

Les techniques de fouille de données appliquées au diagnostic des systèmes électriques

G. Rostaing¹, F. François² et J.P. Rognon³

¹ ENSPG

² LEG

³ ENSIEG

COMMUNICATION DANS LES JOURNEES DU CLUB EEA

« Diagnostic en Electronique, Electrotechnique et Automatique »

17-18 Mars 2005 Lyon

Les techniques de fouille de données appliquées au diagnostic des systèmes électriques.

Gilles Rostaing, Florence François, Jean Pierre Rognon

Gilles Rostaing : Professeur agrégé à l'ENSPG chargé des enseignements d'électronique, instrumentation et informatique industrielle, docteur de l'INPG, chercheur au laboratoire d'électrotechnique de Grenoble (LEG) au sein de l'équipe Syrel sur le thème de la surveillance des systèmes électriques.

Florence François : Ingénieur au laboratoire d'électrotechnique de Grenoble, docteur de l'INPG et chargée des applications liées aux méthodes d'intelligences artificielles au sein de l'équipe Syrel.

Jean Pierre Rognon : Professeur à l'ENSIEG, Chercheur dans l'équipe Syrel sur le thème de la surveillance des systèmes électriques.

Résumé :

La surveillance et le diagnostic nécessitent une bonne connaissance à la fois fonctionnelle et comportementale du système à étudier. De la formalisation de cette connaissance dépend la méthode à employer.

De manière générale, l'expertise du fonctionnement sain est possédée par le constructeur, les données de maintenance sont connues de l'exploitant, mais ces données sont rarement disponibles dans leur intégralité, ce qui rend la conception de systèmes de diagnostic relativement complexe.

La méthode proposée dans cet article vient du constat suivant : « un système vieilli en fonction de ses conditions de fonctionnement ». Il est donc intéressant d'analyser le « vécu » d'un système afin de mettre en évidence les « scénarii » conduisant à une situation de défaillance. Pour cela, il faut organiser l'acquisition de données afin de classer les états de fonctionnement, d'établir un historique et de déterminer l'évolution future de l'état du système. Cette « base de données » des scénarii à risque permet donc d'anticiper l'occurrence des défaillances.

Mots clefs : diagnostic, disponibilité, maintenance prédictive, data mining, fouille de données, réseaux Bayésiens, réseaux de neurones.

1. Introduction

Les systèmes électriques sont des systèmes complexes, l'instrumentation y est peu présente si ce n'est pour des besoins de contrôle. Le diagnostic y est traditionnellement absent car ces systèmes sont, de manière générale, relativement robustes.

On peut donc se demander l'utilité de la surveillance de tels systèmes qui, s'ils sont bien construits et bien utilisés ont des durées de vie importantes. La raison en est simple, ces hypothèses sont fausses dans bien des cas :

- Le constructeur est un assembleur, il est possible que dans ce cas, les aspects économiques ou stratégiques d'un marché fassent que la fiabilité de l'ensemble soit compromise par des composants de moindre qualité, ou par des spécifications moins contraignantes.
- Le changement de technologie chez un fournisseur ou le changement de fournisseur (seconde source) peu engendrer des problèmes de fiabilité (durée de vie de l'équipement modifiée).
- Le constructeur ne maîtrise pas l'exploitation : le matériel est spécifié pour un type d'utilisation or rien ne garanti que l'utilisation sur le site ne se modifiera pas au fil des ans (exemple des armoires électriques et des protections).
- Le constructeur ne maîtrise pas la maintenance qui est réalisée par l'exploitant ou sous-traitée, il ne peut donc pas suivre l'évolution de son matériel et en assurer la disponibilité.
- L'exploitant veut disposer d'heures de disponibilité pour un prix d'achat le plus faible possible, dans un marché concurrentiel : est il possible de concevoir moins cher et peut être moins fiable mais avec une disponibilité maîtrisée grâce à une maintenance prédictive ?

Actuellement, la demande industrielle en terme de système de surveillance est croissante car la disponibilité des processus industriels est devenu un argument tant économique que marketing. La surveillance a pour but, d'une part, d'augmenter la fiabilité des applications critiques, d'autre part, d'assurer la maintenance prédictive et donc planifiée des processus économiquement importants amenant par là une meilleure rentabilité économique. La surveillance, le diagnostic apparaissent donc comme une valeur ajoutée au produit, un service client.

2. Les Objectifs

2.1 . La fiabilité

Une part de la demande industrielle s'exprime donc en terme de **fiabilité**: quelle confiance puis je accorder au fonctionnement de mon système ? Si cette question est depuis longtemps étudiée dans les domaines de la micro électronique, la méthodologie qui est associée encore loin d'être maîtrisée dans les domaines liés à l'électronique de puissance et à l'électrotechnique (voir les travaux du RTP fiabilité)

2.2. La sûreté de fonctionnement

L'autre part de cette demande s'intéresse à la sûreté de fonctionnement, c'est-à-dire à la maîtrise de la disponibilité du procédé. Les objectifs visés peuvent être multiples mais dépendent essentiellement de l'application. Si on passe en revues quelques possibilités de niveau de surveillance on peut distinguer trois cas typiques.

2.2.1. la détection de panne

Le minimum est un outil de détection des pannes permettant d'alerter l'exploitant, de protéger l'application et éventuellement de reconfigurer le procédé pour maintenir le fonctionnement (éventuellement en mode dégradé si il existe une redondance).

2.2.2. La localisation de défaut

L'aide à la maintenance peut être réalisée en utilisant une méthode permettant de localiser le composant élémentaire en cause dans le système.

2.2.3. La maintenance prédictive.

Bien entendu, le niveau ultime de la surveillance, le « saint Graal » est la maintenance prédictive, c'est-à-dire évaluer l'état d'un système de façon à en déduire de manière « optimale » sa disponibilité c'est-à-dire prédire le temps restant avant la première panne. Cette application de la surveillance vise à l'utilisation « optimale » des matériels et système, c'est-à-dire de maîtriser la disponibilité d'un système donné en planifiant sa maintenance évitant le surcoût d'une maintenance préventive (aéronautique) ou la disponibilité aléatoire d'un système soumis à une maintenance corrective.

Dans tous les cas, ces fonctionnalités étant un « plus » apporté à l'exploitant, les intérêts divergent entre celui-ci et le constructeur. Ce dernier souhaite généralement obtenir le maximum d'information avec une instrumentation minimale (généralement celle imposée par l'aspect fonctionnel : le contrôle/ commande ou la tarification dans le domaine de l'énergie). Il existe néanmoins des applications où le service rendu en terme d'aide à la maintenance peut se monnayer et autorise donc des investissements matériels de la part du constructeur.

3. Le contexte applicatif :

La surveillance et le diagnostic nécessitent une connaissance système importante, de la formalisation de cette connaissance dépend la méthode à employer. De manière générale, l'expertise du fonctionnement sain est possédée par le constructeur, les données de maintenance sont connues de l'exploitant. Hors pour une surveillance efficace, il est crucial de connaître non seulement le fonctionnement sain, les conditions de fonctionnement « normal », mais aussi les défaillances potentielles et les processus de vieillissement des éléments critiques. Ces données sont rarement disponibles dans leur intégralité, ce qui rend la conception de systèmes de diagnostic relativement complexe et forcément multidisciplinaire.

4. Le contexte méthodologique :

4.1 Diagnostic par modèle analytique.

Dans la mesure où il existe des **modèles analytiques** « physiques » du procédé à surveiller, si on dispose des « entrées » du modèle, on est capable, en comparant les sorties de modèle et la mesure des sorties du système de déterminer si le fonctionnement est sain. Dans le cas contraire, le fonctionnement est « anormal » et pour en déterminer la cause il est nécessaire d'analyser les erreurs et d'en déduire le comportement défaillant. La confirmation du diagnostic peut être fournie par la mise en place d'un modèle du fonctionnement en défaut (à condition que ce modèle soit connu). La difficulté de ce type de méthode vient du compromis à trouver entre la précision du modèle, sa robustesse aux conditions de fonctionnement et la contrainte de fonctionner en temps réel. Les techniques à base d'observateurs ainsi que les techniques d'identifications de paramètres ont été largement employées par la communauté sur le cas « académique » de la machine asynchrone et ont montré leur difficulté de mise en œuvre [1].

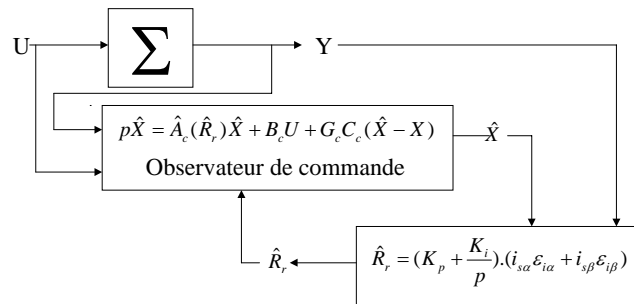


figure 1 : exemple de structure de diagnostic utilisant un observateur.

4.2. L'analyse de signal.

Si l'on se réfère au diagnostic tel qu'il est pratiqué par les équipes de maintenance, l'expertise associée à une instrumentation adaptée permet généralement de détecter et localiser les pannes, dans le contexte d'une automatisation de ces techniques et dans le cadre d'une instrumentation réduite, l'analyse du contenu harmonique des grandeurs mesurées est une technique logique et qui a fait ses preuves. L'analyse de signal est donc, elle aussi, largement utilisée en détection de panne. La difficulté première de cette technique est l'instrumentation qui se doit d'être « proche » du phénomène à détecter sous peine de voir les indicateurs de défaut faussés par la multiplication des harmoniques naturels du système et par l'atténuation apportée par la fonction de transfert de celui-ci : Des défauts mécaniques (roulements ou engrenages) seront bien plus visibles sur des accéléromètres judicieusement positionnés que sur les courants du moteur d'entraînement [2]. L'autre difficulté vient de l'analyse des résultats (signatures) et de la prise de décision qui fait encore une fois appel à l'expertise sous une forme ou une autre.

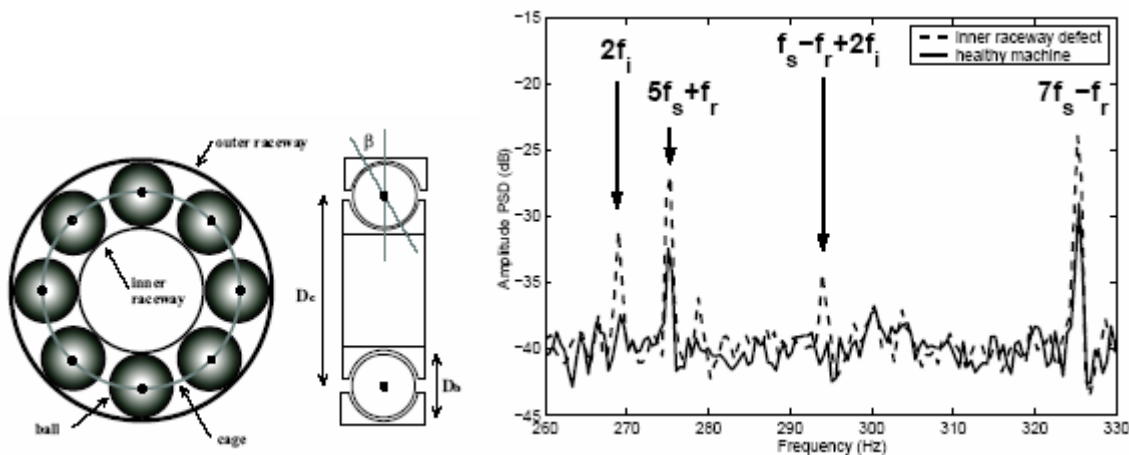


figure 2 : Spectre du courant statorique d'une machine asynchrone avec défaut de roulement.

4.3 Evaluation de l'état d'un système.

Théoriquement l'une et l'autre des types de méthodes présentées au paragraphe précédent et en particulier l'identification de paramètre peu permettre d'accéder à des grandeurs parfois difficilement mesurables et représentatives de certains modes de défaillance. Le suivi de ces paramètres critiques ou des dérives entre modèle et procédé permet d'anticiper l'apparition de certains défauts et de prédire l'occurrence d'une panne ou de suivre son évolution. En pratique, ces méthodes sont encore une fois sensibles à la précision des modèles utilisés et/ou aux conditions de

fonctionnement du système ce qui rend le suivi de variation difficile en terme de précision.

On voit que les méthodes exposées précédemment apportent de l'information sur le fonctionnement du procédé et même, dans le cas de l'identification de paramètre sur les éléments le constituant, mais que pour être exploitable en terme de diagnostic cette information demande à être analysée compte tenu des conditions de fonctionnement. Ces informations sont les symptômes qui permettent de conduire aux défaillances mais on peut se poser la question des causes physiques des défaillances et de leur surveillance. Pour surveiller les causes il est nécessaire de comprendre les phénomènes physiques (voire physico chimiques) conduisant aux défauts. La complexité des processus de dégradation est telle qu'il est souvent difficile de déterminer ou de mesurer de manière pratique les causes de vieillissement.

La solution proposée dans le paragraphe suivant vient du constat suivant : « un système vieilli en fonction de ses conditions de fonctionnement ». On peut donc penser évaluer l'état d'un procédé physique quelconque en sachant ce qu'il a enduré, pour cela, il faut relier de manière causale les conditions de fonctionnement et les défaillances potentielles (la surcharge d'une alimentation linéaire entraîne l'échauffement du régulateur et sa destruction) ce qui fait appel à l'expérience soit du constructeur, soit de l'utilisateur : le retour d'expérience. Malheureusement cette information est généralement partielle et dans ce cas pourquoi ne pas utiliser le système de surveillance à cet effet.

5. Suivi du fonctionnement et capitalisation de l'expérience pour la surveillance.

5.1. L'objectif

Si on établit facilement le lien entre la façon dont on utilise un système et sa durée de vie (exemple de l'automobile) il est en revanche plus difficile d'établir des règles strictes liant un scénario de fonctionnement et une panne car cela nécessite soit la connaissance d'un expert soit une étude expérimentale longue, fastidieuse et difficilement représentative des phénomènes réels. Il est donc intéressant de se doter d'un outil capable d'analyser le « vécu » d'un système afin de mettre en évidence les « scénarii » conduisant à une situation de défaillance. Dans l'analyse du « vécu » il y a à la fois la notion d'état de fonctionnement (classement : nominal, surcharge, défaut par exemple), la notion de répartition dans le temps de ces états successifs (scénario : fonctionnement permanent en deçà du nominal par exemple) et la notion de lien entre un type de scénario et une situation de panne (un système fonctionnant en permanence en deçà du nominal a une durée de vie de 20 ans).

Pour réaliser cette analyse, il faut organiser l'acquisition de données sur le système surveillé afin de classer les états de fonctionnement, d'établir un historique et de déterminer l'évolution future de l'état du système. La connaissance des scénarii conduisant à défaillance peut se faire dans un premier temps à l'aide de l'expertise de l'exploitant construite à partir de systèmes identiques dans des conditions d'exploitations similaires, mais elle peut aussi se construire en capitalisant, dans le système de surveillance les données de maintenance grâce à un apprentissage permanent des scénarios menant aux défauts. Cette « base de donnée » des scénarii à risque permet donc d'anticiper l'occurrence des défaillances.

5.2. Les outils

La grande quantité d'information qui afflue lors de l'acquisition des données en cours de fonctionnement et le traitement qu'elles nécessitent imposent des techniques

adaptées aux flux de données importants, on désire de plus obtenir de cette analyse les liens entre l'historique des données (scénarii) et les disfonctionnements potentiels (défaillances). L'outil classiquement utilisé pour ce type de problème est la fouille de donnée ou Data Mining[3]. De manière classique dans ce type de d'outil d'analyse de données, les informations sont triées, classées dans les différents états du système, historiées et traitées de manière à déterminer l'état futur du système.

Les étapes du processus (2)

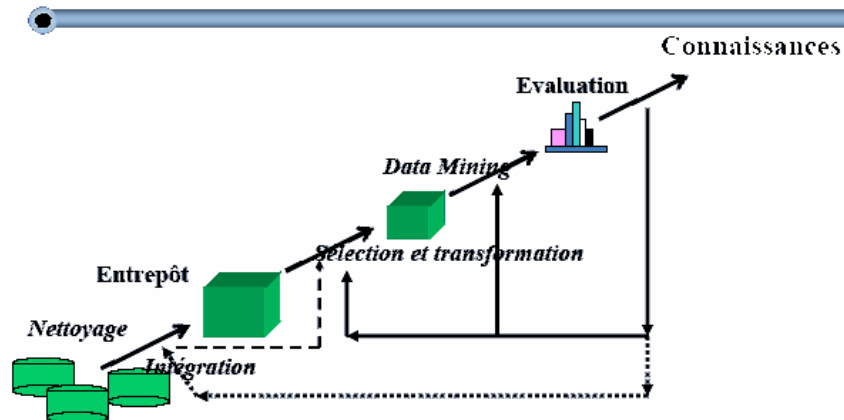
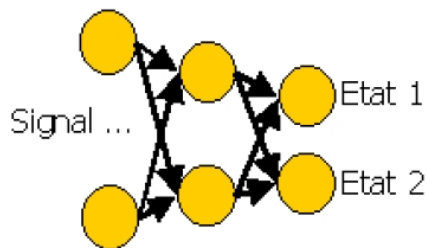


figure 3 : illustration du processus de data mining (cours de Martine Collard)

Parmi les nombreuses techniques liées au Data Mining, la puissance des réseaux de neurones et leur capacité à s'auto déterminer semblent assez bien adaptés au classement des données de cette problématique. Les réseaux bayésiens [4,5] quand à eux seront utiles pour formaliser les relations entre les scénarii ou les différents états de classement : ils permettent de qualifier les relations entre les différents états ainsi que de les quantifier à l'aide de probabilités conditionnelles.

■ Réseaux de neurones



■ Réseaux bayésiens

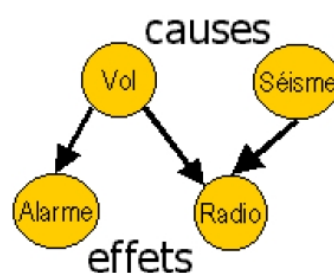


Figure 4 : illustration de réseau Bayésien et réseau de neurones (source Cours de P. Leray)

5.3. Exemples d'applications

Parmi les exemples d'applications, on peut citer les travaux de Roland Casimir au CEGELY concernant le diagnostic de la machine asynchrone par reconnaissance des formes [6].

Concernant les problèmes industriels, on peut noter les travaux de la société ProBayes pour MGE UPS [7], dont l'objectif est d'analyser les défaillances intervenant sur l'ensemble du parc d'onduleurs installés.

Parmi les travaux en cours au LEG on peut noter une étude concernant la surveillance d'un parc de transformateur MT/BT à partir des données des compteurs

client ainsi que la surveillance d'un parc d'armoires électriques à partir des informations issues des protections, des compteurs et de capteurs supplémentaires.

6. Conclusion

Compte tenu de la demande industrielle actuelle, les méthodes utilisées jusqu'à présent dans le domaine du diagnostic des systèmes électriques sont insuffisantes car elles demandent d'une part un modèle précis du système à surveiller et d'autre part une instrumentation adaptée pour discriminer les défaillances du procédé. Dans l'optique de la surveillance d'un système complexe d'une part et d'un parc de systèmes identiques d'autre part, les techniques de fouilles de données semblent être une solution intéressante qu'il importe de valider sur un problème concret. Il est bien entendu qu'en terme de diagnostic toute information est intéressante et il s'agit d'extraire un maximum d'information des données disponibles, en ce sens les techniques à base de modèles analytiques ainsi que les techniques d'analyse des signaux sont indispensables pour extraire des mesures (informations brutes) des indicateurs de défaut représentatif. Ces techniques ne prenant pas en compte la connaissance à priori et l'expertise, elles sont donc complémentaires des techniques d'analyse et de fouille de données.

Bibliographie :

- [1] : thèse INPG Bertrand Raison « Détection et localisation de défaillances sur un entraînement électrique » septembre 2000
- [2] : DEA Martin Blödt, Article IEEE-ISIE04 Ajaccio 2004
- [3] : « Le data mining » René Lefebure, Gilles Venturi, Ed Eyrolles
- [4] : « Réseaux bayésiens » Patrick Naïm, Pierre-Henri Wuillemin, Philippe Leray, Olivier Pourret, Anna Becker, Ed Eyrolles
- [5] : « Apprentissage et diagnostic de systèmes complexes : réseaux de neurones et réseaux bayésiens » Philippe Leray, Thèse de doctorat de l'université Paris 6 – 1998
- [6] : « Diagnostic des défaut des machines asynchrones par reconnaissance des formes » thèse de Roland Casimir de l'école centrale de Lyon décembre 2003
- [7] : Industrie et Technologie n°860 septembre 2004 p80 ou www.probayes.com

Généralités sur les méthodes :

Cours sur le DataMining et les réseaux Bayésiens [Philippe Leray](http://asi.insa-roen.fr/enseignement/siteUV/dm/index.html) <http://asi.insa-roen.fr/enseignement/siteUV/dm/index.html>.

Cours de Martine collard sur le data mining : www.i3s.unice.fr/~mcollard/master/index.html

Exemple d'application au domaine des systèmes électriques :

Machine Asynchrone :

Thèse Roland Casimir « Diagnostic des défaut des machines asynchrones par reconnaissance des formes » au CEGELY : http://cegely.ec-lyon.fr/article.php3?id_article=140

Application aux transformateurs :

Surveillance continue des transformateurs : application des méthodes d'intelligence artificielle

De P. Boss, P. Lorin, ABB Sécheron SA et de A. Schenk, A. Germond, Ecole Polytechnique Fédérale de Lausanne

Apparu dans le Bulletin VSE No. 4, le 3 mars 2000