



Projet IMdR 1/2004

Comment optimiser les collectes de données pour conduire son retour d'expérience ?

Chef de Projet :
Monsieur François BILLY (EDF)

Contractant :
Société SECTOR et l'Université Pierre-Mendès-France de Grenoble

Le projet n° 1/2004 de l'IMdR intitulé « Comment optimiser les collectes de données pour conduire son retour d'expérience » a été réalisé en 2006. Les souscripteurs qui y ont participé sont le CEA/CESTA, la DGA/LRBA, EADS, EDF, l'IMdR et la SNCF.

De nombreuses expériences de retours d'expérience (REX) dans divers secteurs d'activités ont rencontré des difficultés qui ont compromis leur réussite. Par exemple, la collecte de données, qui initie le processus de REX, peut être source d'échecs par diminution progressive des déclarations ou au contraire par augmentation exponentielle des déclarations engorgeant le système d'analyse.

Ce projet a pour objet de proposer des recommandations afin d'assurer des conditions de réussite de la collecte de données. Pour cela, ce projet s'articule autour des quatre tâches suivantes :

- tâche 1 : Recherche bibliographique
- tâche 2 : Prise en compte des problèmes définis par les souscripteurs
- tâche 3 : Optimisation de la collecte
- tâche 4 : Illustration par des exemples

Tâche 1 : Recherche bibliographique

La première tâche de ce projet a consisté à réaliser un état de l'art sur la collecte de données à partir d'une recherche bibliographique parmi les ouvrages de référence et les publications des colloques spécialisés. Treize axes de caractérisation de la collecte de données en ressortent. Ces axes, fortement interdépendants, vont constituer des directions d'optimisation de la collecte. Ces axes sont les suivants :

- définition des objectifs du REX,
- définition de l'échantillon à observer,
- définition des paramètres à collecter,
- prise en compte des opérateurs,
- qualité des données collectées,
- outil de collecte,
- analyse des coûts de collecte,
- REX négatif / positif,
- collecte du facteur humain,
- intégration du REX dans l'entreprise,
- REX comme processus vivant,
- évaluation / validation des informations collectées,
- mise en place de la collecte.

Pour chaque axe, le projet s'est efforcé d'explicitier les apports de la bibliographie consultée dans la perspective de l'optimisation de la collecte de données.

Tâche 2 : Prise en compte des problèmes définis par les souscripteurs

Deux souscripteurs ont proposé leur projet de REX et de collecte de données pour illustrer ce projet :

- EDF-EGD : REX en ligne

L'objectif principal de ce projet de collecte est pour EDF-EGD de suivre la fiabilité des équipements installés vis-à-vis des spécifications contractuelles avec les fournisseurs. L'optimisation porte sur le nombre de régions « panels » qui devront suivre de manière fine les matériels implantés.

- EADS ST : ponts roulants

L'objectif principal de ce projet de collecte est pour EADS ST de suivre la disponibilité d'ateliers, de justifier la fiabilité des freins de ces ponts et d'en déduire les pratiques de maintenance. L'optimisation porte sur l'intérêt que pourrait représenter un REX sur des matériels similaires dans une autre industrie.

Tâche 3 : Optimisation de la collecte

Cette tâche centrale du projet consiste à proposer des recommandations s'appliquant à la définition et à l'évolution du processus de collecte de données dans un but d'optimisation. Les principaux critères d'optimisation sont la simplification de la collecte, la qualité des données collectées, la définition d'un échantillon minimal, l'amélioration de la représentativité des résultats et la diversification des utilisations des données collectées.

L'optimisation d'une collecte ne peut être réalisée qu'au travers de la maîtrise de nombreuses problématiques. Pour la clarté de la démarche, les quatre problématiques suivantes ont été retenues :

- la problématique de définition de la collecte,
- la problématique de conception d'un nouveau système de collecte,
- la problématique d'organisation de la collecte,
- la problématique d'une collecte en vue de la prolongation d'exploitation des matériels.

En termes de définition, l'optimisation de la collecte consiste à suivre les treize axes de caractérisation définis lors de la recherche bibliographique afin de valider que toutes les composantes de la collecte ont été développées. Les principales recommandations qui en ressortent sont :

- l'utilisation des outils de l'analyse de risques,
- la prise en compte des opérateurs de la collecte dès cette phase de définition,
- la définition d'un langage commun,
- la connaissance du patrimoine,
- la responsabilité du management.

La problématique de conception est centrée d'une part sur l'optimisation de la taille de l'échantillon qui fera l'objet de la collecte, et d'autre part sur l'analyse des coûts.

Concernant la taille d'échantillon, deux méthodes se présentent en fonction de l'importance du patrimoine à observer : les méthodes basées sur la maximisation de la vraisemblance et les méthodes bayésiennes.

Dans le cas d'un patrimoine suffisamment important permettant un échantillonnage, la taille minimale d'un échantillon peut être définie par l'utilisation des méthodes statistiques basées sur la maximisation de la vraisemblance ou sur l'algorithme Expectation-Maximisation (EM).

Dans le cas d'un patrimoine faible, l'approche bayésienne d'analyse des données permet la valorisation des données issues de la collecte (en faible quantité) par leur combinaison avec des données extérieures (avis d'experts, autres REX). Cette approche permet la justification de

résultats plus proches de la réalité que par l'analyse « classique » des seules données de la collecte.

Les coûts de la collecte doivent être évalués dans le cadre d'une approche de coût global de possession sur le long terme. Les gains issus du REX sont de différents types (accroissement de fiabilité, de sécurité, diminution des coûts de maintenance, ...).

La problématique de l'organisation de la collecte est traitée autour des trois aspects principaux :

- Les opérateurs
Ils seront généralement chargés de la collecte. Leur rôle est donc primordial pour la réussite de la collecte. Ils doivent être associés dès le début au projet de REX. Ils doivent faire l'objet d'une attention particulière et continue de la part des responsables du REX. Tout particulièrement, ils doivent être régulièrement informés des retombées positives du REX sur les performances de l'entreprise.
- L'implication du management
Le management a la responsabilité de la réalisation du REX. Son implication doit être constante et continue tout au long du projet de REX. Tout particulièrement, le management doit organiser le partage des informations sur les enseignements du REX.
- La communication
Un processus de REX induit, au sein de l'entreprise, de nombreux flux de communication qu'il convient d'animer en permanence.

La problématique de prolongation d'exploitation des matériels peut avoir un impact considérable sur une collecte de données qui ne l'a pas anticipé. En effet, une décision de prolongation d'exploitation de matériels repose sur de nombreuses informations caractérisant aussi bien l'évolution de la fiabilité et de la disponibilité que celles de la maintenance et de ses coûts.

Alors que les matériels « peu fiables » peuvent donner lieu à des modélisations classiques (par exemple l'analyse de Weibull), les matériels « très fiables » nécessitent une collecte orientée vers la connaissance des dégradations et des maintenances, surtout préventives.

L'optimisation de la collecte doit notamment reposer sur l'analyse de criticité des matériels considérés comme importants pour la sécurité ou pour la rentabilité, sur la caractérisation du vieillissement et sur l'identification fine des informations à collecter.

Tâche 4 : Illustration par des exemples

L'exemple proposé par EDF-EGD est le support de l'illustration de la définition de la taille minimale d'un échantillon permettant l'obtention du résultat souhaité (suivi de fiabilité dans ce cas). La modélisation propose, en incrémentant la taille de l'échantillon observé, d'estimer les résultats du REX qui seraient obtenus. L'intérêt de la modélisation présentée est de considérer que des événements peuvent potentiellement être déclarés sur les matériels en dehors des régions panels. Deux méthodes d'analyse sont alors comparées : la méthode « classique » de maximum de vraisemblance qui prend en compte toutes les déclarations sans distinction et la méthode itérative basée sur l'algorithme Expectation-Maximisation (EM) qui permet de considérer que certaines données sont issues en dehors des régions panels. L'approche EM permet la définition d'un échantillon plus petit pour le même niveau de résultat.

D'une manière générale, cet exemple montre que la prise en compte des informations incomplètes issues de l'observation des matériels en dehors des régions panels améliore nettement les résultats de l'analyse des données et permet donc en corollaire de diminuer le nombre de régions panels suivant finement les matériels.

L'exemple proposé par EADS ST est le support de l'illustration de l'approche bayésienne. Les fiabilités objectif des matériels (ponts roulants) exploités par EADS ST et leurs conditions d'exploitation font que la probabilité d'apparition d'une défaillance des freins sur plusieurs années d'observation est extrêmement faible. Dans ces conditions, une approche « classique » de

l'analyse des données n'est pas en mesure de fournir des résultats exploitables : l'estimation du taux de défaillance, par exemple, est soit égale à zéro (non pertinent) soit très largement surestimée.

Des résultats d'un REX sur des matériels similaires existants peuvent être utilisés avec profit par une approche bayésienne. Tout particulièrement une telle approche pourrait permettre de justifier une fiabilité des freins de l'ordre de $10^{-7}/h$ à $10^{-6}/h$ alors que les méthodes classiques limitées au seul REX d'EADS ST conduiraient à un ordre de grandeur de $10^{-5}/h$.

D'une manière générale, cet exemple montre tout l'intérêt de tenir compte du REX sur des matériels similaires dans d'autres industries afin d'optimiser sa propre collecte de données.

Conclusion et perspectives

Les recommandations issues du projet s'appliquent idéalement à la phase de conception d'un processus de collecte de données c'est-à-dire dès la phase de spécification des besoins. Pour autant, ces recommandations peuvent être également prises en compte en cours de vie du processus, tout particulièrement dans toutes ses évolutions.

Les perspectives de poursuite de ce projet sont multiples. Citons de manière limitative :

- la collecte de données concernant spécifiquement le facteur humain et son optimisation.
- le cahier des charges des spécifications d'un outil informatique pour une collecte optimisée.
- l'optimisation de l'analyse de données collectées dans le cadre d'un REX.
- le développement de la modélisation présentée dans ce projet pour le calcul de la taille minimale d'un échantillon.
- l'approche bayésienne (appliquée à l'analyse du facteur humain par exemple).