

Optimal Priority Assignment for Real-Time Systems: A Coevolution-Based Approach

Jaekwon Lee · Seung Yeob Shin ·
Shiva Nejati · Lionel C. Briand

Received: date / Accepted: date

Abstract In real-time systems, priorities assigned to real-time tasks determine the order of task executions, by relying on an underlying task scheduling policy. Assigning optimal priority values to tasks is critical to allow the tasks to complete their executions while maximizing safety margins from their specified deadlines. This enables real-time systems to tolerate unexpected overheads in task executions and still meet their deadlines. In practice, priority assignments result from an interactive process between the development and testing teams. In this article, we propose an automated method that aims to identify the best possible priority assignments in real-time systems, accounting for multiple objectives regarding safety margins and engineering constraints. Our approach is based on a multi-objective, competitive coevolutionary algorithm mimicking the interactive priority assignment process between the development and testing teams. We evaluate our approach by applying it to six industrial systems

Jaekwon Lee
SnT, University of Luxembourg, Luxembourg
University of Ottawa, Canada
E-mail: jaekwon.lee@uni.lu

Seung Yeob Shin
SnT, University of Luxembourg, Luxembourg
E-mail: seungyeob.shin@uni.lu

Shiva Nejati
University of Ottawa, Canada
SnT, University of Luxembourg, Luxembourg
E-mail: snejati@uottawa.ca

Lionel C. Briand
SnT, University of Luxembourg, Luxembourg
University of Ottawa, Canada
E-mail: lionel.briand@uni.lu

from different domains and several synthetic systems. The results indicate that our approach significantly outperforms both our baselines, i.e., random search and sequential search, and solutions defined by practitioners. Our approach scales to complex industrial systems as an offline analysis method that attempts to find near-optimal solutions within acceptable time, i.e., less than 16 hours.

Keywords Priority Assignment, Schedulability Analysis, Real-Time Systems, Coevolutionary Search, Search-Based Software Engineering

1 Introduction

Mission-critical systems are found in many different application domains, such as aerospace, automotive, and healthcare domains. The success of such systems depends on both functional and temporal correctness. For functional correctness, systems are required to provide appropriate outputs in response to the corresponding stimuli. Regarding temporal correctness, systems are supposed to generate outputs within specified time constraints, often referred to as deadlines. The systems that have to comply with such deadlines are known as real-time systems (Liu, 2000). Real-time systems typically run multiple tasks in parallel and rely on a real-time scheduling policy to decide which tasks should have access to processing cores, i.e., CPUs, at any given time.

While developing a real-time system, one of the most common problems that engineers face is the assignment of priorities to real-time tasks in order for the system to meet its deadlines. Based on priorities of real-time tasks, the system’s task scheduler determines a particular order for allocating real-time tasks to processing cores. Hence, a priority assignment that is poorly designed by engineers makes the system scheduler execute tasks in an order that is far from optimal. In addition, the system will likely violate its performance and time constraints, i.e., deadlines, if a poor priority assignment is used.

In real-time systems, the problem of optimally assigning priorities to tasks is important not only to avoid deadline misses but also to maximize *safety margins* from task deadlines and is subject to *engineering constraints*. Tasks may exceed their expected execution times due to unexpected interrupts. For example, it is infeasible to test an aerospace system exhaustively on the ground such that potential environmental uncertainties, e.g., those related to space radiations, are accounted for. Hence, engineers assign optimal priorities to tasks such that the remaining times from tasks’ completion times to their deadlines, i.e., safety margins, are maximized to cope with potential uncertainties. Furthermore, engineers typically have to account for additional engineering constraints, e.g., they assign higher priorities to critical tasks that must always meet their deadlines compared to the tasks that are less critical or non-critical.

A brute force approach to find an optimal priority assignment would have to examine all $n!$ distinct priority assignments, where n denotes the number of tasks. Furthermore, for a given priority assignment, schedulability analysis is, in general, known as a hard problem (Audsley, 2001), which determines

whether or not tasks will always complete their executions within their specified deadlines. Thus, optimizing priority assignments is also a hard problem because the space of all possible system states to explore in order to find optimal priority assignments is very large. Most of the prior works on optimizing priority assignments provide analytical methods (Fineberg and Serlin, 1967; Leung and Whitehead, 1982; Audsley, 1991; Davis and Burns, 2007; Chu and Burns, 2008; Davis and Burns, 2009; Davis and Bertogna, 2012), which rely on well-defined system models and are very restrictive. For example, they assume that tasks are independent, i.e., tasks do not share resources (Davis et al., 2016; Zhao and Zeng, 2017). Industrial systems, however, are typically not compatible with such (simple) system models. In addition, none of the existing work addresses the problem of optimizing priority assignments by simultaneously accounting for multiple objectives, such as safety margins and engineering constraints, as discussed above.

Search-based software engineering (SBSE) has been successfully applied in many application domains, including software testing (Wegener et al., 1997; Wegener and Grochtman, 1998; Lin et al., 2009; Arcuri et al., 2010; Shin et al., 2018), program repair (Weimer et al., 2009; Tan et al., 2016; Abdessalem et al., 2020), and self-adaptation (Andrade and Macêdo, 2013; Chen et al., 2018; Shin et al., 2020), where the search spaces are very large. Despite the success of SBSE, engineering problems in real-time systems have received much less attention in the SBSE community. In the context of real-time systems, there exists limited work on finding stress test scenarios (Briand et al., 2005) and predicting worst-case execution times (Lee et al., 2020b), which complements our work.

In practice, priority assignments result from an interactive process between the development and testing teams. While developing a real-time system, developers assign priorities to real-time tasks in the system and then testers stress the system to check whether or not the system meets its specified deadlines. If testers find a problematic condition under which any of the tasks violates its deadline, developers have to modify the priority assignment to address the problem. The back-and-forth between the development and testing teams continues until a priority assignment that does not lead to any deadline miss is found or the one that yields the least critical deadline misses is identified. The process is, however, not automated.

In this article, we use metaheuristic search algorithms to automate the process of assigning priorities to real-time tasks. To mimic the interactive back-and-forth between the development and testing teams, we use competitive coevolutionary algorithms (Luke, 2013). Coevolutionary algorithms are a specialized class of evolutionary search algorithms. They simultaneously coevolve two populations (also called species) of (candidate) solutions for a given problem. They can be cooperative or competitive. Such competitive coevolution is similar to what happens in nature between predators and preys. For example, faster preys escape predators more easily, and hence they have a higher probability of generating offspring. This impacts the predators, because they need to evolve as well to become faster if they want to feed and survive (Meneghini

et al., 2016). Hence, the two species, i.e., predators and preys, have coevolved competitively. We note that no species has the competing traits of predators and preys simultaneously as such species could not evolve to survive. In our context, priority assignments defined by developers can be seen as preys and stress test scenarios as predators. The priority assignments need to evolve so that stress testing is not able to push the system into breaking its real-time constraints. Dually, stress test scenarios should evolve to be able to break the system when there is a chance to do so.

Contributions. We propose an Optimal Priority Assignment Method for real-time systems (OPAM). Specifically, we apply multi-objective, two-population competitive coevolution (Popovici et al., 2012) to address the problem of finding near-optimal priority assignments, aiming at maximizing the magnitude of safety margins from deadlines and constraint satisfaction. In OPAM, two species relate to priority assignment and stress testing coevolve synchronously, and compete against each other to find the best possible solutions. We evaluated OPAM by applying it to six complex, industrial systems from different domains, including the aerospace, automotive, and avionics domains, and several synthetic systems. Our results show that: (1) OPAM finds significantly better priority assignments compared to our baselines, i.e., random search and sequential search, (2) the execution time of OPAM scales linearly with the number of tasks in a system and the time required to simulate task executions, and (3) OPAM priority assignments significantly outperform those manually defined by engineers based on domain expertise.

We note that OPAM is the first attempt to apply coevolutionary algorithms to address the problem of priority assignment. Further, it enables engineers to explore trade-offs among different priority assignments with respect to two objectives: maximizing safety margins and satisfying engineering constraints. Our full evaluation package is available online (Lee et al., 2021).

Organization. The remainder of this article is structured as follows: Section 2 motivates our work. Section 3 defines our specific problem of priority assignment in practical terms. Section 4 discusses related work. Sections 5 and 6 describe OPAM. Section 7 evaluates OPAM. Section 8 concludes this article.

2 Motivating case study

We motivate our work using an industrial case study from the satellite domain. Our case study concerns a mission-critical real-time satellite, named ESAIL (LuxSpace, 2021), which has been developed by LuxSpace – a leading system integrator for microsatellites and aerospace system. ESAIL tracks vessels’ movements over the entire globe as the satellite orbits the earth. The vessel-tracking service provided by ESAIL requires real-time processing of messages received from vessels in order to ensure that their voyages are safe with the assistance of accurate, prompt route provisions. Also, as ESAIL orbits the planet, it must be oriented in the proper position on time in order to provide

services correctly. Hence, ESAIL's key operations, implemented as real-time tasks, need to be completed within acceptable times, i.e., deadlines.

Engineers at LuxSpace analyze the schedulability of ESAIL across different development stages. At an early design stage, the engineers use a priority assignment method that extends the rate monotonic scheduling policy (Fineberg and Serlin, 1967), which is a theoretical priority assignment algorithm used in real-time systems. At a later development stage, if the engineers found that any real-time task of ESAIL cannot complete its execution within its deadline, the engineers, in our study context, reassign priorities to tasks in order to address the problem of deadline violations.

The rate monotonic policy assigns priorities to tasks that arrive to be executed periodically and must be completed within a certain amount of time, i.e., periodic tasks with hard deadlines. According to the policy, periodic tasks that arrive frequently have higher priorities than those of other tasks that arrive rarely. In ESAIL, for example, if the vessel-tracking task arrives every 100ms and the satellite-position control task arrives every 150ms, the former has a higher priority than the latter. However, the rate monotonic policy does not account for tasks that arrive irregularly and should be completed within a reasonable amount of time, i.e., aperiodic tasks with soft deadlines. ESAIL contains aperiodic tasks with soft deadlines as well, such as a task for updating software. Hence, the engineers extend the rate monotonic policy to assign priorities to all tasks of ESAIL. The extensions are as follows: First, the engineers assign priorities to periodic tasks based on the rate monotonic policy. Second, the engineers assign lower priorities to aperiodic tasks than those of periodic tasks. As aperiodic tasks with soft deadlines are typically considered less critical than periodic tasks with hard deadlines, the engineers aim to ensure that periodic tasks complete their executions within their deadlines by assigning lower priorities to aperiodic tasks while periodic tasks have higher priority. Engineers use a heuristic to assign priorities to aperiodic tasks. They treat aperiodic tasks as (pseudo-)periodic tasks by setting aperiodic tasks' (expected) minimum arrival rates as their fixed arrival periods, making the tasks frequently arrive. The engineers then apply the rate monotonic policy for the aperiodic tasks with the synthetic periods while ensuring that aperiodic tasks have lower priorities than those of periodic tasks.

A priority assignment made at an early design stage keeps changing while developing ESAIL due to various reasons, such as changes in requirements and implementation constraints. At a development stage, instead of relying on the extended rate monotonic policy, the engineers assign priorities based on their domain expertise, manually inspecting schedulability analysis results. Hence, a priority assignment at later development stages often does not follow the extended rate monotonic policy. For example, as aperiodic tasks are also expected to be completed within a reasonable amount of time, some aperiodic tasks may have higher priorities than some periodic tasks as long as they are schedulable.

Engineers at LuxSpace, however, are still faced with the following issues:

- (1) Their priority assignment method, which extends the rate monotonic

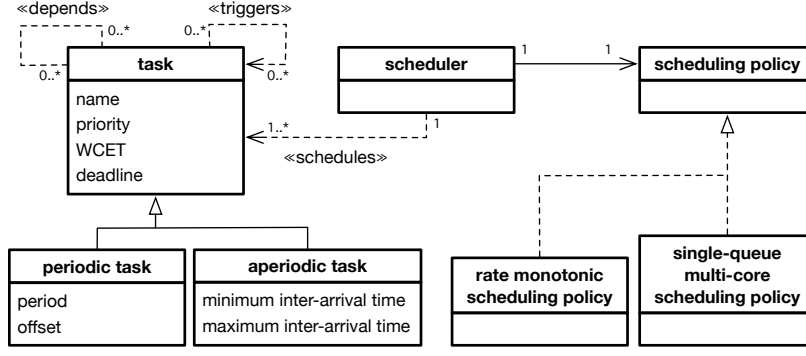


Fig. 1: A conceptual model representing the key abstractions to analyze optimal priority assignments.

scheduling policy, assigns priorities to tasks in order to ensure only that tasks are to be schedulable. However, engineers have a pressing need to understand the quality of priority assignments in detail as they impact ESAIL operations differently. For example, once ESAIL is launched into orbit, the satellite operates in the space environment, which is inherently impossible to be fully tested on the ground. Unexpected space radiations may trigger unusual system interrupts, which hasn't been observed on the ground, resulting in overruns of ESAIL tasks' executions. In such cases, a priority assignment assessed on the ground may not be able to tolerate such unexpected uncertainties. Hence, engineers need a priority assignment that enables ESAIL tasks to tolerate unpredictable uncertainties as much as possible and to be schedulable. (2) Engineers at LuxSpace assign priorities to tasks without any systematic assistance. Instead, they rely on their expertise and the current practices described above to manually assign priorities to ensure that tasks are to be schedulable. To this end, we are collaborating with LuxSpace to develop a solution for addressing these issues in assigning task priority.

3 Problem description

This section defines the task, scheduler, and schedulability concepts, which extend the concepts defined in our previous work (Lee et al., 2020b) by augmenting our previous definitions with the notions of safety margins, constraints in assigning priorities, and relationships between real-time tasks. We then describe the problem of optimizing priority assignments such that we maximize the magnitude of safety margins and the degree of constraint satisfaction. Figure 1 shows an overview of the conceptual model that represents the key abstractions required to analyze optimal priority assignments for real-time systems. The entities in the conceptual model are described below.

Task. We denote by j a real-time task that should complete its execution within a specified deadline after it is activated (or arrived). Every real-time

task j has the following properties: priority denoted by $pr(j)$, deadline denoted by $dl(j)$, and worst-case execution time (WCET) denoted by $wcet(j)$. Task priority pr determines if an execution of a task is preempted by another task. Typically, a task j preempts the execution of a task j' if the priority of j is higher than the priority of j' , i.e., $pr(j) > pr(j')$. The $pr(j)$ priority is a fixed value assigned to task j . Such fixed priorities are determined offline; hence, they are not changed online for any reason. Note that a real-time task scheduler that relies on fixed priorities is applied in all the study subjects in this article (see Section 7.2) and is commonly used in industrial systems (Briand et al., 2005; Guan et al., 2009; Lin et al., 2009; Anssi et al., 2011; Zeng et al., 2014; Di Alesio et al., 2015; Dürr et al., 2019; Lee et al., 2020a).

The $dl(j)$ function determines the deadline of a task j relative to its arrival time. A task deadline can be either *hard* or *soft*. A hard deadline of a task j constrains that j *must* complete its execution within a deadline $dl(j)$ after j is activated. While violations of hard deadlines are not acceptable, depending on the operating context of a system, violating soft deadlines may be to some extent tolerated. Note that we use a metaheuristic search relying on fitness functions quantifying the degrees of deadline misses, safety margins, and constraint satisfaction. Such functions do not depend on the nature of the deadlines. Our approach outputs a set of priority assignments that are Pareto optimal with respect to safety margins and constraint satisfaction. Engineers then perform domain-specific trade-off analysis among Pareto solutions. Hence, in this article, we handle hard and soft deadline tasks in the same manner.

Real-time tasks are either *periodic* or *aperiodic*. Periodic tasks, which are typically triggered by timed events, are invoked at regular intervals specified by their *period*. We denote by $pd(j)$ the period of a periodic task j , i.e., a fixed time interval between subsequent activations (or arrivals) of j . Any task that is not periodic is called aperiodic. Aperiodic tasks have irregular arrival times and are activated by external stimuli which occur irregularly. In real-time analysis, based on domain knowledge, we typically specify a minimum inter-arrival time denoted by $pmin(j)$ and a maximum inter-arrival time denoted by $pmax(j)$ indicating the minimum and maximum time intervals between two consecutive arrivals of an aperiodic task j . In real-time analysis, sporadic tasks are often separately defined as having irregular arrival intervals and hard deadlines (Liu, 2000). In our conceptual definitions, however, we do not introduce new notations for sporadic tasks because the deadline and period concepts defined above sufficiently characterize sporadic tasks. Note that for periodic tasks j , we have $pmin(j) = pmax(j) = pd(j)$. Otherwise, for aperiodic tasks j , we have $pmax(j) > pmin(j)$.

Task relationships. The execution of a task j depends not only on its own parameters described above, e.g., priority $pr(j)$ and period $pd(j)$, but also on its relationships with other tasks. Relationships between tasks are typically determined by task interactions related to accessing shared resources and triggering arrivals of other tasks (Di Alesio et al., 2012). Specifically, if two tasks j and j' access a shared resource r in a mutually exclusive way, j may be blocked from executing for the period during which j' accesses r . We

denote by $dp(j, j')$ the resource-dependency relation between tasks j and j' that holds if j and j' have mutually exclusive access to a shared resource r such that they cannot be executed in parallel or preempt each other, but one can execute only after the other has completed accessing r .

The other type of relationship between tasks is related to a task j triggering the arrival of another task j' . This is a common interaction between tasks (Locke et al., 1990; Anssi et al., 2011; Di Alesio et al., 2015). For example, j may hand over some of its workload to j' due to performance or reliability reasons. We denote by $tr(j, j')$ the triggering relation between tasks j and j' that holds if j triggers the arrival of j' . We note that both relationships are defined at the level of tasks, following prior works (Locke et al., 1990; Anssi et al., 2011; Di Alesio et al., 2015) describing the five industrial case study systems used in our experiments (see Section 7.2).

Scheduler. Let J be a set of tasks to be scheduled by a real-time scheduler. A scheduler then dynamically schedules executions of tasks in J according to the tasks' arrivals and the scheduler's scheduling policy over the scheduling period $\mathbb{T} = [0, \mathbf{T}]$. We denote by $at_k(j)$ the k th arrival time of a task $j \in J$. The first arrival of a periodic task j does not always occur immediately at the system start time (0). Such offset time from the system start time to the first arrival time $at_1(j)$ of j is denoted by $offset(j)$. For a periodic task j , the k th arrival of j within \mathbb{T} is $at_k(j) \leq \mathbf{T}$ and is computed by $at_k(j) = offset(j) + (k-1) \cdot pd(j)$. For an aperiodic task j' , $at_k(j')$ is determined based on the $k-1$ th arrival time of j' and its minimum and maximum arrival times. Specifically, for $k > 1$, $at_k(j') \in [at_{k-1}(j') + pmin(j'), at_{k-1}(j') + pmax(j')]$ and, for $k = 1$, $at_1(j') \in [pmin(j'), pmax(j')]$, where $at_k(j') < \mathbf{T}$.

A scheduler reacts to a task arrival at $at_k(j)$ by scheduling the execution of j . Depending on a scheduling policy (e.g., rate monotonic scheduling policy for single-core systems (Fineberg and Serlin, 1967) and single-queue multi-core scheduling policy (Arpaci-Dusseau and Arpaci-Dusseau, 2018)), an arrived task j may not start its execution at the same time as it arrives when higher priority tasks are executing on all processing cores. Also, task executions may be interrupted due to preemption. We denote by $et_k(j)$ the completion time for the k th arrival of a task j . According to the worst-case execution time of a task j , we have: $et_k(j) \geq at_k(j) + wcet(j)$.

During system operation, a scheduler generates a *schedule scenario* which describes a sequence of task arrivals and their completion time values. We define a schedule scenario as a set S of tuples $(j, at_k(j), et_k(j))$ indicating that a task j has arrived at $at_k(j)$ and completed its execution at $et_k(j)$. Due to a degree of randomness in task execution times and aperiodic task arrivals, a scheduler may generate a different schedule scenario for different runs of a system.

Figure 2 shows two schedule scenarios S (Figure 2a) and S' (Figure 2b) produced by a scheduler over the $[0, 23]$ time period of a system run. Both S and S' describe executions of three tasks, j_1 , j_2 , and j_3 arrived at the same time stamps (see at_i in the figures). In both scenarios, the aperiodic task j_1 is characterized by: $pmin(j_1) = 5$, $pmax(j_1) = 13$, $dl(j_1) = 4$, and $wcet(j_1) = 2$. The

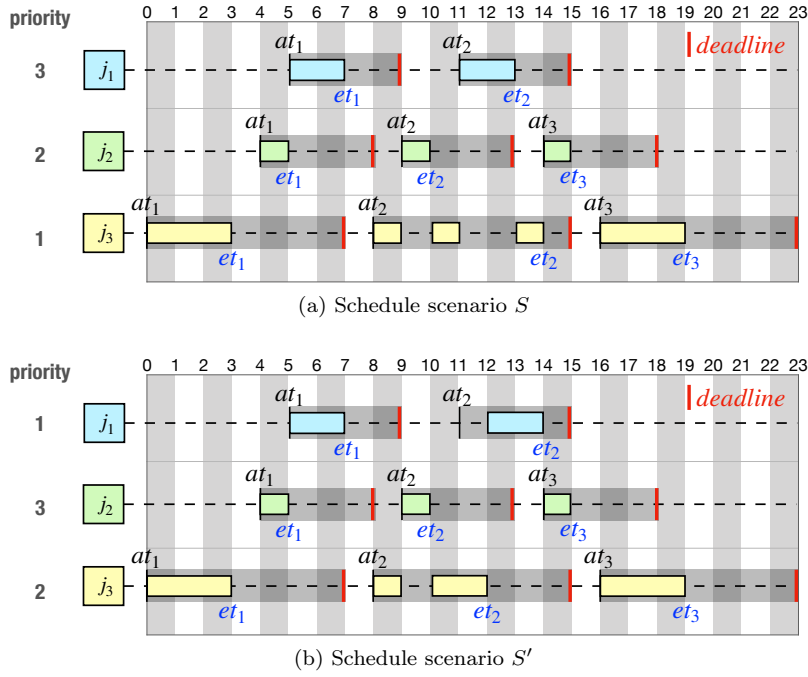


Fig. 2: Example schedule scenarios S and S' of three tasks: j_1 , j_2 , and j_3 . (a) The S schedule scenario is produced when $pr(j_1) = 3$, $pr(j_2) = 2$, and $pr(j_3) = 1$. (b) The S' schedule scenario is produced when $pr(j_1) = 1$, $pr(j_2) = 3$, and $pr(j_3) = 3$.

aperiodic task j_2 is characterized by: $pmin(j_2) = 3$, $pmax(j_2) = 10$, $dl(j_2) = 4$, and $wcet(j_2) = 1$. The periodic task j_3 is characterised by: $pd(j_3) = 8$, $dl(j_3) = 7$, and $wcet(j_3) = 3$. The priorities of the three tasks in S (resp. S') satisfy the following: $pr(j_1) > pr(j_2) > pr(j_3)$ (resp. $pr(j_2) > pr(j_3) > pr(j_1)$). In both scenarios, task executions can be preempted depending on their priorities. Then, S is defined by $S = \{(j_1, 5, 7), \dots, (j_2, 4, 5), \dots, (j_3, 8, 14), (j_3, 16, 19)\}$; and S' is defined by $S' = \{(j_1, 5, 7), \dots, (j_2, 4, 5), \dots, (j_3, 8, 12), (j_3, 16, 19)\}$.

Schedulability. Given a schedule scenario S , a task j is *schedulable* if j completes its execution before its deadline, i.e., for all $et_k(j)$ observed in S , $et_k(j) \leq at_k(j) + dl(j)$. Let J be a set of tasks to be scheduled by a scheduler. A set J of tasks is then schedulable if for every schedule S of J , we have no task $j \in J$ that misses its deadline.

As shown in schedule scenarios S and S' presented in Figures 2a and 2b, respectively, all three tasks, j_1 , j_2 , and j_3 , are schedulable. However, we note that the overall amounts of remaining time, i.e., safety margins, from the tasks' completions to their deadlines observed in S and S' are different (see the second

completion times and deadlines of j_1 , j_2 , and j_3 in S and S') because S and S' are produced by using different priority assignments. Engineers typically desire to assign optimal priorities to real-time tasks that aim at maximizing such safety margins, as discussed below.

Problem. In real-time systems, fixed priorities are typically assigned to tasks (Davis et al., 2016; Lee et al., 2020a). Finding an appropriate priority assignment is important not only for ensuring the schedulability of a system but also for maximizing the safety margins within which a system can tolerate unexpected execution time overheads. For example, if an unpredictable error occurs and triggers check-point mechanisms (Davis and Burns, 2007), which re-execute part or all of a task j , then the execution time of j unexpectedly overruns. Hence, engineers need an optimal priority assignment that maximizes the overall remaining times from task completion times to task deadlines, i.e., safety margins.

While assigning priorities to tasks, engineers also account for constraints, that are often but not always domain-specific. For example, aperiodic tasks' priorities should be lower than those of periodic tasks because periodic tasks are often more critical than aperiodic tasks. Hence, engineers develop a system that prioritizes executions of periodic tasks over aperiodic tasks. Recall from Section 2, this constraint is desirable by engineers. When needed, however, engineers can violate the constraint to some extent in order to ensure that aperiodic tasks complete within a reasonable amount of time while periodic tasks meet their deadlines. Constraints can be either *hard* constraints, which must be satisfied, or *soft* constraints, which are desired to be satisfied. In our study, hard constraints need to be assured while scheduling tasks, e.g., a running task's priority must be higher than a ready task's priority, which are enforced by a scheduler. In the context of optimizing priority assignments, we focus on maximizing the extent of satisfying soft constraints. We refer to a soft constraint as a constraint in this paper.

Our work aims at optimizing priority assignments that maximize the safety margins while satisfying such constraints. Specifically, for a set J of tasks to be analyzed, we define three concepts as follows: (1) a priority assignment for J denoted by \vec{P} , (2) the magnitude of safety margins for a priority assignment \vec{P} denoted by $fs(\vec{P})$, and (3) the degree of constraint satisfaction denoted by $fc(\vec{P})$. We note that Section 6.3 describes how we optimize \vec{P} , and compute $fs(\vec{P})$ and $fc(\vec{P})$ in detail. Our study aims at finding a set \mathbf{B} of best possible priority assignments that are Pareto optimal (Knowles and Corne, 2000) such that a priority assignment $\vec{P} \in \mathbf{B}$ maximizes both $fs(\vec{P})$ and $fc(\vec{P})$, and any other priority assignments in \mathbf{B} are equally viable.

4 Related Work

This section discusses related research strands in the areas of priority assignments, real-time analysis using exhaustive techniques, search-based analysis in real-time systems, and coevolutionary analysis in software engineering.

Table 1: Comparing our work, OPAM, with existing priority assignment techniques with respect to the properties captured in their underlying system models.

| Properties | OPAM | RMPO | DMPO | OPA | OPA-MLD | RPA | FNR-PA | PRPA | OPTA | EPAF |
|-------------------------|------|------|------|-----|---------|-----|--------|------|------|------|
| Periodic task | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Aperiodic task | ○ | | | ○ | ○ | ○ | ○ | ○ | | |
| Resource dependency | ○ | | | | | | | | | |
| Triggering relationship | ○ | | | | | | | | | |
| Multi-core system | ○ | | | ○ | ○ | ○ | ○ | ○ | | |
| Safety margin | ○ | | | | | ○ | ○ | ○ | | |
| Engineering constraint | ○ | | | | ○ | | | | | |

Priority assignment. The problem of optimally assigning priorities to real-time tasks has been widely studied (Fineberg and Serlin, 1967; Liu and Layland, 1973; Leung and Whitehead, 1982; Audsley, 1991; Tindell et al., 1994; George et al., 1996; Audsley, 2001; Davis and Burns, 2007; Chu and Burns, 2008; Davis and Burns, 2009, 2011; Davis and Bertogna, 2012; Davis et al., 2016; Zhao and Zeng, 2017; Hatvani et al., 2018). Fineberg and Serlin (1967) reported early work that relies on a simple system model, assuming, for example, that all tasks arrive periodically, tasks run on a single processing core, tasks’ deadlines are equal to their periods, and task executions are independent from one another. They proposed a priority assignment method, named rate-monotonic priority ordering (RMPO), that assigns higher priorities to the tasks with shorter periods. RMPO can find a feasible priority assignment that guarantees periodic tasks to be schedulable when such priority assignments exist (Liu and Layland, 1973). Leung and Whitehead (1982) extended RMPO to relax one of the underlying assumptions made in RMPO. Specifically, their priority assignment approach, known as deadline-monotonic priority ordering (DMPO), accounts for task deadlines that can be less than or equal to their periods. In contrast to our work, however, these methods are often not applicable to industrial systems that are not compatible with their simplified system models. Recall from Section 3 that a realistic system typically consists of both periodic and aperiodic tasks. Task executions depend on their relationships, i.e., resource dependencies and triggering relationships, with other tasks.

Audsley (2001) designed a priority assignment method, named optimal priority assignment (OPA), that relies on an existing schedulability analysis method M . OPA guarantees to find a feasible priority assignment that is schedulable according to M if such priority assignments exist. OPA is applicable to more complex systems than those supported by the methods mentioned

above, i.e., RMPO and DMPO. Specifically, OPA can find a feasible priority assignment even in the following situations: (1) First arrivals of periodic tasks occur after some offset time (Audsley, 1991). (2) Aperiodic tasks have arbitrary deadlines (Tindell et al., 1994). (3) Task executions are scheduled based on a non-preemptive scheduling policy (George et al., 1996). (4) Tasks run on multiple processing cores (Davis and Burns, 2011). Unlike our approach that accounts for two objectives, safety margins and engineering constraints (see Section 3), OPA attempts to find a feasible priority assignment whose only objective is to make all tasks schedulable. Note that such a feasible priority assignment does not necessarily maximize safety margins as discussed in Section 3. Hence, a feasible priority assignment obtained by OPA is often fragile and sensitive any changes in task executions and unable to accommodate unexpected overheads in task execution times, which are commonly observed in industrial systems (Davis and Burns, 2007).

OPA has been extended by several works (Davis and Burns, 2007; Chu and Burns, 2008; Davis and Burns, 2009; Davis and Bertogna, 2012). Davis and Burns (2007) presented a robust priority assignment method (RPA) with a degree of tolerance for unexpected overruns of task execution times. Chu and Burns (2008) introduced an extended OPA algorithm (OPA-MLD) that minimizes the lexicographical distance between the desired priority assignment and the one obtained by the algorithm. OPA-MLD enables important tasks to have higher priorities. Davis and Bertogna (2012) proposed an RPA extension (FNR-PA) to make RPA work when a system allows task preemption to be deferred for some interval of time. Davis and Burns (2009) developed a probabilistic robust priority assignment method (PRPA) for a real-time system to be less likely to violate its deadlines. Even though the prior works mentioned above improve OPA to some extent, they assume that task executions are independent of one another. In contrast to these existing approaches, OPAM accounts for dependencies among task executions, i.e., resource dependencies and triggering relationships (see our problem description in Section 3).

Some recent priority assignment techniques address scalability. Hatvani et al. (2018) presented an optimal priority and preemption-threshold assignment algorithm (OPTA) that attempts to decrease the computation time for finding a feasible priority assignment. OPTA uses a heuristic to traverse a problem space while pruning infeasible paths to efficiently and effectively explore the problem space. Zhao and Zeng (2017) introduced an effective priority assignment framework (EPAF) that combines a commercial solver for integer linear programs and their problem-specific optimization algorithm. However, these methods rely on simple system models that assume, for example, task executions to be independent and running on a single processing core. Therefore, the applicability of these techniques is limited. In contrast, recall from Sections 2 and 3 that our approach aims at scaling to complex industrial systems while accounting for realistic system characteristics regarding task periods, inter-arrival times, resource dependencies, triggering relationships, and multiple processing cores.

Table 1 compares our work, OPAM, with the other priority assignment techniques mentioned above. As shown in the table, we note that prior works rely on system models that are very restrictive. In particular, existing work assumes that task executions are independent of one another. However, task dependencies such as resource dependencies and triggering relationships are commonly observed in industrial systems. In addition, we note that no existing solution simultaneously accounts for safety margins and engineering constraints. Hence, to our knowledge, OPAM is the first attempt to provide engineers with a set of equally viable priority assignments, allowing trade-off analysis with respect to the two objectives: maximizing safety margins and satisfying engineering constraints.

Real-time analysis using exhaustive techniques. Constraint programming and model checking have been applied to conclusively and exhaustively verify whether or not a system meets its deadlines (Kwiatkowska et al., 2011; Di Alesio et al., 2012; Nejati et al., 2012; Di Alesio et al., 2013). Existing research on priority assignment based on OPA rely on such exhaustive techniques to prove the schedulability of a set of tasks for a given priority assignment. We note that schedulability analysis is, in general, an NP-hard problem (Davis et al., 2016) that cannot be solved in polynomial time. As a result, exhaustive techniques based on model checking and constraint solving are often not amenable to analyze large industrial systems such as ESAIL – our motivating case study system – described in Section 2. To assess if exhaustive techniques could scale to ESAIL, as discussed in Section 7.8, we performed a preliminary experiment using UPPAAL (Behrmann et al., 2004), a model checker for real-time systems. We observed that UPPAAL was not able to verify schedulability of ESAIL tasks for a fixed priority assignment even after letting it run for several days (see Section 7.8 for more details).

Search-based analysis in real-time systems. In real-time systems, most of the existing works that use search-based techniques focus on testing (Wegener et al., 1997; Wegener and Grochtmann, 1998; Briand et al., 2005; Lin et al., 2009; Arcuri et al., 2010). Wegener et al. (1997, 1998) introduced a testing approach based on a genetic algorithm that aims to check computation time, memory usage, and task synchronization by analyzing the control flow of a program. Briand et al. (2005) applied a genetic algorithm to find stress test scenarios for real-time systems. Lin et al. (2009) proposed a search-based approach to check whether a real-time system meets its timing and security constraints. Arcuri et al. (2010) presented a black-box system testing approach based on a genetic algorithm. Beyond testing real-time systems, Nejati et al. (2013, 2014) developed a search-based trade-off analysis technique that helps engineers balance the satisfaction of temporal constraints and keeping the CPU time usage at an acceptable level. Lee et al. (2020b) combined a search algorithm and machine learning to estimate safe ranges of worst-case task execution times within which tasks likely meet their deadlines. In contrast to these prior works, OPAM addresses the problem of optimally assigning priorities to real-time tasks while accounting for multiple objectives regarding safety margins and engineering constraints, thus enabling Pareto (trade-off)

analysis. Further, OPAM uses a multi-objective, competitive coevolutionary search algorithm, which has been rarely applied to date in prior studies of real-time systems, as discussed next.

Coevolutionary analysis in software engineering. Despite the success of search-based software engineering (SBSE) in many application domains including software testing (Wegener et al., 1997; Wegener and Grochtman, 1998; Lin et al., 2009; Arcuri et al., 2010; Shin et al., 2018), program repair (Weimer et al., 2009; Tan et al., 2016; Abdesslem et al., 2020), and self-adaptation (Andrade and Macêdo, 2013; Chen et al., 2018; Shin et al., 2020), coevolutionary algorithms have been applied in only a few prior studies (Wilkerson and Tauritz, 2010; Wilkerson et al., 2012; Boussaa et al., 2013). Wilkerson et al. (2010, 2012) present a coevolution-based approach to automatically correct software. Their work introduced a program representation language to facilitate their automated corrections. Boussaa et al. (2013) developed a code-smells detection approach. The main idea is to evolve two competing populations of code-smell detection rules and artificial code-smells. Unlike these prior works, we study the problem of optimally assigning priorities to tasks in real-time systems. To our knowledge, we are the first to address the priority assignment problem using a multi-objective, competitive coevolutionary search algorithm.

5 Approach Overview

Finding an optimal priority assignment is an inherently interactive process. In practice, once engineers assign priorities to the real-time tasks in a system, testers then stress the system to find a condition, i.e., a particular sequence of task arrivals, in which a task execution violates its deadline. Testers typically use a simulator or hardware equipment to stress the system by triggering plausible worst-case arrivals of tasks that maximize the likelihood of deadline misses. If testers find task arrivals that induce deadline misses, the task arrivals are reported to engineers in order to fix the problem by reassigning priorities. This interactive process of assigning priorities and testing schedulability continues until both engineers and testers ensure that the tasks meet their deadlines.

For such intrinsically interactive problem-solving domains, we conjecture that coevolutionary algorithms are potentially suitable solutions. A coevolutionary algorithm is a search algorithm that mutually adapts one of different species, e.g., in our study, two populations of priority assignments and task-arrival sequences, acting as foils against one another. Specifically, we apply multi-objective, two-population competitive coevolution (Luke, 2013) to address our problem of finding optimal priority assignments (see Section 3). In our approach, the two populations of priority assignments and stress test scenarios, i.e., task-arrival sequences, evolve synchronously, competing with each other in order to search for optimal priority assignments that maximize the magnitude of safety margins from deadlines and the extent of constraint satisfaction. Note that better priority assignments enable a system to achieve

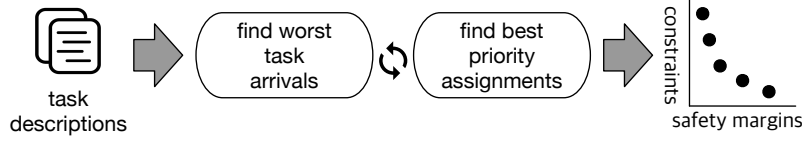


Fig. 3: An overview of our Optimal Priority Assignment Method for real-time systems (OPAM).

larger safety margins. Hence, those priority assignments have a higher chance to pass stress test scenarios. This impacts the stress test scenarios because they need to evolve as well, aiming at inducing deadline misses in the system.

Recall from Section 4 that most of the existing SBSE research relies on search algorithms using a single population (Chen et al., 2018; Abdessalem et al., 2020; Shin et al., 2020). However, such algorithms do not fit the problem of priority assignments targeted here. When (1) two competing traits between task arrivals and priority assignments are encoded together in an individual of a single population and (2) two contradicting fitness functions regarding safety margins and deadline misses, which are exact opposites, assess such individuals, the notion of Pareto optimality is not applicable. In that case, maximizing the magnitude of safety margins necessarily entails minimizing the magnitude of deadline misses. Hence, a single population-based search algorithm cannot make Pareto improvements that maximize safety margins (resp. deadline misses) while not minimizing deadline misses (resp. safety margins). Specifically, the dominance relation over such individuals does not exist because if an individual I is strictly better than another individual I' in one fitness value, I is always worse than I' in the other fitness value. Hence, we are not able to obtain equally viable solutions with respect to the contradicting objectives using such a method.

Figure 3 shows an overview of our proposed solution: Optimal Priority Assignment Method for real-time tasks (OPAM). OPAM requires as input task descriptions defined by engineers, which specify task characteristics and their relationships (see Section 3). Given such input task descriptions, the “find worst task arrivals” and “find best priority assignments” steps aim at generating worst-case sequences of task arrivals and best-case priority assignments, respectively. A worst-case sequence of task arrivals means that the magnitude of deadline misses, i.e., the amounts of time from task deadlines to task completion times, is maximized when tasks arrive as defined in the sequence. Note that if there is no deadline miss, a task-arrival sequence is considered worst-case if tasks complete their executions as close to their deadlines as possible. In contrast, a priority assignment is best-case when the magnitude of safety margins is maximized. Beyond maximizing safety margins, the “find best priority assignments” step accounts for satisfying engineering constraints in assigning priorities to tasks. OPAM evolves two competing populations of task-arrival sequences and priority assignments synchronously generated from the two steps. OPAM then outputs a set of priority assignments that are Pareto

optimal with regards to the magnitude of safety margins and the extent of satisfying constraints. Hence, OPAM allows engineers to perform domain-specific trade-off analysis among Pareto solutions and is useful in practice to support decision making with respect to their task design. For example, suppose engineers develop a weakly hard real-time systems (Bernat et al., 2001) that can tolerate occasional deadline misses. In that case, engineers may consider a few deadline misses as less important (as long as their consequences are negligible) than the overall magnitude of safety margins in their trade-off analysis. Section 6 describes OPAM in detail.

6 Competitive Coevolution

Figure 4 describes the OPAM algorithm for finding optimal priority assignments, which employs multi-objective, two-population competitive coevolution. The algorithm first randomly initializes two populations \mathbf{A} and \mathbf{P} for task-arrival sequences and priority assignments, respectively (lines 13–15). For \mathbf{A} , OPAM randomly varies task arrivals of aperiodic tasks to create ps_a task-arrival sequences, according to the input task descriptions D . Regarding \mathbf{P} , OPAM randomly creates ps_p priority assignments that may include one defined by engineers if available.

The two populations sequentially evolve during the allotted analysis budget (see line 17 in Figure 4). The best priority assignment is the one that makes tasks schedulable and maximizes the magnitude of safety margins, while satisfying engineering constraints for a given worst sequence of task arrivals. Hence, searching for the best priority assignments involves searching for the worst sequences of task arrivals. We create two populations \mathbf{A} and \mathbf{P} searching for the worst arrival sequences and the best priority assignments, respectively. The fitness values of task-arrival sequences in \mathbf{A} are computed based on how well they challenge the priority assignments in \mathbf{P} , i.e., maximizing the magnitude of deadline misses (line 20). Likewise, the priority assignments in \mathbf{P} are evaluated based on how well they perform against the task-arrival sequences in \mathbf{A} , i.e., maximizing the magnitude of safety margins while satisfying constraints (line 25). Once the two populations are assessed against each other, OPAM generates the next populations based on the computed fitness values (lines 21 and 26). OPAM tailors the breeding mechanisms of steady-state genetic algorithms (GA) (Whitley and Kauth, 1988) for \mathbf{A} and NSGAII (Deb et al., 2002) for \mathbf{P} .

OPAM uses two types of fitness functions, namely internal and external fitness evaluations, which play a different and complementary role as described below. The two internal fitness evaluations in lines 20 and 25 of the listing in Figure 4 aim at selecting individuals – task-arrival sequences and priority assignments – for breeding the next \mathbf{A} and \mathbf{P} populations. OPAM evaluates the external fitness for the \mathbf{P} population of priority assignments to find a best Pareto front (lines 28–31). As shown in lines 20 and 25, the internal fitness values of individuals in \mathbf{A} (resp. \mathbf{P}) are computed based on how they perform

```

1  Algorithm Search optimal priority assignments
2  Input  $D$ : task descriptions
3  Input  $n_c$ : number of coevolution cycles //budget
4  Input  $ps_a$ : population size //task-arrival sequences
5  Input  $ps_p$ : population size //priority assignments
6  Input  $cp_a$ : crossover probability //task-arrival sequences
7  Input  $cp_p$ : crossover probability //priority assignments
8  Input  $mp_a$ : mutation probability //task-arrival sequences
9  Input  $mp_p$ : mutation probability //priority assignments
10 Input  $E$ : set of task-arrival sequences //external evaluation
11 Output  $B$ : best Pareto front
12
13 //initialize populations
14  $A \leftarrow \text{randomize\_arrivals}(D, ps_a)$ 
15  $P \leftarrow \text{randomize\_priorities}(D, ps_p)$ 
16
17 for  $n_c$  times do
18     //evolution: find worst-case sequences of task arrivals
19     //objective: deadline misses
20      $\text{evaluate\_internal\_fitness\_arrivals}(A, P)$ 
21      $A \leftarrow \text{breed\_arrivals}(A, P, cp_a, mp_a)$  //GA
22
23     //evolution: find best-case priority assignments
24     //objectives: safety margins and constraints
25      $\text{evaluate\_internal\_fitness\_priorities}(P, A)$ 
26      $P \leftarrow \text{breed\_priorities}(P, A, cp_p, mp_p)$  //NSGAII
27
28     //external fitness evaluation
29     //objectives: safety margins and constraints
30      $\text{evaluate\_external\_fitness}(P, E)$ 
31      $B \leftarrow \text{select\_best}(P \cup B)$ 
32
33 return  $B$ 

```

Fig. 4: Multi-objective two-population competitive coevolution for finding optimal priority assignments.

with respect to individuals in P (resp. A). Hence, an individual's internal fitness is assessed through interactions with competing individuals. For example, a priority assignment in the first generation may have acceptable fitness values regarding safety margins and constraint satisfaction with respect to the first generation of task-arrival sequences, which are likely far from worst-case sequences. However, priority assignment fitness may get worse in later generations as the task-arrival sequences evolve towards larger deadline misses. Thus, if OPAM simply monitors internal fitness, it cannot reliably detect coevolutionary progress as an individual's internal fitness changes according to competing individuals. The problem of monitoring progress in coevolution has been observed in many studies (Ficici, 2004; Popovici et al., 2012). To address it, OPAM computes external fitness values of priority assignments in P based on a set E of task-arrival sequences generated independently from the coevolution process. By doing so, OPAM can observe the monotonic improvement

of external fitness for priority assignments. We note that, in general, if interactions between two competing populations are finite and any interaction can be examined with non-zero probability at any time, monotonicity guarantees that a coevolutionary algorithm converges to a solution (Popovici et al., 2012).

We note that our approach for evolving task-arrival sequences is based on past work (Briand et al., 2005), where a specific genetic algorithm configuration was proposed to find worst-case task-arrival sequences. One significant modification is that OPAM accounts for task relationships – resource-dependency and task triggering relationships – and a multi-core scheduling policy based on simulations to evaluate the magnitude of deadline misses.

Following standard practice (Ralph et al., 2020), the next sections describe OPAM in detail by defining the representations, the scheduler, the fitness functions, and the evolutionary algorithms for coevolving the task-arrival sequences and priority assignments. We then describe the external fitness evaluation of OPAM.

6.1 Representations

OPAM coevolves two populations of task-arrival sequences and priority assignments. A task-arrival sequence is defined by their inter-arrival time characteristics (see Section 3). A priority assignment is defined by a function that maps priorities to tasks.

Task-arrival sequences. Given a set J of tasks to be scheduled, a feasible sequence of task arrivals is a set A of tuples $(j, at_k(j))$ where $j \in J$ and $at_k(j)$ is the k th arrival time of a task j . Thus, a solution A represents a valid sequence of task arrivals of J (see valid $at_k(j)$ computation in Section 3). Let $\mathbb{T} = [0, \mathbf{T}]$ be the time period during which a scheduler receives task arrivals. The size of A is equal to the number of task arrivals over the \mathbb{T} time period. Due to the varying inter-arrival times of aperiodic tasks (Section 3), the size of A will vary across different sequences.

Priority assignments. Given a set J of tasks to be scheduled, a feasible priority assignment is a list \vec{P} of priority $pr(j)$ for each task $j \in J$. OPAM assigns a non-negative integer to a priority $pr(j)$ of j such that priorities are comparable to one another. The size of \vec{P} is equal to the number of tasks in J . Each task in J has a unique priority. Hence, a priority assignment \vec{P} is a permutation of all tasks' priorities. We note that these characteristics of priority assignments are common in many real-time analysis methods (Audsley, 2001; Davis and Burns, 2007; Zhao and Zeng, 2017) and industrial systems (e.g., see our six industrial case study systems described in Section 7.2).

6.2 Simulation

OPAM relies on simulation for analyzing the schedulability of tasks in a scalable way. For instance, an inter-arrival time of a software update task in a

satellite system is approximately at most three months. In such cases, conducting an analysis based on an actual scheduler is prohibitively expensive. Also, applying an exhaustive technique for schedulability analysis typically doesn't scale to an industrial system (e.g., see our experiment results using a model checker described in Section 7.8). Instead, OPAM uses a real-time task scheduling simulator, named OPAMScheduler, which applies a scheduling policy, i.e., single-queue multi-core scheduling policy (Arpaci-Dusseau and Arpaci-Dusseau, 2018), based on discrete simulation time events. Note that we chose the single-queue multi-core scheduling policy for OPAMScheduler since our case study systems (described in Section 7.2) rely on this policy.

OPAMScheduler takes as input a feasible task-arrival sequence A and a priority assignment \vec{P} for scheduling a set J of tasks. It then outputs a schedule scenario as a set S of tuples $(j, at_k(j), et_k(j))$ where $at_k(j)$ and $et_k(j)$ are the k th arrival and end time values of a task j , respectively (see Section 3). For each task j , OPAMScheduler computes $et_k(j)$ based on its WCET and scheduling policy while accounting for task relationships (see the $dp(j, j')$ resource-dependency relationship and the $tr(j, j')$ task triggering relationship in Section 3). To simulate the worst-case executions of tasks, OPAMScheduler assigns tasks' WCETs to their execution times.

OPAMScheduler implements a single-queue multi-core scheduling policy (Arpaci-Dusseau and Arpaci-Dusseau, 2018), which schedules a task j with explicit priority $pr(j)$ and deadline $dl(j)$. When tasks arrive, OPAMScheduler puts them into a single queue that contains tasks to be scheduled. At any simulation time, if there are tasks in the queue and multiple cores are available to execute tasks, OPAMScheduler first fetches a task j from the queue in which j has the highest priority $pr(j)$. OPAMScheduler then allocates task j to any available core. Note that if task j shares a resource with a running task j' in another core, i.e., the $dp(j, j')$ resource-dependency relationship holds, j will be blocked until j' releases the shared resource.

OPAMScheduler works under the assumption that context switching time is negligible, which is also a working assumption in many scheduling analysis methods (Liu and Layland, 1973; Audsley, 2001; Di Alesio et al., 2015). Note that the assumption is practically valid and useful at an early development step in the context of real-time analysis. For instance, our collaborating partner, LuxSpace, accounts for the waiting time of tasks due to context switching between tasks through adding some extra time to WCET estimates at the task design stage. Note that OPAM can be applied with any scheduling policy, including those that account for context switching time and multiple queues.

6.3 Fitness functions

Internal fitness: deadline misses. Given a feasible task-arrival sequence A and a priority assignment \vec{P} , we formulate a function, $fd(A, \vec{P})$, to quantify the degree of deadline misses regarding a set J of tasks to be scheduled. To compute $fd(A, \vec{P})$, OPAM runs OPAMScheduler for A and \vec{P} and obtains

a schedule scenario S . We denote by $dist_k(j)$ the distance between the end time and the deadline of the k th arrival of task j observed in S and define $dist_k(j) = et_k(j) - at_k(j) + dl(j)$ (see Section 3 for the notation end time $et_k(a)$, arrival time $at_k(j)$, and deadline $dl(j)$). We denote by $lk(j)$ the last arrival index of a task j in A . Given a set J of tasks to be scheduled, the $fd(A, \vec{P})$ function is defined as follows:

$$fd(A, \vec{P}) = \sum_{j \in J, k \in [1, lk(j)]} 2^{dist_k(j)}$$

Note that $fd(A, \vec{P})$ is defined as an exponential equation. Hence, when all task executions observed in a schedule scenario S meet their deadlines, $fd(A, \vec{P})$ is a small value as any distance $dist_k(j)$ between the task end time and the deadline of the k th arrival of task j is a negative value. In contrast, deadline misses result in positive values for $dist_k(j)$. In such cases, $fd(A, \vec{P})$ is a large value. The exponential form of $fd(A, \vec{P})$ was precisely selected for this reason, to assign large values for deadline misses but small values when deadlines are met. By doing so, $fd(A, \vec{P})$ prevents an undesirable solution that would result into many task executions meeting deadlines obfuscating a smaller number of deadline misses.

Following the principles of competitive coevolution, individuals in a population \mathbf{A} of task-arrival sequences need to be assessed by pitting them against individuals in the other population \mathbf{P} of priority assignments. We denote by $fd(A, \mathbf{P})$ the internal fitness function that quantifies the overall magnitude of deadline misses across all priority assignment $\vec{P} \in \mathbf{P}$, regarding a set J of tasks to be scheduled. The $fd(A, \mathbf{P})$ fitness is used for breeding the next population of task-arrival sequences. OPAM aims to maximize $fd(A, \mathbf{P})$, defined as follows:

$$fd(A, \mathbf{P}) = \sum_{\vec{P} \in \mathbf{P}} fd(A, \vec{P}) / |\mathbf{P}|$$

Internal fitness: safety margins. Given a feasible priority assignment \vec{P} and a task-arrival sequence A , we denote by $fs(\vec{P}, A)$ the magnitude of safety margins regarding a set J of tasks to be scheduled. The computation of $fs(\vec{P}, A)$ is similar to the computation of $fd(A, \vec{P})$ regarding the use of OPAM-Scheduler, which outputs a schedule scenario S . The difference is that OPAM reverses the sign of $fd(A, \vec{P})$ as OPAM aims at maximizing the magnitude of safety margins. Given a set J of tasks to be scheduled, the $fs(\vec{P}, A)$ function is defined as follows:

$$fs(\vec{P}, A) = \sum_{j \in J, k \in [1, lk(j)]} -2^{dist_k(j)} \quad (\text{i.e., } -fd(A, \vec{P}))$$

Given two populations \mathbf{P} and \mathbf{A} of priority assignments and task-arrival sequences, similar to internal fitness $fd(A, \mathbf{P})$, priority assignments in \mathbf{P} need to be assessed against task-arrival sequences in \mathbf{A} . We formulate an internal fitness function, $fs(\vec{P}, \mathbf{A})$, to quantify the overall magnitude of safety margins

across all task-arrival sequences $A \in \mathbf{A}$, regarding a set J of tasks to be scheduled and a priority assignment \vec{P} . OPAM relies on the $fs(\vec{P}, \mathbf{A})$ function to breed the next population of priority assignments. OPAM aims to maximize $fs(\vec{P}, \mathbf{A})$, which is defined as follows:

$$fs(\vec{P}, \mathbf{A}) = \sum_{A \in \mathbf{A}} fs(\vec{P}, A) / |\mathbf{A}|$$

Internal fitness: constraints. Given a priority assignment \vec{P} , we formulate an internal fitness function, $fc(\vec{P})$, to quantify the degree of satisfaction of soft constraints set by engineers. Such function is required as we recast the satisfaction of such constraints into an optimization problem, in order to minimize constraint violations. Specifically, OPAM accounts for the following constraint: aperiodic tasks should have lower priorities than those of periodic tasks. Recall from Section 2 that engineers consider this constraint to be desirable. We denote by $lp(\vec{P})$ the lowest priority of periodic tasks in \vec{P} . For a set J of tasks to be scheduled, OPAM aims to maximize $fc(\vec{P})$, which is defined as follows:

$$fc(\vec{P}) = \sum_{j \in J} \begin{cases} lp(\vec{P}) - pr(j), & \text{if } j \text{ is an aperiodic task} \\ 0, & \text{otherwise} \end{cases}$$

Greater $pr(j)$ values denote higher priorities. Given a priority assignment \vec{P} , if $pr(j)$ for an aperiodic task j is lower than the priority of any of the periodic tasks, $lp(\vec{P}) - pr(j)$ is a positive value. OPAM measures the difference between priorities of aperiodic and periodic tasks. By doing so, $fc(\vec{P})$ rewards aperiodic tasks that satisfy the above constraint and consistently penalizes those that violate it. Hence, OPAM aims at maximizing $fc(\vec{P})$.

External fitness: safety margins and constraints. To examine the quality of priority assignments and monitor the progress of coevolution, OPAM takes as input a set \mathbf{E} of task-arrival sequences created independently from the coevolution process. Given a set \mathbf{E} of task-arrival sequences and a priority assignment \vec{P} , OPAM utilizes $fs(\vec{P}, \mathbf{E})$ and $fc(\vec{P})$ described above as external fitness functions for quantifying the magnitude of safety margins and the extent of constraint satisfaction, respectively. As \mathbf{E} does not change over the coevolution process, $fs(\vec{P}, \mathbf{E})$ is used for evaluating a priority assignment \vec{P} since it is not impacted by the evolution of task-arrival sequences. Hence, external fitness functions ensure that OPAM monitors the progress of coevolution in a stable manner. Given two populations \mathbf{P} and \mathbf{A} of priority assignments and task-arrival sequences, we recall that the $fd(A, \mathbf{P})$ internal fitness function quantifies the overall magnitude of deadline misses across all priority assignments in \mathbf{P} for the given sequence of task arrivals A . The $fs(\vec{P}, \mathbf{A})$ internal fitness function quantifies the overall magnitude of safety margins across all sequences of task arrivals in \mathbf{A} for the given priority assignments \vec{P} . Hence, the internal fitness of A (resp. \vec{P}) is assessed through interactions with competing individuals in \mathbf{P} (resp. \mathbf{A}). Therefore, if OPAM relies only on the internal fitness functions, it cannot gauge the progress of coevolution in a stable manner as an individual's internal fitness depends on competing individuals.

```

1  Algorithm Task-arrival sequences evolution
2  Input A: population of task-arrival sequences
3  Input P: population of priority assignments
4  Input  $cp_a$ : crossover probability //task-arrival sequences
5  Input  $mp_a$ : mutation probability //task-arrival sequences
6  Output A: population of task-arrival sequences
7
8  //evaluate internal fitness values for A
9  for each  $A_i \in \mathbf{A}$ 
10 |   for each  $\vec{P}_l \in \mathbf{P}$ 
11 |   |    $S \leftarrow \text{simulate}(A_i, \vec{P}_l)$  //OPAMScheduler
12 |   |   //distk(j) is computed based on S
13 |   |    $fd(A_i, \vec{P}_l) = \sum_{j \in J, k \in [1, lk(j)]} 2^{dist_k(j)}$ 
14 |   |    $fd(A_i, \mathbf{P}) = \sum_{\vec{P}_l \in \mathbf{P}} fd(A_i, \vec{P}_l) / |\mathbf{P}|$ 
15 |
16  //breed task-arrival sequences
17  parents  $\leftarrow$  select_arrivals(A)
18  offspring  $\leftarrow$  crossover_arrivals(parents,  $cp_a$ )
19  offspring  $\leftarrow$  mutate_arrivals(offspring,  $mp_a$ )
20  //evaluate internal fitness values for offspring
21  for each  $A_i \in \text{offspring}$ 
22 |   for each  $\vec{P}_l \in \mathbf{P}$ 
23 |   |    $S \leftarrow \text{simulate}(A_i, \vec{P}_l)$  //OPAMScheduler
24 |   |   //distk(j) is computed based on S
25 |   |    $fd(A_i, \vec{P}_l) = \sum_{j \in J, k \in [1, lk(j)]} 2^{dist_k(j)}$ 
26 |   |    $fd(A_i, \mathbf{P}) = \sum_{\vec{P}_l \in \mathbf{P}} fd(A_i, \vec{P}_l) / |\mathbf{P}|$ 
27  A  $\leftarrow$  replace_arrivals(A, offspring)
28
29  return A

```

Fig. 5: A steady-state GA-based algorithm for evolving task-arrival sequences.

We note that soft deadline tasks also require to execute within reasonable execution time, i.e., (soft) deadline. As the above fitness functions return quantified degrees of deadline misses and safety margins, OPAM uses the same fitness functions for both soft and hard deadline tasks.

6.4 Evolution: Worst-case task arrivals

The algorithm in Figure 5 describes in detail the evolution of task-arrival sequences in lines 18–21 of the listing in Figure 4. OPAM adapts a steady-state Genetic Algorithm (GA) (Luke, 2013) for evolving task-arrival sequences. As shown in lines 8–14, OPAM first evaluates each task-arrival sequence in the **A** population against the **P** population of priority assignments. OPAM executes OPAMScheduler to obtain a schedule scenario S for a task-arrival sequence $A_i \in \mathbf{A}$ and a priority assignment $\vec{P}_l \in \mathbf{P}$ (line 11). OPAM then computes the internal fitness $fd(A_i, \mathbf{P})$ capturing the magnitude of deadline misses (lines 12–14). We note that a steady-state GA iteratively breeds offspring, assess their fitness, and then reintroduce them into a population. However, OPAM

computes internal fitness of all task-arrival sequences in \mathbf{A} at every generation. This is because internal fitness is computed in relation to \mathbf{P} , which is coevolving with \mathbf{A} .

Breeding the next population is done by using the following genetic operators: (1) *Selection*: OPAM selects candidate task-arrival sequences using a tournament selection technique, with the tournament size equal to two which is the most common setting (Gendreau and Potvin, 2010) (line 17 in Figure 5). (2) *Crossover*: Selected candidate task-arrival sequences serve as parents to create offspring using a crossover operation (line 18). (3) *Mutation*: The offspring are then mutated (line 19). Below, we describe our crossover and mutation operators.

Crossover. A crossover operator is used to produce offspring by mixing traits of parent solutions. OPAM modifies the standard one-point crossover operator (Luke, 2013) as two parent task-arrival sequences A_p and A_q may have different sizes, i.e., $|A_p| \neq |A_q|$. Let $J = \{j_1, j_2, \dots, j_m\}$ be a set of tasks to be scheduled. Our crossover operator first randomly selects an aperiodic task $j_r \in J$. For all $i \in [1, r]$ and $j_i \in J$, OPAM then swaps all j_i arrivals between the two task-arrival sequences A_p and A_q . Since J is fixed for all solutions, OPAM can cross over two solutions that may have different sizes.

Mutation operator OPAM uses a heuristic mutation algorithm. For a task-arrival sequence A , OPAM mutates the k th task arrival time $at_k(j)$ of an aperiodic task j with a mutation probability. OPAM chooses a new arrival time value of $at_k(j)$ based on the $[pmin(j), pmax(j)]$ inter-arrival time range of j . If such a mutation of the k th arrival time of j does not affect the validity of the $k+1$ th arrival time of j , the mutation operation ends. Specifically, let d be a mutated value of $at_k(j)$. In case $at_{k+1}(j) \in [d + pmin(j), d + pmax(j)]$, OPAM returns the mutated A task-arrival sequence.

After mutating the k th arrival time $at_k(j)$ of a task j in a solution A , if the $k+1$ th arrival becomes invalid, OPAM corrects the remaining arrivals of j . Let o and d be, respectively, the original and mutated k th arrival time of j . For all the arrivals of j after d , OPAM first updates their original arrival time values by adding the difference $d - o$. Let $\mathbb{T} = [0, \mathbf{T}]$ be the scheduling period. OPAM then removes some arrivals of j if they are mutated to arrive after \mathbf{T} or adds new arrivals of j while ensuring that all tasks arrive within \mathbb{T} .

As shown in lines 20–26 in Figure 5, the internal fitness of the generated offspring is computed based on the \mathbf{P} population. OPAM then updates the \mathbf{A} population of task-arrival sequences by comparing the offspring and individuals in \mathbf{A} (line 27).

We note that when a system is only composed of periodic tasks, OPAM will skip evolving for worst-case arrival sequences as arrivals of periodic tasks are deterministic (see Section 3). Nevertheless, OPAM will optimize priority assignments based on given arrivals of periodic tasks. When needed, OPAM can be easily extended to manipulate offset and period values for periodic tasks, in a way identical to how we currently handle inter-arrival times for aperiodic tasks.

```

1  Algorithm Priority assignments evolution
2  Input A: population of task-arrival sequences
3  Input P: population of priority assignments
4  Input  $ps_p$ : population size //priority assignments
5  Input  $cp_p$ : crossover probability //priority assignments
6  Input  $mp_p$ : mutation probability //priority assignments
7  Output P: population of priority assignments
8
9  //evaluate internal fitness values for P
10 for each  $\vec{P}_i \in \mathbf{P}$ 
11   for each  $A_l \in \mathbf{A}$ 
12      $S \leftarrow \text{simulate}(A_l, \vec{P}_i)$  //OPAMScheduler
13     //distk(j) is computed based on S
14      $fs(\vec{P}_i, A_l) = \sum_{j \in J, k \in [1, lk(j)]} -2^{dist_k(j)}$ 
15      $fs(\vec{P}_i, \mathbf{A}) = \sum_{A_l \in \mathbf{A}} fs(\vec{P}_i, A_l) / |\mathbf{A}|$ 
16      $fc(\vec{P}_i) = \sum_{j \in J} \begin{cases} lp(\vec{P}_i) - pr(j), & \text{if } j \text{ is an aperiodic task} \\ 0, & \text{otherwise} \end{cases}$ 
17
18   //breed priority assignments
19    $\vec{R} \leftarrow \text{sort\_non\_dominated\_fronts}(\mathbf{P})$ 
20    $\text{assign\_crowding\_distance}(\vec{R})$ 
21    $\mathbf{P}_\alpha \leftarrow \text{NSGAIIBreed}(\vec{R}, ps_p, cp_p, mp_p)$ 
22   //evaluate internal fitness values for  $\mathbf{P}_\alpha$ 
23   for each  $\vec{P}_i \in \mathbf{P}_\alpha$ 
24     for each  $A_l \in \mathbf{A}$ 
25        $S \leftarrow \text{simulate}(A_l, \vec{P}_i)$  //OPAMScheduler
26       //distk(j) is computed based on S
27        $fs(\vec{P}_i, A_l) = \sum_{j \in J, k \in [1, lk(j)]} -2^{dist_k(j)}$ 
28        $fs(\vec{P}_i, \mathbf{A}) = \sum_{A_l \in \mathbf{A}} fs(\vec{P}_i, A_l) / |\mathbf{A}|$ 
29        $fc(\vec{P}_i) = \sum_{j \in J} \begin{cases} lp(\vec{P}_i) - pr(j), & \text{if } j \text{ is an aperiodic task} \\ 0, & \text{otherwise} \end{cases}$ 
30    $\vec{R} \leftarrow \text{sort\_non\_dominated\_fronts}(\mathbf{P} \cup \mathbf{P}_\alpha)$ 
31    $\text{assign\_crowding\_distance}(\vec{R})$ 
32    $\mathbf{P} \leftarrow \text{select\_archive}(\vec{R}, ps_p)$ 
33
34  return P

```

Fig. 6: An NSGAIIBased algorithm for evolving priority assignments.

6.5 Evolution: Best-case priority assignments

Figure 6 shows the evolution procedure of priority assignments, which refines lines 23–26 in Figure 4. OPAM tailors the Non-dominated Sorting Genetic Algorithm version 2 (NSGAIIB) (Deb et al., 2002) to generate a non-dominating (equally viable) set of priority assignments, representing the best trade-offs found among the given internal fitness functions. This is referred to as a Pareto nondominated front (Knowles and Corne, 2000), where the dominance relation over priority assignments is defined as follows: A priority assignment \vec{P} dominates another priority assignment \vec{P}' if \vec{P} is not worse

than \vec{P}' in all fitness values, and \vec{P} is strictly better than \vec{P}' in at least one fitness value. NSGAII has been applied to many multi-objective optimization problems (Langdon et al., 2010; Shin et al., 2018; Wang et al., 2020).

OPAM maintains a population \mathbf{P} of priority assignments as an archive that contains the best priority assignments discovered during coevolution. Unlike a standard application of NSGAII, in our study, we need to reevaluate the internal fitness values for priority assignments in \mathbf{P} at every generation as the internal fitness values are computed based on the \mathbf{A} population of task-arrival sequences, which coevolves. As shown in lines 9–16 in Figure 6, OPAM first computes the internal fitness functions that measure the magnitude of safety margins and the extent of constraint satisfaction. OPAM then sorts non-dominated Pareto fronts (line 19) and assigns crowding distance (line 20) to introduce diversity among non-dominated priority assignments (Deb et al., 2002).

For breeding the next population of priority assignments (line 21 in Figure 6, OPAM applies the following standard genetic operators (Sivanandam and Deepa, 2008) that have been applied to many similar problems (Islam et al., 2012; Marchetto et al., 2016; Shin et al., 2018): (1) *Selection*. OPAM uses a binary tournament selection based on non-domination ranking and crowding distance. The binary tournament selection has been used in the original implementation of NSGAII (Deb et al., 2002). (2) *Crossover*. OPAM applies a partially mapped crossover (PMX) (Goldberg and Lingle, 1985). PMX ensures that the generated offspring are valid permutations of priorities. (3) *Mutation*. OPAM uses a permutation swap method for mutating a priority assignment. This mutation method interchanges two randomly-selected priorities in a priority assignment according to a given mutation probability.

For the generated population \mathbf{P}_α of priority assignments, OPAM computes the two internal fitness functions (lines 22–29 in Figure 6). OPAM then sorts non-dominated Pareto fronts for the union of the current \mathbf{P} and next \mathbf{P}_α populations (line 30), assign crowding distance (line 31), and select the best archive by accounting for the computed non-domination ranking and crowding distance (line 32).

6.6 External fitness evaluation

Figure 7 shows an algorithm that computes the external fitness functions and finds the best Pareto front, which refines lines 28–31 in Figure 4. To monitor the coevolution progress in a stable manner, OPAM takes as input a set \mathbf{E} of task-arrival sequences that are generated independently from the coevolution process. We use an adaptive random search technique (Chen et al., 2010) to sample task-arrival sequences in order to create \mathbf{E} . The adaptive random search extends the naive random search by maximizing the Euclidean distance between the sampled points such that it maximizes the diversity of task-arrival sequences in \mathbf{E} .

```

1  Algorithm Priority assignments evolution
2  Input E: set of task-arrival sequences //external evaluation
3  Input P: population of priority assignments
4  Input  $ps_p$ : population size //priority assignments
5  Input  $cp_p$ : crossover probability //priority assignments
6  Input  $mp_p$ : mutation probability //priority assignments
7  Output P: population of priority assignments
8
9  //evaluate external fitness values for P
10 for each  $\vec{P}_i \in \mathbf{P}$ 
11   for each  $E_l \in \mathbf{E}$ 
12      $S \leftarrow \text{simulate}(E_l, \vec{P}_i)$  //OPAMScheduler
13     //distk(j) is computed based on S
14      $fs(\vec{P}_i, E_l) = \sum_{j \in J, k \in [1, lk(j)]} -2^{dist_k(j)}$ 
15      $fs(\vec{P}_i, \mathbf{E}) = \sum_{E_l \in \mathbf{E}} fs(\vec{P}_i, E_l) / |\mathbf{E}|$ 
16      $fc(\vec{P}_i) = \sum_{j \in J} \begin{cases} lp(\vec{P}_i) - pr(j), & \text{if } j \text{ is an aperiodic task} \\ 0, & \text{otherwise} \end{cases}$ 
17  $\vec{R} \leftarrow \text{sort\_non\_dominated\_fronts}(\mathbf{P} \cup \mathbf{B})$ 
18  $\text{assign\_crowding\_distance}(\vec{R})$ 
19  $\mathbf{B} \leftarrow \text{select\_best\_front}(\vec{R})$  //|B| ≤ |P|
20
21 return B

```

Fig. 7: An algorithm for evaluating external fitness and finding the best Pareto front.

As shown in lines 9–16 in Figure 7, OPAM computes the two external fitness values for each priority assignment in the **P** population based on a given set **E** of task-arrival sequences. OPAM then sorts non-dominated Pareto fronts for the union of the **P** population and the current best Pareto front (line 17), assigns crowding distance (line 18), and selects the best Pareto front by accounting for the computed non-domination ranking and crowding distance (line 32). OPAM adopts NSGAII in order to maximize the diversity of priority assignments in the best Pareto front.

7 Evaluation

This section describes our evaluation of OPAM through six industrial case studies from different domains and several synthetic subjects. Our full evaluation package is available online (Lee et al., 2021).

7.1 Research questions

RQ1 (Sanity check): *How does OPAM perform compared with Random Search?* For search-based solutions, this RQ is an important *sanity check* to ensure that success is not due to the search problem being easy (Arcuri and

Table 2: Description of the six industrial subject systems: number of periodic and aperiodic tasks, resource dependencies, triggering relations, and platform cores.

| System | Task types | | Relationships | | Platform |
|--------|------------|-----------|---------------|------------|----------|
| | Periodic | Aperiodic | Dependencies | Triggering | Cores |
| ICS | 3 | 3 | 3 | 0 | 3 |
| CCS | 8 | 3 | 3 | 6 | 2 |
| UAV | 12 | 4 | 4 | 0 | 3 |
| GAP | 15 | 8 | 6 | 5 | 2 |
| HPSS | 23 | 9 | 5 | 0 | 1 |
| ESAIL | 11 | 14 | 0 | 0 | 1 |

Briand, 2014). Our conjecture is that a search-based algorithm, although expensive, will significantly outperform naive random search (RS).

RQ2 (Coevolution): *Is competitive coevolution suitable to find best-case priority assignments?* We conjecture that a coevolutionary algorithm is a suitable solution to address the priority assignment problem since it is solved, in practice, through a competing interactive process between the development and testing teams. To answer this RQ, we compare OPAM with a sequential approach that first looks for worst-case sequences of task arrivals and then tries to find best-case priority assignments.

RQ3 (Scalability): *Can OPAM find (near-)optimal solutions for large-scale systems in a reasonable time budget?* In this RQ, we investigate the scalability of OPAM by conducting some experiments with systems of various sizes, including six industrial and several synthetic subjects. We study the relationship between OPAM’s performance measures and the characteristics of study subjects.

RQ4 (Usefulness): *How do priority assignments generated by OPAM compare with priority assignments defined by engineers?* OPAM can be considered useful only when it finds priority assignments that show benefits over those defined (manually) by engineers with domain expertise. This RQ therefore compares the quality of priority assignments generated by OPAM with those defined by engineers. We further discuss the usefulness of OPAM from a practical perspective, based on the feedback received from engineers in LuxSpace.

7.2 Industrial study subjects

To evaluate RQs in realistic and diverse settings, we apply OPAM to six industrial study subjects from different domains such as aerospace, automotive, and avionics domains. Specifically, we obtained one case study subject from our industry partner, LuxSpace. We found the other five industrial study subjects in the literature (Di Alesio et al., 2015), which, consistent with the LuxSpace system, all assume a single-queue, multi-core, fixed-priority scheduling policy.

Note that OPAM uses the same scheduling policy (described in Section 6.2) as in Di Alesio et al.’s work. This policy uses fixed priorities that are determined offline and therefore do not change dynamically. Table 2 summarizes the relevant attributes of these subjects, presenting the number of periodic and aperiodic tasks, resource dependencies, triggering relations, and platform cores. The subjects are characterized by real-time parameters, e.g., periods, deadlines, and priorities, described in Section 3. We note that all the study subjects are deadlock-free systems as they do not have circular resource dependencies. Regarding task priorities, all tasks in the six subjects have fixed priorities, which are defined by experts in their domains. The full task descriptions (including WCET, inter-arrival times, periods, deadlines, priorities, and relationship details) of the subjects are available online (Lee et al., 2021). The main missions of the six subjects are described as follows:

- ICS is an ignition control system that checks the status of an automotive engine and corrects any errors of the engine (Peraldi-Frati and Sorel, 2008). The system was developed by Bosch GmbH¹.
- CCS is a cruise control system that acquires data from vehicle sensors and maintains the specified vehicle speed (Anssi et al., 2011). Continental AG² developed the system.
- UAV is a mini unmanned air vehicle that follows dynamically defined way-points and communicates with a ground station to receive instructions (Traore et al., 2006). The system was developed in collaboration with the University of Poitiers France and ENSMA³.
- GAP is a generic avionics platform for a military aircraft (Locke et al., 1990). The system was designed in a joint project with Carnegie Mellon University, the US Navy, and IBM⁴, aiming at supporting several missions regarding air-to-surface attacks.
- HPSS is a satellite system for two satellites, named Herschel and Planck (Mikučionis et al., 2010). The two satellites share the same computational architecture, although they have different scientific missions. Herschel aims at studying the origin and evolution of stars and galaxies. Planck’s primary mission is the study of the relic radiation from the Big Bang. ESA⁵ carried out the HPSS project.
- ESAIL is a microsatellite for tracking ships worldwide by detecting messages that ships radio-broadcast (see Section 2). Luxspace, our industry partner, developed ESAIL in an ESA project.

¹ Bosch GmbH: <https://www.bosch.com/>

² Continental AG: <https://www.continental.com>

³ ENSMA: <https://www.ensma.fr/>

⁴ IBM: <https://www.ibm.com/>

⁵ ESA: <https://www.esa.int/>

```

1  Algorithm Synthetic task generation
2  Input  $n$ : number of tasks
3  Input  $u_t$ : target utilization
4  Input  $pd_{min}$ : minimum task period
5  Input  $pd_{max}$ : maximum task period
6  Input  $g$ : granularity of task periods
7  Input  $\gamma$ : ratio of aperiodic tasks
8  Input  $\mu$ : range factor to determine maximum inter-arrival times
9  Output  $S$ : set of tasks
10
11  $S \leftarrow \{\}, C \leftarrow \{\}$ 
12 // synthesize a set of periodic tasks
13  $U \leftarrow UUniFast\_discard(n, u_t)$  // task utilizations
14  $I \leftarrow generate\_task\_periods(n, pd_{min}, pd_{max}, g)$  // task periods
15 for each  $j \in [1, n]$ 
16 |  $C \leftarrow C \cup \{U_j \cdot I_j\}$ , where  $U_j \in U$  and  $I_j \in I$  // WCETs
17  $S \leftarrow generate\_task\_set(I, C)$  // set of tasks
18 // convert some periodic tasks to aperiodic tasks
19  $S \leftarrow convert\_aperiodic\_tasks(S, \gamma, \mu)$ 
20
21 return  $S$ 

```

Fig. 8: An algorithm for synthesizing a set of tasks.

7.3 Synthetic study subjects

To investigate RQ3, we use synthetic subjects in order to freely control key parameters in real-time systems. We create a set of tasks by adopting a well-known procedure (Emberson et al., 2010) for synthesizing real-time tasks, which has been applied in many schedulability analysis studies (Davis et al., 2008; Zhang and Burns, 2009; Davis and Burns, 2011; Grass and Nguyen, 2018; Dürr et al., 2019).

Figure 8 describes a procedure that synthesizes a set of real-time tasks. For a given number n of tasks and a target utilization u_t , the procedure first generates a set U of task utilization values by using the UUniFast-Discard algorithm (Davis and Burns, 2011) (line 13). The UUniFast-Discard algorithm is devised to give an unbiased distribution of utilization values, where a utilization $U_j \in U$ is a positive value and $\sum_{U_j \in U} U_j = u_t$.

The procedure then generates a set I of n task periods according to a log-uniform distribution within a range $[pd_{min}, pd_{max}]$, i.e., given a task period (random variable) $I_j \in I$, $\log I_j$ follows a uniform distribution (line 14 in Figure 8). For example, when the minimum and maximum task periods are $pd_{min} = 10\text{ms}$ and $pd_{max} = 1000\text{ms}$, respectively, the procedure generates (approximately) an equal number of tasks in time intervals $[10\text{ms}, 100\text{ms}]$ and $[100\text{ms}, 1000\text{ms}]$. The parameter g is used to choose the granularity of the periods, i.e., task periods are multiples of g . Such a distribution of task periods provides a reasonable degree of realism with respect to what is usually observed in real systems (Baruah et al., 2011).

As shown in lines 15–16 of the procedure in Figure 8, a set \mathbf{C} of task WCETs are computed based on the set \mathbf{U} of task utilization values and the set \mathbf{I} of task periods. Specifically, a task WCET $C_j \in \mathbf{C}$ is computed as $C_j = U_j \cdot I_j$.

As per line 17 of the listing in Figure 8, the procedure synthesizes a set \mathbf{S} of tasks. A task j is characterized by a period I_j and a WCET C_j and it is associated with a deadline $dl(j)$ and a priority $pr(j)$. According to the rate-monotonic scheduling policy (Liu and Layland, 1973), tasks’ deadlines are equal to their periods and tasks with shorter periods are given higher priorities.

To synthesize aperiodic tasks, the procedure converts some periodic tasks to aperiodic tasks according to a given ratio γ of aperiodic tasks among all tasks (see line 19 in Figure 8). A range factor μ is used to determine maximum inter-arrival times of aperiodic tasks. Specifically, for a task j to be converted, the procedure sets the minimum inter-arrival time $pmin(j)$ as $pmin(j) = I_j$. The procedure then selects a uniformly distributed value x from the range $(1, \mu]$ and computes the maximum inter-arrival time $pmax(j)$ as $pmax(j) = x \cdot I_j$.

7.4 Experimental Design

This section describes how we design experiments to answer the RQs described in Section 7.1. We conducted four experiments, EXP1, EXP2, EXP3, and EXP4, as described below.

EXP1. To answer RQ1, EXP1 compares OPAM with our baseline, which relies on random search, to ensure that the effectiveness of OPAM is not due to the search problem being simple. Our baseline, named RS, replaces GA with a random search for finding worst-case sequences of task arrivals and NSGAI with a random search for finding best-case priority assignments. Note that RS uses the same internal and external fitness functions (see Section 6.3) and also maintains the best populations during search; however, it does not employ any genetic operators, i.e., crossover and mutation. In EXP1, we applied OPAM and RS to the six industrial subjects described in Section 7.2.

Recall from Section 6.3 that OPAM uses a set \mathbf{E} of task-arrival sequences that are generated independently from the coevolution process in order to monitor the coevolution progress in a stable manner. As OPAM and RS use the same set \mathbf{E} of task-arrival sequences, EXP1 first compares OPAM and RS based on \mathbf{E} . In addition, EXP1 examines how well the solutions, i.e., priority assignments, found by OPAM and RS perform with other sequences of task arrivals. To do so, we create six sets of sequences of task arrivals for each study subject by varying the method to generate task-arrival sequences and the number of task-arrival sequences. Note that task-arrival sequences generated by different methods are valid with respect to the inter-arrival times defined in each study subject. Below we describe the six sets of task-arrival sequences generated for each subject.

- \mathbf{T}_a^{10} : A set of task-arrival sequences generated by using an adaptive random search technique (Chen et al., 2010) that aims at maximizing the diversity of task-arrival sequences. The \mathbf{T}_a^{10} set contains 10 sequences of task arrivals.

- \mathbf{T}_w^{10} : A set of task-arrival sequences generated by using a stress test case generation method that aims at maximizing the chances of deadline misses in task executions. The stress test case generation method extends prior work (Briand et al., 2005). The extended method uses the fitness function regarding deadline misses and genetic operators that OPAM introduces for evolving worst-case task-arrival sequences (see Section 6). The \mathbf{T}_w^{10} set contains 10 sequences of task arrivals.
- \mathbf{T}_r^{10} : A set of task-arrival sequences generated randomly. The \mathbf{T}_r^{10} set has 10 sequences of task arrivals.
- \mathbf{T}_a^{500} : A set of task-arrival sequences generated by using the adaptive random search technique. The \mathbf{T}_a^{500} set contains 500 sequences of task arrivals.
- \mathbf{T}_w^{500} : A set of task-arrival sequences generated by using the stress test case generation method. The \mathbf{T}_w^{500} set contains 500 sequences of task arrivals.
- \mathbf{T}_r^{500} : A set of task-arrival sequences generated randomly. The \mathbf{T}_r^{500} set has 500 sequences of task arrivals.

EXP2. To answer RQ2, EXP2 compares OPAM with a priority assignment method, named SEQ, that relies on one-population search algorithms. SEQ first finds a set of worst-case sequences of task arrivals using GA with the fitness function that measures the magnitude of deadline misses (see $fd()$ in Section 6.3) and the genetic operators described in Section 6.4. Given a set of worst-case task-arrival sequences obtained from GA, SEQ then aims at finding best-case priority assignments using NSGAII with the fitness functions that quantify the magnitude of safety margins and the degree of constraint satisfaction (see $fs()$ and $fc()$, respectively, in Section 6.3) and the genetic operators described in Section 6.5.

We note that SEQ does not use the external fitness functions as it does not coevolve task-arrival sequences and priority assignments. Hence, the numbers of fitness evaluations of the two methods are not comparable. To fairly compare OPAM and SEQ, we set the same time budget for the two methods. Specifically, we first measure the execution time of OPAM for analyzing each subject. We then split the execution time in half and set each half time as the execution budget of the GA and NSGAII steps in SEQ for the corresponding subject. In order to assess the quality of priority assignments obtained from OPAM and SEQ, we use the sets of task-arrival sequences described in EXP1, i.e., \mathbf{T}_a^{10} , \mathbf{T}_w^{10} , \mathbf{T}_r^{10} , \mathbf{T}_a^{500} , \mathbf{T}_w^{500} , and \mathbf{T}_r^{500} , which are created independently from the two methods.

EXP3. To answer RQ3, EXP3 examines not only the six industrial subjects but also 370 synthetic subjects. We create the synthetic subjects to study correlations between the execution time and memory usage of OPAM and the following parameters: the number of tasks (n), a (part-to-whole) ratio of aperiodic tasks (γ), a range factor for maximum inter-arrival times (μ), and simulation time (T), as described in Sections 7.3 and 6. We note that we chose to control parameters n , γ , and μ because they are the main parameters on which engineers have control to define tasks in real-time systems. Simulation time T obviously impacts the execution time of OPAM as well. But EXP3 aims

at modeling such correlations precisely and providing experimental results. Regarding the other factors that define, for example, task relationships and platform cores, we note significant diversity across the six industrial subjects.

Recall from Section 7.3 that we use the task generation procedure presented in Figure 8 to synthesize tasks. For EXP3, we set some parameter values of the procedure as follows: (1) Target utilization $u_t = 0.7$, which is a common objective in the development of a real-time system in order to guarantee the schedulability of tasks (Fineberg and Serlin, 1967; Dürr et al., 2019). (2) The range of task periods $[pd_{min}, pd_{max}] = [10\text{ms}, 1\text{s}]$, which are common values in many real-time systems (Emberson et al., 2010; Baruah et al., 2011). (3) The granularity of task periods $g = 10\text{ms}$ in order to increase realism as most of the task periods in our industrial subjects are multiples of 10ms. Because of some degree of randomness in the procedure of Figure 8, we create ten synthetic subjects per configuration. Below we further describe how synthetic subjects are created for each controlled experiment.

EXP3.1. To study the correlations between the execution time and memory usage of OPAM with the number of tasks n , we create nine sets of ten synthetic subjects such that no two sets have the same number of tasks. Specifically, we create sets with 10, 15, ..., 50 tasks, respectively. Regarding the ratio of aperiodic tasks, $\gamma = 0.4$ as, on average, the ratio of aperiodic tasks to periodic tasks in our industrial subjects is $2/3$. For the range factor, $\mu = 2$, which is determined based on the inter-arrival times of aperiodic tasks in our industry subjects. We set the simulation time T to 2s in order to ensure that any aperiodic task arrives at least once during that time. We note that, given the maximum task period $pd_{max} = 1\text{s}$ and the range factor $\mu = 2$, the maximum inter-arrival time of an aperiodic task is at most 2s (see Section 7.3).

EXP3.2. To study the correlations between the execution time and memory usage of OPAM with the ratio of aperiodic tasks γ , we create ten sets of synthetic subjects by setting this ratio to the following values: 0.05, 0.10, ..., 0.50. We set the number of tasks to 20 ($n = 20$), which is the average number of tasks in our six industrial subjects. Regarding the other parameters, range factor and simulation time, $\mu = 2$ and $T = 2\text{s}$ are set as discussed in EXP3.1.

EXP3.3. To study the correlations between the execution time and memory usage of OPAM with the range factor μ that is used to determine the maximum inter-arrival times, we create nine sets of synthetic subjects by setting μ to 2, 3, ..., 10. We set the simulation time as follows: $T = 10\text{s}$. This ensures that any aperiodic task arrives at least once during the simulation time when μ is at most 10 (see Section 7.3). The other parameters, the number of tasks and ratio of aperiodic tasks, $n = 20$ and $\gamma = 0.4$ are set as discussed in EXP3.1 and EXP3.2.

EXP3.4. To study the correlations between the execution time and memory usage of OPAM with the simulation time T , we create nine sets of synthetic subjects by setting T to 2s, 3s, ..., 10s. The other parameters, e.g., the number of tasks, the ratio of aperiodic tasks, and the range factor, $n = 20$, $\gamma = 0.4$, and $\mu = 2$, are set as discussed in EXP3.1 and EXP3.2.

EXP4. To answer RQ4, EXP4 compares priority assignments optimized by OPAM and those defined by engineers. We apply OPAM to the six industrial subjects (see Section 7.2) which include priority assignments defined by practitioners. Note that we focus here on the ESAIL subject in collaboration with our industry partner, LuxSpace; The other five subjects are from the literature (Di Alesio et al., 2015) and hence we can only collect feedback from practitioners for ESAIL.

7.5 Evaluation metrics

Multi-objective evaluation metrics. In order to fairly compare the results of search algorithms, based on existing guidelines (Li et al., 2020) for assessing multi-objective search algorithms, we use complementary quality indicators: *Hypervolume* (HV) (Zitzler and Thiele, 1999), *Pareto Compliant Generational Distance* (GD+) (Ishibuchi et al., 2015), and *Spread* (Δ) (Deb et al., 2002). To compute the GD+ and Δ quality indicators, following the usual procedure (Li et al., 2020), we create a reference Pareto front as the union of all the non-dominated solutions obtained from all runs of the algorithms being compared. Identifying the optimal (ideal) Pareto front is typically infeasible for a complex optimization problem (Li et al., 2020). Key features of the three quality indicators are described below.

- HV is defined to measure the volume in the objective space that is covered by members of a Pareto front generated by a search algorithm (Zitzler and Thiele, 1999). The higher the HV values, the more optimal the search outputs.
- GD+ is defined to measure the distance between the points on a Pareto front obtained from a search algorithm and the nearest points on a reference Pareto front (Ishibuchi et al., 2015). GD+ modifies General Distance (GD) (Veldhuizen and Lamont, 1998) to account for the dominance relations when computing the distances. The lower the GD+ values, the more optimal the search outputs.
- Δ is defined to measure the extent of spread among the points on a Pareto front computed by a search algorithm (Deb et al., 2002). We note that OPAM aims at obtaining a wide variety of equally-viable priority assignments on a Pareto front (see Section 6). The lower the Spread values, the more spread out the search outputs.

Interpretable metrics. The two external fitness functions described in Section 6 mainly aim at effectively guiding search. It is, however, difficult for practitioners to interpret the computed fitness values. Since they are not intuitive to practitioners, to assess the usefulness of OPAM from a practitioner perspective, we measure (1) the safety margins from tasks' completion times to their deadlines across our experiments and (2) the number of constraint violations in a priority assignment. In addition, we measure the execution time and memory usage of OPAM.

Statistical comparison metrics. To statistically compare our experiment results, we use the Mann-Whitney U-test (Mann and Whitney, 1947) and Vargha and Delaney’s \hat{A}_{12} effect size (Vargha and Delaney, 2000), which have been frequently applied for evaluating search-based algorithms (Arcuri et al., 2010; Hemmati et al., 2013; Shin et al., 2018). Mann-Whitney U-test determines whether two independent samples are likely or not to belong to the same distribution. We set the level of significance, α , to 0.05. Vargha and Delaney’s \hat{A}_{12} measures probabilistic superiority – effect size – between search algorithms. Two algorithms are considered to be equivalent when the value of \hat{A}_{12} is 0.5.

7.6 Parameter tuning and implementation

Parameters for coevolutionary search. For the coevolutionary search parameters, we set the population size to 10, the crossover rate to 0.8, and the mutation rate to $1/|J|$, where $|J|$ denotes the number of tasks. We apply these parameter values for both the evolution of task-arrival sequences and priority assignments (see Section 6). These values are determined based on existing guidelines (Arcuri and Fraser, 2011; Sayyad et al., 2013) and previous work (Lee et al., 2020b).

We determine the number of coevolution cycles (see Section 6) based on an initial experiment. We applied OPAM to the six industrial subjects and ran OPAM 50 times for each subject. From the experiment results, we observed that there is no notable difference in Pareto fronts generated after 1000 cycles. Hence, we set the number of coevolution cycles to 1000 in our experiments, i.e., EXP1, EXP2, and EXP3 described in Section 7.4.

Parameters for evaluating fitness functions. To evaluate external fitness functions, we use a set of task-arrival sequences that are generated independently from the coevolution process (see Section 6.6). We use an adaptive random search (Chen et al., 2010) to generate a set \mathbf{E} of task-arrival sequences, which varies task arrival times within the specified inter-arrival time ranges of aperiodic tasks. We set the size of \mathbf{E} to 10. From our initial experiment, we observed that this is sufficient to compute the external fitness functions of OPAM under a reasonable time, i.e., less than 15s. We note that \mathbf{E} contains two default sequences of task arrivals as follows: (seq. 1) aperiodic tasks always arrive at their maximum inter-arrival times and (seq. 2) aperiodic tasks always arrive at their minimum inter-arrival times. By having those two sequences of task arrivals as initial elements in \mathbf{E} , the adaptive random search finds other sequences of task arrivals to maximize the diversity of elements in \mathbf{E} .

If a system contains only periodic tasks, the simulation time is often set as the least common multiple (LCM) of their periods to account for all possible arrivals (Peng et al., 1997). However, as the six industrial subjects include aperiodic tasks, this is not applicable. For the experiments with the six industrial subjects, we set the simulation time to the maximum time between the LCM of periodic tasks’ periods and the maximum inter-arrival time among

aperiodic tasks. By doing so, all possible arrival patterns of periodic tasks are examined and any aperiodic task arrives at least once during simulation. Recall from Section 6.4 that OPAM varies arrival times of aperiodic tasks to find worst-case sequences of task arrivals.

We note that the parameters mentioned above can probably be further tuned to improve the performance of our approach. However, since with our current setting, we were able to convincingly and clearly support our conclusions, we do not report further experiments on tuning those values.

Implementation. We implemented OPAM by extending jMetal (Durillo and Nebro, 2011), which is a metaheuristic optimization framework supporting NSGAII and GA. We conducted our experiments using the high-performance computing cluster (Varrette et al., 2014) at the University of Luxembourg. To account for randomness, we repeated each run of OPAM 50 times for all experiments. Each run of OPAM was executed on a different node (equipped with five 2.5GHz cores and 20GB memory) of the cluster, and took less than 16 hours.

7.7 Results

RQ1. Figure 9 shows the best Pareto fronts obtained with 50 runs of OPAM and RS, for the six industrial study subjects described in Section 7.2. The fitness values presented in the figures are computed based on each subject’s set \mathbf{E} of task-arrival sequences (see Section 7.6), which is created independently from OPAM and RS. Figures 9a, 9c, 9d, 9e, and 9f indicate that OPAM finds significantly better solutions than RS for ICS, UAV, GAP, HPSS, and ESAIL. Regarding CCS (see Figure 9b), it is difficult to conclude anything based only on visual inspection. Hence, we compared Pareto fronts obtained by OPAM and RS using the three quality indicators HV, GD+, and Δ , described in Section 7.5.

Figure 10 depicts distributions of HV (Figure 10a), GD+ (Figure 10b), and Δ (Figure 10c) for the six industrial subjects. The boxplots in the figures present the distributions (25%-50%-75%) of the quality values obtained from 50 runs of OPAM and RS. The quality values are computed based on the Pareto fronts obtained by the algorithms and each subject’s set \mathbf{E} of task-arrival sequences (see Section 7.6). In the figures, statistical comparisons of the two corresponding distributions are summarized using p-values and \hat{A}_{12} values, as described in Section 7.5, under each subject name.

As shown in Figures 10a and 10b, OPAM obtains better distributions of HV and GD+ compared to RS for all six subjects. All the differences are statistically significant as the p-values are below 0.05. Regarding Δ , as depicted in Figure 10c, OPAM yields higher diversity in Pareto front solutions than RS for the following subjects: UAV, GAP, and HPSS. For ICS, CCS, and ESAIL, OPAM and RS obtain similar Δ values. From Figures 10a and 10b, and Table 2, we also observe that the higher the number of aperiodic tasks in a subject, the larger the differences in HV and GD+ between OPAM and RS. Hence,

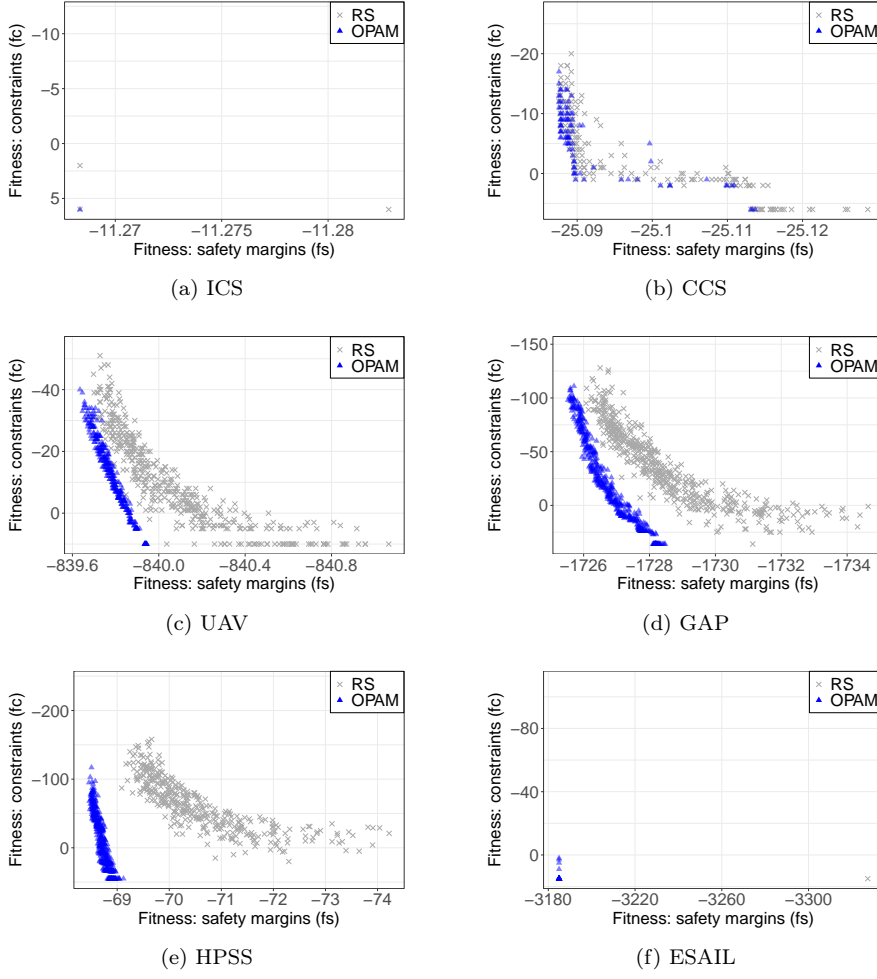


Fig. 9: Pareto fronts obtained by OPAM and RS for the six industrial subjects: (a) ICS, (b) CCS, (c) UAV, (d) GAP, (e) HPSS, and (f) ESAIL. The fitness values are computed based on each subject's set \mathbf{E} of task-arrival sequences (see Section 7.6). The points located closer to the bottom left of each plot are considered to be better priority assignments when compared to points closer to the top right.

for these two quality indicators, OPAM outperforms RS more significantly for more complex search problems. Note that the number of aperiodic tasks is one of the main factors that drives the degree of uncertainty in task arrivals.

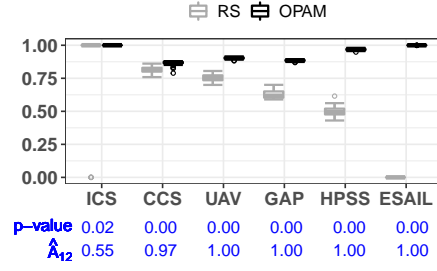
Given the Pareto priority assignments obtained by OPAM and RS, we further assessed the quality values of the solutions by evaluating them with

Table 3: Comparing OPAM and RS using the three quality indicators: HV, GD+, and Δ . Average quality values computed based on 50 runs of OPAM and RS using the different sets of task-arrival sequences (see Section 7.4).

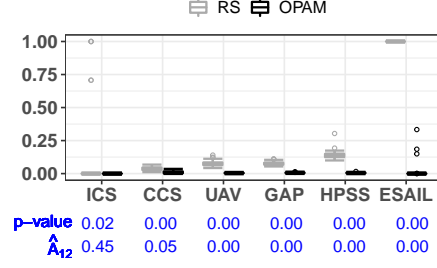
| | | | ICS | CCS | UAV | GAP | HPSS | ESAIL |
|-------------------------------------|----------|------------------|-----------|-----------|-----------|-----------|-----------|-----------|
| T_a^{10} (adaptive, size 10) | HV | OPAM | 1.0000 | 0.7168 | 0.8923 | 0.8864 | 0.9629 | 0.9998 |
| | | RS | 0.9000 | 0.6633 | 0.7488 | 0.6278 | 0.5120 | 0.0000 |
| | | $p \hat{A}_{12}$ | 0.02 0.55 | 0.00 0.80 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 |
| | GD+ | OPAM | 0.0000 | 0.0203 | 0.0068 | 0.0067 | 0.0073 | 0.0135 |
| | | RS | 0.0883 | 0.0472 | 0.0745 | 0.0780 | 0.1380 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.02 0.45 | 0.00 0.04 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 |
| | Δ | OPAM | 1.0000 | 0.7650 | 0.4256 | 0.3631 | 0.5355 | 0.9433 |
| | | RS | 0.9766 | 0.5879 | 0.6112 | 0.6605 | 0.7508 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.16 0.52 | 0.00 0.76 | 0.00 0.15 | 0.00 0.03 | 0.00 0.12 | 0.08 0.47 |
| T_w^{10} (worst, size 10) | HV | OPAM | 0.0000 | 0.7878 | 0.9152 | 0.9280 | 0.9652 | 0.9997 |
| | | RS | 0.0000 | 0.7591 | 0.7782 | 0.6743 | 0.5180 | 0.0000 |
| | | $p \hat{A}_{12}$ | 1.00 0.50 | 0.01 0.65 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 |
| | GD+ | OPAM | 0.0000 | 0.0809 | 0.0053 | 0.0042 | 0.0108 | 0.0135 |
| | | RS | 0.0200 | 0.0866 | 0.0740 | 0.0760 | 0.1405 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.16 0.48 | 0.75 0.52 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 |
| | Δ | OPAM | 1.0000 | 0.7012 | 0.4508 | 0.4009 | 0.4872 | 0.9433 |
| | | RS | 0.9600 | 0.4764 | 0.6032 | 0.7002 | 0.7328 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.16 0.52 | 0.00 0.79 | 0.00 0.22 | 0.00 0.03 | 0.00 0.11 | 0.08 0.47 |
| T_r^{10} (random, size 10) | HV | OPAM | 0.0000 | 0.8976 | 0.9792 | 0.9449 | 0.9837 | 0.9999 |
| | | RS | 0.0000 | 0.8517 | 0.8191 | 0.6879 | 0.5183 | 0.0000 |
| | | $p \hat{A}_{12}$ | 1.00 0.50 | 0.00 0.90 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 |
| | GD+ | OPAM | 0.0000 | 0.0806 | 0.0035 | 0.0043 | 0.0211 | 0.0134 |
| | | RS | 0.0200 | 0.1252 | 0.0912 | 0.0789 | 0.1580 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.16 0.48 | 0.00 0.09 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 |
| | Δ | OPAM | 1.0000 | 0.8662 | 0.4603 | 0.3951 | 0.4728 | 0.9433 |
| | | RS | 0.9600 | 0.6579 | 0.6331 | 0.7035 | 0.7617 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.16 0.52 | 0.00 0.73 | 0.00 0.20 | 0.00 0.02 | 0.00 0.05 | 0.08 0.47 |
| T_a^{500} (adaptive, size 500) | HV | OPAM | 1.0000 | 0.7032 | 0.9424 | 0.9089 | 0.9803 | 0.9999 |
| | | RS | 0.9000 | 0.6518 | 0.7893 | 0.6561 | 0.5167 | 0.0000 |
| | | $p \hat{A}_{12}$ | 0.02 0.55 | 0.00 0.86 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 |
| | GD+ | OPAM | 0.0000 | 0.0159 | 0.0035 | 0.0051 | 0.0064 | 0.0134 |
| | | RS | 0.0883 | 0.0393 | 0.0850 | 0.0746 | 0.1422 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.02 0.45 | 0.00 0.03 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 |
| | Δ | OPAM | 1.0000 | 0.7842 | 0.4715 | 0.3680 | 0.4850 | 0.9433 |
| | | RS | 0.9766 | 0.5354 | 0.6357 | 0.6850 | 0.7565 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.16 0.52 | 0.00 0.84 | 0.00 0.21 | 0.00 0.01 | 0.00 0.09 | 0.08 0.47 |
| T_w^{500} (worst, size 500) | HV | OPAM | 1.0000 | 0.6535 | 0.9223 | 0.9307 | 0.9635 | 0.9997 |
| | | RS | 0.9000 | 0.6050 | 0.7791 | 0.6770 | 0.5032 | 0.0000 |
| | | $p \hat{A}_{12}$ | 0.02 0.55 | 0.00 0.77 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 |
| | GD+ | OPAM | 0.0000 | 0.0302 | 0.0037 | 0.0040 | 0.0054 | 0.0136 |
| | | RS | 0.0883 | 0.0545 | 0.0768 | 0.0763 | 0.1408 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.02 0.45 | 0.00 0.09 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 |
| | Δ | OPAM | 1.0000 | 0.7899 | 0.4640 | 0.4077 | 0.5083 | 0.9433 |
| | | RS | 0.9766 | 0.5910 | 0.6114 | 0.7052 | 0.7448 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.16 0.52 | 0.00 0.84 | 0.00 0.22 | 0.00 0.02 | 0.00 0.11 | 0.08 0.47 |
| T_r^{500} (random, size 500) | HV | OPAM | 1.0000 | 0.6936 | 0.9742 | 0.9481 | 0.9810 | 0.9999 |
| | | RS | 0.9000 | 0.6401 | 0.8138 | 0.6904 | 0.5183 | 0.0000 |
| | | $p \hat{A}_{12}$ | 0.02 0.55 | 0.00 0.85 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 | 0.00 1.00 |
| | GD+ | OPAM | 0.0000 | 0.0169 | 0.0031 | 0.0041 | 0.0062 | 0.0134 |
| | | RS | 0.0883 | 0.0394 | 0.0914 | 0.0794 | 0.1420 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.02 0.45 | 0.00 0.03 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 | 0.00 0.00 |
| | Δ | OPAM | 1.0000 | 0.7415 | 0.4637 | 0.4077 | 0.4854 | 0.9433 |
| | | RS | 0.9766 | 0.5251 | 0.6358 | 0.7042 | 0.7535 | 1.0000 |
| | | $p \hat{A}_{12}$ | 0.16 0.52 | 0.00 0.80 | 0.00 0.20 | 0.00 0.03 | 0.00 0.09 | 0.08 0.47 |

n.nnnn : OPAM outperforms RS

n.nnnn : RS outperforms OPAM



(a) HV



(b) GD+

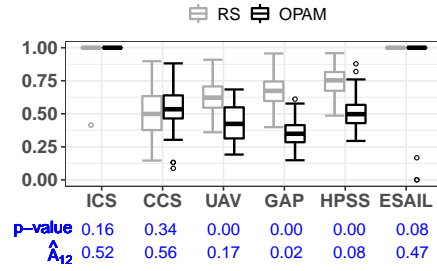
(c) Δ

Fig. 10: Comparing OPAM and RS using the three quality indicators: (a) HV, (b) GD+, and (c) Δ . The boxplots (25%-50%-75%) show the quality values obtained from 50 runs of OPAM and RS. The quality values are computed based on the Pareto fronts obtained by the algorithms and each subject's set \mathbf{E} of task-arrival sequences (see Section 7.6).

different sets of task-arrival sequences. As described in Section 7.4, we created six test sets of task-arrival sequences for each subject by varying the sequence generation methods and the number of task-arrival sequences in a set (see \mathbf{T}_a^{10} , \mathbf{T}_w^{10} , \mathbf{T}_r^{10} , \mathbf{T}_a^{500} , \mathbf{T}_w^{500} , and \mathbf{T}_r^{500} described in Section 7.4). Table 3 reports the average quality values measured by HV, GD+, and Δ based on 50 runs of OPAM and RS with the different test sets of task-arrival sequences. The

results indicate that OPAM significantly outperforms RS in most comparison cases. Specifically, out of a total of 108 comparisons, OPAM outperforms RS 87 times (see the blue-colored cells related to OPAM in Table 3). Regarding Δ , RS outperforms OPAM for the CCS subject (see the gray-colored cells related to RS in Table 3). As shown in Table 2, CCS has only 3 aperiodic tasks and RS was therefore able to find better solutions with respect to Δ for such a simple subject.

*The answer to **RQ1** is that OPAM significantly outperforms RS with respect to HV and GD+. In particular, OPAM performs considerably better than RS when more aperiodic tasks are involved.*

RQ2. To compare OPAM and SEQ, we first visually inspect the best Pareto fronts obtained from 50 runs of OPAM and SEQ for the six study systems described in Section 7.2 by varying the test sets of task-arrival sequences for each subject (see \mathbf{T}_a^{10} , \mathbf{T}_w^{10} , \mathbf{T}_r^{10} , \mathbf{T}_a^{500} , \mathbf{T}_w^{500} , and \mathbf{T}_r^{500} described in Section 7.4), which are created independently from OPAM and SEQ. Overall, we observed that OPAM finds significantly better priority assignments in most cases. For example, Figure 11 depicts the best Pareto fronts obtained by OPAM and SEQ when the fitness values are computed based on each subject's test set \mathbf{T}_a^{500} of 500 task-arrival sequences, which are generated with adaptive random search. The results clearly show that OPAM outperforms SEQ with respect to producing more optimal Pareto fronts for ICS, CCS, UAV, HPSS, and ESAIL. For GAP, the visual inspection is not sufficient to provide any conclusions. Hence, we further compare OPAM and SEQ based on the quality indicators described in Section 7.5.

Table 4 compares the quality values measured by HV, GD+, and Δ for the six study subjects. To fairly compare the priority assignments obtained by OPAM and SEQ, we assess them with the test sets of task-arrival sequences for each subject (see \mathbf{T}_a^{10} , \mathbf{T}_w^{10} , \mathbf{T}_r^{10} , \mathbf{T}_a^{500} , \mathbf{T}_w^{500} , and \mathbf{T}_r^{500} described in Section 7.4). Table 4 reports the average quality values computed based on 50 runs of OPAM and SEQ. In Table 4, the statistical comparison of the two corresponding distributions are reported using p-values and \hat{A}_{12} values.

As shown in Table 4, we compared OPAM and SEQ 108 times by varying the study subjects, the quality indicators, the number of task-arrival sequences, and the task-arrival sequence generation methods. Out of 108 comparisons, OPAM significantly outperforms SEQ 63 times. Specifically, out of 36 HV comparisons, OPAM obtains better HV values than SEQ 28 times. For ICS (6 HV comparisons), the differences in HV values between OPAM and SEQ are not statistically significant. In only one HV comparison for CCS, SEQ outperforms OPAM (see the gray-colored cell related to HV and CCS in Table 4). To interpret these results, one must recall from Table 2 that ICS and CCS have only three aperiodic tasks that impact the degree of uncertainty in task arrivals and therefore represent simple cases. Out of 36 GD+ comparisons, OPAM outperforms SEQ 32 times. SEQ outperforms OPAM only two times for CCS. Hence, overall, the results indicate that OPAM outperforms SEQ, in terms of generating more optimal Pareto fronts, when the subjects

Table 4: Comparing OPAM and SEQ using the three quality indicators: HV, GD+, and Δ . Average quality values computed based on 50 runs of OPAM and SEQ using the different sets of task-arrival sequences (see Section 7.4).

| | | | ICS | CCS | UAV | GAP | HPSS | ESAIL |
|-----------------------------------|-------------------------------------|------------------|------------------|---------------|---------------|---------------|---------------|---------------|
| T_a^{10} (adaptive, size 10) | HV | OPAM | 0.0000 | 0.6052 | 0.6011 | 0.6088 | 0.6290 | 0.9808 |
| | | SEQ | 0.0000 | 0.4172 | 0.5354 | 0.5868 | 0.6086 | 0.4470 |
| | | $p \hat{A}_{12}$ | 1.00 0.50 | 0.00 1.00 | 0.00 0.95 | 0.00 0.76 | 0.02 0.63 | 0.00 1.00 |
| | GD+ | OPAM | 0.0000 | 0.0244 | 0.0175 | 0.0148 | 0.0529 | 0.0249 |
| | | SEQ | 0.2191 | 0.0835 | 0.0350 | 0.0201 | 0.0625 | 0.1887 |
| | | $p \hat{A}_{12}$ | 0.00 0.01 | 0.00 0.00 | 0.00 0.01 | 0.00 0.25 | 0.00 0.26 | 0.00 0.03 |
| | Δ | OPAM | 1.0000 | 0.7653 | 0.4239 | 0.3343 | 0.5297 | 0.9444 |
| | | SEQ | 0.0200 | 0.5656 | 0.3628 | 0.2875 | 0.5706 | 0.8285 |
| | | $p \hat{A}_{12}$ | 0.00 0.99 | 0.00 0.81 | 0.01 0.64 | 0.01 0.65 | 0.33 0.44 | 0.00 0.75 |
| | T_w^{10} (worst, size 10) | HV | OPAM | 0.0000 | 0.7345 | 0.6258 | 0.6290 | 0.7460 |
| | | | SEQ | 0.0000 | 0.6794 | 0.5933 | 0.5928 | 0.5046 |
| | | | $p \hat{A}_{12}$ | 1.00 0.50 | 0.00 0.82 | 0.00 0.82 | 0.00 0.88 | 0.00 1.00 |
| | | GD+ | OPAM | 0.0000 | 0.0912 | 0.0191 | 0.0131 | 0.0340 |
| | | | SEQ | 0.0000 | 0.0695 | 0.0272 | 0.0211 | 0.0667 |
| | | | $p \hat{A}_{12}$ | 1.00 0.50 | 0.00 0.86 | 0.00 0.12 | 0.00 0.14 | 0.00 0.03 |
| | T_r^{10} (random, size 10) | HV | OPAM | 1.0000 | 0.7009 | 0.4835 | 0.3616 | 0.9470 |
| | | | SEQ | 1.0000 | 0.5376 | 0.3111 | 0.3054 | 0.7547 |
| | | | $p \hat{A}_{12}$ | 1.00 0.50 | 0.00 0.74 | 0.00 0.83 | 0.01 0.66 | 0.01 0.35 |
| | | GD+ | OPAM | 0.0000 | 0.0911 | 0.0205 | 0.0160 | 0.0472 |
| | | | SEQ | 0.0000 | 0.1358 | 0.0882 | 0.0277 | 0.0646 |
| | | | $p \hat{A}_{12}$ | 1.00 0.50 | 0.00 0.01 | 0.00 0.00 | 0.00 0.10 | 0.00 0.19 |
| | T_a^{500} (adaptive, size 500) | HV | OPAM | 0.0000 | 0.6781 | 0.7134 | 0.6261 | 0.7332 |
| | | | SEQ | 0.0000 | 0.4854 | 0.6179 | 0.5981 | 0.7056 |
| | | | $p \hat{A}_{12}$ | 1.00 0.50 | 0.00 1.00 | 0.00 1.00 | 0.00 0.83 | 0.00 0.73 |
| | | GD+ | OPAM | 0.0000 | 0.0174 | 0.0140 | 0.0134 | 0.0320 |
| | | | SEQ | 0.2191 | 0.0727 | 0.0549 | 0.0197 | 0.0565 |
| | | | $p \hat{A}_{12}$ | 0.00 0.01 | 0.00 0.00 | 0.00 0.00 | 0.00 0.20 | 0.00 0.08 |
| | T_w^{500} (worst, size 500) | HV | OPAM | 1.0000 | 0.7833 | 0.4964 | 0.3588 | 0.9442 |
| | | | SEQ | 0.0200 | 0.7319 | 0.4002 | 0.3315 | 0.8554 |
| | | | $p \hat{A}_{12}$ | 0.00 0.99 | 0.23 0.57 | 0.00 0.72 | 0.07 0.60 | 0.02 0.36 |
| | | GD+ | OPAM | 0.0000 | 0.0511 | 0.0141 | 0.0135 | 0.0258 |
| | | | SEQ | 0.2191 | 0.0343 | 0.0267 | 0.0226 | 0.0336 |
| | | | $p \hat{A}_{12}$ | 0.00 0.01 | 0.00 0.96 | 0.00 0.05 | 0.00 0.11 | 0.00 0.24 |
| | T_r^{500} (random, size 500) | HV | OPAM | 1.0000 | 0.7569 | 0.4950 | 0.3751 | 0.9469 |
| | | | SEQ | 0.0200 | 0.7259 | 0.3315 | 0.3139 | 0.8957 |
| | | | $p \hat{A}_{12}$ | 0.00 0.99 | 0.43 0.55 | 0.00 0.82 | 0.01 0.66 | 0.20 0.57 |
| | | GD+ | OPAM | 0.0000 | 0.6646 | 0.8446 | 0.6321 | 0.7087 |
| | | | SEQ | 0.0000 | 0.4876 | 0.7242 | 0.5839 | 0.6786 |
| | | | $p \hat{A}_{12}$ | 1.00 0.50 | 0.00 1.00 | 0.00 1.00 | 0.00 0.93 | 0.00 0.72 |
| | T_a^{500} (adaptive, size 500) | HV | OPAM | 0.0000 | 0.6646 | 0.8446 | 0.6321 | 0.7087 |
| | | | SEQ | 0.0000 | 0.4876 | 0.7242 | 0.5839 | 0.6786 |
| | | | $p \hat{A}_{12}$ | 1.00 0.50 | 0.00 1.00 | 0.00 1.00 | 0.00 0.93 | 0.00 0.72 |
| | | GD+ | OPAM | 0.0000 | 0.0184 | 0.0172 | 0.0165 | 0.0327 |
| | | | SEQ | 0.2191 | 0.0684 | 0.0791 | 0.0285 | 0.0580 |
| | | | $p \hat{A}_{12}$ | 0.00 0.01 | 0.00 0.00 | 0.00 0.00 | 0.00 0.09 | 0.00 0.06 |
| | T_w^{500} (worst, size 500) | HV | OPAM | 1.0000 | 0.7449 | 0.5059 | 0.3960 | 0.9472 |
| | | | SEQ | 0.0200 | 0.6798 | 0.4156 | 0.3341 | 0.8546 |
| | | | $p \hat{A}_{12}$ | 0.00 0.99 | 0.19 0.58 | 0.00 0.71 | 0.01 0.66 | 0.03 0.38 |
| | | GD+ | OPAM | 0.0000 | 0.6646 | 0.8446 | 0.6321 | 0.7087 |
| | | | SEQ | 0.0000 | 0.4876 | 0.7242 | 0.5839 | 0.6786 |
| | | | $p \hat{A}_{12}$ | 1.00 0.50 | 0.00 1.00 | 0.00 1.00 | 0.00 0.93 | 0.00 0.72 |

n.nnnn : OPAM outperforms SEQ

n.nnnn : SEQ outperforms OPAM

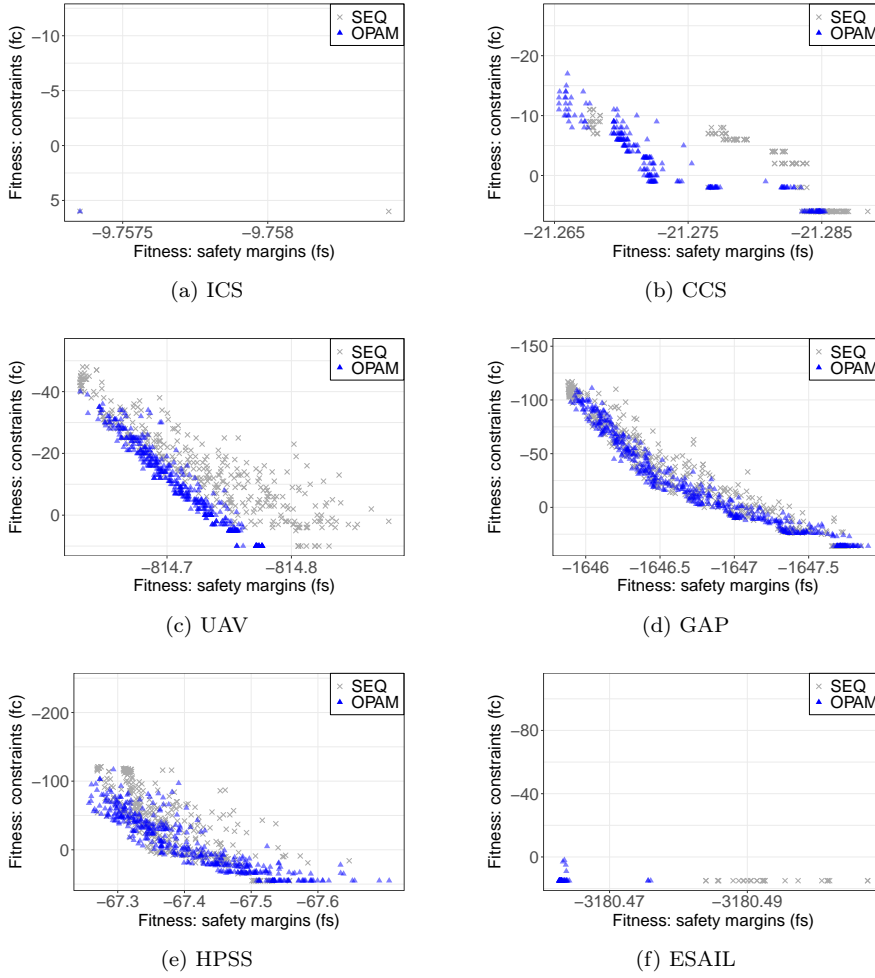


Fig. 11: Pareto fronts obtained by OPAM and SEQ for the six industrial subjects: (a) ICS, (b) CCS, (c) UAV, (d) GAP, (e) HPSS, and (f) ESAIL. The fitness values are computed based on each subject's set \mathbf{T}_a^{500} of task-arrival sequences (see Section 7.4). The points located closer to the bottom left of each plot are considered to be better priority assignments when compared to points closer to the top right.

feature a considerable degree of uncertainty in task arrivals and therefore make our search problem more complex. Otherwise differences are not statistically or practically significant. Regarding Δ , which focuses on the diversity of solutions on the Pareto front, SEQ outperforms OPAM 24 times out of 36 comparisons (see the gray-colored cells related to Δ in Table 4). However, since OPAM pro-

Table 5: Execution times and memory usage required to run OPAM for the six industrial subjects. Average values computed based on 50 runs of OPAM are reported.

| Subject | Execution time (s) | Memory usage (MB) |
|---------|--------------------|-------------------|
| ICS | 104.34 | 89.97 |
| CCS | 165.50 | 111.85 |
| UAV | 1455.35 | 312.85 |
| GAP | 2819.03 | 730.29 |
| HPSS | 226.98 | 127.77 |
| ESAIL | 55844.23 | 2879.79 |

duces enough alternative priority assignments spreading across Pareto fronts (as visible from the solutions obtained by OPAM in Figure 11), these differences in Δ have limited implications in practice.

*The answer to **RQ2** is that OPAM significantly outperforms SEQ with respect to HV and GD+ when in the presence of more than a few aperiodic tasks and therefore higher uncertainty in terms of task arrivals. OPAM therefore generate solutions on a Pareto front that is closer to the unknown, optimal one. In other words, coevolution is a suitable and successful strategy for finding better priority assignments in complex systems.*

RQ3. Table 5 reports the average execution times and memory usage required to run OPAM for the six industrial subjects, over 50 runs. As shown in Table 5, finding optimal priority assignments for ESAIL requires the largest execution time (≈ 15.5 h) and memory usage (≈ 2.9 GB), compared to the other subjects. We note that such execution time and memory usage are acceptable as OPAM can be executed offline in practice.

Figures 12 and 13 show, respectively, the execution times and memory usage from EXP3.1 (a), EXP3.2 (b), EXP3.3 (c), and EXP3.4 (d), described in Section 7.4. The boxplots in the figures show distributions (25%-50%-75%) obtained from 50×10 runs of OPAM for a set of 10 synthetic subjects, which are created with the same experimental setting. Regarding the execution time of OPAM, Figures 12a and 12d show that the execution time of OPAM is linear both in the number of tasks and simulation time. As for the memory usage of OPAM, results in Figures 13a and 13d indicate that memory usage is linear both in the number of tasks and in the simulation time. However, the results depicted in Figures 12b, 12c, 13b, and 13c indicate that there are no correlations between OPAM execution time and memory usage and the following two parameters: ratio of aperiodic tasks and range factor. Therefore, we expect OPAM to scale well as the numbers of tasks and simulation time increase.

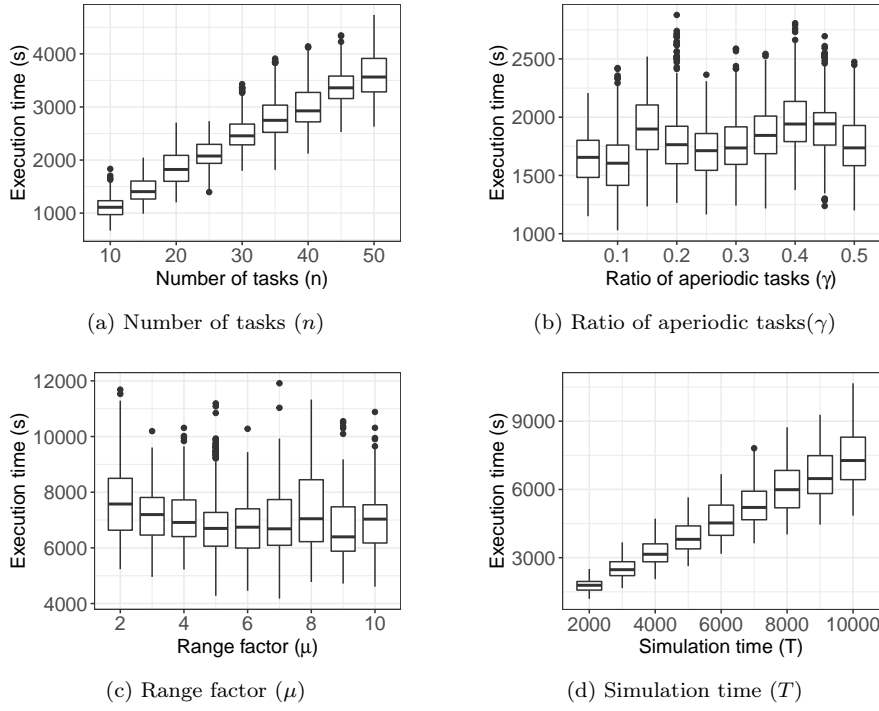


Fig. 12: Execution times of OPAM when varying the values of the following parameters: (a) number of tasks n , (b) ratio of aperiodic tasks γ , (c) range factor μ , and (d) simulation time T . The boxplots (25%-50%-75%) show the execution times obtained from 500 runs of OPAM, i.e., 50 runs for each of the 10 synthetic subjects with the same configuration.

*The answer to **RQ3** is that the execution time and memory usage of OPAM are linear in the number of tasks and simulation time, thus scaling to industrial systems. Further, across our experiments, OPAM takes at most 15.5h using 2.9GB of memory to optimize priority assignments, an acceptable result since this is done offline.*

RQ4. Figure 14 compares, with respect to external fitness (see the $fs()$ and $fc()$ fitness functions and the set \mathbf{E} of sequences of task arrivals described in Section 6.6), the Pareto solutions obtained by OPAM against the priority assignments defined by engineers for the six industrial subjects: ICS (Figure 14a), CCS (Figure 14b), UAV (Figure 14c), GAP (Figure 14d), HPSS (Figure 14e), and ESAIL (Figure 14f).

As shown in the figure, the solutions obtained by OPAM clearly outperform the priority assignments defined by engineers regarding the two external ob-

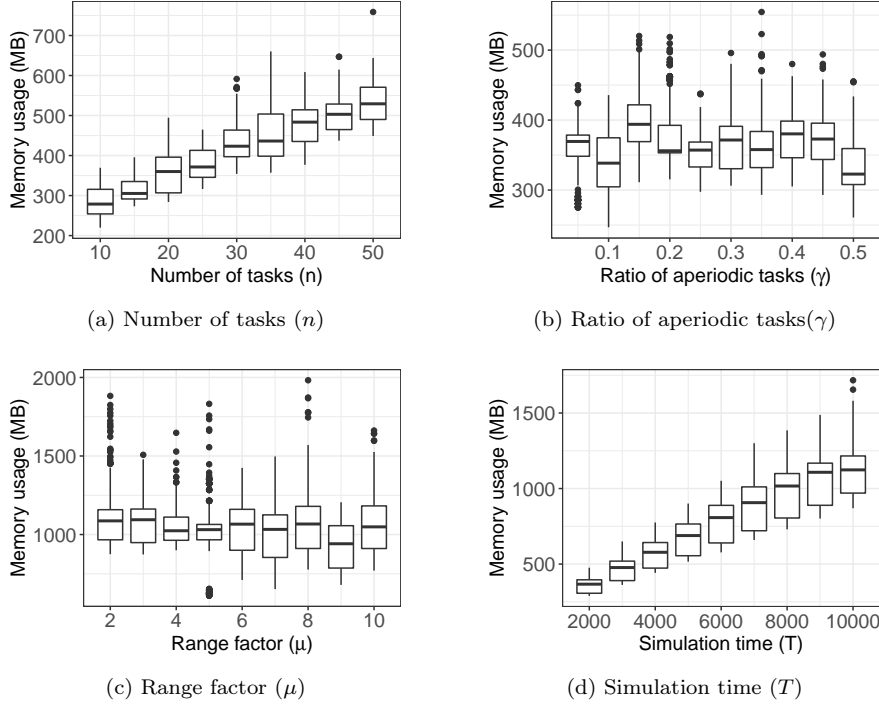


Fig. 13: Memory usage of OPAM when varying the values of the following parameters: (a) number of tasks n , (b) ratio of aperiodic tasks γ , (c) range factor μ , and (d) simulation time T . The boxplots (25%-50%-75%) show the memory usage obtained from 500 runs of OPAM, i.e., 50 runs for each of the synthetic subjects with the same configuration.

jectives: the magnitude of safety margins and the extent to which constraints are satisfied.

Table 6 summarizes safety margins from the task executions of ESAIL when using one of our priority assignments optimized by OPAM and the one defined by engineers at LuxSpace. Note that we focus on ESAIL as it is not possible to access the engineers who developed the other five industrial subjects reported in the literature (Locke et al., 1990; Traore et al., 2006; Peraldi-Frati and Sorel, 2008; Mikučionis et al., 2010; Anssi et al., 2011). For comparison, we chose the bottom-left solution in Figure 14f since it is optimal for the constraint fitness, which is the same as the fitness value of the priority assignment defined by engineers, and the differences in safety margin fitness among our solutions are negligible.

As shown in Table 6, our optimized priority assignment significantly outperforms the one of engineers. Our solution increases safety margins, on av-

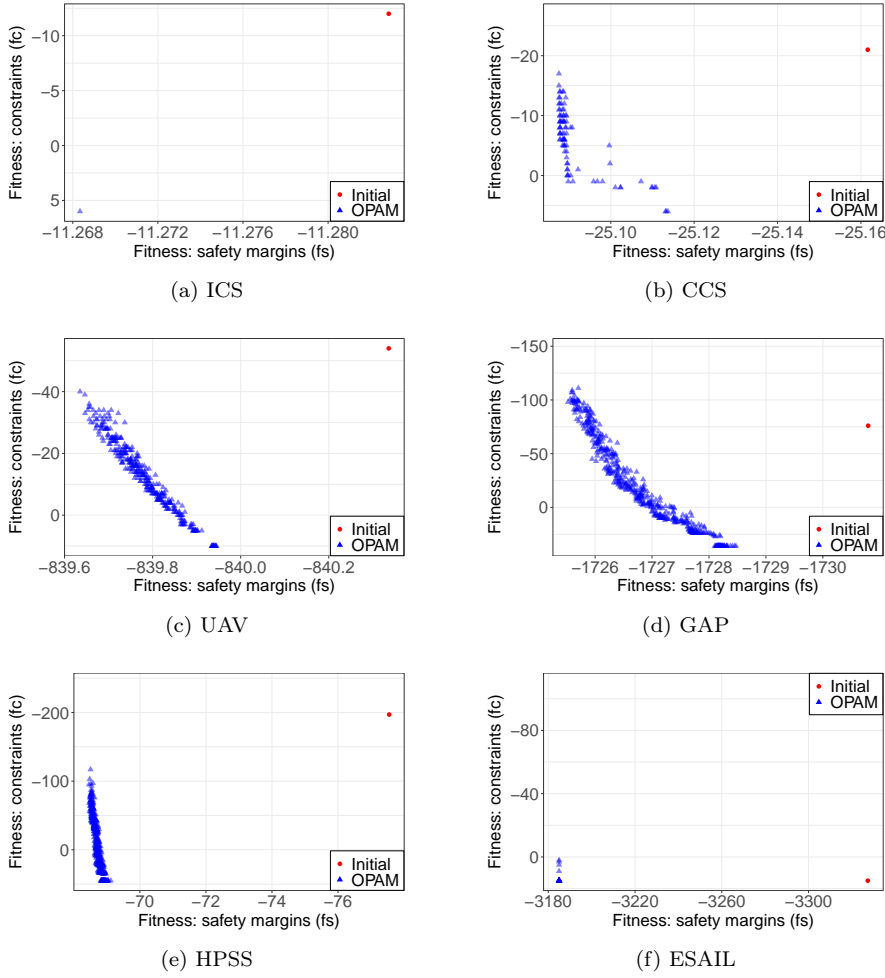


Fig. 14: Comparing Pareto solutions obtained by OPAM and priority assignments defined by engineers for the six industrial subjects: (a) ICS, (b) CCS, (c) UAV, (d) GAP, (e) HPSS, and (f) ESAIL. The points located closer to the bottom left of each plot are considered to be better priority assignments when compared to points closer to the top right.

erage, by 5.33% compared to the engineers' solution. For aperiodic tasks, our solution decreases safety margins by 0.01% (4.2ms difference) when the safety margins being compared are the maximum margins observed in both solutions (see the maximum safety margins, 59710.3ms obtained by engineers' solution and 59707.2ms obtained by OPAM, in Table 6). Such a small decrease is however negligible in the context of ESAIL as the maximum safety margin

Table 6: Comparing safety margins from the task executions of ESAIL when using our optimized priority assignment and the one defined by engineers.

| | | Periodic tasks | Aperiodic tasks | All tasks |
|--------------|--------|----------------|-----------------|-----------|
| Engineer | Min | -44.5 | 9.4 | -44.5 |
| | Max | 1879.7 | 59710.3 | 59710.3 |
| | Avg. | 126.6 | 52.6 | 78.1 |
| | Median | 82.1 | 9.4 | 48.1 |
| OPAM | Min | 48.1 | 9.4 | 9.4 |
| | Max | 1879.7 | 59707.2 | 59707.2 |
| | Avg. | 129.8 | 57.2 | 82.3 |
| | Median | 85.7 | 9.4 | 48.1 |
| % Difference | Min | 208.09% | 0.00% | 121.12% |
| | Max | 0.00% | -0.01% | -0.01% |
| | Avg. | 2.53% | 8.89% | 5.33% |
| | Median | 4.38% | 0.00% | 0.00% |

* Unit of time: ms

obtained by our solution is still large, i.e., $\approx 1\text{m}$. For periodic tasks, we note that our solution increases safety margins by 208.09% when the safety margins being compared are the minimum margins observed in both solutions (see the minimum safety margins, -44.5ms obtained by engineers' solution and 48.1ms obtained by OPAM, in Table 6). Note that the minimum safety margin of -44.5ms obtained with the engineers' solution indicates that a task violates its deadline. In the context of ESAIL, which is a mission-critical system, such gain in safety margins in the executions of periodic tasks is important because the hard deadlines of periodic tasks are more critical than the soft deadlines of aperiodic tasks.

Investigating practitioners' perceptions of the benefits of OPAM is necessary to adopt OPAM in practice. To do so, we draw on the qualitative reflections of three software engineers at LuxSpace, with whom we have been collaborating on this research. They have had four to seven years of experience developing satellite systems at LuxSpace, with more than 50 years of collective experience in companies. All the reflections are based on observations made throughout our interactions. The engineers at LuxSpace deemed OPAM to be an improvement over their current practice as it allows them to perform domain-specific trade-off analysis among Pareto solutions and is useful in practice to support decision making with respect to their task design. Encouraged by the promising results, we are now applying OPAM to new systems in collaboration with LuxSpace.

*The answer to **RQ4** is that OPAM helps optimize priority assignments such that they outperform those manually defined by engineers based on domain expertise. Our results show that OPAM, compared to current practice, increases safety margins, on average, by 5.33%.*

7.8 Threats to Validity

To mitigate the main threats that arise from not accounting for random variation, we compared OPAM against RS under identical parameter settings. We present all the underlying parameters and provide the full package of our experiments to facilitate replication. Also, we ran OPAM 50 times for each study subject and compared results using statistical analysis, i.e., Mann-Whitney U-test and Vargha and Delaney’s \hat{A}_{12} .

We note that there are prior studies that aim at optimizing priority assignments such as OPA (Audsley, 1991) and RPA (Davis and Burns, 2007). However, to our knowledge, none of the existing works offer ways to analyze trade-offs among equally viable priority assignments with respect to safety margins and the satisfaction of constraints. Nevertheless, we attempted to compare OPAM with an extension of an existing method, e.g., RPA (Davis and Burns, 2007). To do so, we first applied an exhaustive schedulability analysis technique to the ESAIL subject – our motivating case study – in order to verify whether the ESAIL tasks are schedulable for a given priority assignment. Note that existing priority assignment techniques are built on such schedulability analysis methods, which are therefore a prerequisite. We chose UPPAAL (Behrmann et al., 2004), a model checker, for schedulability analysis as it has been used in real-time system studies (Mikučionis et al., 2010; Yu et al., 2010; Yalcinkaya et al., 2019). However, our experiment results using UPPAAL for ESAIL showed that it was not able to complete the analysis task, even after 5 days of execution, for a single priority assignment. We were therefore not able to perform experimental comparisons with existing priority assignment methods. Since this evaluation is not the main focus of this article, we point the reader to the UPPAAL specification of ESAIL available online (Lee et al., 2021).

Recall from Section 6.2 that OPAM assigns tasks’ WCETs to their execution times when it simulates the worst-case executions of tasks while varying task arrival times. In many real-time systems studies (Briand et al., 2005; Guan et al., 2009; Lin et al., 2009; Anssi et al., 2011; Zeng et al., 2014; Di Alesio et al., 2015; Dürr et al., 2019), static WCETs are often used instead of varying task execution times for the purpose of real-time analysis. For example, practitioners typically use WCETs to estimate the lowest bound of CPU utilization required to properly apply the rate monotonic scheduling policy (Fineberg and Serlin, 1967) to their systems. Similarly, OPAM assumes that near-worst-case schedule scenarios can be simulated by assigning tasks’ WCETs to their execution times and varying tasks’ arrival times using search. A near-worst-case schedule scenario entails that the magnitude of deadline misses is maximized when tasks execute as per this scenario. Under this working assumption, we were able to empirically evaluate the sanity, coevolution, scalability, and usefulness aspects of OPAM (see Section 7). The results indicate that OPAM is a promising and useful tool. However, the formal proof of whether or not the WCET assumption holds in the system model described in Section 3 requires complex analysis, accounting for varying task arrival times, triggering relation-

ships, resource dependencies, and multiple cores. When task execution times need to be varied during simulation, engineers can adapt OPAM by utilizing Monte-Carlo simulation (Kroese et al., 2014) to account for such variations.

The main threat to external validity is that our results may not generalize to other systems. We mitigate potential biases and errors in our experiments by drawing on real industrial subjects from different domains and several synthetic subjects. Specifically, we selected two subjects from the aerospace domain, two from the automotive domain, and two from the avionics domain. The positive feedback obtained from LuxSpace and the encouraging results from our industrial case studies indicate that OPAM is a scalable and practical solution. Furthermore, we believe OPAM introduces a promising avenue for addressing the problem of priority assignment by applying coevolutionary algorithms, even for systems that use other scheduling policies, e.g., priority inheritance. In order for OPAM to support different scheduling policies, the main requirement is to replace the existing simulator (described in Section 6) with a new simulator supporting the desired scheduling policy. In our approach, the coevolution part of OPAM is separated from the scheduling policy, which is contained in the simulator. Hence, we deem the expected changes for the coevolution part of OPAM to be minimal. Future studies are nevertheless necessary to investigate how OPAM can be adapted to find near-optimal priority assignments for other real-time systems in different contexts.

8 Conclusion

We developed OPAM, a priority assignment method for real-time systems, that aims to find equally viable priority assignments that maximize the magnitude of safety margins and the extent to which engineering constraints are satisfied. OPAM uses a novel approach, based on multi-objective, competitive coevolutionary search, that simultaneously evolves different species, i.e., populations of priority assignments and stress test scenarios, that compete with one another with opposite objectives, the former trying to minimize chances of deadline misses while the latter attempts to maximize them. We evaluated OPAM on a number of synthetic systems as well as six industrial systems from different domains. The results indicate that OPAM is able to find significantly better solutions than both those manually defined by engineers based on expert knowledge and those obtained by our baselines: random search and sequential search. Further, OPAM scales linearly with the number of tasks in a system and the time required to simulate task executions. Execution times on our industrial systems are practically acceptable.

In the future, we will continue to study the problem of optimal priority assignment by accounting for (1) priority assignments that change dynamically, (2) WCET value ranges that account for non-deterministic computation times, (3) interrupt handling routines that execute differently compared to real-time tasks, and (4) hybrid scheduling policies that combine multiple standard policies. We also plan to develop a real-time task modeling language to

specify task characteristics such as resource dependencies, triggering relationships, engineering constraints, and behaviors of real-time tasks and to facilitate real-time system analysis, e.g., optimal priority assignment and schedulability analysis. In addition, we would like to incorporate additional analysis capabilities into OPAM in order to verify whether or not a system satisfies the required properties, e.g., schedulability of tasks and absence of deadlocks, for a given priority assignment. For example, statistical model checking (Legay et al., 2010) may allow us to verify whether tasks meet their deadlines for a given priority assignment with a probabilistic guarantee. In the long term, we plan to more conclusively validate the usefulness of OPAM by applying it to additional case studies in different application domains.

Acknowledgements We thank Yago Isasi Parache, LuxSpace, for his support in conducting our industrial case study. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 694277), and NSERC of Canada under the Discovery and CRC programs. The experiments presented in this paper were carried out using the HPC facilities of the University of Luxembourg (Varrette et al., 2014) – see hpc.uni.lu.

References

- Abdessalem RB, Panichella A, Nejati S, Briand LC, Stifter T (2020) Automated repair of feature interaction failures in automated driving systems. In: Proceedings of the 29th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA’20), pp 88–100
- Andrade SS, Macêdo RJdA (2013) A search-based approach for architectural design of feedback control concerns in self-adaptive systems. In: Proceedings of the 2013 IEEE 7th International Conference on Self-Adaptive and Self-Organizing Systems (SASO’13), pp 61–70
- Anssi S, Tucci-Piergiovanni S, Kuntz S, Gérard S, Terrier F (2011) Enabling scheduling analysis for AUTOSAR systems. In: Proceedings of the 14th IEEE International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing (ISORC’11), pp 152–159
- Arcuri A, Briand LC (2014) A hitchhiker’s guide to statistical tests for assessing randomized algorithms in software engineering. *Software Testing, Verification and Reliability* 24(3):219–250
- Arcuri A, Fraser G (2011) On parameter tuning in search based software engineering. In: Proceedings of the 3th International Symposium on Search Based Software Engineering (SSBSE’11), pp 33–47
- Arcuri A, Iqbal MZ, Briand LC (2010) Black-box system testing of real-time embedded systems using random and search-based testing. In: Proceedings of the IFIP International Conference on Testing Software and Systems (ICTSS’10), vol 6435, pp 95–110
- Arpaci-Dusseau RH, Arpaci-Dusseau AC (2018) *Operating Systems: Three Easy Pieces*, 1st edn. Arpaci-Dusseau Books

- Audsley NC (1991) Optimal priority assignment and feasibility of static priority tasks with arbitrary start times. Tech. rep., Dept. Computer Science, University of York
- Audsley NC (2001) On priority assignment in fixed priority scheduling. *Information Processing Letters* 79(1):39–44
- Baruah SK, Burns A, Davis RI (2011) Response-time analysis for mixed criticality systems. In: *Proceedings of the 2011 IEEE 32nd Real-Time Systems Symposium (RTSS'11)*, pp 34–43
- Behrmann G, David A, Larsen KG (2004) A tutorial on UPPAAL. In: *Formal Methods for the Design of Real-Time Systems: International School on Formal Methods for the Design of Computer, Communication, and Software Systems*, pp 200–236
- Bernat G, Burns A, Llamas A (2001) Weakly hard real-time systems. *IEEE Transactions on Computers* 50(4):308–321
- Boussaa M, Kessentini W, Kessentini M, Bechikh S, Ben Chikha S (2013) Competitive coevolutionary code-smells detection. In: *Proceedings of the 5th International Symposium on Search Based Software Engineering (SS-BSE'13)*, pp 50–65
- Briand LC, Labiche Y, Shousha M (2005) Stress testing real-time systems with genetic algorithms. In: *Proceedings of the 7th Annual Conference on Genetic and evolutionary Computation (GECCO'05)*, pp 1021–1028
- Chen T, Li K, Bahsoon R, Yao X (2018) FEMOSAA: Feature-guided and knee-driven multi-objective optimization for self-adaptive software. *ACM Transactions on Software Engineering and Methodology* 27(2):1–50
- Chen TY, Kuo FC, Merkel RG, Tse TH (2010) Adaptive Random Testing: The ART of test case diversity. *Journal of Systems and Software* 83(1):60–66
- Chu Y, Burns A (2008) Flexible hard real-time scheduling for deliberative AI systems. *Real-Time Systems* 40(3):241–263
- Davis RI, Bertogna M (2012) Optimal fixed priority scheduling with deferred pre-emption. In: *Proceedings of the 2012 IEEE 33rd Real-Time Systems Symposium (RTSS'12)*, pp 39–50
- Davis RI, Burns A (2007) Robust priority assignment for fixed priority real-time systems. In: *Proceedings of the 28th IEEE International Real-Time Systems Symposium (RTSS'07)*, pp 3–14
- Davis RI, Burns A (2009) Robust priority assignment for messages on Controller Area Network (CAN). *Real-Time Systems* 41(2):152–180
- Davis RI, Burns A (2011) Improved priority assignment for global fixed priority pre-emptive scheduling in multiprocessor real-time systems. *Real-Time Systems* 47(1):1–40
- Davis RI, Zabus A, Burns A (2008) Efficient exact schedulability tests for fixed priority real-time systems. *IEEE Transactions on Computers* 57(9):1261–1276
- Davis RI, Cucu-Grosjean L, Bertogna M, Burns A (2016) A review of priority assignment in real-time systems. *Journal of Systems Architecture* 65:64–82

- Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6(2):182–197
- Di Alesio S, Gotlieb A, Nejati S, Briand LC (2012) Testing deadline misses for real-time systems using constraint optimization techniques. In: *Proceedings of the 2012 IEEE 5th International Conference on Software Testing, Verification and Validation (ICST'12)*, pp 764–769
- Di Alesio S, Nejati S, Briand LC, Gotlieb A (2013) Stress testing of task deadlines: A constraint programming approach. In: *Proceedings of the IEEE 24th International Symposium on Software Reliability Engineering (ISSRE'13)*, pp 158–167
- Di Alesio S, Briand LC, Nejati S, Gotlieb A (2015) Combining genetic algorithms and constraint programming to support stress testing of task deadlines. *ACM Transactions on Software Engineering and Methodology* 25(1):1–37
- Durillo JJ, Nebro AJ (2011) JMetal: A java framework for multi-objective optimization. *Advances in Engineering Software* 42(10):760–771
- Dürr M, Brüggén GVD, Chen KH, Chen JJ (2019) End-to-end timing analysis of sporadic cause-effect chains in distributed systems. *ACM Transactions on Embedded Computing Systems* 18(5s):1–24
- Emberson P, Stafford R, Davis RI (2010) Techniques for the synthesis of multiprocessor tasksets. In: *Proceedings of the 1st International Workshop on Analysis Tools and Methodologies for Embedded and Real-time Systems (WATERS'10)*, pp 6–11
- Ficici SG (2004) Solution concepts in coevolutionary algorithms. Ph.d. thesis, Brandeis University, Department of Computer Science, Waltham, MA
- Fineberg MS, Serlin O (1967) Multiprogramming for hybrid computation. In: *Proceedings of the AFIPS Fall Joint Computing Conference (AFIPS'67)*, pp 1–13
- Gendreau M, Potvin JY (2010) *Handbook of Metaheuristics*. Springer
- George L, Rivierre N, Spuri M (1996) Preemptive and non-preemptive real-time uniprocessor scheduling. Research Report RR-2966, INRIA, projet RE-FLECS
- Goldberg DE, Lingle R (1985) Alleles and the traveling salesman problem. In: *Proceedings of the 1st International Conference on Genetic Algorithms*, pp 154–159
- Grass W, Nguyen THC (2018) Improved response-time bounds in fixed priority scheduling with arbitrary deadlines. *Real-Time Systems* 54(1):1–30
- Guan N, Stigge M, Yi W, Yu G (2009) New response time bounds for fixed priority multiprocessor scheduling. In: *Proceedings of the 2009 30th IEEE International Real-Time Systems Symposium (RTSS'09)*, pp 387–397
- Hatvani L, Afshar S, Bril RJ (2018) Optimal priority and threshold assignment for fixed-priority preemption threshold scheduling. *ACM SIGBED Review* 15(1):43–49

- Hemmati H, Arcuri A, Briand LC (2013) Achieving scalable model-based testing through test case diversity. *ACM Transactions on Software Engineering and Methodology* 22(1):1–42
- Ishibuchi H, Masuda H, Tanigaki Y, Nojima Y (2015) Modified distance calculation in generational distance and inverted generational distance. In: *Proceedings of the 8th International Conference on Evolutionary Multi-Criterion Optimization (EMO'15)*, pp 110–125
- Islam MM, Marchetto A, Susi A, Scanniello G (2012) A multi-objective technique to prioritize test cases based on latent semantic indexing. In: *Proceedings of the 2012 16th European Conference on Software Maintenance and Reengineering (CSMR'12)*, pp 21–30
- Knowles JD, Corne DW (2000) Approximating the nondominated front using the pareto archived evolution strategy. *Evolutionary Computation* 8(2):149–172
- Kroese DP, Brereton TJ, Taimre T, Botev ZI (2014) Why the Monte Carlo method is so important today. *Wiley Interdisciplinary Reviews: Computational Statistics* 6:386–392
- Kwiatkowska M, Norman G, Parker D (2011) PRISM 4.0: Verification of probabilistic real-time systems. In: *Proceedings of the 23rd International Conference on Computer Aided Verification (CAV'11)*, pp 585–591
- Langdon WB, Harman M, Jia Y (2010) Efficient multi-objective higher order mutation testing with genetic programming. *Journal of Systems and Software* 83(12):2416–2430
- Lee H, Lee J, Yeom I, Woo H (2020a) Panda: Reinforcement learning-based priority assignment for multi-processor real-time scheduling. *IEEE Access* 8:185570–185583
- Lee J, Shin SY, Nejati S, Briand LC, Parache YI (2020b) Schedulability analysis of real-time systems with uncertain worst-case execution times. *CoRR abs/2007.10490*
- Lee J, Shin SY, Nejati S, Briand LC (2021) [Evaluation package] Optimal priority assignment method for real-time systems. <https://github.com/SNTSVV/OPAM>
- Legay A, Delahaye B, Bensalem S (2010) Statistical model checking: An overview. In: *Proceedings of the International Conference on Runtime Verification (RV'10)*, pp 122–135
- Leung JYT, Whitehead J (1982) On the complexity of fixed-priority scheduling of periodic, real-time tasks. *Performance Evaluation* 2(4):237–250
- Li M, Chen T, Yao X (2020) How to evaluate solutions in Pareto-based search-based software engineering? A critical review and methodological guidance. *IEEE Transactions on Software Engineering*
- Lin M, Xu L, Yang LT, Qin X, Zheng N, Wu Z, Qiu M (2009) Static security optimization for real-time systems. *IEEE Transactions on Industrial Informatics* 5(1):22–37
- Liu CL, Layland JW (1973) Scheduling algorithms for multiprogramming in a hard-real-time environment. *Journal of the ACM* 20(1):46–61
- Liu JWS (2000) *Real-Time Systems*, 1st edn. Prentice Hall PTR

- Locke CD, Vogel DR, Lucas L, Goodenough JB (1990) Generic avionics software specification. Tech. rep., DTIC Document
- Luke S (2013) Essentials of Metaheuristics, 2nd edn. Lulu, available for free at <http://cs.gmu.edu/~sean/book/metaheuristics/>
- LuxSpace (2021) ESAIL. <https://luxspace.lu/triton-2/>
- Mann HB, Whitney DR (1947) On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics* 18(1):50–60
- Marchetto A, Islam MM, Asghar W, Susi A, Scanniello G (2016) A multi-objective technique to prioritize test cases. *IEEE Transactions on Software Engineering* 42(10):918–940
- Meneghini IR, Guimarães FG, Gaspar-Cunha A (2016) Competitive coevolutionary algorithm for robust multi-objective optimization: The worst case minimization. In: *Proceedings of the 2016 IEEE Congress on Evolutionary Computation (CEC'16)*, pp 586–593
- Mikučionis M, Larsen KG, Rasmussen JI, Nielsen B, Skou A, Palm SU, Pedersen JS, Hougaard P (2010) Schedulability analysis using UPPAAL: Herschel-Planck case study. In: *Proceedings of the International Symposium On Leveraging Applications of Formal Methods, Verification and Validation (ISoLA'10)*, pp 175–190
- Nejati S, Briand LC (2014) Identifying optimal trade-offs between CPU time usage and temporal constraints using search. In: *Proceedings of the 2014 International Symposium on Software Testing and Analysis (ISSTA'14)*, pp 351–361
- Nejati S, Di Alesio S, Sabetzadeh M, Briand LC (2012) Modeling and analysis of CPU usage in safety-critical embedded systems to support stress testing. In: *Proceedings of the 15th International Conference of Model Driven Engineering Languages and Systems (MODELS'12)*, vol 7590, pp 759–775
- Nejati S, Adedjouma M, Briand LC, Hellebaut J, Begey J, Clement Y (2013) Minimizing CPU time shortage risks in integrated embedded software. In: *Proceedings of the 28th IEEE/ACM International Conference on Automated Software Engineering (ASE'13)*, pp 529–539
- Peng D, Shin KG, Abdelzaher TF (1997) Assignment and scheduling communicating periodic tasks in distributed real-time systems. *IEEE Transactions on Software Engineering* 23(12):745–758
- Peraldi-Frati MA, Sorel Y (2008) From high-level modelling of time in MARTE to real-time scheduling analysis. In: *Proceedings of the MODELS'08 Workshop on Model Based Architecting and Construction of Embedded Systems (ACES-MB)*, vol 503, pp 129–144
- Popovici E, Bucci A, Wiegand RP, De Jong ED (2012) Coevolutionary principles. In: *Handbook of Natural Computing*, Springer, pp 987–1033
- Ralph P, bin Ali N, Baltés S, Bianculli D, Diaz J, Dittrich Y, Ernst N, Felderer M, Feldt R, Filieri A, de França BBN, Furia CA, Gay G, Gold N, Graziotin D, He P, Hoda R, Juristo N, Kitchenham B, Lenarduzzi V, Martínez J, Melegati J, Mendez D, Menzies T, Molleri J, Pfahl D, Robbes R, Russo D, Saarimäki N, Sarro F, Taibi D, Siegmund J, Spinellis D, Staron M, Stol K,

- Storey MA, Taibi D, Tamburri D, Torchiano M, Treude C, Turhan B, Wang X, Vegas S (2020) Empirical standards for software engineering research. 2010.03525
- Sayyad AS, Goseva-Popstojanova K, Menzies T, Ammar H (2013) On parameter tuning in search based software engineering: A replicated empirical study. In: Proceedings of the 2013 3rd International Workshop on Replication in Empirical Software Engineering Research (RESER'13), pp 84–90
- Shin SY, Nejati S, Sabetzadeh M, Briand LC, Zimmer F (2018) Test case prioritization for acceptance testing of cyber physical systems: A multi-objective search-based approach. In: Proceedings of the 27th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA'18), pp 49–60
- Shin SY, Nejati S, Sabetzadeh M, Briand LC, Arora C, Zimmer F (2020) Dynamic adaptation of software-defined networks for IoT systems: A search-based approach. In: Proceedings of the IEEE/ACM 15th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS'20), pp 137–148
- Sivanandam SN, Deepa SN (2008) Introduction to Genetic Algorithms, 1st edn. Springer-Verlag
- Tan SH, Yoshida H, Prasad MR, Roychoudhury A (2016) Anti-patterns in search-based program repair. In: Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering (FSE'16), pp 727–738
- Tindell KW, Burns RA, Wellings A (1994) An extendible approach for analyzing fixed priority hard real-time tasks. *Real-Time Systems* 6(2):133–151
- Traore K, Grolleau E, Cottet F (2006) Simpler analysis of serial transactions using reverse transactions. In: Proceedings of the International Conference on Autonomic and Autonomous Systems (ICAS'06), p 11
- Vargha A, Delaney H (2000) A critique and improvement of the "CL" common language effect size statistics of McGraw and Wong. *Journal of Educational and Behavioral Statistics* 25(2):101–132
- Varrette S, Bouvry P, Cartiaux H, Georgatos F (2014) Management of an academic HPC cluster: The UL experience. In: Proceedings of the 2014 International Conference on High Performance Computing & Simulation (HPCS'14), pp 959–967
- Veldhuizen DAV, Lamont GB (1998) Multiobjective evolutionary algorithm research : A history and analysis. Tech. Rep. TR-98-03, Air Force Institute of Technology, Wright-Patterson AFB
- Wang P, Huang J, Cui Z, Xie L, Chen J (2020) A gaussian error correction multi-objective positioning model with NSGA-II. *Concurrency and Computation: Practice and Experience* 32(5):1–16
- Wegener J, Grochtmann M (1998) Verifying timing constraints of real-time systems by means of evolutionary testing. *Real-Time Systems* 15(3):275–298
- Wegener J, Sthamer H, Jones BF, Eyres DE (1997) Testing real-time systems using genetic algorithms. *Software Quality Journal* 6(2):127–135

- Weimer W, Nguyen T, Le Goues C, Forrest S (2009) Automatically finding patches using genetic programming. In: Proceedings of the 31st International Conference on Software Engineering (ICSE'09), pp 364–374
- Whitley D, Kauth J (1988) GENITOR: A different genetic algorithm. In: Proceedings of the 1988 Rocky Mountain Conference on Artificial Intelligence, pp 118–130
- Wilkerson JL, Tauritz D (2010) Coevolutionary automated software correction. In: Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation (GECCO'10), pp 1391–1392
- Wilkerson JL, Tauritz DR, Bridges JM (2012) Multi-objective coevolutionary automated software correction. In: Proceedings of the 14th Annual Conference on Genetic and Evolutionary Computation (GECCO'12), pp 1229–1236
- Yalcinkaya B, Nasri M, Brandenburg BB (2019) An exact schedulability test for non-preemptive self-suspending real-time tasks. In: Proceedings of the 2019 Design, Automation & Test in Europe Conference & Exhibition (DATE'19), pp 1228–1233
- Yu F, Li G, Xiong N (2010) Schedulability analysis of multi-processor real-time systems using UPPAAL. In: Proceedings of the 2nd International Conference on Information Science and Engineering (ICISE'10), pp 1–6
- Zeng H, Natale MD, Zhu Q (2014) Minimizing stack and communication memory usage in real-time embedded applications. *ACM Transactions on Embedded Computing Systems* 13(5s):1–25
- Zhang F, Burns A (2009) Schedulability analysis for real-time systems with EDF scheduling. *IEEE Transactions on Computers* 58(9):1250–1258
- Zhao Y, Zeng H (2017) The virtual deadline based optimization algorithm for priority assignment in fixed-priority scheduling. In: Proceedings of the 2017 IEEE Real-Time Systems Symposium (RTSS'17), pp 116–127
- Zitzler E, Thiele L (1999) Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach. *IEEE Transactions on Evolutionary Computation* 3(4):257–271



Jaekwon Lee is a Ph.D. candidate at the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg. He received his M.S. in Computer Engineering from Chungbuk National University (South Korea) in 2015. His research interests are in search-based software engineering, empirical software engineering, and machine learning.



Seung Yeob Shin is a Research Scientist at the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg. He received his Ph.D. degree from the Laboratory for Advanced Software Engineering Research (LASER) in the College of Information and Computer Sciences at the University of Massachusetts Amherst in 2016. His research interests are in software engineering, focusing on model-driven software development, search-based software engineering, empirical software engineering, and analysis of complex systems.



Shiva Nejati is an Associate Professor at the School of Electrical Engineering and Computer Science of the University of Ottawa and a part-time Research Scientist at the SnT Centre for Security, Reliability, and Trust, University of Luxembourg. From 2009 to 2012, she was a researcher at Simula Research Laboratory in Norway. She received her Ph.D. degree from the University of Toronto, Canada in 2008. Nejati's research interests are in software engineering, focussing on model-based development, software testing, analysis of cyber-physical systems, search-based software engineering and formal and empirical software engineering methods. Nejati has coauthored over 50 journal and conference papers, and regularly serves on the program committees of international conferences in the area of software engineering. She has for the past ten years been conducting her research in close collaboration with industry partners in telecommunication, maritime, energy, automotive and aerospace sectors.



Lionel C. Briand is professor of software engineering and has shared appointments between (1) The University of Ottawa, Canada and (2) The SnT centre for Security, Reliability, and Trust, University of Luxembourg. In collaboration with colleagues, over 25 years, he has run many collaborative research projects with companies in the automotive, satellite, aerospace, energy, financial, and legal domains. Lionel has held various engineering, academic, and leading positions in six countries. He was one of the founders of the ICST conference (IEEE Int. Conf. on Software Testing, Verification, and Validation, a CORE A event) and its first general chair. He was also EiC of Empirical Software Engineering (Springer) for 13 years and led, in collaboration with first Victor Basili and then Tom Zimmermann, the journal to the top tier of the very best publication venues in software engineering.

Lionel was elevated to the grades of IEEE Fellow and ACM Fellow for his work on software testing and verification. He was granted the IEEE Computer Society Harlan Mills award and the IEEE Reliability Society engineer-of-the-year award for his work on model-

based verification and testing, respectively in 2012 and 2013. He received an ERC Advanced grant in 2016 – on the topic of modelling and testing cyber-physical systems – which is the most prestigious individual research award in the European Union. Most recently, he was awarded a Canada Research Chair (Tier 1) on "Intelligent Software Dependability and Compliance". His research interests include: software testing and verification, applications of AI in software engineering, model-driven software development, requirements engineering, and empirical software engineering.