

## Proposition sujet de thèse

### Processus 4.0 pour l'analyse des défaillances en production microélectronique

**UMR de rattachement :** LIMOS, UMR CNRS 6158.

**Localisation :** Mines Saint-Etienne, Institut FAYOL

**Directeur de thèse :** Pr. Mireille Batton-Hubert, EMSE – Institut FAYOL - LIMOS.

**Co-directeur :** Pr. Xavier Boucher, EMSE – Institut FAYOL - LIMOS.

**Ecole Doctorale :** Ecole Doctorale SIS, Ecole des Mines de Saint-Etienne et Université Jean Monnet.

**Date de démarrage proposée :** Octobre 2020

**Collaboration industrielle :** STMicroelectronics, Projet Européen FA4.0.

**Tuteur industriel :** Pascal Gounet, Grenoble Reliability & Analysis Laboratory

**Mots clés :** Data-mining, Machine Learning, Process-mining, Théorie de la décision, Artificial Intelligence, Industrie du Futur, Industrie micro-électronique

#### Contexte scientifique et industriel

Cette thèse se développe dans le cadre du projet européen FA4.0, en collaboration avec les équipes de STMicroelectronics sur Grenoble (Grenoble Reliability & Analysis Laboratory). Plus précisément, le projet vise à déployer des solutions d'Intelligence Artificielle développée dans une perspective 'Industrie du Futur' pour transformer et améliorer les pratiques industrielles au sein d'un service particulier de l'entreprise : le laboratoire d'analyse des défaillances. Ce laboratoire est un maillon essentiel de la production industrielle, en charge d'analyser et de diagnostiquer les problèmes de qualité et de défaillance qui apparaissent dans les processus de production de wafers micro-électroniques (fabrication de micro-puces) et de produits complexes.

L'analyse des défaillances tout au long de la chaîne de valeur à partir d'une puce, est une condition préalable à l'amélioration de la fiabilité et de la qualité et donc de la compétitivité des dispositifs électroniques, en particulier sur le marché de l'automobile et de l'industrie pour les applications exigeantes en matière de fiabilité et de sécurité. La forte progression de la numérisation et de l'automatisation, offre un formidable potentiel pour mieux maîtriser la fiabilité des produits de haute technologie qui sont basés sur des systèmes électroniques de plus en plus complexes.

Les nouveaux concepts comme la conduite autonome représentent le défi technologique le plus élevé pour les composants électroniques, qui malgré leur complexité extrême, doivent fonctionner quotidiennement de manière fiable et sécurisée. La garantie de fiabilité est donc extrêmement critique, au niveau des composants électroniques et du système. Les risques de fiabilité ne peuvent être maîtrisés que par une caractérisation et une analyse efficiente des défaillances liées au processus et par un contrôle global des incertitudes le long de la chaîne couvrant la conception, la fabrication et la gestion de la qualité des composants.

Grâce à l'apport du digital, l'objectif du projet est donc de transformer les pratiques actuelles du laboratoire d'analyse de défaillance, en mobilisant les techniques d'intelligence artificielle du domaine du machine learning. Ces techniques permettent d'enrichir la capacité explicative des modèles statistiques, par une capacité renforcée à discriminer les causes possibles de non fiabilité et à prendre en compte leurs interactions.

#### Problématique et objectifs de recherche

La spécificité d'un laboratoire d'analyse des défaillances est d'opérer un certain nombre de tests électriques, physiques et chimiques, qui utilisent largement des méthodes d'exploration et d'analyse des supports micro-électroniques qui contiennent les composants électroniques susceptibles d'être atteint par un défaut (défaillance). Lors de cette analyse, la succession des tâches d'exploration dépend des résultats observés à chaque étape  $t$  ; la

décision pour le choix de l'étape de test suivant  $t+1$  dépend des résultats obtenus à la phase  $t$  voire  $t-1$ , mais surtout de l'expertise humaine. Ce diagnostic, appliqué à un support électronique donné faisant apparaître un défaut observé, peut être vu comme un parcours possible sur un graphe de décision où chaque nœud est un choix décisionnel parmi un ensemble d'alternatives possibles (alternatives : quelle est l'étape d'analyse suivante à réaliser ?). Les résultats des analyses et ou des tests se présentent sous la forme de données diverses et variées (notamment d'importantes bases de données d'imagerie de microscopies diverses, de signaux physiques, de mesures, de données textuelles ...) qui sont utilisées par l'analyste pour la construction de son diagnostic au cours du processus d'analyse du défaut. Les choix d'analyses à réaliser sont nombreux, d'où un nombre important de combinaisons de parcours ; le choix n'est pas forcément unique ni mono-objectif et n'est pas lié à un modèle de causalité déterministe mais plutôt d'ordre décisionnel (notion d'*utilité espérée*). Les résultats des analyses à une étape et aux étapes précédentes, orientent la sélection de l'étape suivante à réaliser.

Il s'agit alors de développer des **méthodes adaptées de diagnostic basées sur l'analyse des profils de décision et du flot de données, afin d'orienter la stratégie optimale (ou meilleures stratégies) pour sélectionner, progressivement, les étapes d'analyse de défaillance pertinentes dans un cas donné**. Les décisions de choix d'étapes d'analyse permettent ainsi de **construire progressivement l'ensemble du parcours d'analyse de défaillance**. De plus, cela doit permettre de limiter les mauvais choix lors du passage d'une étape non-destructive à une étape destructive.

L'enjeu est d'exploiter des motifs dans les données, le flux de données et les choix décisionnels dans un cadre de grande dimension, de données hétérogènes où l'analyste est une source d'information incontournable, en utilisant des techniques avancées du *machine learning* et du *process mining*. D'un point de vue industriel, ces travaux visent à **accroître significativement l'efficacité du processus d'analyse de défaillances** (notions de temps de cycle et de taux de succès), notamment par une réactivité améliorée, une réduction des phases d'analyses n'apportant pas de valeur ajoutée, une fiabilisation des éléments de diagnostic fournis.

### Domaine scientifique et Résultats attendus

- Domaine scientifique concerné: Approches par apprentissage non-supervisé ; Théorie de la décision statistique ; arbre de décision
- Mots clés : *Process mining*, graphe décisionnel probabiliste, apprentissage et partitionnement
- Mots clés : Couplage entre apprentissage automatique et expertise, Similarité et induction sur les 'cas de défauts'. Par ailleurs une analyse des mécanismes de décision inter-étapes est nécessaire, en vue d'aide à la décision couplant apprentissage probabilistique et d'autres formes de raisonnements décisionnels.

Les résultats attendus : Il s'agit de contribuer à l'*Automation du processus* global 'Failure Analysis' en fournissant une aide à la décision entre chaque étape d'analyse, destinée à optimiser la construction progressive de l'ensemble du parcours d'analyse. A chaque étape, une sélection des tâches d'analyse pour les étapes suivantes est alors proposée pour répondre aux besoins diagnostic du défaut concerné.

L'approche développée devra permettre de :

- d'extraire des bases de données existantes, la connaissance sur la structure implicite des 'Parcours décisionnels (= 'Failure Analysis flows') ;
- de faire l'analyse des 'parcours décisionnels' archivés afin d'identifier, sur l'ensemble de la base, s'il existe une typologie (classification) des 'parcours décisionnels', ou bien des motifs récurrents au sein de ces parcours ;
- de construire un 'failure analysis flow' pour un nouveau 'cas de défaut détecté' à chaque étape ;
- et de répondre à certaines questions scientifiques : Quel est le modèle de représentation de l'ensemble de l'espace décisionnel probabiliste correspond le mieux au 'Failure analysis flow' sa nature (processus markovien ou non, etc. ) ?

## Procédure de candidature

Dossier incluant:

- CV
- Lettre de motivation
- Relevés des notes des 2 dernières années de formation.
- Résumé du projet recherche de master
- Lettres de recommandations.

Le dossier est à transmettre en version écrite à Mireille Batton-Hubert, et en copie électronique à Xavier Boucher : { [batton@emse.fr](mailto:batton@emse.fr); [boucher@emse.fr](mailto:boucher@emse.fr) }

Mireille Batton-Hubert Ecole des Mines de Saint Etienne, Institut FAYOL  
158 Cours Fauriel  
CS 62362  
42023 Saint Etienne cedex 2

## Profil du candidat

Master français ou européen de recherche ou bien Ingénieur (Généraliste ou Sciences des données) avec possibilité d'équivalence Master recherche (première expérience de projet recherche). Nous recherchons à couvrir une ou bien plusieurs des compétences suivantes :

- Mathématiques appliquées, orientée vers l'analyse des données et l'aide à la décision
- Science des données, Machine learning, Process mining
- Recherche opérationnelle, Aide à la décision, Optimisation multicritères.
- Culture en génie industriel, Process industriel, Systèmes de production

Bien sûr un intérêt ou des expériences complémentaires en lien avec l'industrie de la microélectronique seront également appréciées.

## Références

W. M. P. van der Aalst, *Process Mining: Data Science in Action*. Springer, 2016.

W. M. P. van der Aalst and A. J. M. M. Weijters, "Process mining: A research agenda," *Comput. Ind.*, vol. 53, no. 3, pp. 231-244, 2004.

[Trevor Hastie](#), [Robert Tibshirani](#) et Jerome Friedman, *The Elements of Statistical Learning : Data Mining, Inference, and Prediction*, 2009, 2<sup>e</sup> éd.

Liens utiles

**Institut FAYOL** : <https://www.mines-stetienne.fr/recherche/5-centres-de-formation-et-de-recherche/institut-henri-fayol/>

**Publications FAYOL** : <https://www.mines-stetienne.fr/recherche/departements/departement-genie-de-l-environnement-et-des-organisations/publications/>

**IT'm Factory** : <https://www.mines-stetienne.fr/entreprise/itm-factory/>