



Centre de Recherche : Institut Fayol - ENMSE

Titre :

- Knowledge models to support an agile and optimised information system in Microelectronics Industry – Case study applied to reactive maintenance.
- Modèles de connaissances contribuant à l'agilité et à l'optimisation du système d'information en industrie Microélectronique – Application au pilotage de la maintenance.

Directeur de thèse : Xavier Boucher, Prof.

Co-directeur : Rodolphe Leriche, HDR

Co-encadrant : Philippe Beaune

Ecole Doctorale : Ecole Doctorale SIS, de l'ENSMSE et de l'Université Jean Monnet.

Collaboration, rattachement à un programme : Collaboration Industrielle ST Microelectronics

Financement de la thèse : Convention CIFRE – St Microelectronics

Mots clés : Système d'information adaptatif, ontologie, fouille de données, advanced process control, semi-conducteurs.

Contexte Industriel

Les systèmes de production du secteur microélectroniques sont caractérisés par un haut degré de technologie et d'automatisation, mais également par une confrontation quasi permanente à la gestion du changement. Dans ce contexte, le système d'information est un vecteur majeur de création de valeur : il fournit des volumes sans cesse croissant d'informations utiles à la supervision, au pilotage et à l'adaptation des processus industriels ; il constitue une clé de voute de l'agilité du système productif. Dans ce contexte, la thèse proposée aura pour enjeux (i) de contribuer à un net accroissement de l'usage et de la valeur ajoutée de l'information disponible et de (ii) mettre en place des mécanismes de traitement de l'information permettant l'adaptation continue du SI aux évolutions, dans une perspective (à plus long terme) de système d'information auto-adaptatif.

Les énormes quantités d'information dont disposent aujourd'hui les sites de production de STMicroelectronics se trouvent réparties au sein de multiples bases de données et composants du système d'information global. Ces informations contribuent au système décisionnel de pilotage et de contrôle de la fabrication, et peuvent être captées et archivées pour des usages très locaux ou plus globaux. Ces données concernent, entre autres, le pilotage et la gestion de la production, le contrôle de la qualité, la surveillance des équipements, le pilotage réactif des équipements en fonction des lots produits, la maintenance des équipements, etc...

La valeur ajoutée de ces masses d'information est actuellement faiblement exploitée. En effet, d'un point de vue industriel, plusieurs problématiques contribuent à la complexité de gérer un tel système d'information :

- La répartition de l'information au sein d'un système très distribué pose des problèmes spécifiques d'interopérabilité qui limitent fortement l'usage des données disponibles.
- La connaissance sur le système d'information en lui-même (méta-connaissance) est soumise à des phénomènes multiples d'obsolescence, de perte, de manque de mise à jour qui induisent une large perte de performance et qui réduisent l'adaptabilité du système.
- Les mécanismes de traitement d'information inclus dans le système décisionnels n'ont pas intégré à ce jour les principes permettant l'adaptabilité de la collecte et du traitement des données selon les changements de contexte, ni les principes permettant de gérer l'évolutivité du système.

Cette problématique industrielle est liée à des verrous scientifiques soulignés ci-après. Ce travail de thèse propose d'aborder ces enjeux par l'association entre (i) la création de modèles de connaissances implémentés sous la forme d'une ontologie de système d'information et (ii) la mise au point de mécanismes de fouille de données dans l'ensemble du système d'information, qui répondront à la nécessité d'adaptabilité au contexte et d'évolutivité du système soulignée ci-dessus. Le champ d'application plus spécifique de ce travail se focalisera sur l'optimisation de la maintenance des équipements à partir d'historiques hétérogènes de données.

Sujet et objectifs

Les verrous scientifiques de ce travail peuvent être regroupés essentiellement sous deux aspects :

- La création et les mécanismes de maintien à jour d'un modèle de connaissance sur le système d'information susceptible de résoudre des problématiques d'interopérabilité des données et des systèmes, d'interprétation contextuelle de l'information, et de maîtrise de l'évolutivité du système. Nous proposons de nous appuyer sur le domaine des ontologies pour traiter ces aspects.
- L'intégration, au sein d'une ontologie de système d'information, de mécanismes de fouilles de données répondant à des principes d'adaptabilité du traitement des données au contexte et de prise en compte de l'évolutivité des données disponibles.

Ontologie & modèle de connaissances : dans le domaine des ontologies, cette recherche bénéficiera comme point de départ des acquis de l'ENSMSE dans le cadre de la thèse CIFRE de Rahee Ghuburn soutenue en 2008, en collaboration avec un autre site de production de STMicroelectronics. Ces apports fourniront une base initiale de structuration ontologique des connaissances qui permettront d'approfondir plus avant les mécanismes dynamiques associés à ce modèle de connaissance : formalisation de la notion de contexte et mécanismes d'interprétation en fonction de ce contexte ; mécanismes permettant des chemins alternatifs d'accès à l'information en fonction des besoins ; mécanismes de gestion sémantique permettant d'automatiser l'évolution de l'ontologie en fonction des changements.

Ainsi, tout en apportant des solutions d'ordre sémantique à la gestion de l'interopérabilité des informations et au maintien d'une méta-connaissance sur le système d'information, l'ontologie devra également être enrichie par le couplage avec les techniques de fouilles de données. Ces techniques d'analyses statistiques offrent notamment l'opportunité de créer des méta-informations (informations construites à partir d'autres informations de base) permettant d'augmenter l'utilisation des données disponibles au sein du système décisionnel de l'entreprise. Ces méta-informations utilisables dans les processus de pilotage et de prise de décision, devront trouver leur place dans l'ontologie. Notamment elles devront être traçables (comment ont elles été produites, à partir de quelles données, quelle précision peut être associée ?), et pouvoir être construites à la demande (ce qui est lié à l'existence d'une sémantique).

Fouille de données : en complément au modèle de connaissance ontologique, les techniques de fouille de données seront utilisées d'une part pour généraliser l'utilisation intelligente et efficace des giga-octets d'informations enregistrées et d'autre part afin de spécifier des mécanismes de traitement de l'information et de prise de décision adaptables à différents besoins.

De manière générique, une des problématique clé consiste à reconstruire une information utile à partir de données basiques dispersées dans l'ensemble des bases. Il s'agit de construire une fonction f à partir de données $\{X\}$ choisies parmi tous les X . $f(\{X\})$ devra remplacer une information enregistrée par le passé ou construite pour l'occasion. Les usages potentiels de $f(\{X\})$ sont nombreux : contrôle plan adaptatif, création d'alarmes pour la maintenance, mesures de remplacement, archivage des données.

D'un point de vue scientifique la construction de ce type de fonction repose sur la démarche de sélection des données, sur la spécification des méthodes de fouilles de données pertinentes (machine à support vectoriel, krigeage, réseau de neurones...), ainsi que sur la gestion de la qualité des prévisions et de la traçabilité des incertitudes. Deux difficultés semblent inhérentes à l'application aux semi-conducteurs. Tout d'abord, les quantités de données à explorer sont énormes. Ainsi, une indication par un expert des $X^{(i)}$'s candidats est a priori nécessaire. Ensuite, chaque $X^{(i)}$ est enregistrée à des fréquences temporelles très variables. Il faudra donc ajouter à la fouille de données des méthodes de mise en concordance temporelle, typiquement de la compression de données. Dans le cadre de cette thèse, nous ne proposerons pas une méthode unique de fouille de données, mais un ensemble de méthodes utilisables selon le contexte. L'accent sera mis sur une description générique des méthodes de fouilles de données et leur inclusion dans l'ontologie. Le travail se focalisera donc sur l'évolution des données et des méthodes de fouilles de données.

Domaine d'application : Une application en maintenance servira de guide aux développements. Elle permettra de délimiter un domaine de travail facilitant la spécification de l'ontologie du SI. Elle servira de champ applicatif pour la mise aux points des méthodes de fouille de données : l'objectif consistera à améliorer la performance des algorithmes de maintenance et du système de collecte de données associé.

Contexte de la collaboration

Cette proposition de thèse CIFRE fait suite à différentes collaborations de recherche d'ores et déjà développées entre l'Institut Fayol et St Microelectronics. Outre les travaux sur les ontologies cités précédemment, une autre thèse CIFRE a démarré en mai 2010, sur l'aide à la décision pour anticiper l'impact des scénarios d'évolution sur la capacité du système d'information.

Profil du candidat

Diplômé d'un master recherche ou Grande École en Gestion des systèmes d'information, Informatique ou Génie industriel.

Idéalement nous cherchons une double compétence en système d'information et analyse statistique (fouille de données).

Certains acquis parmi les compétences suivantes seront appréciés :

- Architecture et modélisation des systèmes d'information
- Modèle de connaissances et ontologies
- Analyse statistique, fouille de données
- Modèles statistiques de maintenance

Contact

Candidatures à adresser conjointement à

M. Xavier BOUCHER

ECOLE NATIONALE SUPERIEURE DES MINES

Institut Fayol

158 Cours Fauriel

42023 SAINT ETIENNE Cedex 2

04.77.42.01.33

boucher@emse.fr

M. David ROZIER

STMicroelectronics SA (Site de Crolles),

850 rue Jean Monnet

F-38926 Crolles cedex

david.rozier@st.com