ISSN:2320-0790

COMPUSOFT
An International Journal of Advanced Computer Technology

Estimation of Ready Queue Processing Time using Factor-Type (F-T) Estimator in Multiprocessor Environment

Anjali Jain[1], Diwakar Shukla[2]

[1]DCSA, Dr. H.S. Gour Central University, Sagar M.P., INDIA
[2]Department of Mathematics and Statistics, Dr. H.S.Gour Central University, Sagar, M.P., INDIA

Abstract: The ready queue processing estimation problem appears when many processes remain in the ready queue after the sudden failure. The system manager has to decide immediately how much further time is required to process all the remaining jobs in the ready queue. In lottery scheduling, this prediction is possible with the help of sampling techniques. Ratio method, existing in sampling literature, was previously used by authors to predict the time required for remaining jobs to finish after failure, provided that highly correlated source of auxiliary information provides better processing time prediction.

This paper proposes two new estimators T_A and T_B which are compared with previously defined ratio estimator in terms of total processing time. Under large sample approximation the bias and m.s.e of proposed estimators have been obtained in the set up of lottery scheduling. The confidence intervals are calculated for the numerical support to the theoretical findings.

Keywords: Lottery Scheduling, Factor Type Estimator, Bias, Mean Squared Error (M.S.E), Variance, Confidence Intervals.

I. INTRODUCTION

Suppose that there are k processors in a multiprocessor and multi-user environment and a large number of processes, say N, are in waiting queue. The scheduler adopts lottery scheduling procedure to choose randomly any n processes from waiting queue (n<N) and allocates to k processors in sequential manner. Lottery scheduling is different from basic scheduling algorithm where each process is allocated a number of lottery tickets determining the possibility of process when to use the CPU. At each schedule point, a lottery is held and the process in the ready queue with the winning ticket gets the CPU utilization. Unlike priority scheduling every job has equal chance of being represented to the processors. Lottery scheduling does not suffer from starvation.

The technical problem appears when N is very large, the congestion in processing occurs, many processes have to wait until they are called in random manner. If suddenly the system collapse due to failure of power supply, maintenance problem, technical faults or any other, the system manager has to rely on backup management. His problem at this juncture is to know how much time requires finishing the remaining processes. These predictions are uncertain and require probability mechanism to resolve. This paper takes such problem and presents an estimation method for predicting the possible time interval required for processing.

Shukla and Jain [12] discussed multiprocessor environment and innovate usual lottery scheduling and discussed a procedure to obtain ready queue time estimate. A method of estimation is suggested by authors and they computed the predicted time intervals. Shukla et al. [8] study similar problem using systematic lottery scheduling scheme in order to improve prediction of ready queue processing time. Shukla et al. [11] discussed similar problem when processes are grouped according to some criteria in different queue and total ready queue processing time and confidence intervals are predicted. Shukla et al. [10] introduce size based priority scheme for the ready queue time length prediction and proved that it is better than usual lottery scheduling in prediction of confidence interval for total time estimation of ready queue.

II. A REVIEW

Cochran [2] contains an introduction to the methods of sampling theory with application over multiple data. David [3] extended lottery scheduling, a proportional share resource management algorithm to provide the performance assurance present in traditional non real time process scheduler. Dynamic tickets were incorporated into a lottery scheduler to improve the interactive response time and to reduce kernel lock contention. Raz et al. [4] presented procedure of deciding priorities among jobs by maintaining fairness n selection procedure. Shukla and Jain [7] [9] tackled Marhov Chain based study of transitions in multilevel queue scheduling. Shukla and Jain [13] performed analysis of thread scheduling and Deficit Round Robin Alternated (DRRA) scheduling algorithm using Markov Chain Model approach.

Shukla and Jain [5] studied a stochastic model approach for reaching probabilities of message flow in space-division switches. Shukla and Ojha [6] performed analysis of multilevel queue with the effect of data model approach. Waldspurger [1] proposed that lottery scheduling ticket/currency framework can accommodate scheduling mechanism other than the probabilistic lottery algorithm and discussed the proportional share resource management technique in lottery scheduling.

Yiping [19] developed a queuing theory model to predict system behavior and CPU queue length in Microsoft NT, Windows 2000 and device fair share scheduling which guarantees application performance by explicitly allocating share of system resources among competing workloads. Some other useful contributions are [14], [15], [16] [17], [18].

III. PROBLEM DEFINATION

It is common and well known idea that often the more input information provides better prediction subject to condition if information is related. Based on this thought factor type estimation technique in this paper has been introduced in order to get more precise confidence intervals compared to Shukla et al. [14].

IV. PROCESSOR STRUCTURE AND NOTATIONS

Let $Q_1, Q_2, Q_3 \ldots\ldots\ldots Q_k$ be k processors who take intake from the ready queue containing $P_1, P_2, P_3 \ldots\ldots\ldots P_N$ processes (n<N). Processes are related to long, medium and short term scheduling queues prepared for transferring to ready queue. When a process is blocked or suspended, it is back to respective queue. The figure 1 shows the diagram of scheduling process structure with k processors. Let Y_i be the CPU burst time of each process $(i = 1,2,3 \ldots N)$ and X be one auxiliary variable like size of process or process priority or processes response time (time interval between arrival time and processor entering time). The mean of CPU burst time of N processes in the ready queue are:

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \quad \bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$$

IV- A. MODIFIED MULTIPROCESSOR LOTTERY SCHEDULING [As per Shukla, Jain and Choudhary[12])

Step I: When a process enters into ready queue, it is allotted a random number (in specified range).

Step II: Each processor $Q_1, Q_2, Q_3 \ldots\ldots\ldots Q_K$ generates unique and uncommon random number in similar specified range stated in Step I.

Step III: Matching of both random numbers takes place between process and processor. If both random numbers are same for a process in ready queue, process is assigned to that processor.

Step IV: Processor either blocks or processes the job. It selects another process by random manner as stated above.

Step V: When one job processed completely or partially processors generate time consumed in processing as y_i (Time by j^{th} processor) where $(j = 1,2,3 \ldots\ldots K)$ other information available auxiliary variables X_i .

V. ESTIMATION OF READY QUEUE PROCESSING TIME (as derived in [12])

Ratio estimator: $\bar{y}_r = \bar{y}\bar{X} / \bar{x}$

$$B(\bar{y}_r) = \bar{Y}(V_{02} - V_{11}), \qquad \ldots.\ (1)$$

$$M(\bar{y}_r) = \bar{Y}^2(V_{20} - 2V_{11} + V_{02}), \qquad \ldots.\ (2)$$

In above expressions

$$V_{ij} = E\left(\bar{y} - \bar{Y}\right)^i \left(\bar{x} - \bar{X}\right)^j / \bar{Y}^i \bar{X}^j ; \bar{Y}, \bar{X} \neq 0. \qquad \text{.... (3)}$$

Consider following notations;

$f = n/N,$

$A = (d-1)(d-2),$

$B = (d-1)(d-4),$

$C = (d-2)(d-3)(d-4).$

Where d is a constant $(0 < d < \infty)$.

The class of factor-type estimator proposed by [14]

$$T_d = \bar{y}\left[\frac{(A+C)\bar{X} + fB\bar{x}}{(A+fB)\bar{X} + C\bar{x}}\right]$$

where d > 0 is a non-negative constant.

It gives $T_1 = \bar{y}_r, T_2 = \bar{y}_p, T_3 = \bar{y}_a$ and $T_4 = \bar{y}$ [as Shukla et.al[14]].

V-A. BIAS AND M.S.E of T_A, T_B [As per [14]]

Taking large sample approximations, $\bar{y} = \bar{Y}(1+e_1)$ and $\bar{x} = \bar{X}(1+e_2)$ such that $E\left(e_1^i e_2^j\right) = V_{ij}$ we get

$$E(T_d) = E\left[\bar{Y}(1+e_1)\left\{\frac{(A+C) + fB(1+e_1)}{(A+fB) + C(1+e_2)}\right\}\right]$$

As it is obvious that

$$\left|\frac{Ce_2}{A+fB+C}\right| < 1$$

For all choices of A, B and C, the bias up to the second order moments, is

$$B(T_d) = \frac{(C-fB)\bar{Y}}{(A+fB+C)}\left\{\frac{C}{A+fB+C}V_{02} - V_{11}\right\}$$

Similarly the expression for M.S.E of T_d is given by

$$M(T_d) = E\left(T_d - \bar{Y}\right)^2$$
$$= \bar{Y}^2\left[V_{20} + P^2 V_{02} + 2PV_{11}\right]$$

Where $P = (fB - C)/(A + fB + C)$

SPECIAL CASE

At $d = 5$, we choose one estimator in class T_d.

$$T_A = \bar{y}\left[\frac{9\bar{X} + 2f\bar{x}}{(6+2f)\bar{X} + 3\bar{x}}\right]$$

$$B(T_A) = \frac{(3-2f)\bar{Y}}{(9+2f)}\left\{\frac{3}{9+2f}V_{02} - V_{11}\right\}$$

$$M(T_A) = \bar{Y}^2\left[V_{20} + \frac{4f^2 - 12f + 9}{81 + 36f + 4f^2}V_{02} + \frac{4f-6}{18+4f}V_{11}\right],$$

Similarly at $d = 6$, we choose one estimator from class T_d.

$$T_B = \bar{y}\left[\frac{22\bar{X} + 5f\bar{x}}{(10+5f)\bar{X} + 12\bar{x}}\right]$$

$$B(T_B) = \frac{(12-5f)\bar{Y}}{(22+5f)}\left\{\frac{12}{22+5f}V_{02} - V_{11}\right\}$$

$$M(T_B) = \bar{Y}^2\left[V_{20} + \frac{25f^2 - 120f + 144}{484 + 220f + 25f^2}V_{02} + \frac{10f - 24}{44+10f}V_{11}\right]$$

The 95% confidence interval of the estimate using T_A and T_B are:

$$P[T_A \pm 1.96\sqrt{V(T_A)}] = 0.95$$
$$P[T_B \pm 1.96\sqrt{V(T_B)}] = 0.95$$

Where

$$V(T_A) = \left[M(T_A) - \left(B(T_A)^2\right)\right]$$
$$V(T_B) = \left[M(T_B) - \left(B(T_B)^2\right)\right]$$

More explicitly one can write for confidence intervals

$$P[T_A \pm 1.96\sqrt{V(T_A)} \leq \bar{Y} \leq T_A \pm 1.96\sqrt{V(T_A)}] = 0.95$$
$$P[T_B \pm 1.96\sqrt{V(T_B)} \leq \bar{Y} \leq T_B \pm 1.96\sqrt{V(T_B)}] = 0.95$$

where T_A, T_B are the sample based estimates of population parameter $\bar{Y}$. These estimates T_A, T_B are predictors for average time required to complete a process by processors. Suppose out of N processes, n are processed (n<N) and remaining (N-n) are still in the system when sudden collapse occurs. Now, the predicted total time required for remaining jobs;

$$t_A = (N-n)T_A \qquad \text{.... (4)}$$
$$t_B = (N-n)T_B \qquad \text{.... (5)}$$

VI. NUMERICAL DATA ANALYSIS

Consider 30 processes in ready queue at a time whose size measure X is also given in terms of bytes. If we assume that all the processes are processed completely in the ready queue, the CPU burst time Y is mentioned against them.

Table1: Presents Ready Queue Process size measures.

Process ID	Process Size Parameter (X_i)	CPU Burst Time (Y_i)
1	210	30
2	897	20
3	312	112
4	171	40
5	461	59
6	290	60
7	379	30
8	220	43
9	470	101
10	636	69
11	455	138
12	682	43
13	952	109
14	574	26
15	536	74
16	416	89
17	788	123
18	902	67
19	623	58
20	563	84
21	111	143
22	341	29
23	775	147
24	913	94
25	745	131
26	130	79
27	877	46
28	927	59
29	424	72
30	356	22

Table 2: Sample of n processed from ready queue.

	$N=30$	$n=5$	$f=.1666$
$V=0.13361$	**At**	$d=5$	$d=6$
$v_{02}=0.03949$	$BIAS(T_A,T_B)$	0.16974	0.59156
$v_{20}=0.04568$	$m.s.e(T_A,T_B)$	246.086	267.944
$v_{11}=0.005276$	$Var(\bar{y})(T_A,T_B)$	246.057	267.594

Estimated Confidence Interval lengths		
$\hat{\bar{Y}}$	*Confidence Intervals* *at* $d=5$	*Confidence Intervals* *at* $d=6$
72.16	(41.41-102.90)	(40.74-104.22)

Processes Parameters	**Randomly selected Processes from Ready Queue**				
Processes	P_9	P_{18}	P_{30}	P_{24}	P_{13}
CPU Burst Time Y	101	67	22	94	109
Processes Size X	470	902	356	913	952

Table 3: Processed processes estimated time length.

VII. CONCLUSION

The use of Factor-type estimator play important role in the prediction of the possible processing time of ready queue. It considerable reduced the length of confidence intervals comparatively with ratio estimator in the setup of lottery scheduling. Two estimators T_A and T_B suggested for estimation of average ready queue remaining time. The estimate by T_A = 65.9 while T_B = 62.08 (when n=5). Both are close to true value. The M.S.E of T_A (=246.08) is lesser to the M.S.E of T_B (=267.9) for n=5. T_A is uniformly efficient over T_B for all n=5, 10, 15.... (n-1) due to lower M.S.E. The true values of CPU burst time lies within the range of confidence interval. It is recommended to prefer T_A over T_B in setup of lottery scheduling for estimation purpose.

REFERENCES

[1] Carl, A. Waldspurger and William, E. Weihl 1994. Lottery Scheduling a flexible proportional-share resource management. In Proceedings of the 1st USENIX Symposium on Operating Systems Design and Implementation (OSDI).

[2] Cochran. 2005. Sampling Technique, Wiley Eastern Publication, New Delhi.

[3] David, P., Garth, A. Gibson and John, W. M. 1999. Implementing Lottery Scheduling: Matching the specializations in Traditional Schedulers. In Proceedings of the USENIX Annual Technical Conference USA.

[4] Raz, D., Itzahak, B. and Levy, H. 2004. Classes, Priorities and Fairness in Queuing Systems. Research report. Rutgers University.

[5] Shukla, D. and Jain, S. 2010. A Stochastic Model Approach for Reaching Probabilities of Message Flow in Space-Division Switches. International Journal of Computer Networks.

[6] Shukla, D., Jain, S. and Ojha, S. 2010. Effect of Data Model Approach for the Analysis of Multi-Level Queue Scheduling. International Journal of Advanced Networking and Applications.

[7] Shukla, D. and Jain, S. 2007. Deadlock state study in security based multilevel queue scheduling scheme in operating system. In Proceedings of National Conference on Network Security and Management, NCNSM-07.

[8] Shukla, D. and Jain, A. 2010. Estimation of ready queue processing time under SL scheduling scheme in multiprocessor environment**.** International Journal of Computer Science and Security (IJCSS).

[9] Shukla, D. and Jain, S. 2009. Analysis of Thread scheduling with multiple processors under a Markov chain model. Journal of Computer Science.

[10] Shukla, D. and Jain, A. 2011. Analysis of Ready Queue Processing Time under PPS-LS and SRS-LS scheme in Multiprocessing Environment. GESJ: Computer Science and Telecommunication.

[11] Shukla, D., Jain, A. and Choudhary, A. 2010. Estimation of ready queue processing time under Usual Group Lottery Scheduling (GLS) in Multiprocessor Environment**.** International Journal of Computer and Applications (IJCA).

[12] Shukla, D., Jain, A. and Choudhary, A. 2011. Estimation of ready queue processing time under Usual Lottery Scheduling (ULS) in Multiprocessor Environment**.** Journal of Applied Computer Science and Mathematics (JACSM).

[13] Shukla, D., Jain, S. and Singh, U. 2008. A Markov chain model for Deficit Round Robin Alternated (DRRA) scheduling algorithm. In Proceedings of the International Conference on Mathematics and Computer Science, ICMCS-08.

[14] Singh, V.K. and Shukla, D. 1987. An efficient one-parameter family of factor-type estimator in sample surveys, METRON International Journal of Statistics.

[15] Silberschatz, A. and Galvin, P. 1999. Operating System Concepts. Ed.5, John Wiley and Sons (Asia), Inc.

[16] Singh, D. and Choudhary, F.S. 1986. Theory and Analysis of Sample Survey and Designs, Wiley Eastern Limited, New Delhi.

[17] Stalling, W. 2004. Operating Systems, Ed.5, Pearson Education, Singapore, Indian Edition, New Delhi.

[18] Tanenbaum, A. 2000. Operating system, Ed. 8, Prentice Hall of India, New Delhi.

[19] Yiping, D. and William, F. 2000. Interpreting Windows NT Processor Queue Length Measurements. In Proceedings of the 31st Computer Measurement Group Conference.