Optimal grouping and scheduling of preventive maintenance activities

P. Viveros

Departamento de Ingeniería Industrial, Universidad Técnica Federico Santa María, Av. Santa María 6400, Santiago, Chile. Email: pablo.viveros@usm.cl

R. Mena

Departamento de Ingeniería Industrial, Universidad Técnica Federico Santa María, Av. Santa María 6400, Santiago, Chile. Email: rodrigo.mena@usm.cl

E. Zio

MINES ParisTech, PSL Research University, CRC, Sophia Antipolis, France.

Energy Department, Politecnico di Milano, Via La Masa 34, 20156 Milan, Italy.

Eminant Scholar, Department of Nuclear Engineering, College of Engineering, Kyung Hee University, Republic of Korea.

Email: enrico.zio@polimi.it

S. Campos

Departamento de Ingeniería Industrial, Universidad Técnica Federico Santa María, Av. Santa María 6400, Santiago, Chile. Email: sebastian.campos.12@sansano.usm.cl

Reliability Centered Maintenance (RCM) is a well-known methodology for the optimized design of the maintenance plans in industry, whose fundamental elements are the functional analysis of the systems and their reliability quantification. From its application, a list of maintenance activities of different type is obtained (e.g., preventive or corrective), also with recommendations for the time at which the activities should be executed. Subsequently, this information is passed to management for the planning of the maintenance equipment, scheduling and for grouping the activities in feasible packages. The grouping must seek simultaneous performing opportunism, so that, the window of the time requested to the production area (equipment possession planning) as many activities as possible are executed. Several researchers have studied this problem and proposed optimization models. However, due to the complexity of the interactions of the activities and the planning horizon, it has been necessary to introduce simplification criteria. Typically, the studies focus on determining the optimal maintenance plan, seeking to minimize total maintenance costs with respect to the desired level of reliability in a multicomponent/equipment system. This paper presents an optimization model that seeks to schedule and group a plan of preventive maintenance activities for a single machine, minimizing its unavailability and respecting tolerance criteria in the periodicity of execution (advance or delay). The proposed model is continuous and solved within the mixed-integer linear programming paradigm.

Keywords: Preventive maintenance, opportunistic maintenance grouping, maintenance activities planning, mixed-integer linear programming.

1. Introduction

One of the most recognized methodologies in maintenance engineering, and widely applied in the industry, is Reliability Centered Maintenance (RCM). Its main objective is the definition of appropriate maintenance policies so that the assets deliver the required levels of functionality.
Indirectly, RCM also seeks to reduce maintenance costs focusing on the system and avoiding unnecessary maintenance operations. RCM proposes an effective way to address the problem of unplanned detentions, pursuing cost reduction, and providing crucial information for planning activities. It is essential to mention that both RCM and other methodologies are not responsible for the planning and programming processes of the maintenance activities, i.e., do not necessarily integrate the designing and scheduling of the list of activities to execute.

The design of a maintenance plan includes different tasks that seek to manage the resources that the organization owns in the best way, determining when and how maintenance activities should be executed. Expert opinions, historical data, benchmarking, recommendations of manufacturers, among other elements, often guide this. The design must be understood as a process of continuous improvement that is flexible and adaptable to dynamic demands that the productive system demands. The question then arises as to how to use the RCM output (or equivalent) for planning and programming maintenance activities looking for opportunities to simultaneously perform them in order to reduce system interruptions and consequently inefficiency costs. This problem is trivial when the execution frequencies of preventive activities are multiples of each other, resulting in simple programming of activities where the activities are grouped into the least common multiple of their periodicities. However, when this multiplicity in the frequencies of execution is not available, programming becomes complex and unsystematic.

To address this problem and take advantage of the opportunities of grouping activities, maintenance plans that consider certain tolerance in the execution of activities can be developed, that is, with a delay or advance at the execution time. In this way, activities can be grouped without the need for their frequencies to be multiples of each other. However, grouping possibility also depends on the ability to program the simultaneous execution of the activities, i.e., their feasibility of being simultaneously performed, and also, the sufficient availability of resources. For example, in Fig. 1 activities 1 and 2 are detailed and executed with execution periodicities of $T_1$ and $T_2$, respectively. It is detailed how activity 1 must be delayed in order to be performed simultaneously with activity 2, deducing that the first grouping of both activities will occur in $T_2$. However, it is considered that there is a limit of delay and/or advancement of activities given by performance tolerances. The execution of grouped activities seeks to generate as much as possible packages to improve production efficiency, which means reducing process interruptions and thus minimize unavailability.

![Fig. 1. Two activities performed with different periodicities, being able to be delayed and/or advanced.](image)

Many researchers have addressed this problem by developing optimization models. Because of the number of activities and the planning horizon, the proposed models turn out to be highly complex due to their combinatorial nature. For this reason, assumptions for the problem simplification must be made. For example, Doostparast, Kolahan & Doostparast (2014) focus on determining an optimal maintenance plan minimizing total maintenance costs looking for a certain level of reliability in a multi-component system with some constraints such as available resources and system reliability. Maatouk, Jarkass, Châtelet & Chebbo (2017) address the problem using a GA, then apply a local search algorithm (LS) in order to increase the searching speed of the optimal local solution building 4 different algorithms to compare their efficiencies in obtaining the optimal intervals to execute PM actions in a series-parallel system. On the other hand, Dahia, Z., Bellaouar, A., & Billel, S. (2018) prefer to use a model that considers a deterioration in the components of the system. The assumption that the machines remain “As Good As New” when performing PM activities is known as “perfect maintenance” and aims to reduce the complexity of the problem. However, authors such as Xia, Xi, Zhou & Lee (2012) emphasize that this assumption goes too far with reality. Considering this, Sachdeva, Kumar & Kumar (2008) developed a multi-objective model that seeks to maximize availability and minimize costs simultaneously for a paper production system through a GA and Simulink (Matlab). In Zhou, Xi & Lee (2009), seek to be more realistic, incorporating the duration of activities and the assumption of imperfect maintenance. Tambe et al. (2013) consider imperfect maintenance in a productive system, taking into account a limit time for performing maintenance. The authors demand certain availability in the system, adding...
This research proposes a model for planning a set of preventive maintenance activities for a single machine within a defined time horizon. The activities can be delayed or advanced in their executions to take advantage of the opportunity to make groupings or packages of activities. The model is more effective in its resolution than other models found in the literature, allowing grouping a higher amount of activities on a longer horizon. This highlight is achieved by formulating a continuous-time Mixed-Integer Linear Programming (MILP) model that allows finding the instants of time in which the activities are executed, maximizing the number of groups (or clusters) to be formed without having to discretize the entire time horizon.

2. Problem definition

There is a set of preventive maintenance activities to be performed on a machine. The execution instant of each activity is established by a previously defined frequency or periodicity, for instance, from the RCM application. It is necessary to plan the execution of the activities on the machine, taking into account energy production levels, seeking to minimize maintenance costs, transportation costs, duration of activities, and a penalty for overtime use of the human resource. Mishra, Shrivastava, and Vrat (2019) present a metaheuristic called “Jaya algorithm” that obtain optimal instants of execution of preventive and corrective maintenance activities in several components/equipment in series, looking for opportunistic groupings in order to reduce possible inefficiency costs. Another perspective, authors such as Hosseini, Kalam, Barker & Ramirez-Marquez (2019) consider that activities can be delayed and/or advanced in their executions in order to make groupings for reducing unproductive costs. The authors wish to know what activities should be delayed and/or advanced, taking into account a certain periodicity of execution from RCM or other tools. To do this, they seek to minimize costs, which consist of equipment inefficiency costs, penalty costs for the delay and/or advance of the activities, and costs for execution of activities. The authors propose a Greedy Heuristic-Based Local Search Algorithm (GHLSA) to find the global solution. Plus, more variables can be added to the problem as in Aquino, Chagas & Souza (2018), in which the authors add the technical team responsible for executing the maintenance in order to execute the largest number of PM activities minimizing the number of technical work teams to use.
be advanced and activity 2 delayed to achieve groups that reduce the total amount of interventions in the machine. In this way, if the activities are not grouped, there would be 4 interventions; however, the number of interventions is reduced to 2 when activities are grouped by their tolerances. There is a change in the decisions of future execution of activities in the planning horizon by delaying or advancing these tasks; therefore, the problem has the characteristic of being dynamic through time. For the problem formulation, the following notation is used:

**Parameters:**

- \( T \): Planning horizon for maintenance activities, finite number.
- \( I \): Number of activities to execute on a machine.
- \( J_i \): Number of executions of activity \( i \) on the horizon \( T \), \( \forall i \in \{1,2,\ldots,I\} \).
- \( N_i \): Conjunto de intervalos de tiempo posibles para ejecutar la actividad \( i \) en el horizonte \( T \), \( \forall i \in \{1,2,\ldots,I\} \).
- \( N \): Set of possible groups of activities on the horizon \( T \).
- \( T_i \): Periodicity of the activity \( i \), \( \forall i \in \{1,2,\ldots,I\} \).
- \( M \): Large enough number.
- \( e_i \): Tolerance execution of the activity \( i \), \( \forall i \in \{1,2,\ldots,I\} \).

**Decision variables:**

- \( t_{j,i} \): Instant of the \( j \)th execution of the activity \( i \), \( \forall j \in \{0,1,2,\ldots,J_i\}, \forall i \in \{1,2,\ldots,I\} \).
- \( s_{j,i,o,n} = 1 \) if the \( j \)th execution of the activity \( i \) is part of the grouping by the nth execution of the activity n.
- \( = 0 \) otherwise.
- \( s_{j,i,o,n} = 1 \) if the \( j \)th execution of the activity \( i \) is part of the grouping. \( \forall (j,i,o,n) \in N \).
- \( = 0 \) otherwise.

It is considered a machine with \( i \in I \) maintenance activities. Each activity needs to be executed with a fixed periodicity \( T_i \). Each activity has a \( e_i \) tolerance for being able to be advanced or delayed. In this way, every time it is required to execute the activity \( i \), there is a time window in which it can be executed within \([T_i - e_i; T_i + e_i]\), as seen in the Fig. 2. The execution of activity can only be advanced or delayed when it is possible to form a grouping with another activity (or activities); otherwise, it must be executed as its periodicity \( T_i \) indicates it. It is not necessarily an advance or delay in the execution of events for being grouped. Each activity is performed times during the planning horizon; therefore, the number of performances of the activity \( i \) is given by equation (1).

\[
J_i = 1,2,\ldots,\lfloor T_i/T_i \rfloor \forall i \in I \quad (1)
\]

Due to the possible delays or advances in the activities execution, the planning may differ from the planning horizon \( T \); however, the variation in the duration of this planning will depend fundamentally on the tolerance \( e_i \). In Fig. 3, \( J_i \) executions of the activity \( i \) are observed during the horizon \( T \), with \( J_i = 4 \).

If activity \( i \) is delayed due to grouping with some other activity, the planning horizon should be extended depending mainly on \( J_i \) and \( e_i \), represented in Fig. 4. In the figure, the activity \( i \) is delayed, therefore, grouped with some other, causing that the activity \( i \) to perform beyond what is stipulated by horizon \( T \).

In this way, the specific problem corresponds to the optimal planning of the executions of preventive maintenance activities in order to minimize inefficiency within an operational horizon \( T \). Each activity has a certain execution tolerance \( e_i \); therefore, the moments in which the activities are executed must be determined considering the groups formation. For the groupings to be justified, the costs associated with inefficiency must be significant. Due to the problem complexity, a series of assumptions will be made in order to simplify the number of variables involved and be able to solve it in reasonable computational times. Thus, we assume that there are sufficient resources to carry out independent activities, and can be grouped in any way. In the same way, duration of activities are not considered in the model in order to simplify the complexity of computational resolution; however, they can be analyzed after results. On the other hand, we consider a machine that operates under a regime of 24 hours a day, so that the machine must stop its operations to execute any activity. Several models in the literature discretize the planning horizon \( T \) since the combinatorial nature problem. So they look at each point of the discretize horizon by making thorough
comparisons. That is why the optimization problem is modeled through a linear-mixed (MILP) full-time continuous programming model, in such a way to find the instances of time in which activities are executed without having to go through each point discretized horizon. The continuous-time model is limited to looking only at the possible intervals of executions of the activities; in this way, groups of a higher number of activities are determined. However, despite the increase in comparative efficiency, the "combinatorial explosion" that occurs as the number of activities continues. In this context, sets $N_i$ and $N$ are defined and used to further improve the way of comparing instances of execution between activities and thus increase the number of activities for which the model can be solved in a reasonable amount of time.

The set $N_i$ contains the intervals of all possible executions of activity $i$, which is defined in equation (2). In this way, the ends of the intervals contain the maximum arrears and advances of the executions of the activity $i$, given by the respective tolerance considered $e_i$.

$$N_i = \{ [(T_i - e_i), (T_i + e_i)], [2(T_i - e_i), 2(T_i + e_i)], \ldots, [|J_i|T_i - |J_i|e_i], [|J_i|T_i + |J_i|e_i] \} \quad \forall i \in I \quad (2)$$

For example, if activity $i$ is overdue, it will be executed in $T_i - e_i$. If the next execution is overdue again, it will be executed in $2T_i - 2e_i$, and so on, obtaining the analogous when advancing. So then, the activity $i$ will be executed in a mandatory way within theses intervals since that higher delays or advances are not allowed, as seen in Fig. 5.

![Fig. 5. $N$ set construction](image)

As shown in Fig. 5. The first execution of activity $i$ will be between time instants $T_i - e_i$ and $T_i + e_i$. Then, since this execution of activity $i$ can be advanced to $T_i - e_i$, the second possible execution can be advanced, at most, to $2T_i - 2e_i$, and so on until $|J_i|$-nth execution. It should be noted that each time activity $i$ is executed, the next interval in set $N$ will have a greater amplitude. There is an analogous situation for the case in which the executions of the activity $i$ are delayed.

4. Problem formulation

Given a set $I$ of preventive maintenance activities with periodicities $T_i$ and tolerances $e_i$, where $i = 1, 2, \ldots, |I|$. The objective is the planning of activities on the $T$ horizon, minimizing interventions, and, consequently, the inefficiency of the machine. For this, the following assumptions are considered:

a. Duration of activities is not considered, that is, the machine is repaired instantly. This assumption is used to simplify the number of variables involved in solving the problem.

b. There are enough resources to carry out the activities individually and grouped in any way, so there is no limiting factor in performing any specific activities grouping. This assumption facilitates activity grouping. Besides, there are no infeasible solutions due to insufficient resources.

c. Maximizing the number of grouping maintenance activities is equivalent to minimizing the inefficiency of interventions. It is considered that, with the grouping of a higher number of activities at a given moment, the machine must be less intervened favoring a decrease in the unproductivity of the machine. The validity of this assumption depends fundamentally on the fulfillment of the assumption about the duration of activities.

The formulation of the MILP proposed model is presented as follows:

$$\max \sum_{j \in J} \sum_{i \in I} h_{i,j} x_{i,j} \quad (4)$$

$$t_{i,j} = 0 \quad \forall j \in J \quad (5)$$

$$t_{i,j} \leq t_{i,j-1} + T_i + e_i + f_i \quad \forall j \in J, i \in I \quad (6)$$

$$t_{i,j} \geq t_{i,j-1} + T_i - e_i - f_i \quad \forall j \in J, i \in I \quad (7)$$

$$t_{i,j} \leq t_{i,|J_i|} + H + \sum_{g=1}^{G_{i,j}} \sum_{k \in K} \sum_{l=1}^{L_{g,k,l}} \sum_{n=1}^{N_{1,l}} x_{i,j,g,k,l,n} + \sum_{k \in K} \sum_{l=1}^{L_{g,k,l}} \sum_{n=1}^{N_{2,l}} x_{i,j,g,k,l,n} + \sum_{k \in K} \sum_{l=1}^{L_{g,k,l}} \sum_{n=1}^{N_{3,l}} x_{i,j,g,k,l,n} \quad \forall j \in J, i \in I \quad (8)$$

$$t_{i,j} \geq t_{i,|J_i|} + H - \sum_{g=1}^{G_{i,j}} \sum_{k \in K} \sum_{l=1}^{L_{g,k,l}} \sum_{n=1}^{N_{4,l}} x_{i,j,g,k,l,n} + \sum_{k \in K} \sum_{l=1}^{L_{g,k,l}} \sum_{n=1}^{N_{5,l}} x_{i,j,g,k,l,n} + \sum_{k \in K} \sum_{l=1}^{L_{g,k,l}} \sum_{n=1}^{N_{6,l}} x_{i,j,g,k,l,n} \quad \forall j \in J, i \in I \quad (9)$$

$$\sum_{g=1}^{G_{i,j}} \sum_{k \in K} \sum_{l=1}^{L_{g,k,l}} \sum_{n=1}^{N_{3,l}} x_{i,j,g,k,l,n} \leq n_i \quad \forall j \in J, i \in I \quad (10)$$

$$t_{i,j} \leq t_{i,|J_i|} + M(1 - x_{i,j}) \quad \forall j \in J, i \in I \quad (11)$$

$$t_{i,j} \geq t_{i,|J_i|} + M(1 - x_{i,j}) \quad \forall j \in J, i \in I \quad (12)$$

$$t_{i,j} \leq t_{i,|J_i|} - M(1 - x_{i,j}) \quad \forall j \in J, i \in I \quad (13)$$

$$t_{i,j} \geq t_{i,|J_i|} - M(1 - x_{i,j}) \quad \forall j \in J, i \in I \quad (14)$$

$$x_{i,j} \in \{0, 1\} \quad \forall j \in J, i \in I \quad (15)$$

The objective function (4) seeks to maximize the number of elements that can be grouped, and consequently, the number of groups. The constraint (5) is a border constraint which
indicates that no activity has been executed; hence, all activities will be executed at the instant equal to 0. (6), (7), (8) and (9) detail the mobility of the activity execution, that is, in case of no grouping. The activity execution will depend only on its frequency; otherwise, the activity may move within its tolerance. (10) Restricts the set of activities that will be grouped with a particular activity; so there is an activity that forms the grouping while other activities belong to it. (11) details that in order for an activity to form a grouping, there must be a specific activity grouped into it. (12) And (13) show that the activities belonging to a group must be performed simultaneously. Finally (13) and (14) declare the nature of the decision variables.

Results

The resolution of the MILP model presented in the previous section was implemented in the Python programming language using the Pyomo library. The solver used corresponds to Gurobi, and the case studies run on a computer of @ 2.7 GHz 2.9 GHz with 16 GB of RAM. The resolution performance of the model depends on the planning horizon, the tolerances considered, and the activity with less periodicity of execution. This last factor is more shocking on the resolution times since it is a higher number of executions of an activity, the greater will be the number of comparisons made to find the optimal solution.

A series of experiments were performed to test the model and analyze the impact on the reduction of inefficiency. The experiments focus on finding the optimal semiannual planning of 5 preventive maintenance activities (PM) on a continuously functioning machine (it works 24/7) by looking for time instants in which they must be executed to stop the equipment operation as little as possible. The experiments consider 3 possible cases:

Case I: The execution periodicities of the activities are multiples of each other.
Case II: The execution periodicities of the activities are not multiples of each other.
Case III: The combination of the previous cases.

These cases seek to simulate output information after some methodology application as RCM in order to use it as an input for the model. The model seeks to reach the scenario in which the intervals of the execution of the activities are multiples of each other, as it occurs in case I. In Fig. 6 it is observed how Case I is optimized since inefficiency cannot be more reduced because all activities have multiplicity between their periodicities. The model allows to obtain benefits in the most realistic case (case III) as the best case (case II); however, this decrease is affecting the tolerance considered. This tolerance was defined given a percentage of the periodicity of the activities, but it can be defined for any positive real number. It should be noted that tolerance can be incorporated that takes different values in each execution of the activities; that is, tolerance can be different in each instant of time.

Conclusions

Undoubtedly, having adequate maintenance management has become one of the fundamental tasks of organizations with a high endowment of assets; these organizations are continuously seeking better performance at the lowest possible cost. In this sense, the correct design of a maintenance plan for its consequent process of planning, programming and execution, should be understood as a strategic task within the maintenance management and the assets of the companies. Different methodologies such as RCM support the design process of maintenance policies; however, they do not necessarily address the problem of future planning and programming of recommended maintenance activities. Besides, some criteria allow the
creation of a maintenance plan, such as the recommendation of the manufacturers, benchmarking, experience of the technical suppliers, etc. On the other hand, there are criteria defined in the scientific literature incorporated through mathematical models that allow for optimal maintenance plans, such as maintaining some levels of reliability and availability required in the systems or reducing costs due to the unproductivity of the components. These models are not trivial to implement or solve due to the number of variables involved and the problem complexity to be defined. This paper aims to develop an optimal maintenance plan using as criteria a certain level of tolerance for the execution of activities already defined in the maintenance policy (input) through the use of a mathematical optimization model. The model considers a set of PM activities from the RCM application or other methodologies in order to generate work packages or groupings of tasks in order to reduce the number of stops on the machine. For this, a MILP was developed to achieve rapid solutions considering only three input parameters: planning horizon, periodicities, and tolerances of the activities.

References


