ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

Survey on Profit and Penalty Aware Scheduling for Time Dependent Applications

Madhuri Bedase[1], D. A. Phalke[2]

M. E. Student, Dept. of Computer Engineering, DYPCOE, Savitribai Phule Pune University, Pune, MH, India[1]

Assistant Professor, Dept. of Computer Engineering, DYPCOE, Savitribai Phule Pune University, Pune, MH, India[2]

ABSTRACT: Scheduling algorithms are the most important part of real time systems, based on uniprocessor and multiprocessor environment as per applications, have their own challenges and complexities, such as earliest deadline first, rate monotonic etc. As computer and internet technology continue to growing, real time online services becomes a huge market. The goal of traditional real time applications is to complete the task before deadline. Different from this, the objective of online provider, is to maximizing profit with timely services. For this type of applications with aperiodic task set, the time dependent functions provides two distinct features: accrued profit- if task finished upto deadline else accrued penalty. This paper presents various available scheduling algorithms to increase the profit and minimize the penalty of online service providerswith comparative analysis and also show that, how PP-aware scheduling performs better than other existing algorithm in terms of accrued profit and reducing processing time.

KEYWORDS: Time dependent applications; RTOS; profit and penalty aware; scheduling algorithm; SLA

I. INTRODUCTION

Scheduling plays very important role in daily life applications. The various applications includes concept of scheduling like, logistic planning, work-flow systems, space mission planning, entertainment, medical systems, traffic control systems, military applications, nuclear power plant control system and so on. A real time system is usually one that controls and/or monitors a physical (real-world) process. In real time systems, users requesting to service provider for getting some services. These multiple requests are scheduled by processor using scheduling algorithms and these requests to service provider are treated as tasks. The examples of real time online systems are: Internet booking engine, airline reservation systems, travel planning service provider and so on. The real time systems are divided into three types. First is hard real time systems. In this, the task is requested to get service from a service provider and having some time limit also called as a deadline. If the time limit is finished before completing task request, the task has failed. This failure leads to a total collapse of the system. In soft real time system, missing of deadline only leads to degradation in performance of the entire system but not a complete failure. Third one is firm real time system, where missing of task deadline has no value. Some tasks are periodic and some are aperiodic. Some tasks are preemptive and some are non-preemptive. Preemption means switching from low priority task to high priority task. That is, during the execution of any task, if the high priority task is arriving then currently executing task is suspended and processor is assigned to newly arrived high priority task. After completion of the high priority task, the execution of low priority task is resumed from that point where it is stopped. Pre-emptive tasks can be pre-empted at any time, but in non-preemptive, switching between tasks is not allowed, even if the high priority task is arrived [18].

Scheduling is a strategy by which the system decides which task should be executed at a given time on which processor. The problem of real time scheduling is different from multiprogramming time sharing scheduling because of the role of timing constraints is taken into consideration in the evaluation of the system performance. Normal multiprogramming time sharing systems are processing multiple job streams simultaneously on different processors, therefore the scheduling of these tasks has the goals of maximizing rate of processing request and maintaining quality. The primary performance of real time systems is to perform critical operations within a set of critical time constraints assign by user. When a system can meet all its tasks' deadlines, then this task set is schedulable and the system is called as under loaded. On the other hand, if at least one task cannot finish by its deadline, then the system is called as overloaded. There are some goals, which should be achieve by scheduler in real-time system such as, maintaining the deadline restriction, avoiding simultaneous access to shared resources and devices, obtain high degree of utilization

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

when satisfying the timing constraints of the system, reducing the cost of preemption, reducing the communication cost of real time systems. Tasks having some parameters, such as release time, deadline, run time, minimum delay, maximum delay, worst case execution time, best case execution time, the response time and priority.

This paper mainly focused on various algorithm used for scheduling of real time online services. In literature most of the scheduling algorithms was based on Earliest Deadline First (EDF) algorithm. Various forms of EDF are available, such as EDF-FM, DQDF, EDF-ACO, DO-EDF, DR-EDF etc. Some algorithms are based on genetic algorithm, rate monotonic, and deadline monotonic concept. In Recent, some authors are working of new model of task scheduling algorithm: Profit and Penalty (PP-aware) scheduling. The main aim of this scheduler is to increase the total profit and reducing the penalty of service providers [16], [17].

This concept of profit and penalty between service provider and client is based on Service Level Agreement (SLA). This SLA is agrees by both sides. Service provider agrees this SLA by providing guaranteed services within a client's deadline and client agree this SLA by paying to the service provider after getting services within a deadline. If the service providers not achieving a deadline, then according to SLA penalty is applied to them [19].

II. RELATED WORK

In [2], author proposed Dynamic queue deadline first scheduling (DQDF) algorithm. This algorithm is used to handle the scheduling of dynamic multiple tasks in real time systems. It minimizing the overhead and percentage of deadline miss ratio to improve the system utilization and fairness. It uses static priority approach. First uses the EDF for scheduling tasks. Calculating the deadline miss ratio for every processor. If deadline miss ratio is greater than threshold of processors then run time dynamic queue is created and remaining tasks of that process move to this run time queue which having highest priority. So that the maximum task can meet their deadline. To reducing the number of run time queues, proposed algorithm generate the shared run time dynamic queue which is shared by all processors.

Research Gap:

1. System is not applicable of multiprocessor environment.
2. Approach can be useful for hard real time operating systems by eliminating deadline misses.

Apurva Shah, KetanKotecha combines Earliest Deadline First (EDF) and Aunt Colony Optimization (ACO) algorithm in [3]. Adaptive algorithm combines the advantages of both and overcomes the limitation of EDF and ACO algorithms. EDF is fast but reducing the performance during overload, and ACO better performs during overload but execution speed is slow compared to EDF. Adaptive algorithm is start with EDF. When two consecutive jobs miss their deadline, then assume that system goes into overload condition. In this case of overloading adaptive algorithm automatically switch to ACO and continue to scheduling of jobs. During ACO if ten jobs successfully achieve the deadline continuously, adaptive algorithm consider that the system is under loaded and system will switching from ACO to EDF.

Research Gap: System can be extended for loosely coupled multiprocessor or distributed systems.

Marko Bertogna and SanjoyBaruah [6] have proposed a limited-preemption scheduling technique. This technique combines the benefits of both Preemptive and Non-preemptive scheduling. This system avoids the unnecessary preemptions, so that the runtime overheads is very low. Runtime overhead causes worst case execution of a task. System also prove that, any system with speriodic tasks supporting preemptive EDF is also supporting to limited-preemption version EDF. So, as compared to preemptive EDF, limited-preemption EDF has minimum schedulability overhead. Limited-preemption EDF offers low runtime overhead as compared to non-preemptive EDF. System also offers efficient support for mutual exclusion.

Research Gap:

1. Reduce the preemption overhead of each task.
2. Proper Selection of WCET bound to avoid missing of deadline.

In efficient scheduling algorithm for real time distributed systems [4], Kotecha K. modify the EDF algorithm and proposed D-O-EDF and D-R-EDF algorithms. These algorithms are applicable for client-server distributed system with soft timing constraints. EDF having the disadvantage like domino effect during overload system. Also in client server distributed system, server is jam because of increasing flow of requests at www. To overcome these issues, D-O-EDF and D-R-EDF are developed. During underload D-O-EDF scheduling is exactly similar to EDF. During overload, D-O-

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

EDF uses the static priority 0 and 1 for scheduling the job. If priority is 0 then job discarding otherwise not. D-R-EDF combines both static and dynamic priority approach and automatic switching among them. During underload, D-R-EDF working same as EDF and during overload working same as Rate Monotonic (RM). In RM task with earliest period is scheduled first. Scheduling start with EDF, when two jobs continue missing deadline, overload condition occurs and algorithm switch to RM. After 5 jobs continuously meet deadline, underload situation occurs and switch to EDF.
Research Gap: Not applicable for hard timing constraint in distributed systems.

RUN is an efficient scheduler and reduces the multiprocessor problem to a series of uniprocessor problem, developed in [9] by P. Regnier. Finally uniprocessor schedule return to original multi- processor problem. It reduces the overhead due to preemptions and migration and improve processor utilization. It produces the multiprocessor schedule for a set of fixed rate task with few preemptions per jobs.
Research Gap:
1. Consider fault tolerance, energy consumption, and adaptability while considering the uniprocessor scheduling to solve the multiprocessor problems.
2. Extend this model for sporadic tasks with constrained deadlines.

Shuhui Li et al has proposed, Profit and Penalty (PP) aware scheduling in [1]. Author has introduced two different Time Utility Functions (TUF): profit TUF and penalty TUF associated with each task. Profit TUF represent the gain on completing a task before deadline and penalty TUF is the loss by crossing deadline or aborting the tasks. System scheduling the tasks which arriving a-periodically. PP- aware scheduler process these tasks either by preemptively or by non-preemptively. The main aim of the system is, to complete the as many as tasks before deadline and increase the systems total accrued profit. This system introduced the VEP task execution model with gain and loss function. According to this model, task execution time is not known priori, but task execution probable time is known. This paper also introduced a Risk factor of task, use to define the risk of processing task. This factor reducing the high penalty cost.
Research Gap:
1. There is a need of optimal algorithm for scheduling, because present algorithm is highly heuristic.
2. Identify the optimal risk factor of system with different task.
3. Identify more complicated profit and penalty function.
4. Identify the negative function of loss.
5. Calculation of preemption point using simple periodic sampling approach.

In [7], author present non-preemptive utility accrued scheduling approaches. This approach is based on opportunity cost concept and a speculated execution order for expected utility. Then constrained preemptive approach is proposed. This scheduling algorithms carefully selects highly profitable tasks to execute. System also aggressively removes non profitable tasks, which lead large penalty. The system proves that the proposed algorithms successfully performs better than traditional EDF, utility accrued and heuristic approach based on profit and penalty aware task model.
Research Gap: Need to improve the preemption point calculation.

R. Sharma proposed, task migration with EDF-RM scheduling algorithms for distributed system in [8]. This algorithm makes simultaneous use of EDF and RM algorithm. Initially global scheduler is used to maintain global task queue. Distribute these task in global queue to different processors independently using RM. Then processors receives these task and store into local task queue and schedule these tasks by using EDF.
Research Gap: Domino effect should be resolve by using duplication of tasks based on deadline.

Shuhui Li, Miao Song Zheng Li, ShangpingRen and Gang Quan proposed two scheduling algorithms in [5], prediction-based highest gain density first (PHGDF) and iterative PHGDF (iPHGDF). These algorithms are used to maximize the systems accrued profit. Both algorithms works non-preemptively. The executing task selection criteria in prediction-based highest gain density first is that, scheduler selects task with highest unit time potential gain. If the selected task taking maximum time for execution and making other task non profitable, then scheduler will remove the current executing task. The iterative version of prediction-based highest gain density first is used to make the prediction

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

as early as possible. To achieve this, the pending queue is updated that is sorted every time by the task scheduling algorithm. Hence a better system profit is obtained.

Research Gap: Reduce the cost overhead occurs due to re-order of task pending queue.

Omniscient and Active Local Deadline Assignment (OLDA and ALDA) algorithms are proposed in [10]. These algorithms are based on local deadline assignment problem and highly efficient to achieve end to end job deadline in distributed real time systems. It provide feasible solution for job deadline assignment problem on multiple processor locally. It reduced the slack time among sub jobs on the processor with removing the overhead of global clock synchronization.

Research Gap:

1. To validate the algorithm, this system can be deploy in real time OS with real world applications.
2. These algorithms can be improved with appropriate bus-based network platform by solving the problem of communication delays and with some other quality of service criteria.

The deadline guaranteed algorithm for multi cloud online systems is implemented in [11] based on Service Level Agreement (SLA). This paper proposed the parallel deadline guaranteed scheme (*PDG*), for load balancing among overloaded and under loaded cloud servers. Also enhance this scheme with priority based data reallocation algorithm. There is another system for work load balancing in cloud environment is proposed in [13], with Idle-Server Monitoring algorithm.

Research gap: To check the real performance, apply this system on real cloud storage systems.

The technique to identify the deadline miss probability for periodic task set is introduced in [12], based on resource reservation algorithm. Modelling approach used for reducing the complexity of probability computation. The metric for calculation of end to end deadline missing probability at each local processing unit is proposed in [15]. This approach is applicable for soft real time distributed applications. The delay-impact-based (DIB) local deadline assignment algorithm is also proposed based on this metric. With this algorithm, each processor can be able to schedule and execute their job independent of other processor. It will increase the deadline meet ratio in real time applications.

Research gap:

1. Investigate the connection between quality of service and probabilistic deadlines in various applications domain.
2. Apply the scheme on those applications with multiple tasks arrival.

The multiple workflow scheduling problem under dynamic environment is solve in [14] by introducing two algorithms namely, OMPHC-PCPR and OPHC-TR. These algorithms are based on the constraints of deadline and privacy. It utilizes ranking method for supporting to prioritized job scheduling during resource allocation.

Research gap: Improve this system with discrete optimization technique to minimize the cost of scheduling.

The comparative analysis of some important techniques are described in Table I. Table represent the title, year of publication and author of paper, approach used to develop the system, performance metrics and summary of system.

III. IDENTIFIED FUTURE DIRECTIONS

Following are some future directions to design and implement a good quality of scheduler to increase the profit of real time systems by completing the tasks within deadline:

1. Select proper WCET bounds to avoid deadline missing
2. Proper preemption point placement
3. Scheduler should support various application models
4. Domino effect should be resolved
5. It should be cost effective, reliable and efficient
6. Optimal scheduler algorithm which can work in both uniprocessor and multiprocessor environment.

Implement and apply the scheduler in real world applications to get actual benefit.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

IV. CONCLUSION

The simulation results showed that the proposed algorithm performs better with the total transmission energy metric than the maximum number of hops metric. The proposed algorithm provides energy efficient path for data transmission and maximizes the lifetime of entire network.As the performance of the proposed algorithm is analyzed between two metrics in future with some modifications in design considerations the performance of the proposed algorithm can be compared with other energy efficient algorithm. We have used very small network of 5 nodes, as number of nodes increases the complexity will increase. We can increase the number of nodes and analyze the performance.

TABLE I
COMPARATIVE ANALYSIS OF VARIOUSAPPROACHES

Paper Title / Year	Author	Approach	Performance Metrics	Summary
Profit and penalty aware scheduling for real time online services, IEEE Transaction, Feb 2012.	Shuhui Li, Li Wang	Preemptive and Non-Preemptive PP-aware scheduling	Systems total accrued profit, Average number of preemptions	Makes use of task admission test and Preemption Threshold.
Dynamic queue deadline first Scheduling algorithm for soft real time systems, Sept 2005.	Arshad Iqbal, Asia Zafar	DQDF With EDF and DQS algorithm	Deadline miss ratio	Achieve tasks deadline by using General run time queue.
Adaptive scheduling algorithm for real time multiprocessor real time multiprocessor real time systems, 2009.	Apurva Shah, KetanKotecha	Adaptive scheduling combination of EDF and ACO	Number of Preemption, Success Ratio, CPU utilization	Start with EDF, according to systems load switching among EDF and ACO.
Efficient scheduling algorithms for real time distributed systems, 2010.	Kotecha K.	D-O-EDFandD-R-EDF	Success Ratio, Effective CPU Utilization	Modified EDF algorithm for client/server distributed system with soft timing constraints.
Maximizing Online Service Profit for Time-Dependent Applications, 2013.	Shuhui Li,GangQuan	PHGDFandiPHGDF	Systems accrued profit	Make the prediction earlier, the pending queue is reorderedevery time by the task scheduling algorithm after simulating arequest dispatch
Limited preemption EDF scheduling of sporadic task systems, 2009.	Marko Bertogna,SanjoyBaruah	Preemptive and Nonpreemptive EDF	Limited Preemption	The preemptions that are not needed are avoided thus, runtime overheads that cause worst case execution of a task are very low.
On-Line Real-Time Service-Oriented Task Scheduling Using TUF, 2012.	Shuo Liu	Scheduling with opportunity cost and speculated execution order	Reducing number of preemptions and increasing profit	Non preemptive utility accrued scheduling approaches upon a metric developed according to the opportunity cost concept and a speculation-based metric for expected utility

REFERENCES

1. S. R. Shuhui Li, 'Profit and penalty aware scheduling for real time online services', IEEE Transactions on Industrial Informatics, VOL. 8, NO. 1, pp. 78-79, 2012.
2. A. Iqbal, 'Dynamic queue deadline first scheduling algorithm for soft real time systems', IEEE International Conference on Engineering Technologies, pp. 346-351, 2005.
3. K. K. Apurva Shah, 'Adaptive scheduling algorithm for real time multiprocessor systems', IEEE International Advance Computing Conference, pp. 35-39, 2009.
4. A. Shah, 'Efficient scheduling algorithms for real-time distributed systems', IEEE International Conference on Parallel, Distributed and Grid Computing, pp. 44-48, 2010.
5. Li, Shuhui, 'Maximizing online service profit for time-dependent applications', IEEE 19th International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA), pp. 342-245, 2013.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 5, May 2016

6. Bertogna, Marko, SanjoyBaruah, ‘Limited preemption EDF scheduling of sporadic task systems’, IEEE Transactions on Industrial Informatics, pp. 579-591, 2010.
7. Shuo Liu, ‘On-Line Real-Time Service-Oriented Task Scheduling Using TUF’, ISRN Software Engineering, Article ID 681985, 2012.
8. R. Sharma, ‘Task migration with EDF-RM scheduling algorithms in distributed system’, IEEE International Conference on Advances in Computing and Communications, (ICACC), pp. 182-185, 2012.
9. P. Regnier, ‘Run: Optimal multiprocessor real time scheduling via reduction to uniprocessor’, IEEE Real Time Systems Symposium (RTSS), pp. 104-115, 2011.
10. Hong S, Chantem T, Hu XS, ‘Local-Deadline Assignment for Distributed Real-Time Systems’, IEEE Transactions on Computers, pp. 1983-1997, 2015.
11. Liu G, Shen H., ‘Deadline guaranteed service for multi-tenant cloud storage’, IEEE International Conference on InPeer-to-Peer Computing (P2P), pp. 1-10, 2015.
12. Luigi Palopoli, ‘An Analytical Solution for Probabilistic Guarantees of Reservation Based Soft Real–Time Systems’, IEEE Transactions on Parallel and Distributed Systems, 2015.
13. Rajeshwari, B. S., M. Dakshayini, ‘Optimized service level agreement based workload balancing strategy for cloud environment’, IEEE International Advance Computing Conference (IACC), pp. 160-165, 2015.
14. Sharif, Shaghayegh, ‘Online Multiple Workflow Scheduling under Privacy and Deadline in Hybrid Cloud Environment’, IEEE 6th International Conference on Cloud Computing Technology and Science (CloudCom), pp. 455-462, 2014.
15. Song M, Li S, Ren S, QuanG., ‘Delay-impact-based local deadline assignment for online scheduling of distributed soft real-time applications’, IEEE International Conference In Performance Computing and Communications (IPCCC), pp. 1-8, 2014.
16. S. Mohammadi, ‘Scheduling algorithms for real time systems’, Queens University, Technical Report, 2005.
17. L. Zhang, ‘Scheduling algorithms for multiprocessor real-time systems’, IEEE International Conference onInformation, Communications and Signal Processing, volume 3, pp. 1470-1474, 1997.
18. N. D. Thai, ‘Real-time scheduling in distributed systems’,IEEE Proceedings of the International Conference on Parallel Computing in Electrical Engineering, pp. 165-170, 2002.
19. P. Kittipipattanathaworn, N. Nupairoj, ‘SLA guarantee real-time monitoring system with soft deadline constraint’, *11th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pp. 52-57, 2014.

BIOGRAPHY

Madhuri Ravindra Bedase received the BE in Computer Engineering from S. S. V. P. S. college of engineering which is under North Maharashtra University,Jalgon, in 2011. Currently pursuing ME in Computer engineering from D. Y.Patil College of engineering, Akurdi, Pune, India.

D. A. Phalke received the ME in Computer Engineering from D. Y. Patil College of Engineering, Akurdi through University of Pune and she is working as Assistant professor in the Department of Computer Engineering at D. Y. Patil College of Engineering and has 14 year of experience in the field. Currently pursuing PhD in department of technology, SavitribaiPhule Pune University.