ABSTRACT

Scheduling periodic hard real-time tasks is of great importance to many real-time applications, such as open-architecture machine tool controllers and avionics systems. The rate-monotonic scheduling algorithm has been proven to provide an optimal static-priority assignment under idealized conditions. However, some of these conditions are not met in a real computer system. In particular, the release times of tasks can deviate from the specified time instants because of operating system software timer unpredictability. In this paper, we investigate the timer behaviors in three commercial real-time operating systems, VxWorks, QNX and pSOSystem. Based on our findings, we propose an empirical task schedulability model, called RMTU (Rate-Monotonic in the presence of Timing Unpredictability), to augment the rate-monotonic scheduling theory in order to handle timing unpredictability. We then design an approach to systematically derive the model parameters by measurements. With RMTU, task deadlines can be empirically guaranteed. The validity of RMTU is supported by our measurement data. These results are useful not only to control application developers, but also to real-time practitioners at large.

1 Introduction

Scheduling periodic hard real-time tasks is of great importance to many real-time applications [20], such as open-architecture machine tool controllers [16] and avionics systems. For a set of independent, preemptive, and periodic tasks, whose hard deadlines are the same as their respective periods, Liu and Layland [14] proved that the rate-monotonic (RM) priority assignment is an optimal static-priority scheduling algorithm. A scheduling algorithm is said to be static if the priorities are not changed once they are assigned to the tasks; otherwise, it is dynamic.

However, this result was derived under idealized conditions, some of which do not hold in a real computer system. In particular, we observe that the intervals of periodic tasks may vary because the timers that realize the periodicity are not perfect [25]. The precise times of operating system (OS) timer firings can deviate from the nominal time instants, thus introducing variation in actual task periods. This can adversely affect the performance of scheduling algorithms in terms of deadline miss ratios.

To address this problem, we propose an empirical task schedulability model, called RMTU (Rate-Monotonic in the presence of Timing Unpredictability), to augment the original RM scheduling algorithm to handle timing unpredictability. In particular, we introduce two parameters to capture the timer delay and the CPU utilization by the OS. These model parameters are determined empirically and systematically for each target system by measuring the execution of a set of predetermined tasks. The model is empirical in the sense that its parameters are derived from measurement data and validated experimentally. We characterize timer behaviors on three commercial real-time operating systems (RTOSs), VxWorks, QNX and pSOSystem, and validate RMTU in a VxWorks testbed.

The remainder of the paper is organized as follows. Sections 2 describes our findings of timer characteristics in three commercial RTOSs—VxWorks, QNX and pSOSystem. In Sections 3 and 4, we describe RMTU and a systematic approach to derive model parameters, respectively. We discuss the results in Section 5. Finally, we present related work and conclusions in Sections 6 and 7, respectively.

2 Software timer characteristics

The RM scheduling theory [14] is very simple and elegant. However, not all its assumptions hold in real-world computer systems. In particular, we notice that, in a multi-tasking environment, periodic tasks rely on the OS timer to realize the periodicity. The timer service typically consists of one or more functions, such as the POSIX [4]
functions `timer_create()`, `timer_delete()`, `timer_gettime()` and `timer_settime()`. A task uses a software timer to generate an event periodically. After each invocation of the task is completed, the task goes into sleep until the timer wakes it up with the next timer event.

Clearly, the characteristics of the OS timer service have potentially significant impact on the performance of periodic tasks. We now present the experiments to measure the timer behaviors of three commercial RTOSs—VxWorks, QNX and pSOS System.

2.1 VxWorks timer

Our first set of timer experiments is conducted on a Motorola MVME147 board. This is a VMEbus-based processor board with a Motorola 30 MHz 68030 CPU and 4 MB of RAM. It runs VxWorks (version 5.1.1) [24]. VxWorks is a development and execution environment for real-time and embedded applications on a wide variety of target processors. It includes a high-performance scalable RTOS which executes on a target processor. The system clock resolution of the VxWorks kernel is 1 \( ms \) (1000 ticks/second).

The VME StopWatch [3] is used for timing measurement. It is a board that is plugged into the VME chassis and time-stamps bus events. These events are reads or writes to specific VME extended addresses. In our experiments, a simple inline function call is used to generate the events. The internal clock resolution of the VME StopWatch is 25 \( ns \).

We use the POSIX timer to periodically trigger a simple task \( \tau \) which consists of two consecutive function calls that generate events \( E_1 \) and \( E_2 \), respectively. There is no other code between the function calls. Therefore, the elapsed time between \( E_1 \) and \( E_2 \) is the execution time of one function call, while the elapsed time between two consecutive \( E_1 \) (or \( E_2 \)) events is the task interval (or timer interval, i.e., the interval between two consecutive timer firings). Table I lists the statistics of the task \( \tau \) with a 10 \( ms \) period using the POSIX timer, while Figures 1 and 2 show the histograms of the measured timer intervals and event-generating function execution times, respectively.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Execution Time</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>963</td>
<td>1.90 ( \mu s )</td>
<td>9999.9 ( \mu s )</td>
</tr>
<tr>
<td>962</td>
<td>0.0383 ( \mu s )</td>
<td>202.7 ( \mu s )</td>
</tr>
<tr>
<td>1.90 ( \mu s )</td>
<td>2.20 ( \mu s )</td>
<td>11964.9 ( \mu s )</td>
</tr>
</tbody>
</table>

Table I. VxWorks POSIX timer statistics.

From the measurements, we have two important observations about timer characteristics. First, timer intervals vary around the nominal period (see Fig. 1). This is because there are other OS activities besides the user task that consume CPU cycles. Many OS tasks run at higher priorities than user tasks. For example, the VxWorks process manager `tExcTask` runs at the highest priority. The hardware interrupts generated by devices (e.g., network devices) are also serviced at higher priorities by the OS kernel. All these inevitably cause fluctuations in the timer firings as well as task execution times (see Fig. 2).

Second, timers have “memory” behavior. By examining the actual measurement data (not shown here) [27], we observe that whenever an interval deviates from the nominal period by at least one or two percentage points, the next interval almost always swings in the opposite direction by roughly the same amount. For example, the longest interval of 11,964.9 \( \mu s \) in Table I is followed immediately by the shortest interval of 7,716.3 \( \mu s \), which is in turn followed by a slightly longer interval of 10,316.9 \( \mu s \). This can be explained by the use of absolute time inside the OS kernel. What happens is that one timer firing is late (hence the longer interval) because of other higher-priority system activities. The RTOS still tries to fire the timer at its next originally scheduled time. Therefore, the interval immediately following the longer one is shorter than the nominal period. In this example, the second timer firing is slightly early, which further shortens the second interval. The third firing is on time. The average of these three intervals is 9,999.4 \( \mu s \), which is very close to the nominal period of 10 \( ms \).

The nature of these two timer characteristics, variation and “memory” behavior, suggests that they are generic phenomena of RTOSs. To validate this, we con-
duct similar timer experiments using two other RTOSs—QNX and pSOSystem.

2.2 QNX timer

QNX is a commercial, micro-kernel, POSIX-compliant RTOS [17]. It uses a priority-based, preemptive kernel scheduler. Our experiments are conducted on a XYCOM XVME-674/16 board running QNX version 4.22. This board is a VMEbus PC/AT processor module with an Intel 66 MHz 80486DX2, 32 Mbytes dual-access DRAM, SVGA and IDE controllers. The QNX system clock resolution is set to 50 µs and the VME StopWatch is used for the timing measurement.

For the simple task \( \tau \) with a 10 ms period (same as that in the VxWorks experiments) using POSIX timer functions, Table II gives measured task execution time and interval statistics. The QNX POSIX timer also exhibits interval variation. The timer “memory” behavior is present as well, which can be illustrated by the following sequence of six consecutive timer intervals extracted from the measurement data: 9974.3, 9976.9, 10257.5, 9782.0, 9966.4, and 10005.0 µs. The first two intervals are close to the nominal period. The third interval is abnormally long, which is compensated by the next abnormally short interval. The fifth and sixth intervals are again close to the nominal period.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Execution Time</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950</td>
<td>1949</td>
<td></td>
</tr>
</tbody>
</table>

Table II. QNX POSIX timer statistics.

2.3 pSOSystem timer

The pSOSSystem is another commercial modular high-performance RTOS designed specifically for embedded microprocessors [5]. It includes a real-time multi-tasking kernel pSOS*. Though the pSOSSystem (version 1.1) is not POSIX-compliant, it has system calls similar to POSIX timer functions. We substitute the POSIX timer functions with pSOSSystem-specific functions in our experiments.

We use a setup similar to that for VxWorks and QNX, where pSOS* runs on a VMEbus-based processor board—Ironics IV3207 (Motorola 25 MHz 68040 with 4 MB RAM). The system clock resolution of pSOS* is 10 ms (100 ticks/second). Again, timer interval variation (Table III) and “memory” behavior are observed in the measurement data.

2.4 Timer delay factor

Suppose there is a set of \( m \) independent preemptive periodic tasks, \( \tau_1, \tau_2, \ldots, \tau_m \), with periods \( T_1, T_2, \ldots, T_m \), deadlines \( D_1, D_2, \ldots, D_m \), and worst-case nominal execution times \( C_1, C_2, \ldots, C_m \), and initial release times \( I_1, I_2, \ldots, I_m \), respectively. For each task \( \tau_i, i=1,2,\ldots,m \), its dead-

\[
	ext{Table III. pSOSystem timer statistics.}
\]

<table>
<thead>
<tr>
<th>Mean (µs)</th>
<th>Standard Deviation (µs)</th>
<th>Min (µs)</th>
<th>Max (µs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.65</td>
<td>9764.2</td>
<td>0.78</td>
<td>10244.8</td>
</tr>
</tbody>
</table>

\[
\text{Figure 3. Task execution with timer variation.}
\]

line is assumed to be equal to its period, i.e., \( D_i=T_i \). Without loss of generality, the task periods are assumed to be sorted in a non-decreasing order: \( T_1 \leq T_2 \leq \ldots \leq T_m \).

Figure 3 illustrates the effect of the timer “memory” behavior on task schedulability. The initial timer firing is on time at time \( I \), the second is late by \( v \) time units at \( I+T+v \), and the third is again on time at \( I+2T \). The net effect is a longer interval of \( T+v \), followed by a shorter interval of \( T-v \). We call the parameter \( v \) timer delay factor. If \( v > T-C \), the job in that period will miss its deadline.

Considering that the timer delay is caused by higher-priority OS activities, it should be independent of nominal timer intervals. To verify this, we measure the variation of timers with different nominal intervals. We use the VxWorks setup to run 8 periodic tasks, one at a time, and measure their periods. The nominal periods of these tasks range from 5 ms to 200 ms. The typical range of control task periods is from 1 ms to 100 ms. Since the tick size is 1 ms, it is inappropriate to run tasks with a period that is very close to, or less than, 1 ms.

Table IV summarizes the measurement results. The mean, minimum and maximum periods in the table are normalized by their respective nominal task periods so that it is easier to compare the measurements of tasks with different periods. For example, for the 10 ms task (the second row in the table), the actual mean, minimum and maximum intervals are 10000.38, 8391 and 11487 µs, respectively. We can see that the worst-case timer deviations are between one and two milliseconds and have no linear relationship with nominal timer periods.

\[
\text{Table IV. Statistics of timers with different nominal periods.}
\]

<table>
<thead>
<tr>
<th>Task Period (ms)</th>
<th>Sample Size</th>
<th>Interval Mean (µs)</th>
<th>Standard Deviation (µs)</th>
<th>Interval Min (µs)</th>
<th>Interval Max (µs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1213</td>
<td>1.26</td>
<td>0.0011</td>
<td>-1669</td>
<td>1544</td>
</tr>
<tr>
<td>10</td>
<td>802</td>
<td>0.38</td>
<td>175</td>
<td>-1669</td>
<td>1487</td>
</tr>
<tr>
<td>20</td>
<td>607</td>
<td>-0.33</td>
<td>196</td>
<td>-1653</td>
<td>1385</td>
</tr>
<tr>
<td>30</td>
<td>539</td>
<td>-0.71</td>
<td>247</td>
<td>-1373</td>
<td>1393</td>
</tr>
<tr>
<td>50</td>
<td>394</td>
<td>0.42</td>
<td>254</td>
<td>-1060</td>
<td>1156</td>
</tr>
<tr>
<td>70</td>
<td>522</td>
<td>1.40</td>
<td>224</td>
<td>-1180</td>
<td>1284</td>
</tr>
<tr>
<td>100</td>
<td>417</td>
<td>3.54</td>
<td>150</td>
<td>-1425</td>
<td>1608</td>
</tr>
<tr>
<td>200</td>
<td>510</td>
<td>1.06</td>
<td>133</td>
<td>-1565</td>
<td>1540</td>
</tr>
</tbody>
</table>
3 RMTU: an empirical schedulability model

As described in the previous section, we have observed the timer variation and “memory” behavior in three popular commercial RTOSs. The nature of such timing unpredictability leads us to believe that they are generic phenomena. In [25] and [26], we further show that timing unpredictability has an adverse impact on scheduling algorithm performance in terms of deadline miss ratios. For example, some task deadline guarantees that would have been provided by RM under idealized conditions may no longer be possible due to timing unpredictability. Therefore, it is important to take timing unpredictability into account when checking the schedulability of tasks. Specifically, we propose an empirical task schedulability model, called RMTU (Rate-Monotonic in the presence of Timing Unpredictability), which extends the original RM scheduling theory to handle timing unpredictability. By empirical, we mean that the model parameters reflect some specific properties of the target system and thus are experimentally derived from the measurement data of the target system.

RMTU quantifies the effect of timing unpredictability on task schedulability by extending the RM scheduling theory with the following sufficient condition:

\[
\forall \tau, 1 \leq i \leq m, U_s + \frac{C_i}{T_i} + \ldots + \frac{C_i}{T_i} + v \leq i \left(2^i - 1\right) \quad \text{(EQ 1)}
\]

where \(U_s\) is a constant representing the CPU utilization of OS activities, such as the OS kernel scheduling and interrupt handling; and \(v\) is the worse-case task delay factor, which is a constant, independent of task periods \(T_i\). RMTU uses \(U_s\) and \(v\) to capture the unpredictability of RTOS. These two parameters can be determined empirically as described in the next section.

Once we obtain the parameters \(U_s\) and \(v\) and know the task characteristics \(C_i\) and \(T_i\), we can use EQ 1 to determine if all the hard deadlines can be met. This model results in a tighter upper bound of CPU utilization for RM task scheduling in the presence of RTOS unpredictability than the original RM.

4 Derivation of model parameters

Now that we have introduced the empirical task schedulability model RMTU, we will describe a systematic approach to deriving model parameters.

4.1 Assumptions and derivation

To apply RMTU, we need to know all the parameters in EQ 1. \(T_1, T_2, \ldots, T_m\) are the task periods, which are realized using OS timers in applications. In an actual implementation of tasks, the periods could differ from the specified values. Since timers can deliver mean periods very close (typically, within 0.1%) to their respective task periods (see also the timer measurement data in Section 2), we will use the task periods as given in our schedulability analysis.

Theoretically, the worst-case nominal task execution times \(C_1, C_2, \ldots, C_m\) may be obtained by analyzing the task code, assuming that the task execution times are bounded. In a real system, however, the same OS activities that cause timer glitches can affect task execution times as well. Any CPU time used by OS activities during the execution of a user task increases the task response time. Since this increase is not caused by other user tasks, it effectively increases the task’s execution time. Therefore, the measured worst-case execution times reflect more accurately the actual system performance than those obtained from code analysis and should be used in determining task schedulability. Finally, the worst-case timer delay factor \(v\) can only be obtained experimentally.

Our systematic approach to parameter derivation includes experiments using a single task, in which case RMTU can be written as EQ 2. For simplicity, the task subscript and the less-than sign are dropped. This simplification is acceptable because the derived model parameters will be plugged back into EQ 1.

\[
C = (1 - U_s)T - v \quad \text{(EQ 2)}
\]

We define the available CPU utilization (\(U_{avail}\)) as the percentage of CPU cycles that is available to user tasks, which is equal to 1 - \(U_s\). We further define the achievable CPU utilization (\(U_{achiev}\)) as the summation of the worst-case execution time to period ratios of individual tasks, which implies the maximal CPU utilization by user tasks without missing any deadlines. \(U_{achiev}\) has the same value as the CPU utilization in the original RM theory.

In EQ 2, the task period \(T\) is given, while the worst-case execution time \(C\) can be measured. The available CPU utilization 1 - \(U_s\) and the timer delay factor \(v\) can be derived by using the approach we will describe next.

4.2 Systematic approach to parameter derivation

Based on EQ 2, we design the following systematic approach to determining the model parameters empirically:

1. Run a task with a fixed period and an adjustable execution time at the highest priority allowed by the RTOS.

2. Measure the start and completion time of each task invocation to check if it misses its deadline. The number of measured task invocations should be as large as possible and no less than a few hundreds for a given task execution time.\(^1\)

3. If the task misses its deadline, decrease the task execution time and repeat Step 2; otherwise, increase it. Adjust the task execution time until the task does not miss its deadline \(and\) a small\(^2\) increase in its execution time would cause it to miss its deadline.

4. Record the worst-case execution time\(^3\) \(C\) and the

---

1. Density estimation techniques may be used to determine the number of measurement and confidence level of the results [21].
2. In our experiments, a small change is typically less than 1% of total task execution time.
corresponding task period $T$.

5. Repeat Steps 1-4 for several (e.g., five or more) different task periods in the range of the target application task periods.\(^4\)

6. After obtaining $(C, T)$ value pairs from Steps 1-5, use linear regression to derive the available CPU utilization $(I-U_s)$ and the timer delay factor $v$.

### 4.3 Experimental results

We conducted the experiments described above on the computer system running VxWorks described in Section 2.1. Fig. 4 shows the relationship between the measured achievable CPU utilization and the task period. While the RM scheduling theory predicts that the achievable CPU utilization should be 1 regardless of different periods in the single-task case, Fig. 4 clearly shows an upward trend in achievable CPU utilization as the task period increases (as indicated by EQ 2). This is because of the period-independent timer delay factor in RMTU.

![Figure 4. Single-task achievable CPU utilization versus task period.](image)

Table V lists the measurement and computation results. The first eight rows of the table show the data for the tasks whose periods range from 5 to 200 ms. The last three rows are the results of linear regression. Using the measured maximum execution times, linear regression produces a timer delay factor $v$ of 1.802 ms and an available CPU utilization $(I-U_s)$ of 1.0016 (fourth column in the table). We will discuss why $U\text{avail}$ could be greater than 1 in Section 5. The correlation coefficient is very close to 1, indicating a strong positive correlation between the task period and measured worst-case execution time.

### 4.4 Model validation

To validate the empirical model, we have run several sets of three and five independent tasks in the same

3. This is actually the worst-case response time, i.e., the elapsed time between the start and the finish of the task. With our approach, the effect of OS activities on task execution time is also factored into our empirical schedulability model.

4. This is intended to improve the precision of linear regression.

### 5 Discussion

There are a few important observations from our experimental measurements in Section 4. First, RMTU gives a tighter upper bound of the achievable CPU utilization than the original RM scheduling theory. Second, the tasks can typically reach better achievable CPU utilization than predicted by either RM or RMTU, because most task executions take less time than the worst case.

Third, the available CPU utilization $(I-U_s)$ could be greater than one. At the first glance, it appears inappropriate that the available CPU utilization can be greater than one. However, the available CPU utilization is not the actual CPU utilization. When the worst-case task execution times are used in EQ 2 to derive model parameters, the available CPU utilization can be over-estimated (thus resulting in a value larger than one). This is because only a very small number of task invocations actually take that much time to complete.

RMTU is premised on that the system characteristics revealed by single-task experiments are applicable to multiple tasks. However, we found that the presence of other tasks does have an effect on the variation of task execution times, even if the task has the highest priority [27]. It was also observed that a single application can receive periodic callbacks or timer notifications every 1 ms, for instance, with a high degree of precision. But when two or more applications try to do this simultaneously, there are outliers lasting tens of milliseconds that never appear in the single application case. This happens even though the system is essentially completely idle [7]. While RMTU (or

5. Measurement results [27] are omitted due to space limit.
any empirical models) cannot theoretically provide hard guarantees, it works well for practical purposes as shown by our experiments. For applications that require deadline guarantees but not necessarily hard, we instead propose probabilistic deadline guarantees [27].

6 Related work

RM has been extended for the cases with $D_i \leq T_i$ [8, 10, 11, 12], with deadlines being an integral multiple of periods [11, 13], and with $D_i > T_i$ [19]. Rajkumar et al. [18] and Lortz and Shin [15] studied scheduling issues for periodic tasks with blocking due to non-preemptive critical task sections or mutually exclusive access to shared resources. All of the above extensions of the RM scheduling algorithm assume ideal timers.

Kettler, Katcher and Stronsider [9] proposed an engineering methodology that allows users to accurately model and evaluate RTOSs, thus providing a framework to account for implementation costs in real-time scheduling theory. Their approach requires a high degree of expertise to create a valid model for any given RTOS. Such expertise may be scarce among users. Furthermore, as RTOSs become more sophisticated and new releases come out once or twice a year, modeling them becomes more difficult. Instead, we design a set of systematic experiments to characterize the underlying RTOS and provide a simple empirical sufficient condition for applying the RM scheduling theory to practical systems.

Audsley et al. [1, 22] extended static-priority scheduling techniques to address timer jitter. Their solution is to use a non-preemptive approach trying to avoid or at least minimize the problem. Baruah et al. [2] provided a formal feasibility analysis of dynamic-priority systems exhibiting jitter. However, we observe that the overhead of run-time priority change can be prohibitively expensive (e.g., QNX’s setpriot() system call takes more than 100 μs). A custom-built scheduling system is not always a viable alternative either. We instead address the problem by incorporating jitter effects in static-priority systems using commercial RTOSs.

Jeffay and Stone [6] studied the feasibility and schedulability problems for periodic tasks in the presence of interrupt handlers. While their solutions are computationally feasible, some of their assumptions are overly-simplistic. For example, they assume that interrupt handlers are strictly periodic, which is rarely the case in practice.

7 Conclusions

Many scheduling algorithms, e.g., rate-monotonic, may not work well in practice because some of their idealized assumptions do not hold in a real computer system. In particular, we observed timer period variations in three commercial RTOSs—VxWorks, QNX and pSOS. In order to handle such timing unpredictability, we proposed an empirical model—RMTU—to extend rate-monotonic. We further designed a set of systematic experiments, from which the model parameters can be empirically derived. Our experiments verified the validity of RMTU.

References