Probabilistic Analysis of Multi-processor Scheduling of Tasks with Uncertain Parameters

Amare Leulseged and Nimal Nissanke

School of Computing, Information Systems and Mathematics
South Bank University, 103 Borough Road, London SE1 0AA, UK

Abstract. A new approach is proposed for the probabilistic assessment of schedulability of periodic tasks with uncertain characteristics in dynamic multi–processor scheduling. It is aimed at non–critical real–time applications such as multimedia, which allow some leeway with respect to compliance with timing requirements, provided that certain minimum Quality of Service (QoS) requirements are met. Uncertainties are taken into account through random variables at the task arrival times and by characterising subsequent task characteristics in probabilistic terms. By examining each pair of possible computation time and deadline of a given task at each time unit in relation to the same of other tasks, an execution pattern is derived. This forms the basis for computing various QoS attributes such as probability of successful execution, latency in response time, jitter, etc. Illustrative examples address, amongst others, the performance of two particular algorithms, EDF and LLF, in the presence of uncertainties in task characteristics.

1 Introduction

It is a common practice in real–time scheduling algorithms to assume that task characteristics such as computation time and deadline are known precisely, sometimes in advance, and remain constant throughout the life time of the task. However, this is rarely the case in practice and the lack of precise prior knowledge about task characteristics remains a major concern in scheduling. This applies especially to non–critical real–time applications such as multimedia systems, computer vision, real–time tracking based on radar or sonar. Computational tasks in them tend to vary widely in execution times depending on the complexity of the specific task instance being handled. In addition, tasks may or may not arrive at fixed periodic intervals. Experiments in [12] show deviations of actual periods from the nominal ones and a tendency for them to alternate between short and long periods in consecutive instances. In the face of such unpredictabilities, task deadlines too are subject to change in order to indirectly account for uncertainties in task execution times and request times.

A common approach to dealing with uncertainties so arising is to adopt a worst–case strategy and to assign an extreme value to the computation time, regardless of its frequency relative to its other possible values and its representativeness. This is an acceptable solution in critical applications but is an overly

J. Chen and S. Hong (Eds.): RTCSA 2003, LNCS 2968, pp. 103–122, 2004.

demanding one in non–critical applications. In applications such as multimedia user perception is not affected often by such variation to the same degree, while in other applications there are more tolerant alternative ways to dealing with occasional failures. Under such circumstances under–utilisation of computing resources resulting from worst–case considerations could seriously undermine the cost–effectiveness of such applications. This underlines the importance of arriving at an acceptable balance between the Quality of Service (QoS) and the overall system performance, such as throughput and resource utilisation in terms of processor workloads. However, uncertainties in task characteristics must still be dealt with, because missing deadlines result in both QoS violations and wastage of computing resources, to the detriment of the balance between both factors mentioned above.

For tasks with uncertain parameters, on–line and off–line scheduling algorithms and schedulability analysis have been proposed in the literature. Zhou et. al [12] propose a modified rate–monotonic schedulability analysis, incorporating two new experimentally determined parameters to account for uncertainties in operating system overheads, namely, a constant representing the CPU utilisation of operating system activities and a worst–case timer delay factor. A Statistical Rate-Monotonic Scheduling approach [1], with an implementation described in [2], allows scheduling of periodic tasks with highly variable execution times expressed through a probability density function. It also allows the consideration of statistical QoS requirements defined in terms of the probability of a random instance of the task chosen from an arbitrarily long execution history meeting its deadline. Recently, Manolache et al [9] have presented an approach to performance analysis of periodic task sets with their execution times specified as a continuous probability distribution. Although it is a non preemptable and is confined to single processor environments, the approach is elegant and shares the same objective as this paper. Dealing with the so–called monotone processes, i.e., those where the quality of the result keep improving after surpassing a minimum threshold computation time, Chung et. al [3] propose an imprecise computational model that involves a mandatory initial part and an optional follow–on part. Mandatory parts of all tasks are to be executed within the deadline of each task, while the optional part is left free to execute longer, if it can be accommodated, thus refining and improving the result. Hamann et. al [7] extends the imprecise computational model by incorporating an additional minimum reservation time for each task that assures a certain probability of successfully completing a given percentage of its optional parts. In assessing computational times of real–time tasks, there have been several attempts such as [11] based on deterministic code analysis. Recognition of their inappropriateness is evident from works such as [8] devoted to an estimation of execution times statistically from past observations. As is demonstrated in [6] using Gumbel distribution for estimating the worst–case execution time (WCET), statistical models are likely to result in a more realistic assessment of execution times.

The works devoted to uncertainties in task characteristics are extensive. The above are a small selection illustrating a range of approaches addressing, in differ-