Degree of Schedulability of Mixed-Criticality Real-time Systems with Probabilistic Sporadic Tasks

A.Jalil Boudjadar, Alexandre David, Jin Hyun Kim, Kim G. Larsen, Marius Mikuciuonis, Ulrik Nyman, Arne Skou
Institute of Computer Science, Aalborg University, Denmark

Abstract—We present the concept of degree of schedulability for mixed-criticality scheduling systems. This concept is given in terms of the two factors 1) Percentage of Missed Deadlines (PoMD); and 2) Degradation of the Quality of Service (DoQoS). The novel aspect is that we consider task arrival patterns that follow user-defined continuous probability distributions. We determine the degree of schedulability of a single scheduling component which can contain both periodic and sporadic tasks using statistical model checking in the form of UPPAAL SMC. We support uniform, exponential, Gaussian and any user-defined probability distribution.

I. INTRODUCTION

When constructing an embedded software system engineers could be interested, not only, in whether or not the system always meets its requirements, but also how it behaves with insufficient resources. For mixed-criticality systems in can be of special importance to know how the quality of service is degraded for the soft-real time parts if they are provided with less resources than needed. A certain level of degradation may be acceptable in a given setting and we thus consider it important to answer questions regarding schedulability with estimates of the quality instead of just providing a yes/no answer.

This paper presents the degree of schedulability as a new concept to measure the degradation of the quality of service of a given system. The degree of schedulability is given in terms of two metrics; the Percentage of Missed Deadlines and the average delay per missed deadline, called Degradation of Quality of Service.

A hierarchical scheduling system [10] is a component-based system encompassing global resources shared between the system components. The system workload consists of a set of tasks declared with a set of timing attributes such as period, deadline and execution time.

The sporadic task model [3] [15] has received research attention [22] [21] over the years because of its usefulness in modeling recurring processes for hard-real-time systems.

It is obvious that it is not possible to guarantee the schedulability of real-time systems where sporadic tasks can occur arbitrarily frequently [3]. To alleviate this the sporadic task model operates with a minimal inter-arrival time. In many mixed-critical systems like tracking systems and automotive info-tainment systems, arrival times are adequately described by specific distribution functions. In this paper, we model real world events separately from the tasks. Events are modeled according to stochastic arrival patterns with an inter-arrival time that could be zero.

We consider a framework where the triggering mechanism of sporadic tasks is located outside the hierarchical scheduling system as illustrated in Fig. 1. Each sporadic task has its individual arrival pattern, which is characterized by a continuous probability distribution. In this paper, we are focusing on modeling and analyzing a single component at a time, with both sporadic and periodic real-time tasks. This analysis method fits within a larger compositional analysis framework [4], which analyzes a complete hierarchical scheduling system. Our main contributions are:

- We introduce continuous probability distributions to model sporadic events that trigger the execution of sporadic tasks.
- We study the system schedulability and determine the degree of schedulability (Sched°) in terms of the Percentage of Missed Deadlines (PoMD); and average delay per missed deadline, called Degradation of Quality of Service (DoQoS).
- We provide a framework including explicit environment models, which allows us to model and analyze mixed-criticality hierarchical scheduling systems.

Compared to treating the minimal inter-arrival time as a period, our method aims at providing both more realistic and optimistic resource estimates for sporadic tasks. The rest of the paper is structured as follows: Section II cites relevant related work. Section III introduces the compositional analysis framework. In Sections IV, V and VI we introduce respectively continuous sporadic tasks, the models used to analyze mixed-criticality hierarchical scheduling systems.

II. RELATED WORK

In this section we present related work with a specific focus on sporadic tasks. To the best of our knowledge, there is no
previous related work which uses continuous probabilities to characterize the arrival patterns of sporadic tasks. The sporadic task model [3], [15], which is an extension of an earlier task model known as the Liu and Layland (LL) [12] task model has received immense research attention over the years. In [22], the authors propose a framework for the schedulability analysis of real-time systems, where they define a generalized model for sporadic tasks to characterize more precisely the task arrival times. The authors characterize each task by two constraints: \textit{higher instantaneous arrival rate} which bounds the maximum number of task arrivals during some small time interval; \textit{lower average arrival rate} which is used to specify the maximum number of arrivals over some longer time interval. In [21], the authors propose a method to control the preemptive behavior of real-time sporadic task systems by the use of CPU frequency scaling. They introduced a new sporadic task model in which the task arrival may deviate, according to a \textit{discrete} time probability distribution, from the minimum inter-arrival time. In fact, a task arrival \( T \) may deviate with a delay \( t \) if the probability of \( T \) to occur at instant \( t \) is greater than a certain threshold. Based on the probability of arrivals, the authors propose an on-line algorithm computing CPU frequencies that guarantee non-preemptiveness of task behavior while preserving system schedulable. In [3], the authors propose an exact schedulability analysis by providing some necessary and sufficient conditions for a sporadic task system to be schedulable. In fact, the authors consider sporadic tasks with minimum inter-arrival time as periodic tasks, then define the set of legal requests that a task may perform. Based on such a function, they analyze the system schedulability regardless of the schedulability policy. However, considering sporadic tasks with known minimum inter-arrival times as periodic tasks may lead the schedulability analysis to be pessimistic and seriously overestimates the number of task arrivals. Our work differs by modeling probabilistic inter-arrival times and quantifying the system schedulability according to hard and soft real-time requirements. A concept similar to PoMD as introduced in this paper is given in [14]. The term “degree of schedulability” was first introduced in [16] to characterize the sum of response time delays from the individual deadlines. We define the concept DoQoS in a similar way, but focus on the total amount of time by which deadlines are missed. We define our notion of \textit{degree of schedulability} (Sched\textsuperscript{o}) by combining PoMD and DoQoS into one measure.

III. \textbf{COMPOSITIONAL ANALYSIS FRAMEWORK FOR Sched\textsuperscript{o}}

A hierarchical scheduling system [1] consists of a set of concurrent real-time components sharing a set of resources according to a scheduling policy. Each component can again be internally organized as a set of components, giving the organization of the system a tree like structure. The use of temporal partitioning [18] between the components is motivated by the fact that it provides reduction of complexity, separation of concerns, confinement of failure modes, and temporal isolation among system applications. One obvious partitioning of the components in a mixed criticality system is to group them according to their criticality. Such a grouping enables easier certification of the safety critical components when they have minimal communication with the non safety critical parts [17].

In this paper we focus on the schedulability analysis of one component inside a hierarchical scheduling system. Formally, a hierarchical scheduling system \( S = (C, R, A) \) is given by a set of hierarchical components \( C \), a set of resources \( R \) and a scheduling algorithm \( A \). A component, in turn, can be either a hierarchical unit \( \{(C_1, \ldots, C_n), A \} \) of other components \( C_i \), or a basic composition \((W, A)\) of a workload \( W \), together with a scheduling policy \( A \). The workload \( W \) is a set of real-time tasks having time constraints like deadline, execution time and next arrival. The real-time interface \( I \) [20] of a component \((W, A)\) specifies the collective resource requirements that the workloads \( W \) performs under the scheduling policy \( A \). \( I \) is simply given by a period \( p \) and a budget \( b \) in our framework.

In a compositional schedulability analysis framework [7], [4], a hierarchical system is said to be schedulable if each component is schedulable.

The analytical approaches [11], [3], [13] compute whether or not a system is schedulable, according to a scheduling policy, by giving a firm response to the following question: is the demand bound function \( \text{dbf} \) of each component workload \( W \), over a time interval \( t \), lower or equal to the supply bound function \( \text{sbf} \) of a resource according to interface \( I \), over the same time interval, i.e., \( \forall t > 0 \text{ dbf}(W, t) \leq \text{sbf}(I) \). If such an equation is satisfied, the component is said to be schedulable. In the same way, in a model-based setting [21], [2], [9], [4] a system is said to be schedulable if the error locations, stating the deadline violation, are unreachable.

In contrast to the mentioned techniques, we do not only consider if a system is schedulable or not, but we provide the \textit{degree of schedulability} (Sched\textsuperscript{o}) as a way to measure how schedulable a system is. We define the Sched\textsuperscript{o} of an entity (system, component or task) by the two concepts: \textit{Percentage of Missed Deadlines} (PoMD) and \textit{Degradation of Quality of Service} (DoQoS). Each of these concepts can be computed for either a task, a component or a complete embedded system. They should be measured or simulated over a sufficiently large time bounded run and a sufficiently large number of runs in order to obtain usable values.

The PoMD is computed over a set of traces as the number of missed deadlines divided by the number of triggerings of the associated tasks. The DoQoS is defined as 0 if no deadline is missed, otherwise it is the accumulation of overruns divided by the number of missed deadlines. Thus, representing the average overrun when a deadline is missed.

IV. \textbf{CONTINUOUS PROBABILITY BASED SPORADIC TASKS}

In this section, we introduce the characteristics of the probability-based sporadic tasks. Our framework models both a fixed inter-arrival time and a probability distribution. Obviously, a task cannot arrive before the inter-arrival time, and the inter-arrival time can potentially be set to zero. After the inter-arrival time, the arrival of a given task delays with \( \delta \) according to a continuous probability distribution, such as Gaussian \( N = (\mu, \sigma^2) \) with a mean value \( \mu \) and a variance \( \sigma^2 \) (Fig. 2(a)). Fig. 2 shows the three specific probability
distributions we consider in our setting: Gaussian, exponential and uniform. As the probability distribution is a parameter of the sporadic tasks in our framework, any user defined probability distribution can be used.

A. Probability Distributions

We have implemented the continuous probability distributions we consider via a set of UPPAAL embedded functions over the time domain. An example of a Gaussian normalized curve, generated by UPPAAL SMC, is depicted in Fig. 2(a) where the x axis represents continuous time from 0 to 100, $\mu=50$, and $\sigma=50$. Fig. 2(b) shows an exponential probability distribution, with the rate of exponential $\lambda$ being $1/800$. The smaller $\lambda$ is, the more spread out the distribution is. In contrast to the two previous probability distributions, the uniform distribution (Fig. 2(c)) has a equal probability for all time instances up to a maximum time where the probability drops to zero.

B. Conceptual Sporadic Task Model

Our conceptual event model is shown in Fig. 3(a). When the delay has elapsed, the event triggers the corresponding task and moves to the location InterArrivalWait waiting for one inter-arrival time $I$ before starting a new round. The conceptual task model (Fig. 3(a)) starts in location Wait waiting for the triggering event (trigger?) and then moves to the location Run. Depending on whether or not the deadline of the task is missed, the task moves either directly to location Wait or to location MissedDeadline. If the task finishes its current execution after missing a deadline, the overrun will be measured (used for estimating the DoQoS) before moving to the location Wait. The conceptual models shown in Fig. 3 are instantiated as UPPAAL models. Because of space limitations the UPPAAL models are not presented in this paper.

V. Analysis Models for the Degree of Schedulability

For our compositional analysis framework, the hierarchical scheduling systems and their analysis elements consist of environment models, scheduling models, resource model, and task models.

We are using UPPAAL SMC to perform a formalized statistical simulation of our models, known as Statistical Model Checking (SMC). SMC enables quantitative performance measurements instead of the Boolean (true, false) evaluation that symbolic model checking techniques provide. We can summarize the main features of UPPAAL SMC in the following:

- Stopwatches \cite{8} are clocks that can be stopped and resumed without a reset. They are very practical to measure the execution time of preemptive tasks.
- Simulation and estimation of the expected minimum or maximum value of expressions over a set of runs, $E[bound](\min: expr)$ and $E[bound](\max: expr)$, for a given simulation time and/or number of runs specified by bound.
- Probability evaluation $Pr[bound] (P)$ for a property $P$ to be satisfied within a given simulation time and/or number of runs specified by bound. $P$ is specified using either LTL or tMITL logic.

The disadvantage of using statistical model checking is that it will not provide complete certainty that a property is satisfied, but only verify it up to a specific confidence level, given as an analysis parameter \cite{6}.

A. Resource model

The resource model of this paper is a periodic one, which provides a specific amount of resources to a set of tasks or components \cite{19}. The resource model is given by a stochastic supplier, which supplies a resource allocation non-deterministically over supplier’s period.

Fig. 4 shows supplier’s supplying and tasks’ running of $T_p^3$ and $T_4^3$. In this setting, $T_4^3$ has priority over $T_p^3$, and
Fig. 4. Supplier and Task Execution.

Fig. 5. Budgets causing deadline miss

Fig. 6. PoMD calculator

Fig. 7. Targeting component

cntMissDline[tid], and DoQoS[tid] are used to calculate PoMD and DoQoS with PoMD calculator in Fig. 6 and the following queries:

\[
\begin{align*}
E[g\text{Clock} <= \text{simTime}; \text{simNum}] (\text{max: PoMD[tid]}) \\
E[g\text{Clock} <= \text{simTime}; \text{simNum}] (\text{max: DoQoS[tid]})
\end{align*}
\]

Where tid is the id of the task to be analyzed.

For every simulation (simNum) of which time is up to sim-Time, one PoMD[tid] is obtained by calculating the percentage of the accumulated number of missed deadlines of cntMissDline[tid], and the count of task executions of cntExecution[tid]. Finally, PoMD is calculated from the average of PoMD[tid]s of all traces. DoQoS[tid] measures delay after a task misses a deadline, and the maximum DoQoS[tid] is selected from one trace. Finally, DoQoS is determined by calculating the average of the maximum DoQoS[tid]s of each individual simulation trace.

VI. ANALYSIS OF THE DEGREE OF SCHEDULABILITY

We use as a running example in this section the Targeting component of Fig. 7. The workload is characterized by a periodic task \( T_{p3} \) and a sporadic task \( T_{s4} \). In our setting, \( T_{s4} \) has priority over \( T_{p3} \). Both tasks are scheduled according to the fixed priority scheduling (FPS) policy. The sporadic task \( T_{s4} \) follows the uniform probability distribution between 0 and 20 time units.

The analysis is performed in the following steps:

1) Estimate a budget for the component as described in Section V.

2) Analyze the Sched\(^\circ\) for the estimated and lower budgets.

To estimate the budget of a component, we use the budget estimation technique described in [4]. As a result, we found that 23 time units every 40 time units is a good candidate as a sufficient budget for both tasks. In order to have valid results, in the next analysis section we perform experiments where we analyze the same system with a varying amount of traces and simulation time. When reaching more than 1,000 traces and a simulation time of more than 100,000 time units we see that the results stabilize.

A. Analysis Results

In Table I, we show that \( T_3 \) and \( T_4 \) are schedulable under the budget (40, 23) even if \( T_4 \) is treated as a periodic task with a period equal to the minimal inter-arrival time. This is classical worst-case budget estimation, and our analysis also
confirms that tasks miss exactly 0% of their deadlines and have a DoQoS of 0. Throughout the running example, we use FPS scheduling but our framework supports other scheduling policies. For a deficit budget (40, 18), Table II shows the degree of schedulability when the sporadic task $T^4_s$ is assumed to follow an exponential probability distribution with different rates of exponential. Table III shows the Sched ² for the two tasks given a uniform probability distribution for triggering the sporadic task $T^4_s$. Table IV shows the results of our analysis when using different Gaussian distributions, all with a mean value $\mu$ of 10 and different deviations $\sigma$.

VII. CONCLUSION

We have presented a compositional method for analyzing the degree of schedulability of hierarchical real-time systems. The system is modeled in terms of components containing periodic and sporadic tasks. In order to characterize more accurately the arrival time of sporadic tasks, we introduced continuous probability distributions. Given hard and soft real-time requirements, our approach provides probabilistic guarantees on the system schedulability. The Sched ² is defined by the two factors: 1) PoMD and 2) DoQoS. These concepts are helpful when analyzing systems or components with insufficient budgets to meet all deadlines. UPPAAL SMC is used to perform statistical model checking, in order to compute the DoQoS and PoMD. A future work could be the application of the current framework to analyze the energy efficiency [5] of hierarchical scheduling systems.

REFERENCES


TABLE I

<table>
<thead>
<tr>
<th>Component (i[(0, 23)], FPS)</th>
<th>PoMD</th>
<th>DoQoS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T^3_s(40, 4)$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$T^3_s(40, 2)$</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

TABLE II

<table>
<thead>
<tr>
<th>Component (i[(40, 18), FP])</th>
<th>Rate of Exp. of $T^3_s$</th>
<th>PoMD</th>
<th>DoQoS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T^0_s(40, 4)$</td>
<td>1/100,000</td>
<td>0.350 ± 0.435</td>
<td>3.630 ± 0.672</td>
</tr>
<tr>
<td>$T^0_s(10, 1)$</td>
<td>1/1000</td>
<td>0.363 ± 0.434</td>
<td>4.950 ± 0.840</td>
</tr>
<tr>
<td>$T^0_s(1)$</td>
<td>1/10</td>
<td>0.488 ± 0.049</td>
<td>8.404 ± 1.186</td>
</tr>
<tr>
<td>$T^0_s(2)$ (Exp., 40, 2, 40)</td>
<td>1/100,000</td>
<td>0.233 ± 0.028</td>
<td>0.022 ± 0.040</td>
</tr>
<tr>
<td>$T^0_s(10, 2)$</td>
<td>1/1000</td>
<td>0.326 ± 0.032</td>
<td>2.167 ± 0.405</td>
</tr>
<tr>
<td>$T^0_s(10, 4)$</td>
<td>1/1000</td>
<td>0.072 ± 0.016</td>
<td>6.766 ± 0.636</td>
</tr>
</tbody>
</table>

TABLE III

<table>
<thead>
<tr>
<th>Component (i[(40, 18), FP])</th>
<th>PoMD</th>
<th>DoQoS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T^3_s(40, 4)$</td>
<td>0.497 ± 0.035</td>
<td>7.564 ± 1.126</td>
</tr>
<tr>
<td>$T^3_s(40, 2, 40)$</td>
<td>0.052 ± 0.015</td>
<td>6.515 ± 0.688</td>
</tr>
</tbody>
</table>

TABLE IV

<table>
<thead>
<tr>
<th>Component (i[(40, 18), FP])</th>
<th>PoMD</th>
<th>DoQoS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T^3_s(40, 4)$</td>
<td>2.000 ± 0.653</td>
<td>3.677 ± 2.411</td>
</tr>
<tr>
<td>$T^3_s(40, 2, 40)$</td>
<td>1.440 ± 0.539</td>
<td>3.024 ± 1.914</td>
</tr>
<tr>
<td>$T^3_s(40, 1)$</td>
<td>1.640 ± 0.608</td>
<td>3.200 ± 1.966</td>
</tr>
<tr>
<td>$T^3_s(8)$</td>
<td>1.400 ± 0.579</td>
<td>3.983 ± 2.611</td>
</tr>
<tr>
<td>$T^3_s(8)$</td>
<td>0.350 ± 0.255</td>
<td>1.687 ± 1.359</td>
</tr>
<tr>
<td>$T^3_s(8)$</td>
<td>0.490 ± 0.326</td>
<td>1.335 ± 1.076</td>
</tr>
<tr>
<td>$T^3_s(8)$</td>
<td>0.390 ± 0.249</td>
<td>0.551 ± 0.672</td>
</tr>
<tr>
<td>$T^3_s(10)$</td>
<td>0.600 ± 0.324</td>
<td>0.704 ± 0.768</td>
</tr>
</tbody>
</table>