

OPTIMAL PREDICTIVE MAINTENANCE POLICY FOR MULTI-COMPONENT SYSTEMS

Tiffany Cherchi ^{1,2}, Camille Baysse ¹, Benoîte de Saporta ² & François Dufour³

¹ *Thales, France. E-mail : tiffany.cherchi@umontpellier.fr ; camille.baysse@fr.thalesgroup.com*

² *IMAG, Univ Montpellier, CNRS, Montpellier, France. E-mail : benoite.de-saporta@umontpellier.fr*

³ *INRIA CQFD, IMB, Univ Bordeaux, Bordeaux INP, CNRS, France. E-mail : francois.dufour@math.u-bordeaux.fr*

Résumé. Nous présentons un problème d’optimisation pour la maintenance d’un système multi-composants sujet à des détériorations ou défaillances aléatoires de ses composants, entraînant l’évolution de l’état général du système. Le système peut être requis pour effectuer des missions de fréquences et durées déterministes. Notre objectif à long terme est de mettre en place une politique de maintenance optimisée afin de garantir le bon déroulement des missions tout en minimisant les coûts de maintenance. L’idée principale de ce travail est de proposer un modèle mathématique pour l’évolution du système en utilisant le formalisme d’un Processus Markovien Décisionnel (MDP). Par simulations de Monte Carlo, nous comparons les performances de plusieurs politiques de référence.

Mots-clés. Optimisation de maintenance, système multi-composants, Processus Markovien décisionnel, Simulations de Monte Carlo, Politique de maintenance . . .

Abstract. We present an optimization problem for the maintenance of a multi-component system subject to random deteriorations or failures of its components, resulting in the evolution of the overall system’s state. The system can be required to perform missions with deterministic occurrences and durations. Our long term objective is to set up an optimized maintenance policy to ensure the smooth running of missions while minimizing maintenance costs. The main idea of this work is to propose a mathematical model for the evolution of the system by using the formalism of a Markov Decision Processes (MDP). Through Monte Carlo simulations, we compare the performances of several reference policies.

Keywords. Maintenance Optimization, Multi-component system, Markov decision process, Monte Carlo simulation, Maintenance policy. . .

1 Introduction

Cet article étudie un problème d’optimisation pour la maintenance d’un système multi-composant sujet à des détériorations ou pannes aléatoires de ses composants. Plus spécifiquement, nous étudions la dynamique d’un système à trois composants pouvant être requis pour effectuer des missions de fréquences et durées déterministes. Chaque composant est sujet à des détériorations ou défaillances aléatoires entraînant l’évolution de l’état global du système. Notre objectif à long terme est de calculer une politique de maintenance optimisée pour assurer le bon déroulement des missions tout en minimisant les coûts de maintenance.

Nous modélisons ce problème d’optimisation de maintenance via le formalisme d’un processus markovien décisionnel (MDP). Cette famille de processus stochastiques est bien adaptée à

la modélisation de problèmes de décision séquentiels dans l'incertain, voir [HLL96], [Put94] ou [BR11]. Typiquement, à chaque pas de temps (*p. ex.* quotidien ou à chaque mission), étant donné l'état actuel du processus (*p. ex.* état des composants, durée de vie restante des composants, ...), le contrôleur choisit une action. Cette action a une influence sur la transition vers l'état suivant et génère un coût (*p. ex.* coût des opérations de maintenance effectuées, pénalité d'indisponibilité, pénalité pour les missions échouées, ...) en fonction de l'état initial et de l'action choisie. Une suite d'actions s'appelle une politique. Chaque politique génère un coût total pour toute la trajectoire jusqu'à l'horizon de contrôle. Les principales quantités d'intérêt sont la *fonction valeur* correspondant au coût minimal sur toutes les stratégies admissibles et la *stratégie optimale* réalisant ce minimum. Dans le cas où l'espace d'états est fini (et pas trop grand) et que l'intégration contre le noyau de transition du MDP est réalisable analytiquement, l'algorithme de programmation dynamique fournit à la fois la fonction valeur et une stratégie optimale.

Le présent article porte sur la simulation du MDP correspondant à notre problème de maintenance et sur l'évaluation du coût généré par les politiques de référence. Le calcul de la fonction valeur et la présentation d'une stratégie optimale ne sont pas possibles dans le cas que nous étudions, car l'espace d'états est continu et le noyau de transition n'est pas explicite. Ce sera l'objet de futurs travaux.

Concernant la maintenance de systèmes industriels, il existe une littérature assez importante et basée sur des modèles aléatoires. Deux familles de problèmes sont principalement étudiées.

La première famille concerne l'optimisation de familles spécifiques de politiques paramétrées. Par exemple, on peut vouloir optimiser une durée fixe entre les inspections, voir [ZFA17] ou un seuil pour l'accumulation de dommages avant de déclencher une intervention, voir [MC19].

La deuxième famille, dans laquelle s'inscrit l'objectif à long terme de ce papier, concerne l'optimisation globale de toute une classe de règles de maintenance non paramétriques, voir *p. ex.* [dSDZE12]. Dans ce dernier article, les auteurs proposent une méthode numérique pour optimiser la maintenance d'une structure métallique soumise à la corrosion. La classe de stratégies de maintenance qu'ils considèrent sont les stratégies dans lesquelles une seule intervention peut avoir lieu et la structure est ensuite remplacée par une nouvelle. Dans ce papier, nous considérons un problème plus général car nous autorisons plusieurs interventions sur le même équipement sans le remplacer par un nouveau.

Ce papier est organisé comme suit. La section 2 est consacrée à la description et à la modélisation du système physique étudié. Dans la section 3, nous comparons les performances de plusieurs politiques de référence et commentons les résultats numériques obtenus. Enfin, dans la section 4, une conclusion et des perspectives sont présentées.

2 Problème d'optimisation de maintenance

Dans cette section, nous présentons la modélisation du problème de maintenance considéré. Nous décrivons d'abord le système physique et les hypothèses liées aux phénomènes de dégradation et de maintenance. Ensuite, nous expliquons comment ce système physique peut être modélisé dans le formalisme des MDP.

2.1 Modélisation du système physique et hypothèses

Nous étudions la dynamique d'un système formé de trois composants pouvant être requis pour effectuer des missions de fréquences déterministes T_f et de durées T_m sur un horizon de temps fini N . Nous considérons que les trois composants du système sont non-identiques, avec des processus de détérioration ou de défaillance indépendants, et tous les trois sont nécessaires au bon fonctionnement du système.

Chaque composant du système est sujet à des détériorations ou pannes aléatoires. Les composants peuvent être dans trois états possibles (ou régimes, ou modes) : stable, dégradé ou en panne. Les modes de chaque composant déterminent l'état global du système. Nous dirons que le système est dans un état stable si tous ses composants sont dans l'état stable, en panne si au moins un de ses composants est en panne, et dégradé sinon.

Au cours du temps, un composant i passe d'un état stable à un état dégradé selon une distribution de Weibull de paramètres (β_i, α_i) . Ensuite, il passe d'un état dégradé à un état de panne selon une distribution exponentielle de paramètre λ_i . Lorsque le système n'est pas en fonctionnement, nous dirons qu'il est au repos et qu'il ne peut ni se dégrader ni tomber en panne. Le système peut fonctionner dans un état dégradé, mais il est arrêté dès la première panne, de sorte que seuls 20 états sont possibles pour le système. Ainsi, en l'absence d'action de maintenance, le système passe à l'état de panne et y reste.

Pour chaque composant, deux actions de maintenance sont possibles pour prolonger sa durée de vie. Ces interventions sont soit un entretien, soit un remplacement. Dans les états stable et dégradé, des actions d'entretien et de remplacement peuvent être effectuées sur les composants, mais en cas de panne, l'action d'entretien n'est pas possible. Dans les deux cas, les composants sont considérés comme neufs à la suite d'une intervention et leurs temps de fonctionnement sont réinitialisés à 0.

Nous considérons qu'une visite à l'atelier nécessite l'immobilisation de tout le système, générant un coût fixe c_w , auquel s'ajoute un coût de remplacement c_r et/ou d'entretien c_e en fonction des actions effectuées sur chaque composant du système. On suppose que les opérations de maintenance peuvent être effectuées en parallèle. Le temps passé dans l'atelier T_w est donc fixé quelles que soient les actions de maintenance effectuées. Le système est donc indisponible pendant ce temps. De plus, une pénalité d'indisponibilité c_u doit être payée pour chaque jour que l'équipement passe dans l'état de panne, ainsi qu'une pénalité c_f pour les missions échouées.

Une politique de maintenance est une suite d'actions prises à des moments discrets (par exemple, quotidiennement ou à chaque mission) et tenant compte de l'état actuel du système (par exemple, l'état des composants, la durée de vie restante des composants, etc.). Nous ne tiendrons compte que des politiques dans lesquelles chaque visite à l'atelier ramène l'équipement à l'état stable. Chaque politique génère un coût correspondant aux actions de maintenance effectuées, une pénalité d'indisponibilité lorsque le système est en panne, une pénalité pour les missions ayant échoué. L'objectif à long terme de ce travail est de calculer une politique de maintenance optimisée afin d'assurer le bon déroulement des missions tout en optimisant le coût global sur un horizon de temps fini. L'objectif à court terme est de simuler l'évolution du système et d'évaluer le coût généré par certaines politiques de référence.

2.2 Processus Markovien Décisionnel

Pour résoudre notre problème de maintenance, nous proposons de modéliser notre système par un processus markovien décisionnel (MDP). Cette famille de processus stochastiques est bien adaptée à la modélisation de problèmes de décision séquentiels dans l'incertain, voir [HLL96]. Considérons un système qui peut être contrôlé par des décisions séquentielles. Un MDP est la suite de variables aléatoires décrivant l'évolution stochastique de l'état du système. Bien entendu, la distribution des états dépend des actions choisies.

Un processus markovien décisionnel est défini par les paramètres suivants :

- Un espace d'états \mathbb{X} qui décrit les états possibles du système que l'on cherche à modéliser ;
- Un espace d'action \mathbb{A} ainsi qu'une famille de sous ensemble $\mathbb{A}(x)$ de \mathbb{A} représentant l'ensemble des actions réalisables lorsque le système se trouve dans l'état x ;
- Un noyau de transition markovien $Q(dy|x, a)$ qui permet de calculer les transitions aléatoires du système lorsque celui-ci se trouve dans l'état x et que l'action a a été choisie ;
- Une fonction coût c à valeurs réelles dépendant du couple état, action (x, a) ;

A partir de ces paramètres, la dynamique contrôlée du système est alors définie de la façon suivante. Partant d'un point initial x_0 de \mathbb{X} :

- (1) le contrôleur observe l'état du système $x_0 \in \mathbb{X}$ et choisit une action $a_0 \in \mathbb{A}(x_0)$;
- (2) cela génère un coût $c(x_0, a_0)$;
- (3) une transition aléatoire est induite pour le système vers un nouvel état $x_1 \in \mathbb{X}$ selon la distribution de probabilité donnée par le noyau stochastique $Q(dy|x_0, a_0)$.

Cette procédure est ensuite répétée de manière itérative pour générer une trajectoire du processus en termes d'une suite d'états et d'actions (x_n, a_n) . L'objectif du contrôleur est alors de trouver des actions optimales en fonction d'un critère d'optimalité donné, basé sur la fonction de coût c . Typiquement, une politique π de contrôle est une suite $\pi = \{f_n\}_{n \geq 0}$ de fonctions $f_n : \mathbb{X} \rightarrow \mathbb{A}$ permettant d'associer au moment $n \geq 0$ une action à exécuter lorsque le système est dans l'état $x \in \mathbb{X}$ et tel que $f_n(x) \in \mathbb{A}(x)$ pour tout n . L'ensemble de toutes les politiques admissibles est noté Π .

Dans notre cas, l'objectif est d'optimiser un critère d'optimalité basé sur la fonction coût c . On suppose ici que l'agent doit contrôler le système en N étapes, avec N fini. Le critère à horizon fini est intuitivement l'espérance de la somme des coûts générés en partant de l'état initial $x \in \mathbb{X}$ et suivant la politique $\pi \in \Pi$:

$$V_N(\pi, x) = \mathbb{E}_x^\pi \left[\sum_{n=0}^N c(x_n, a_n) \right].$$

Le problème de contrôle optimal associé à un processus markovien décisionnel consiste à minimiser la fonction $\pi \rightarrow V_N(\pi, x)$ sur Π . L'optimum est appelé *fonction valeur* et est donné par

$$V(x) = \inf_{\pi \in \Pi} V_N(\pi; x).$$

Une stratégie $\pi^* \in \Pi$ est dite optimale si elle vérifie

$$V_N(\pi^*, x) = V(x).$$

3 Résultats numériques

Cet article se concentre sur la simulation d'un MDP correspondant à notre problème de maintenance et sur l'évaluation du coût engendré par des politiques de référence. L'objectif de cette section est de fournir les paramètres du modèle considéré ainsi que les valeurs numériques de la fonction de coût associée à différentes stratégies. Plus spécifiquement, nous étudions la dynamique d'un système à trois composants.

3.1 Modèle MDP du système

Étant données les hypothèses de modélisation de la section 2, les paramètres du système de notre MDP sont les suivants :

- Un espace d'état $\mathbb{X} = \{x = (e_i, r_i); e_i \in \{\text{stable, dégradé, panne}\}, r_i \in \mathbb{R}^+\}$ où e_i est une variable discrète décrivant l'état du composant i et r_i est une variable continue donnant la durée de vie utile restante du composant i jusqu'à son prochain changement d'état.
- L'espace d'action donnant l'ensemble des contrôles $\mathbb{A} = \{a = (a_1, a_2, a_3)\}$, où a_i dans $\{\text{rien, entretenir, remplacer}\}$. L'ensemble $\mathbb{A}(x)$ est tel qu'un composant en panne ne peut pas être réparé et qu'après une intervention en atelier, tous les composants sont dans un état stable.
- Le noyau de transition $Q(dy|x, a)$ n'est pas explicite analytiquement, il est seulement simulable selon la procédure donnée dans la section 2. Les valeurs des différents paramètres de fiabilité sont données dans la Table 1.
- La fonction de coût c dépend des actions de maintenance éventuellement effectuées et des pénalités potentielles si le système est en état de panne. On suppose que $c_r < c_c$ puisque le remplacement correctif est généralement plus complexe et plus coûteux que l'entretien préventif, avec $c_r = 50$ et $c_c = 150$. De plus, on suppose que $c_u < c_f$ puisque l'échec d'une mission cause plus de désagréments industriels que l'indisponibilité, avec $c_u = 500$ et $c_f = 750$.

Le calcul de la fonction valeur $V(x)$ et d'une stratégie optimale ne sont pas simples ici car l'espace d'état \mathbb{X} est continu et le noyau de transition n'est pas explicite. Ce sera l'objet de futurs travaux.

La première étape pour résoudre notre problème consiste à implémenter un simulateur de ce processus. Les paramètres du modèle sont fixés arbitrairement comme suit. L'horizon temporel est de $N = 50$ jours, la fréquence des missions est quotidienne $T_f = 1$ et leur durée est de $T_m = 20$ heures. De plus, le délai nécessaire pour envoyer le système à l'atelier est de $T_d = 1$ jour et le temps passé dans l'atelier est de $T_w = 4$ jours.

α_1	160	α_2	140	α_3	175
β_1	1.4	β_2	1.9	β_3	1.7
λ_1	1/45	λ_2	1/40	λ_3	1/35

TABLE 1 – Paramètres de fiabilité

3.2 Politiques de références

Nous introduisons trois politiques de référence.

1. *Politique sans intervention* π_1 : On ne fait aucune intervention de maintenance (ni remplacement, ni entretien) pendant toute la période étudiée.
2. *Politique de maintenance corrective* π_2 : Après 1 jour passé en état de panne, le système est envoyé à l'atelier. Chaque composant en état de panne est remplacé, et chaque composant dégradé est entretenu.
3. *Politique de maintenance préventive* π_3 : Après 1 jour passé en état de panne ou dégradé, le système est envoyé à l'atelier. Chaque composant en état de panne est remplacé, et chaque composant en état dégradé est entretenu.

Les trois politiques peuvent être simulées de façon exacte.

3.3 Résultats numériques

Grâce à des simulations exactes, nous comparons les performances de ces politiques de référence. Leur coût a été évalué au moyen de 10^5 simulations Monte Carlo et les résultats numériques sont donnés dans la Table 2.

Politique	coût	95% IC
π_1	22892	[22884, 22900]
π_2	18134	[18121, 18147]
π_3	15435	[15423, 15447]

TABLE 2 – Coûts des politiques de référence

Comme attendu, une stratégie de maintenance préventive π_3 réduit efficacement les coûts de maintenance en intervenant sur le système avant la défaillance. Cela génère un gain relatif par rapport à la stratégie non contrôlée π_1 de 33 % et un gain relatif de 15 % par rapport à la stratégie corrective π_2 .

4 Conclusions et perspectives

Dans ce papier, nous avons considéré un problème de maintenance pour un système à plusieurs composants. La dynamique du système a été modélisée par un processus markovien décisionnel. Nous avons simulé le processus sous différentes stratégies de maintenance corrective ou préventive de référence et comparé leurs coûts. Par simulations Monte-Carlo, nous avons montré qu'une politique de maintenance préventive réduisait efficacement les coûts de maintenance en intervenant sur le système avant la défaillance.

La prochaine étape de ce travail est de calculer le coût et la stratégie optimale sur l'ensemble Π des stratégies admissibles. Il s'agit d'un problème d'optimisation non standard dans le cadre des MDP car le noyau de transition n'est pas explicite analytiquement, il est seulement simulable et l'espace d'état n'est pas fini. Ainsi, les techniques d'optimisation standard pour les MDP telles que la programmation dynamique et ses extensions comme l'algorithme d'itération de la stratégie et l'algorithme d'itération de la valeur ne s'appliquent pas. L'étape suivante vers la résolution du problème d'optimisation global consistera à discrétiser l'espace d'états. Les coûts des politiques de référence seront utilisés pour évaluer l'impact de la discrétisation sur le calcul des coûts. Il doit y avoir un compromis entre précision, conduisant à un espace d'états avec une cardinalité élevée et une complexité numérique conduisant à un espace d'états le plus petit possible. Sur cet espace d'états finis, le noyau de transition ne sera toujours pas explicite, et nous utiliserons des algorithmes d'optimisation basés sur la simulation pour résoudre ce problème. Ce sera l'objet de futurs travaux.

Références

- [BR11] Nicole Bauerle and Ulrich Rieder. *Markov Decision Processes with Applications to Finance*. Universitext. Springer, Heidelberg., 01 2011.
- [dSDZE12] Benoîte de Saporta, François Dufour, Huilong Zhang, and Charles Elegbede. Optimal stopping for the predictive maintenance of a structure subject to corrosion. *Proceedings of the Institution of Mechanical Engineers, Part O : Journal of Risk and Reliability*, 226(2) :169–181, 2012.
- [HLL96] Onésimo Hernández-Lerma and Jean Bernard Lasserre. *Discrete-Time Markov Control Processes : Basic Optimality Criteria*. volume 30 de Applications of Mathematics. New-York : Springer-Verlaga, 1996.
- [MC19] Sophie Mercier and Inmaculada T. Castro. Stochastic comparisons of imperfect maintenance models for a gamma deteriorating system. *European Journal of Operational Research*, 273(1) :237–248, 2019.
- [Put94] Martin L. Puterman. *Markov Decision Processes : Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York, NY, USA, 1st edition, 1994.
- [ZFA17] Nan Zhang, Mitra Fouladirad, and Barros Anne. Maintenance analysis of a two-component load-sharing system. *Reliability Engineering & System Safety*, 167 :67 – 74, 2017.