, 16]. Many various sampling methodologies exist [10, 13, 14] who are comparatively better over another. The best method of selection among them [15] is always possible for precise prediction of unknown parameter. For missing data, the imputation techniques are popular who to replace the non-responding units [19, 20, 21] by known values. Some of most popular imputation methods are mean imputation, deductive imputation, mean imputation within classes, deductive imputation within classes, hot deck imputation, cold deck imputation etc. ([22, 23, 24, 25]). The content of this paper follows idea of [5] and [4] and uses them as input sources in order to resolve the issue of remaining time estimation in presence of sudden breakdown of the system. The contribution in [26] has opened up avenues to think for the use of size measure of processes.

I. Remaining Time Estimation Problem

Let there are finite number of N processes in a ready queue and n (n < N) have been processed completely before the system breakdown, obviously (N-n) are still in waiting to get signal for resource allocation. One can assume that n processes are just like a random sample selected from ready queue of size N using lottery scheduling. If θ is mean time obtained through sample then remaining total time estimate is $\Delta = [(N-n)\ \theta]$ which is an unknown quantity. For numbers 'c' and 'd', if Δ is predicted as $\Delta \in (c, d)$ who is an interval containing Δ with very high probability, then $\Delta_1 = [(N-n)\ c]$ is lowest, $\Delta_2 = [(N-n)\ d]$ is upper expected remaining time. If highest expected time is precisely estimated then it could be used for backup management during system failure. The efficient estimation of this expected range is a problem which is chosen in this paper for strategy formation in the multiprocessor setup with the consideration of multiple real life possibilities.

II. Confidence Interval (CI)

A confidence interval is a kind of predictive range for catching of unknown parameter. The feature of a confidence interval is that it contains the true value with 95% precision. Let P[A] denotes the probability of happening of an event A. In statistical theory, contains for any two real numbers a', b', the 95% confidence interval is defined as P [a' < true unknown value < b'] = 0.95. It could be interpreted as chance of being true value within a', b' is 95 percent. The length of confidence interval is a tool for measure of betterment. It is a difference of lower limit and upper limit. Let there are m different confidence intervals of length ($l_1, l_2, l_3, l_4 \ldots l_m$) who all catch the true value than an efficiency measure is: Best Confidence Interval = min [$l_1, l_2, l_3, l_4 \ldots l_m$]

III. Simulated Cost Aspect

Let C_0 be the fixed cost and C_1 be the cost per unit predicted time. If δ_1 is the minimum and δ_2 is the maximum remaining time after the occurrence of breakdown than

(a) Linear cost function is total cost $(T_c)_{1A} = C_0 + C_1 * \delta_1$ and $(T_c)_{2A} = C_0 + C_1 * \delta_2$
(b) Square root cost function $(T_c)_{1B} = C_0 + C_1 \sqrt{\delta_1}$ and $(T_c)_{2B} = C_0 + C_1 \sqrt{\delta_2}$

(c) Squared cost function is $(T_c)_{1C} = C_0 + C_1 * \delta_1^2$ and $(T_c)_{2C} = C_0 + C_1 * \delta_2^2$

Overall average cost = [Linear cost + Square root cost + Squared cost] / 3

The average cost is likely to incur in the recovery management of resources after the system breakdown. Averaging over linear, squared function and square-root function is taken to control the sampling fluctuations due to lottery scheduling sample.

IV. Sample based Estimation Method

Let $(Y_1, X_1), (Y_2, X_2), (Y_3, X_3)......... (Y_N, X_N)$ be the data of totality of size N where Y is variable of main interest and X is the support correlated information. For example, the Y may be expenditure of army officers in a country while x is income data which is known from the salary register of organization/head quarter. The mean of population is $\overline{Y} = (1/N) \sum Y_i$ and $\overline{X} = (1/N) \sum X_i$

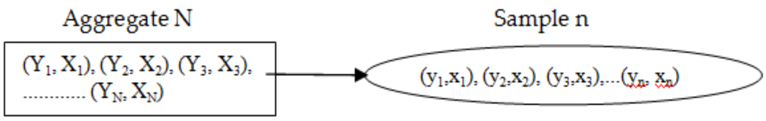


Figure 4: *Sample selection from Aggregate (n<N)*

A sample of size n (n<N) is drawn randomly from N by simple random sampling without replacement method. Sample values are $(y_1, x_1), (y_2, x_2), (y_3, x_3) ... (y_n, x_n)$.

Sample mean are $\overline{y} = (1/n) \sum y_i$ and $\overline{x} = (1/n) \sum x_i$

The objective is to estimate unknown parameter $\overline{Y}$ using known $\overline{X}$ along with sample means $\overline{y}$ and $\overline{x}$. Some well known estimators are:

- Sample mean estimator: $\overline{y}$
- Ratio-estimator: $\overline{y_r} = \overline{y}\ (\overline{X}/\overline{x})$
- Difference estimator: $\overline{y_d} = \overline{y} + d\ (\overline{X} - \overline{x})$

III. Motivation

Earlier contributions (specially [4], [5]) were under assumption that processes who exist in a multiprocessors system are completed before sudden failure. But this is not a practical reality. Since some jobs may complete, some may partially processed and some may blocked by the processors [see figure 4]. The processed and unprocessed case was considered in [4] [see figure (6)]. This paper extends the approach of [4] and [26] by applying the tools of random imputation method against the blocked processes.

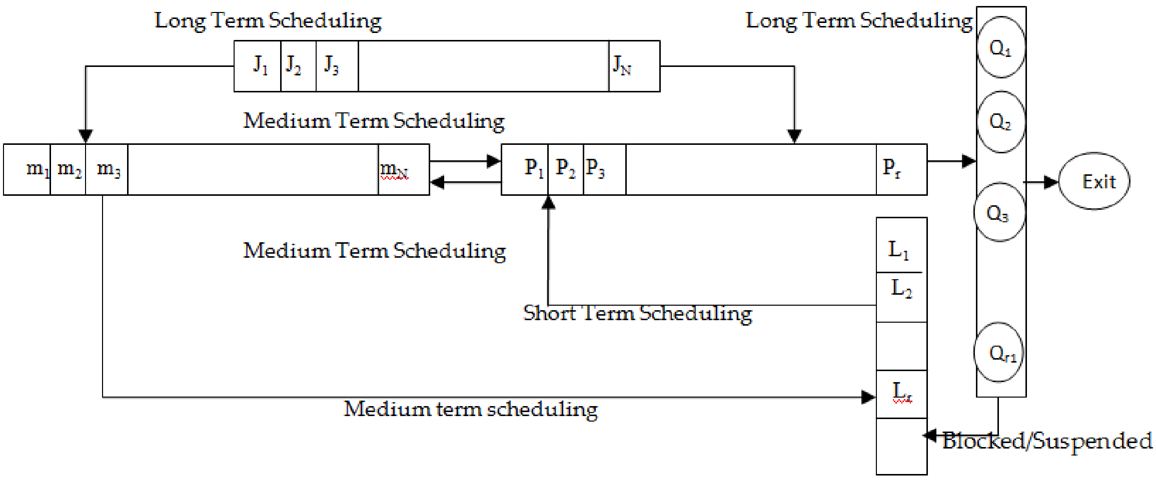


Figure 5: *Ready Queue Processing under Lottery Scheduling (due to [6])*

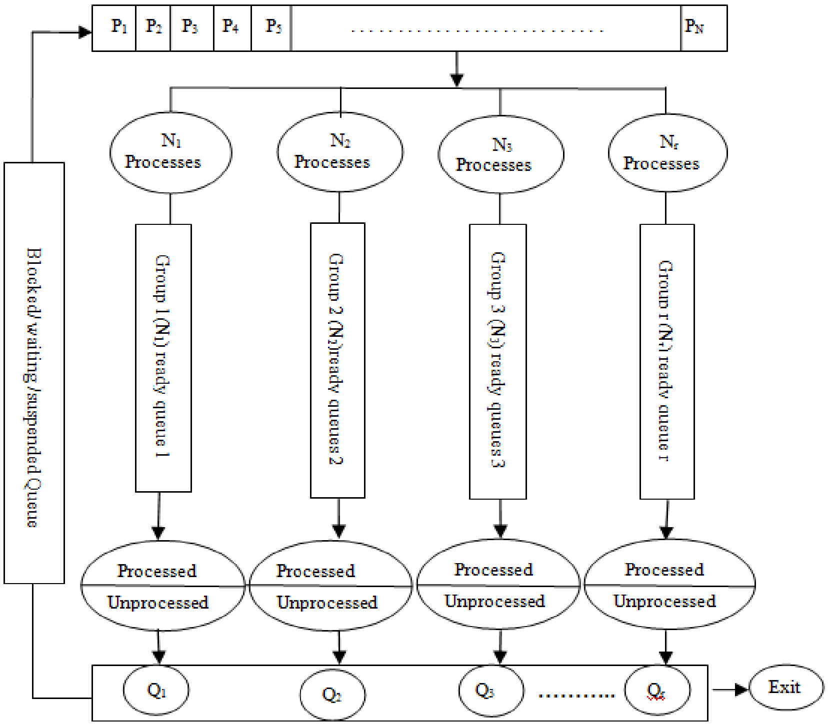


Figure 6: *Setup of ready queue and multiprocessor environment (due to[23])*

IV. Proposed Generalized Computational Setup

Assume the existence a virtual sampled ready queue in a computer system having multiprocessors environment. Some jobs are randomly selected using lottery scheduling from the ready queue and placed in the sampled ready queue from top to bottom in the sequential manner of their selection. Processors are assigned processes in the ordered manner from top to bottom of the virtual sampled ready queue. Figure 5 shows basic setup of this approach but without the size measure while figure 5 shows the earlier approaches [4], [5], [6], [7]. Moreover, figure 6 reveals the special case when all sample units processed before the occurrence of breakdown.

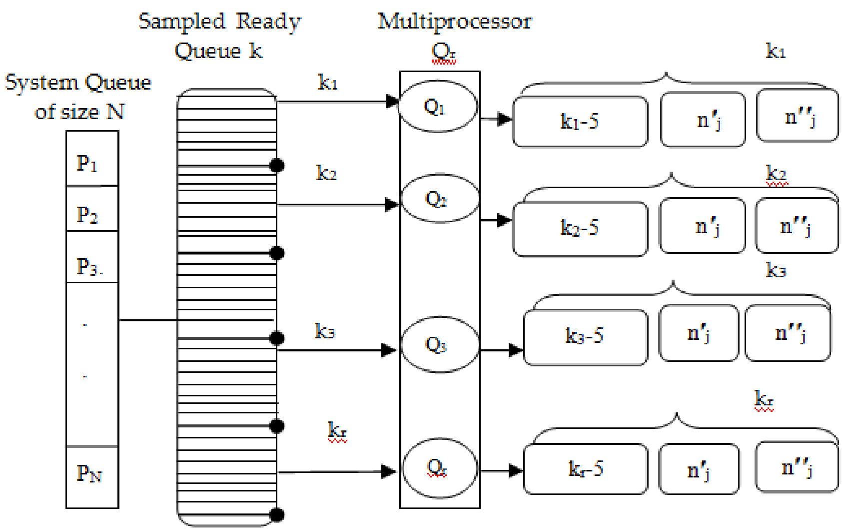


Figure 7: *Sampled Ready Queue Processing Time Estimation setup without size measure*

V. Generalized Assumption and Model

As per figure 7, let the selection of processes is according to lottery scheduling. The process who selects first is placed at the top of the virtual queue who is segment or group of processes likely to allocate to the multi-processors.

1. Assume r processors and a ready queue of N processes in a system like denoted as [P_1, P_2, P_3……..P_N] waiting for allocation of resources.
2. The selection of process for resource allocation is on priority basis using lottery scheduling.
3. If all N are processed completely, time consumed are [t_1, t_2, t_3 ….t_N] who has known size measure [x_1, x_2, x_3 ….x_N].
4. Overall ready queue mean time $\bar{t}=\frac{1}{N}\sum_{i=1}^{N} t_i$, mean size measure $\bar{x}=\frac{1}{N}\sum_{i=1}^{N} x_i$ mean squares $S_t^2=\frac{1}{N-1}\sum_{i=1}^{N}(t_i-\bar{t})^2$, $S_x^2=\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2$.
5. The P_i of known size X_i consumes time t_i(i = 1,2,3,……N) when all assumed processed.
6. Consider r multiprocessors Q_1, Q_2, Q_3…..Q_r, (r < N) and time consumed by the i^{th} process in the j^{th} processor is t_{ij} with corresponding size measures x_{ij} (j = 1,2,3,……r)
7. The unknown total completion time of ready queue is $N\bar{t}$, which is an unknown quantity. This paper is focused to estimate such using sampling methodology. Lottery scheduling is a tool for such estimation where process P_i has a bunch of token numbers and Q_j generates a random number. A process who receives the random number gets the desired resource from Q_j. This scheduling produces a random sample.
8. A virtual ready queue of size k (k < N, k>3r) exists to store sequentially the records of randomly selected k processes from N. The j^{th} segment of virtual sampled queue is k_j($k=\sum_{j=1}^{r} k_j$), who is allocated to the j^{th} processor Q_j in sequential manner.
9. In sample, let sx_{jl} denotes the file size measure and st_{jl} denotes time consumed by i^{th} process in Q_j (l = 1,2,3,...k_j) when all processed completely who are included in the sample of size k.
   - Sample mean of time $\bar{st}=\frac{1}{k}\sum_{j=1}^{r}\sum_{l=1}^{k_j} st_{jl}$
   - Sample mean square of time, $(es)_t^2=\frac{1}{k-1}\sum_{j=1}^{r}\sum_{l=1}^{k_j}(st_{jl}-\bar{st})^2$
   - Sample mean of size, $(\bar{sx})=\frac{1}{K-1}\sum_{j=1}^{k_j}\sum_{l=1}^{k_j}(sx_{jl})$
   - Sample mean square of size, $(es)_x^2=\frac{1}{k-1}\sum_{j=1}^{r}\sum_{l=1}^{k_j}(sx_{jl}-\bar{sx})^2$

i. The term $\bar{st}$, $\bar{sx}$, $(es)_t^2$, $(es)_x^2$ hold when system runs without failure.

10. Assume system breakdown occurs at the time instant T and there are (k_j – n'_j – n''_j) processes completed in Q_j, but n'_j remain partially processed and n''_j remain unprocessed (blocked). This is an assumed generalized model shown in figure 7. Define $g=\sum_{j=1}^{r} n'_j$ and $u=\sum_{j=1}^{r} n''_j$
11. Let $(st')_{jl}$ is time consumed by the l^{th} process in the processor Q_j [l =1, 2, 3... (k_j – n'_j – n''_j)],who is among those processed completely before the occurrence of T.
12. Some sample mean related measures are:
   - Sample mean of (k_j – n'_j – n''_j) process, $(\bar{st}')_j=\frac{1}{(k_j-n'_j-n''_j)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(st'_{jl})$
   - Sample mean square, $(es')_j^2=\frac{1}{(k_j-n'_j-n''_j-1)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(st'_{jl}-(\bar{st}')_j)^2$
   - Similar is for size measure also as (sx'_{jl}) represents size of l^{th} process who is in Q_j before T.
   - Sample mean, $(\bar{sx}')_j=\frac{1}{(k_j-n'_j-n''_j)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(sx'_{jl})$
   - $(\bar{sx})_j=\frac{1}{(k_j-n'_j-n''_j)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(sx'_{jl})$is sample mean of all k_j known values related to x in j^{th} segment of ready queue.
   - Sample mean square, $(ex')_j^2=\frac{1}{(k_j-n'_j-n''_j-1)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(sx'_{jl}-(\bar{sx}')_j)^2$
   - Sample Covariance, $(es'x')_j=\frac{1}{(k_j-n'_j-n''_j-1)}\sum_{l=1}^{(k_j-n'_j-n''_j)}(st'_{jl}-(\bar{st}')_j)(sx'_{jl}-(\bar{sx}')_j)$

13. Assume t_m^* is partially processed time of a process in Q_j (j = m =1,2,3....r) whose sample mean under T is
14. $(\bar{t}^*/T) = \frac{1}{r}\sum_{m=1}^{r} t_m^*$, Variance $(\bar{t}^*/T) = V(\bar{t}^*/T) = \left(\frac{1}{g} - \frac{1}{N-k+g}\right) S_T^2$, where S_T^2 is the conditional ready queue mean square of the remaining un-sampled part [N- k + g] expressed as: $S_T^2 = \frac{1}{(N-k+g-1)}\sum_{i=i}^{N-k+g}(t_i - \bar{t}_T)^2$ where $\bar{t}_T = \frac{1}{N-k+g}\sum_{i=1}^{N-k+g}(t_i)$ where $g = \sum_{j=1}^{r} n'_j$
15. Herein to mention that S_T^2 and $\bar{t}_T$ contain time t_i only from non-sampled processes (N-k) of the main ready queue with the addition of those g who partially processed. For such, the size converts from N into (N - k + g) and only those processes are the part of $\bar{t}_T$ and S_T^2 who are in (N – k + g).
16. The u blocked processes are imputed by Random Imputation Method using random selection of a process among $(k_j - n'_j - n''_j)$ relating to Q_j. Let from Q_j all random imputed time are denoted as t_m^{**}.
    - Sample mean of all random imputed time, $\bar{t}^{**} = \frac{1}{u}\sum_{m=1}^{u} t_m^{**}$
    - Variance of imputation under T, $V(\bar{t}^{**}/T) = \left(\frac{1}{u} - \frac{1}{k}\right)(es)^2$, u < k.
17. Sample based estimate of $(es)^2$ can be obtained by using all k values of time consumption in sample including the partially processed time t_m^* and imputed time value t_m^{**}. It is denoted as $(es^*)^2$ and mathematically expressed as $(es^*)^2 = \frac{1}{k-1}\sum_{j=1}^{r}\sum_{l=1}^{k_j}(st^*_{jl} - \overline{st}^*)^2$ where (st^*_{jl}) and $\overline{st}^*$ include completely processed time st^*_{ij}, partially processed t_m^* and imputed t_m^{**}.
18. The sample estimate of S_T^2 is $(es')^2 = \frac{1}{g-1}\left[\sum_{m=1}^{g}(t_m^* - \bar{t}^*)^2\right]$
19. Bias of estimation strategy is assumed negligible wherever appears and applicable in mathematical expressions

I. Computational Set-up

Aim is to compute the remaining ready queue processing time after occurrence of sudden failure of system at time instant T. This is subject to condition that r processes are partially processed, r is unprocessed (blocked) and remaining fully completed. Blocked and partially processed are n_j' and n_j'' from every Q_j and known size measures are the part of computation. Some frequently used symbols for process time t and process size measure X are as under:

$$\bar{t} = \frac{1}{N}\sum_{i=1}^{N} t_i = \frac{1}{N}\sum\sum t_{ij} \quad (1)$$

$$\bar{t}^* = \frac{1}{g}\sum_{m=1}^{g} t_m^* \quad (2)$$

$$\bar{t}^{**} = \frac{1}{u}\sum_{m=1}^{u} t_m^{**} \quad (3)$$

$$(\overline{st}')_j = \frac{1}{(k_j - n'_j - n''_j - 1)}\sum_{j=1}^{(k_j - n'_j - n''_j)}(st'_{jl}) \quad (4)$$

$$(\overline{sx}')_j = \frac{1}{(k_j - n'_j - n''_j - 1)}\sum_{j=1}^{(k_j - n'_j - n''_j)}(sx'_{jl}) \quad (5)$$

$$(\overline{sx})_j = \frac{1}{(k_j)}\sum_{l=1}^{k_j}(sx'_{jl}) \quad (6)$$

$$(es')_j^2 = 1/(k_j - n'_j - n''_j - 1)\sum_{l=1}^{(k_j - n'_j - n''_j)}(st'_{jl} - (\overline{st}')_j)^2 \quad (7)$$

$$(ex')_j^2 = 1/(k_j - n'_j - n''_j - 1)\sum_{l=1}^{(k_j - n'_j - n''_j)}(sx'_{jl} - (\overline{sx}')_j)^2 \quad (8)$$

$$(es'x')_j = \frac{1}{(k_j - n'_j - n''_j - 1)}\sum_{l=1}^{(k_j - n'_j - n''_j)}\left(st'_{jl} - (\overline{st}')_j\right)\left(sx'_{jl} - (\overline{sx}')\right)_j \quad (9)$$

$$(es^*)^2 = \frac{1}{k-1}\sum_{j=1}^{r}\sum_{l=1}^{k_j}(st^*_{jl} - \overline{st}^*)^2 \quad (10)$$

$$\bar{t}_{rj} = [(\overline{st}')_j + d_j\{(\overline{sx})_j - (\overline{sx}')_j\}], \ d_j \text{ being constant, } (0 < d_j < \infty) \quad (11)$$

Note: The $\bar{t}_{rj}$ is a Difference type estimator as stated in subsection IV of section II.

II. Estimation Strategy

The sample based proposed estimation strategy for mean time is:

$$(t_{mean}/T) = €_1 [\sum_{j=1}^{r} w_j \ (\bar{t}_{rj}/T)] + €_2 \ (\bar{t}^{*}/ T) + \ (1- €_1 – €_2) \ (\bar{t}^{**}/T) \tag{12}$$

with condition that $\sum_{p=1}^{3} €_p = 1$ and $€_p$ denotes constants to be determine suitability and $w_j = (k_j/k)$ is known weight ($\sum w_j = 1$). With the help of Cochran [16; see page 166, page 27, 29] for t_{mean}, the expected value E[.] is expressed as:

$$E\,[t_{mean}/T] = E[\, €_1 [\sum_{j=1}^{r} w_j \ (\bar{t}_{rj}/T)] + €_2 \ (\bar{t}^{*}/T) + (1- €_1 – €_2) \ (\bar{t}^{**}/T)]$$

$$= €_1 [\sum_{j=1}^{r} W_j E \ (\bar{t}_{rj}/T)] + €_2 \, E\, (\bar{t}^{*}/T) + (1- €_1 – €_2) E\, (\bar{t}^{**}/T)] \tag{13}$$

$\neq \bar{t}$ which shows estimator (t_{mean}/T) is biased.

III. Mean Squared Error

Let MSE (.), V (.) and B (.) denote mean squared error, variance and bias respectively. One can express

MSE (t_{mean}/T) = Variance (t_{mean}/T) + [Bias $(t_{mean}/T)]^2$ which holds in general. Assume the bias is small, therefore negligible (as in assumption no. 16)

$$MSE\,(t_{mean}/T) = Variance\,(t_{mean}/T) = €_1^2 [\sum_{j=1}^{r} w_j^2 \, V(\bar{t}_{rj}/T)] + €_2^2 V\,(\bar{t}^{*}/T) + (1- €_1 – €_2) V\,(\bar{t}^{**}/T)]$$

$$= €_1^2 \Big[\sum_{j=1}^{r} \Big(\frac{1}{(k_j - n'_j - n''_j)} - \frac{1}{k}\Big) w_j^2 \{(es')_j^2 + d_j^2 (ex')_j^2 - 2d_j (es'x')_j\}\Big] + €_2^2 \Big[\Big(\frac{1}{g} - \frac{1}{N-k+g}\Big) s_T^2\Big] + (1- €_1 - €_2)^2 \sum_{j=1}^{r} \Big(1 - \frac{1}{(k_j - n'_j - n''_j)}\Big) w_j (es')_j^2 \Big]$$

(as per Cochran[12] page 24, page 29 and page 164) (14)

The expressions P, Q, R are in the sample based estimate form of population parameters

Let $P = \sum_{j=1}^{r} \Big(\frac{1}{(k_j - n'_j - n''_j)} - \frac{1}{k}\Big) w_j^2 \{(es')_j^2 + d_j^2 \, (ex')_j^2 - 2d_j \, (es'x')_j\}$,

$Q = \Big(\frac{1}{g} - \frac{1}{N-k+g}\Big) s_T^2$

$R = \sum_{j=1}^{r} \Big(1 - \frac{1}{(k_j - n'_j - n''_j)}\Big) w_j^2 (es')_j^2$

The above expression is re-written as:

$V[t_{mean/T}] = [€_1^2 \, P + €_2^2 \, Q + (1- €_1 - €_2)^2 R\,]$ ignoring the covariance terms due to independency. For optimum variance, differentiate $V[t_{mean}/T]$ with respect to $€_1$ and $€_2$ and equate to zero, one gets

$$(€_1)_{opt} = (QR) / [PQ+PR+QR] = QM \tag{15}$$

$$(€_2)_{opt} = PQ/ [PQ+PR+QR] = PM \text{ where } M = R/ [PQ+PR+QR] \tag{16}$$

One can differentiate the variance expression by d_j also to get optimum value which is $(d_j)_{opt} = [(es'x')_j/(ex')_j^2]$ Substituting optimum choices in expression, the optimum variance is:

$$V[t_{mean}/T]_{opt} = (€_1)^2_{opt} \, P + (€_2)^2_{opt} \, Q + (1- (€_1)_{opt} - (€_2)_{opt})^2 R] \text{ with } (d_j)_{opt} \tag{17}$$

VI. Numerical Illustration

Consider the 150 processes with processed CPU time whose details are in table 1 with assumption that all 150 processes have been completed.

Table 1: *System Ready Queue Processes with time (N = 150)*

Process	J_{1}	J_{2}	J_{3}	J_{4}	J_{5}	J_{6}	J_{7}	J_{8}	J_{9}	J_{10}	J_{11}	J_{12}	J_{13}	J_{14}	J_{15}
CPU Time	30	20	42	45	59	35	25	48	50	60	32	55	62	47	69
Process Size	41	71	103	142	316	82	199	163	220	127	76	192	251	52	133
Process	J_{16}	J_{17}	J_{18}	J_{19}	J_{20}	J_{21}	J_{22}	J_{23}	J_{24}	J_{25}	J_{26}	J_{27}	J_{28}	J_{29}	J_{30}
CPU Time	34	24	44	70	57	65	38	84	101	66	80	90	92	111	85
Process Size	318	202	106	181	242	148	46	252	136	222	261	97	109	271	116
Process	J_{31}	J_{32}	J_{33}	J_{34}	J_{35}	J_{36}	J_{37}	J_{38}	J_{39}	J_{40}	J_{41}	J_{42}	J_{43}	J_{44}	J_{45}
CPU Time	61	52	72	75	89	67	51	78	80	91	63	86	93	77	99
Process Size	172	243	253	262	83	203	183	166	219	193	223	272	281	301	289
Process	J_{46}	J_{47}	J_{48}	J_{49}	J_{50}	J_{51}	J_{52}	J_{53}	J_{54}	J_{55}	J_{56}	J_{57}	J_{58}	J_{59}	J_{60}
CPU Time	64	54	74	100	87	95	68	114	131	96	110	123	122	141	49
Process Size	205	244	223	254	146	263	53	218	273	139	282	302	173	309	290
Process	J_{61}	J_{62}	J_{63}	J_{64}	J_{65}	J_{66}	J_{67}	J_{68}	J_{69}	J_{70}	J_{71}	J_{72}	J_{73}	J_{74}	J_{75}
CPU Time	118	81	102	105	119	97	88	108	110	121	240	113	122	107	129
Process Size	313	194	153	255	225	169	206	264	58	274	283	303	184	291	216
Process	J_{76}	J_{77}	J_{78}	J_{79}	J_{80}	J_{81}	J_{82}	J_{83}	J_{84}	J_{85}	J_{86}	J_{87}	J_{88}	J_{89}	J_{90}
CPU Time	94	73	104	130	117	234	98	237	161	126	143	236	152	171	233
Process Size	207	246	228	360	256	275	217	265	226	195	284	292	304	300	280
Process	J_{91}	J_{92}	J_{93}	J_{94}	J_{95}	J_{96}	J_{97}	J_{98}	J_{99}	J_{100}	J_{101}	J_{102}	J_{103}	J_{104}	J_{105}
CPU Time	120	112	132	135	149	125	115	138	140	150	122	232	152	137	159
Process Size	247	79	208	276	285	257	56	293	266	187	305	178	310	299	215
Process	J_{106}	J_{107}	J_{108}	J_{109}	J_{110}	J_{111}	J_{112}	J_{113}	J_{114}	J_{115}	J_{116}	J_{117}	J_{118}	J_{119}	J_{120}
CPU Time	124	114	134	160	147	155	128	174	191	156	170	180	182	201	175
Process Size	277	286	211	248	227	294	157	258	229	267	196	298	188	306	270
Process	J_{121}	J_{122}	J_{123}	J_{124}	J_{125}	J_{126}	J_{127}	J_{128}	J_{129}	J_{130}	J_{131}	J_{132}	J_{133}	J_{134}	J_{135}
CPU Time	235	142	162	165	179	151	145	168	171	238	152	175	189	167	241
Process Size	287	278	295	197	249	307	268	311	213	350	112	314	259	297	230
Process	J_{136}	J_{137}	J_{138}	J_{139}	J_{140}	J_{141}	J_{142}	J_{143}	J_{144}	J_{145}	J_{146}	J_{147}	J_{148}	J_{149}	J_{150}
CPU Time	154	144	164	190	177	185	158	204	221	186	200	210	212	231	209
Process Size	214	250	260	279	288	296	308	269	312	245	317	198	319	315	239

Table 2: *Descriptive Statistics of Table 1*

S. No.	Parameters Name	Calculated value
1	Number of Processes N	150
2	Mean time ($\bar{t}$)	122.56

I. Case-I: where each sample size k=40, and $d_j = 0$ ($d_1 = 0$, $d_2 = 0$, $d_3 = 0$)

Table 3: *Calculation for Sample No. 1*

k_1:16	**k_2:13**	**k_3:11**
{(J_{01}),(30),(41)},{(J_{31}),(61),(172)}, {(J_{61}),(118),(313)},{(J_{91}),(120),(247)}, {(J_{121},(235),287)},{(J_{63}),(102),(153)}, {(J_{32}),(52),(243)},{(J_{62}),(81),(194)}, {(J_{92}),(112),(79)},{(J_{122}),(142),(278)}, {(J_{3}),(42),103)},{(J_{33}),(72),(253)}, {(J_{141}),(185),(296)},{(J_{21}),(65),(148)}, {(J_{86}),(143),(284)},{(J_{100}),(150),(187)}	{(J_{49}),(100),(254)},{(J_{34}),(75),(262)}, {(J_{64}),(105),(255)},{(J_{94}),(135),(276)}, {(J_{124}),(165),(197)},{(J_{135}),(241),(230)} {(J_{35}),(89),(83)},{(J_{65}),(119),(225)}, {(J_{95}),(149),(285)},{(J_{150}),(209),(239)}, {(J_{99}),(140),(266)},{(J_{143}),(204),(269)}, {(J_{116}),(170),(196)}	{(J_{29}),(111),(271)},{(J_{59}),(141),(309)} {(J_{28}),(92),(109)},{(J_{96}),(125),(257)} {(J_{119})(201)(306)},{(J_{149})(231)(315)}, {(J_{142}),(158),(308)},{(J_{97}),(115),(56)}, {(J_{108}),(134),(211)},{(J_{112})(128)(157)}, {(J_{120}), (175), (270)}
$n_i' = 2$, $n_i'' = 3$	$n_i' = 2$, $n_i'' = 2$	$n_i' = 2$, $n_i'' = 3$
Partial Processed ={(J_{33}),(72)(253)}{(J_{141}),(185),(296)} (Processed=50 unprocessed=22) (Processed=90 unprocessed=95) Blocked = {(J_{21}),(65),(148)}, {(J_{86}),(143),(284)},{(J_{100}),(150),(187)} Blocked replaced α_1' = {(J_{91}),(120),(247)} α_2' ={(J_{32}),(52),(243)} α_3' = {(J_{01}),(30),(41)} [$\bar{st}_1'$ = 99.54, from eq.(4.4), $(es')_1^2$ = 3330.87, from eq.(4.7)], [$\bar{s}x_1$=3583/16 = 223.94, from eq.(4.5) $\bar{s}x_1'$=2110/11=191.81 from eq. (4.6)], [$(ex')_1^2$ =8210.36, from eq.(4.8)] [$(es'x')_1$ = 3230.60, from eq.(4.9)]	Partial Processed={(J_{150})(209)(239)} {(J_{99}),(140),(266)} (Processed=120, unprocessed=89) (Processed=90, unprocessed=50), Blocked={(J_{143}),(204),(269)}, {(J_{116}),(170),(196)} Blocked replaced β_1' ={(J_{64}),(105),(255)} β_2' ={(J_{135}),(241),(230)} [$\bar{st}_2'$ = 130.88, from eq. (4.4), $(es')_2^2$ = 2534.61 from eq.(4.7)] [$\bar{s}x_2$ =3149/13 =242.23, from eq.(4.5), $\bar{s}x_2' = \frac{2067}{9}$ = 229.66, from eq. (4.6), $(ex')_2^2$ =3761,from eq.(4.8)] [$(es'x')_2$ = 387.45, from eq.(4.9)]	Partial Processed = {(J_{142})(158)(308)} {(J_{97})(115)(56)} (Processed=110unprocessed=48), (Processed=65 unprocessed=55), Blocked={(J_{108}),(134),(211)}, {(J_{112})(128)(157)},{(J_{120})(175)(270)} Blocked replaced γ_1' = {(J_{119})(201)(306)} γ_2' = {(J_{59}),(141),(309)} γ_3' ={(J_{29}),(111),(271)} [$\bar{st}_3'$ = 150.16, from eq.(4.4), $(es')_3^2$ = 2950.56 from eq.(4.7)] [$\bar{s}x_3$ =2641/11=240.09,from eq.(4.5) $\bar{s}x_3' = \frac{1567}{6}$ = 261.16, from eq. (4.6) [$(ex')_3^2$ = 6092.96, from eq.(4.8)] [$(es'x')_3$ = 2952.56, from eq.(4.9)]

$\bar{t}^*$ = (50+90+120+90+110+65)/6 = 87.5

$\bar{t}^{**}$ = (α' + β' + γ')/8 = (120+52+30+105+241+201+141+111) / 8 = 125.13

Estimated [s_T^2 = 2,204.16] (using point 15) S_T^2 is $(es')^2 = \frac{1}{g-1}[\sum_{m=1}^{g}(t_m^*-\bar{t}^*)^2]$

[(50-**87.5**)^2 +(90-**87.5**)^2 +(190-**87.5**)^2 +(110-**87.5**)^2 +(140-**87.5**)^2+(95-**87.5**)^2]/5

=[4,333.58+1,167.58+4,117.78+250.58+200.78+950.48]= 2,204.16

Let $P = \sum_{j=1}^{r}\left(\frac{1}{(k_j - n'_j - n''_j)} - \frac{1}{k}\right) w_j^2 \{(es')_j^2 + d_j^2 (ex')_j^2 - 2d_j (es'x')_j\}$, $Q = \left(\frac{1}{g} - \frac{1}{N-k+g}\right)s_T^2$

$R = \sum_{j=1}^{r}\left(1 - \frac{1}{(k_j - n'_j - n''_j)}\right) w_j^2 (es')_j^2$

$P = \left(\frac{1}{16-2-3} - \frac{1}{40}\right)(0.4)^2 * \{3330.87\} + \left(\frac{1}{13-2-2} - \frac{1}{40}\right)(0.33)^2\{2534.61\} + \left(\frac{1}{11-2-3} - \frac{1}{40}\right)(0.28)^2\{2950.56\}$

= 0.0659 *0.16*3330.87+ 0.0861*0.1089*2534.61+0.1416*0.0784*2950.56= 91.64

$Q = (\frac{1}{3} - \frac{1}{150-40+3})$ 2,204.16 = 0.3245 *2,204.16 = 715.25

$R = (1 - \frac{1}{16-2-3})$ (0.4)²*3330.87+ $\left(1-\frac{1}{13-2-2}\right)$ (0.33)² *2534.61 +$\left(1-\frac{1}{11-2-3}\right)$ (0.28)² *2950.56

= 0.9091 *0.16*3330.87+ 0.8889*0.1089*2534.61 +0.8334*0.0784*2950.56 = 922.63

Calculation of mean and Variance V[t_{mean}/T] at d_j = 0 (for all j = 1,2,3)

$(€_1)_{opt}$= (QR) / [PQ+PR+QR] = QM = 715.25*922.63/[91.64*715.25+91.64*922.63+715.25*922.63]

= 659911.1075/810006.4307 = 0.8147

$(€_2)_{opt}$= PQ/ [PQ+PR+QR] = PM = 91.64*715.25/[91.64*715.25+91.64*870.50+715.25*870.50]

= 65545.51/ 810006.4307= 0.0809

$(t_{mean}/T) = (€_1)_{opt} [\sum_{j=1}^{r} w_j \bar{t}_{rj}] + (€_2)_{opt} (\bar{t}^*) + (1- (€_1)_{opt} - (€_2)_{opt}) (\bar{t}^{**})$

$\bar{t}_{rj} = [(\overline{st}')_j + d_j\{(\overline{sx})_j - (\overline{sx}')_j\}]$,

$\bar{t}_{rj}$= [0.4*99.54+0*(223.94-191.81)]+[0.33*130.88+0*(242.45-229.66)] +[0.28*150.16+0*(240.09-261.16)]

= 39.82+43.19+42.04 = 125.05

(t_{mean}/T) = 0.8147*125.05+0.0809***87.5**+0.1044***125.13 = 122.02**

$V[t_{mean}/T] = (€_1)^2_{opt} P + (€_2)^2_{opt} Q + (1- (€_1)_{opt} - (€_2)_{opt})^2 R]$

$V[t_{mean}/T]$ = [(0.8147)² *91.64+ (0.0809)²*715.25 + 0.0108*922.63] = 60.82+4.68+9.96 = 75.46

The 95% confidence intervals for $\bar{t}$, P [$(t_{mean}/T) \pm 1.96\sqrt{[V(t_{mean}/T)]}$] = 0.95

= 122.02± $1.96\sqrt{75.46}$ = 122.02 ± 17.02 = (104.99, 139.04)

Table 4: *Estimated Sample Mean, Variance and Confidence Interval (CI) of Ten Random Samples*

Case-I: At $(€_1)_{opt}$, $(€_2)_{opt}$, d_j = 0 (d_1 = 0, d_2 = 0, d_3 = 0) where True mean = 122.51				
S.No.	Estimated Sample Mean	V[t_{mean}/T]	95% Confidence Interval (CI)	CI Length
1	122.02	75.46	(104.99, 139.04)	34.05
2	134.58	64.83	(118.79, 150.36)	31.57
3	117.56	74.36	(100.66, 134.46)	33.80
4	113.89	48.45	(100.25, 127.53)	27.28
5	127.00	85.37	(108.89, 145.11)	36.22
6	119.27	46.42	(105.92, 132.62)	26.70
7	123.39	45.41	(110.18, 136.60)	26.42
8	113.12	97.36	(93.78, 132.46)	38.68
9	115.01	53.05	(100.73, 129.28)	28.55
10	120.21	60.91	(104.91, 135.51)	30.60
			Average Length (3138/10) = 31.38	

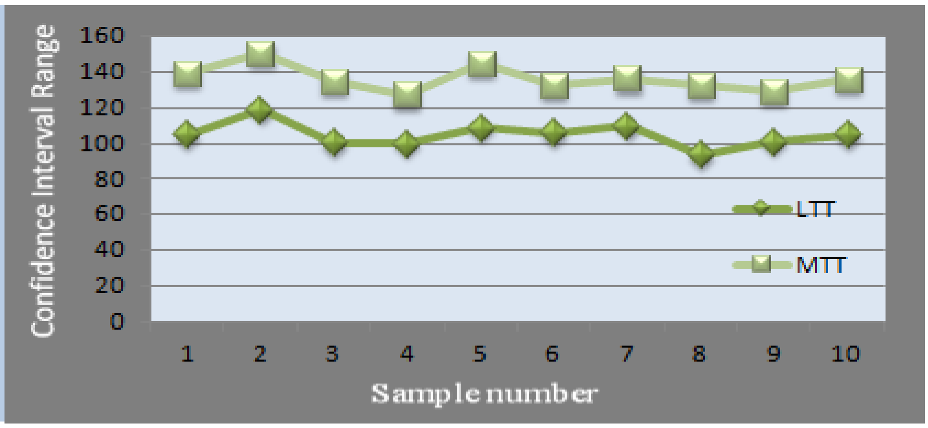


Figure 8: *Graphical representation of Confidence Interval range of Ten Random Samples for Case-I of Table 4 (X-axis has sample number as shown in table 4)*

II. Case-II: where each sample size k=40, and $(d_{opt})_j = (es'x')_j / (ex')_j^2$

Table 5: *Calculation for Sample No. 1*

k_1:16	k_2:13	k_3:11
{(J_{01}),(30),(41)},{(J_{31}),(61),(172)}, {(J_{61}),(118),(313)},{(J_{91}),(120),(247)}, {(J_{121},(235),287)},{(J_{63}),(102),(153)}, {(J_{32}),(52),(243)},{(J_{62}),(81),(194)}, {(J_{92}),(112),(79)},{(J_{122}),(142),(278)}, {(J_3),(42),103)},{(J_{33}),(72),(253)}, {(J_{141}),(185),(296)},{(J_{21}),(65),(148)}, {(J_{86}),(143),(284)},{(J_{100}),(150),(187)}	{(J_{49}),(100),(254)},{(J_{34}),(75),(262)}, {(J_{64}),(105),(255)},{(J_{94}),(135),(276)}, {(J_{124}),(165),(197)},{(J_{135}),(241),(230)} {(J_{35}),(89),(83)},{(J_{65}),(119),(225)}, {(J_{95}),(149),(285)},{(J_{150}),(209),(239)}, {(J_{99}),(140),(266)},{(J_{143}),(204),(269)}, {(J_{116}),(170),(196)}	{(J_{29}),(111),(271)},{(J_{59}),(141),(309)} {(J_{28}),(92),(109)},{(J_{96}),(125),(257)} {(J_{119})(201)(306)},{(J_{149})(231)(315)}, {(J_{142}),(158),(308)},{(J_{97}),(115),(56)}, {(J_{108}),(134),(211)},{(J_{112})(128)(157)}, {(J_{120}), (175), (270)}
$n_i' = 2$, $n_i'' = 3$	$n_i' = 2$, $n_i'' = 2$	$n_i' = 2$, $n_i'' = 3$
Partial Processed = {(J_{33}),(72),(253)}, {(J_{141}),(185),(296)} (Processed=50, unprocessed=22), (Processed=90, unprocessed=95),	Partial Processed={(J_{150}),(209),(239)}, {(J_{99}),(140),(266)} (Processed=120, unprocessed=89) (Processed=90, unprocessed=50),	Partial Processed = {(J_{142}),(158),(308)},{(J_{97})(115)(56)} (Processed=110, unprocessed=48), (Processed=65, unprocessed=55),
Blocked = {(J_{21}),(65),(148)}, {(J_{86}),(143),(284)},{(J_{100}),(150),(187)} Blocked replaced α_1'={(J_{91}),(120),(247)},α_2'={(J_{32}),(52),(243)} α_3' = {(J_{01}),(30),(41)}	Blocked={(J_{143}),(204),(269)}, {(J_{116}),(170),(196)} Blocked replaced β_1' ={(J_{64}),(105),(255)} β_2' ={(J_{135}),(241),(230)}	Blocked={(J_{108}),(134),(211)}, {(J_{112})(128)(157)},{(J_{120})(175)(270)} Blocked replaced γ_1'={(J_{119})(201)(306)},γ_2'={(J_{59}),(141),(309)} γ_3' ={(J_{29}),(111),(271)}
[$\bar{st}_1'$ = 99.54, from eq.(4.4), $(es')_1^2$ = 3330.87, from eq.(4.7)],[$\bar{sx}_1$=3583/16 = 223.94, from eq.(4.5), $\bar{sx}_1'$=2110/11=191.81from eq. (4.6)], [$(ex')_1^2$ =8210.36, from eq.(4.8)] ,[$(es'x')_1$ = 3230.60, from eq.(4.9)],$(d_{opt})_1 = (es'x')_1 / (ex')_1^2$ = 3230.60/8210.36 =0.3935	[$\bar{st}_2'$ = 130.88, from eq. (4.4), $(es')_2^2$ = 2534.61 from eq.(4.7)],[$\bar{sx}_2$ =3149/13 =242.23, from eq.(4.5),$\bar{sx}_2'$= $\frac{2067}{9}$ = 229.66, from eq. (4.6), $(ex')_2^2$ =3761,from eq.(4.8)],[$(es'x')_2$ = 387.45, from eq.(4.9)],$(d_{opt})_2$= $(es'x')_2 / (ex')_2^2$ = 387.45/3761 = 0.1030	[$\bar{st}_3'$ = 150.16, from eq.(4.4),$(es')_3^2$ = 2950.56 from eq.(4.7)] ,[$\bar{sx}_3$ =2641/11=240.09,from eq.(4.5),$\bar{sx}_3'$= $\frac{1567}{6}$ = 261.16, from eq. (4.6),[$(ex')_3^2$ = 6092.96, from eq.(4.8)],[$(es'x')_3$ = 2952.56, from eq.(4.9)],$(d_{opt})_3$= $(es'x')_3 / (ex')_3^2$ =2952.56/6092.96 = 0.48

$\bar{t}^*$ = (50+90+120+90+110+65)/6 = 87.5

$\bar{t}^{**} = (\alpha' + \beta' + \gamma')/8$ = (120+52+30+105+241+201+141+111) / 8 = 125.13

Estimated [s_T^2 = 2,204.16] (using point 15) S_T^2 is $(es')^2 = \frac{1}{g-1}[\sum_{m=1}^{g}(t_m^*-\bar{t}^*)^2]$

[(50-**87.5**)^2 +(90-**87.5**)^2 +(190-**87.5**)^2 +(110-**87.5**)^2 +(140-**87.5**)^2+(95-**87.5**)^2]/5

=[4,333.58+1,167.58+4,117.78+250.58+200.78+950.48]= 2,204.16

Let $P = \sum_{j=1}^{r}\left(\frac{1}{(k_j - n'_j - n''_j)}-\frac{1}{k}\right) w_j^2\{(es')_j^2+d_j^2 (ex')_j^2 -2d_j (es'x')_j\}$, $Q = \left(\frac{1}{g}-\frac{1}{N-k+g}\right)s_T^2$

$R = \sum_{j=1}^{r}\left(1-\frac{1}{(k_j - n'_j - n''_j)}\right) w_j^2 (es')_j^2$

$P = \left(\frac{1}{16-2-3}-\frac{1}{40}\right)(0.4)^2$*{3330.87+0.39*0.39*8210.36 - 2*0.39*3230.60}+ $\left(\frac{1}{13-2-2}-\frac{1}{40}\right)(0.33)^2$ {2534.61+0.10*0.10*3761- 2*0.10*387.45}+ $\left(\frac{1}{11-2-3}-\frac{1}{40}\right)(0.28)^2$ {2950.56 +0.48*0.48*6092.96- 2*0.48*2952.56}

= 0.0659 *0.16*2059.79+ 0.0861*0.1089*2494.73+0.1416*0.0784*1519.92 = 61.98

$Q = \left(\frac{1}{3}-\frac{1}{150-40+3}\right)$ 2,204.16 = 0.3245 *2,204.16 = 715.25

R = $\left(1-\frac{1}{16-2-3}\right)(0.4)^2$*3330.87+ $\left(1-\frac{1}{13-2-2}\right)(0.33)^2$*2534.61 +$\left(1-\frac{1}{11-2-3}\right)(0.28)^2$*2950.56

= 0.9091 *0.16*3330.87+ 0.8889*0.1089*2534.61 +0.8334*0.0784*2950.56 = 922.63

Calculation of mean and Variance V[t_{mean}/T] at $d_j = (d_{opt})_j$

$(€_1)_{opt}$= (QR) / [PQ+PR+QR] = QM = 715.25*922.63/[61.98*715.25+61.98*922.63+715.25*922.63]

= 659911.1075/761426.9099= 0.8666

$(€_2)_{opt}$= PQ/ [PQ+PR+QR] = PM = 61.98*715.25/[61.98*715.25+61.98*870.50+715.25*870.50]

= 44331.195/ 761426.9099 = 0.0582

(t_{mean}/T) = $(€_1)_{opt}$ [$\sum_{j=1}^{r} w_j\ \bar{t}_{rj}$] + $(€_2)_{opt}$ ($\bar{t}^*$) + (1- $(€_1)_{opt}$ – $(€_2)_{opt}$) ($\bar{t}^{**}$)

$\bar{t}_{rj}$= [$(\overline{st}')_j + d_j\{(\overline{sx})_j - (\overline{sx}')_j\}$],

$\bar{t}_{rj}$= [0.4*99.54+0.39*(223.94-191.81)]+[0.33*130.88+0.10*(242.45-229.66)] +[0.28*150.16+0.48*(240.09-261.16)] = [0.4*99.54+12.64]+[0.33*130.88+2.11] +[0.28*163.33-45.50] = 52.45+44.47+31.93 = 128.85

(t_{mean}/T) = 0.8666*128.85+0.0582***87.5**+0.0752***125.13 = 126.16**

V[t_{mean}/T] = $(€_1)^2_{opt}$ P + $(€_2)^2_{opt}$ Q + (1- $(€_1)_{opt}$ – $(€_2)_{opt}$)^{2}R]

V[t_{mean}/T] = [$(0.8666)^2$ *61.98+ $(0.0582)^2$*715.25 + 0.0056*922.63] = 46.54+2.42+5.17 = 54.13

The 95% confidence intervals for $\bar{t}$, P [(t_{mean}/T) ± 1.96$\sqrt{[V(t_{mean}/T)]}$ = 0.95

= **126.16** ± 1.96$\sqrt{54.13}$ = **126.16** ± 14.42 = (111.74, 140.58)

Table 6: *Estimated Sample Mean, Variance and Confidence Interval (CI) of Ten Random Samples*

Case-II: At $(€_1)_{opt}$, $(€_2)_{opt}$, $(d_{opt})_j$ = $(es'x')_j$ / $(ex')_j^2$ where True mean = 122.51				
S.No.	Estimated Sample Mean	V[t_{mean}/T]	95% Confidence Interval (CI)	CI Length
1	**126.16**	54.13	(111.74, 140.58)	28.84
2	130.78	39.24	(118.50, 143.06)	24.56
3	125.24	48.98	(111.52, 138.96)	27.44
4	124.84	45	(111.70, 137.99)	26.29
5	128.89	53.58	(114.54, 143.24)	28.7
6	140.30	100.86	(120.62, 159.98)	39.36
7	125.99	29.81	(115.29, 136.69)	21.4
8	110.79	77.25	(93.56, 128.02)	34.46
9	128.36	50.01	(114.50, 142.22)	27.72
10	128.07	38.42	(115.92, 140.22)	24.3
Average Length (28307/10) = 28.30				

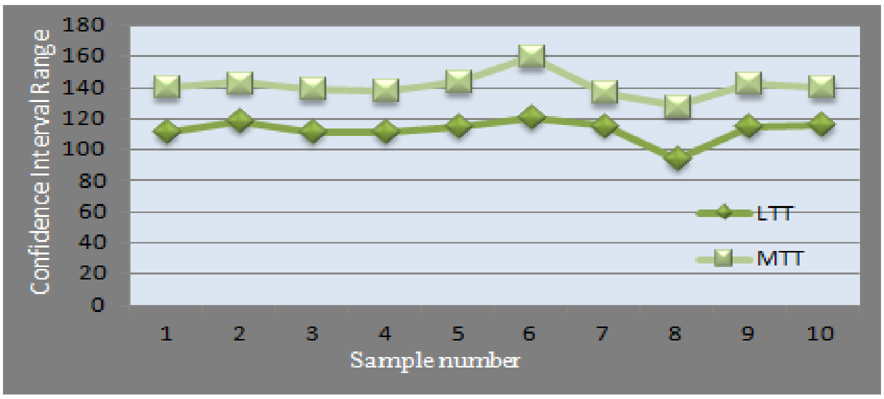


Figure 9: *Graphical representation of Confidence Interval range of Ten Random Samples for Case-II of Table 6 (X-axis has sample number as shown in table 6)*

Table 7: *Comparison between Case-I and Case-II*

S. NO	CASE-I $d_j = 0$ ($d_1 = 0$, $d_2 = 0$, $d_3 = 0$)		CASE-II $(d)_j = (d_{opt})_j$	
	95% Confidence Interval	**Length**	**95% Confidence Interval**	**Length**
1.	(104.99, 139.04)	34.05	(111.74, 140.58)	28.84
2.	(118.79, 150.36)	31.57	(118.50, 143.06)	24.56
3.	(100.66, 134.46)	33.8	(111.52, 138.96)	27.44
4.	(100.25, 127.53)	27.28	(111.70, 137.99)	26.29
5.	(108.89, 145.11)	36.22	(114.54, 143.24)	28.7
6.	(105.92, 132.62)	26.7	(120.62, 159.98)	39.36
7.	(110.18, 136.60)	26.42	(115.29, 136.69)	21.4
8	(93.78, 132.46)	38.68	(93.56, 128.02)	34.46
9.	(100.73, 129.28)	28.55	(114.50, 142.22)	27.72
10.	(104.91, 135.51)	30.6	(115.92, 140.22)	24.3
Average Length (3138/10)		**31.38**	**Average Length (2830/10)**	**28.30**

Table 8: *Case-I: Cost aspect when C_0 = 100 units, C_1 = 10 units*

S. NO	C.I Lower Limit	C I Upper Limit	δ_1	δ_2	Total cost			Total cost		
			$\delta 1$	$\delta 2$	$(T_c)_{1A}$	$(T_c)_{1B}$	$(T_c)_{1C}$	$(T_c)_{2A}$	$(T_c)_{2B}$	$(T_c)_{2C}$
1	104.99	139.04	11,548.90	15,294.40	115589	1174.65	1333771012	153044	1336.7053	2339186814
2	118.79	150.36	13,066.90	16,539.60	130769	1243.10	1707438856	165496	1386.063762	2735583782
3	100.66	134.46	11,072.60	14,790.60	110826	1152.26	1226024808	148006	1316.166107	2187618584
4	100.25	127.53	11,027.50	14,028.30	110375	1150.11	1216057663	140383	1284.411246	1967932109
5	108.89	145.11	11,977.90	15,962.10	119879	1194.43	1434700984	159721	1363.412047	2547886464
6	105.92	132.62	11,651.20	14,588.20	116612	1179.40	1357504714	145982	1307.816211	2128155892
7	110.18	136.6	12,119.80	15,026.00	121298	1200.89	1468895620	150360	1325.805857	2257806860
8	93.78	132.46	10,315.80	14,570.60	103258	1115.66	1064157396	145806	1307.087404	2123023944
9	100.73	129.28	11,080.30	14,220.80	110903	1152.63	1227730581	142308	1292.509958	2022311626
10	104.91	135.51	11,540.10	14,906.10	115501	1174.24	1331739180	149161	1320.905402	2221918272
	Average				115501	1173.74	1336802081	150026.7	1324.088329	2253142435

NOTE 8.1: Overall average cost by lower limit = (115501 +1173.743546+ 1336802081)/3
= 445639585.25 units

NOTE 8.2: Overall average cost by upper limit = (150026.7 + 1324.088329+ 2253142435)/3
= 751097928.59 units

Table 9: *Case- II: Cost aspect when C_0 = 100 units, C_1 = 10 units*

S. NO	C.I Lower Limit	C I Upper Limit	δ_1	δ_2	Total cost $(T_c)_{1A}$	$(T_c)_{1B}$	$(T_c)_{1C}$	Total cost $(T_c)_{2A}$	$(T_c)_{2B}$	$(T_c)_{2C}$
1	111.74	140.58	12,291.40	15,463.80	123014	1208.665865	1510785240	154738	1343.535283	2391291204
2	118.5	143.06	13,035.00	15,736.60	130450	1241.709245	1699112350	157466	1354.456057	2476405896
3	111.52	138.96	12,267.20	15,285.60	122772	1207.573925	1504842058	152956	1336.349465	2336495774
4	111.7	137.99	12,287.00	15,178.90	122970	1208.46741	1509703790	151889	1332.026785	2303990152
5	114.54	143.24	12,599.40	15,756.40	126094	1222.47049	1587448904	157664	1355.244996	2482641510
6	120.62	159.98	13,268.20	17,597.80	132782	1251.87673	1760451412	176078	1426.566998	3096825748
7	115.29	136.69	12,681.90	15,035.90	126919	1226.139423	1608305976	150459	1326.209607	2260782988
8	93.56	128.02	10,291.60	14,082.20	103016	1114.475234	1059170406	140922	1286.684457	1983083668
9	114.5	142.22	12,595.00	15,644.20	126050	1222.274476	1586340350	156542	1350.767764	2447410036
10	115.92	140.22	12,751.20	15,424.20	127612	1229.212115	1625931114	154342	1341.942028	2379059556
			Average value		124167.9	1213.286491	1545209160	155305.6	1345.378344	2415798653

NOTE 9.1: Overall average cost by lower limit = (124167.9+1213.286491+ 1545209160)/3
= 515111513.72 units

NOTE 9.2: Overall average cost by upper limit = (155305.6+ 1345.378344+ 2415798653)/3
= 805318434.65 units

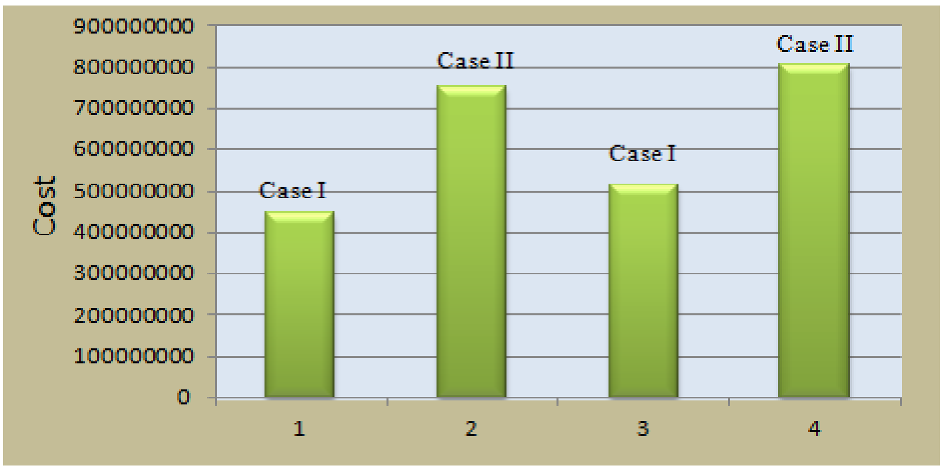


Figure 10: *Pair of graph lines for Case-I and Case-II*

VII. Discussion

In Section VI the data description is in Table 1 and 2 where 150 processes are presented assuming all finished before T. Their total processing time and size process measures are noted. The proposed estimate t_{mean} has unknown constants $€_1$, $€_2$ and d whose suitable values need to be obtained for obtaining a best estimate. Two cases are considered herein as

Case I: $€_1 = (€_1)_{opt}, €_2 = (€_2)_{opt,}$ and $d_1 = 0$, $d_2 = 0$, $d_3 = 0$.

This case indicates for no use of size measure in the estimation strategy at the optimum choice of $€_1$ and $€_2$. The average confidence interval length, under Case-I is 31.38 as evident form table 3. The lowest predicted total remaining time is 11540.1 units while highest is 14991.9 units (table 10). Average cost consumption for lowest estimated time is 445639585.25 units and at highest time level it is 751097928.59 units (table 8).

Case II: $€_1 = (€_1)_{opt}, €_2 = (€_2)_{opt,}$ and $d_1 = (d_{opt})_1$, $d_2 = (d_{opt})_2$, $d_3 = (d_{opt})_3$

This case contains choice of all constants at the optimum level and size measure information x has also been used. The impact of using the support information seems positive since the average length reduced to 30.53 in this case with respect to Case-I while simulated over 10 samples. Figure 9 also reveals for more condensed pair of graph lines for Case-II. Lowest predicted remaining time is 12406.9 units and highest is 15521 units (Table 10). Average cost likely to consume is 515111513.72 units as minimum whereas 805318434.65 units as highest (Table 9).

The percentage relative efficiency of Case-II with respect to Case-I is 9.82 % which supports the use of size measure in estimation (Table 2). The highest cost by Case-I and lowest by Case-II are the recommended cost required for infrastructure creation for backup management (Figure 10).

Table 10: *Ten Sample average Confidence Interval and estimated total Remaining time of processing for Recovery Management*

	Case-I (Without size measure)	**Case-II (With size measure)**	**True Value**
Average Interval (Over 10 samples)	(104.91 - 136.29)	(112.79 - 141.10)	122.51
CI Length	31.38	28.30	
Lowest Predicted Remaining time	(N-k)* 104.91 = 11540.1 units	(N-k)* 112.79 = 12406.9 units	------
Highest Predicted Remaining time	(N-k)* 136.29= 14991.9 units	(N-k)* 141.10 = 15521 units	

$$\textbf{Percentage Relative Efficiency (PRE)} = \left[\frac{[\text{Length of CI of case}-\text{I}]-[\text{Length of CI of other cases}]}{\text{Length of CI of case}-\text{I}} \right] \text{X } 100$$

Table 11: *Percentage Relative Efficiency (PRE)*

Case-II with respect to Case-I
PRE = 9.82 %

VIII. Conclusion

In case when the sudden breakdown occurs in a multiprocessor computer system this paper represents an idea of calculating the ready queue remaining processing time. The paper assumes that $(k_j - n_j' - n_j'')$ processes are completely finished before breakdown, n_j' are partially processed and n_j'' are blocked by j^{th} processor. Under this an estimation strategy is proposed for estimating the total remaining time of jobs to be processed in waiting ready queue. The proposed generalized strategy contains constants whose optimum values are derived and used. Two cases are compared where the first case is having no consideration of size measure of jobs in waiting queue whereas the second case considers the additional features of size measure of processes. The confidence interval is used as a tool for predicting about the unknown with 95% accuracy. Three cost functions are suggested for predicting about the backup infrastructure cost needed for recovery management after system breakdown. The proposed methodology under Case-II performs better than Case-I by comparing the length of confidence intervals. The highest predicted remaining time under ten considered samples is 15521 units, under Case-II. Moreover, the Case-II is 9.82 % more efficient than Case-I. The average cost required for recovery after occurrence of failure is also lower in Case-II. Overall it is found that the suggested estimation strategy is effective for predicting the remaining total time with high efficiency. The suggested is a new methodological approach for predicting the unknown using sampling methodology in the multiprocessor environment. Proposed advocates for the use of size measure of processes, if available for predicting unknown parameters like remaining time of a ready queue.

References

[1] More Sarla, and Shukla Diwakar, (2020). Some new methods for ready queue processing time estimation problem in a multiprocessor environment, *Social Networking and Computational Intelligence, Lecture notes in Networks and Systems, Springer, Singapore*, available at doi.org/10.1007/978-981-15-2071-6_54, Vol. 100, pp 661-670.

[2] More, Sarla and Shukla Diwakar, (2019). Analysis, and extension of methods in ready queue processing time Estimation in Multiprocessor Environment, *Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity university Rajasthan, Jaipur-India*, available at SSRN: https://ssrn.com/ abstract = 3356312 or https:// dx.doi.org/ 10.2139/ SSRN 3356312, pp 1558-1563.

[3] More, Sarla and Shukla Diwakar, (2018). A review on ready queue processing time estimation problem and methodologies used in multiprocessor environment, *International Journal of Commuter Science and Engineering,* available at https://doi.org/10.26438/ijcse/v6i5.11511155, Vol.6, Issue 5, pp 1186-1191.

[4] Diwakar Shukla and Sarla More, (2020). Modified group lottery scheduling algorithm for ready queue mean time estimation in multiprocessor environment, *Reliability: Theory & Applications (RT&A),* Vol. 15,No 4(59), pp 69-85.

[5] Shukla Diwakar, Jain Anjali and Choudhary Amita, (2010). Prediction of ready queue processing time in multiprocessor environment using lottery scheduling (ULS), *International Journal of Commuter Internet and Management*, Vol.18, No.3, pp 58-65.

[6] Shukla Diwakar, Jain Anjali and Choudhary Amita, (2010). Estimation of ready queue Processing time under usual group lottery scheduling (GLS) in multiprocessor environment, *International Journal of Commuter Applications*, Vol.8, No.14, pp 39-45.

[7] Shukla Diwakar, Jain Anjali, and Choudhary Amita, (2010). Estimation of ready queue processing time under SL scheduling scheme in multiprocessors environment, *International Journal of Computer Science and Security*, Vol. 4, Issue 1, pp 74-81.

[8] Carl A. Waldspurger and E William Weihl, (1994). Lottery Scheduling: Flexible proportional share resource management, *The 1994 Operating systems design and implementation conference (OSDI '94), Monterey, California*.

[9] Johnnie Daniel, (2011). Sampling Essentials: Practical Guidelines for Making Sampling Choices, *Sage Publication*.

[10] Paul S. Levy and Stanley Lemeshow, (2008). Sampling of Populations: Methods and Applications, *Wiley Series in Survey Methodology,* Volume 543.

[11] Sampath, S., (2005). Sampling Theory and Methods, *Alpha Science International Publication.*

[12] Cochran, W.G, (2005). Sampling Technique, *Wiley Eastern publication, New Delhi*.

[13] Poduri S. R. S. Rao, (2000). Sampling Methodologies with Applications, *Texts in Statistical Science, Chapman and Hall/CRC Press.*

[14] Ranjan K. Som, (1995). Practical Sampling Techniques, *Second Edition Statistics: A Series of Textbooks and Monographs, CRC Press.*

[15] Steven K. Thompson, (1992). Sampling, *Wiley Series in Probability and Statistics*, Volume 272.

[16] Nurbek Saparkhojayev, Yermek Nugmanov, Amy Apon, Mamyrbek Beysenbi, (2013). Dynamic Lottery Scheduling, *AWERProcedia Information Technology & Computer Science,* Vol 03 3rd World Conference on Information Technology (WCIT-2012), pp1310-1318.

[17] J. Prassanna and Neelanarayanan Venkataraman (2019). Adaptive regressive holt–winters workload prediction and firefly optimized lottery scheduling for load balancing in cloud, *Springer Science+Business Media, LLC, part of Springer Nature, Wireless Networks* https://doi.org/10.1007/s11276-019-02090-8

[18] Hala ElAarag, David Bauschlicher, and Steven Bauschlicher, (2011). Simulation-Based Comparison Of Scheduling Techniques In Multiprogramming Operating Systems on Single and Multi-Core Processors, *The Journal of Computing Sciences in Colleges*, Volume 27, Number 2 Papers of the Twentieth Annual CCSC Rocky Mountain Conference October 14-15, Utah Valley University Orem, Utah

[19] Emily Berg, Jae-Kwang Kim, Chris Skinner, (2016). Imputation Under Informative Sampling, *Journal of Survey Statistics and Methodology*, Volume 4, Issue 4, https://doi.org/10.1093/jssam/smw032, pp 436–462.

[20] Graham Kalton & Leslie Kish, (2007). Some efficient random imputation methods, *Communications in Statistics - Theory and Methods* , Volume 13, - Issue 16, pp 1919-1939.

[21] David A. Binder and Weimin Sun, Frequency Valid Multiple Imputation for Surveys with a Complex Design Statistics Canada Business Survey Methods Division, Statistics Canada, Ottawa, ON, Canada K IA 0T6

[22] J. K. Kim, S. Yang (2017). A note on multiple imputation under complex sampling, *Biometrika*, Volume 104, Issue 1, pp 221–228,

[23] Michael R. Elliott (2021). Weighted Dirichlet Process Mixture Models to Accommodate Complex Sample Designs for Linear and Quantile Regression, *Journal of official Statistics*, http://dx.doi.org/10.2478/JOS-2021-0004, Vol. 37, No. 1, pp. 71–95

[24] Dimitris Bertsimas, Colin Pawlowski, and Ying Daisy Zhuo, (2018). From Predictive Methods to Missing Data Imputation: An Optimization Approach, *Journal of Machine Learning Research* 18, 1-39.

[25] B. K. Singh and Upasana Gogoi, (2017). Estimation of Population mean using Exponential Dual to Ratio Type Compromised Imputation for Missing data in Survey Sampling, *Journal of Statistics Applications & Probability An International Journal*, Vol. 6, No. 3, 515-522

[26] Sarla More and Diwakar Shukla (2021). Sampled Ready queue processing time estimation using size measure information in multiprocessor environment, *Reliability: Theory and application (RT&A)*, No 3(63), vol.16, pp 63-80.