



Développement du projet de recherche CIFRE

22.09.2019

WorldCast Systems

20 Avenue Neil Armstrong,
33700 Mérignac.

<https://www.worldcastsystems.com/fr/>

Vue d'ensemble

Ce document décrit le projet de recherche en partenariat entre l'**IMT Atlantique** et l'entreprise **Worldcast Systems** en vue d'une thèse CIFRE sur la thématique de l'intelligence artificielle appliquée à la maintenance prédictive. Le contexte de la recherche est décrit suivi des objectifs visés et de l'organisation prévisionnelle de la thèse .

Mots-clés : Intelligence Artificielle, Industrie 4.0, Maintenance Prédictive, Apprentissage Machine, SVM, Radiodiffusion

Contexte

L'entreprise WorldCast Systems conçoit et développe une large gamme d'équipements de diffusion et de surveillance pour les marchés de la radio et de la télévision.

Elle s'appuie pour cela sur une expertise acquise au cours de plus de soixante ans d'expérience dans le domaine pour fournir aux ingénieurs en radiodiffusions des produits hautement fiables, efficaces, et innovants.

La stratégie de l'entreprise est en effet d'investir massivement dans la recherche et le développement afin de proposer des technologies de pointe à ses clients et se différencier ainsi de ses concurrents. Un chiffre dénote cette volonté : le département de Recherche et Développement représente 40% du nombre d'employés de l'entreprise. C'est dans ce contexte que l'entreprise WorldCast Systems souhaite investir et se placer dans la thématique de l'**Intelligence Artificielle** (IA). Elle s'inscrit en cela dans le plan Villani pour le développement de l'IA à travers le renforcement des collaborations entre le monde de la recherche académique et celui de la recherche industrielle [1].

Dans une première thèse réalisée en collaboration avec l'Ecole Nationale d'Ingénieurs de Tarbes entre les années 2016 et 2019, des premiers travaux ont été réalisés en vue d'extraire et d'organiser les données extraites des différents produits vendus par l'entreprise WorldCast Systems, puis de les analyser dans le but de réaliser des opérations de **Maintenance Prédictive** (ou PHM, Predictive Health Management [2]). Ces travaux ont porté sur une gamme d'équipements produits par l'entreprise. La structure interne et les composants varie peu au sein de cette gamme, ce qui permet facilement de transférer les résultats obtenus de l'un à l'autre des appareils.

Dans un premier temps, une analyse de type AMDEC (Analyse des Modes de Défaillance, de leurs Effets et de leur Criticités) fut réalisée sur ces équipements afin d'analyser les sources de pannes à partir des composants internes des équipements. Cette analyse s'est nourrie de l'expertise des concepteurs mais également des données de retour d'expérience fournie par le service après vente. A la suite de cette analyse, les données de surveillance pertinentes ont été identifiées. Ces données, remontées de l'ensemble des équipements considérés dans l'étude, ont ensuite été traitées et uniformisées afin de créer une base de données. Celle-ci fut ensuite prétraitée pour obtenir une base de données fiables. Les méthodologies originales développées lors de la thèse pour rendre ces données exploitables sont décrite dans [3][4].

Une fois cette base de données fiables constituée, la deuxième partie du travail fut la proposition d'une méthodologie d'organisation et de classifications des données provenant des appareils en exploitation, puis la définition de critères de défaillance. Les méthodologies utilisées furent de réduire les dimensions du problème (de 14 dimensions

vers 2 dimensions) par l'utilisation de méthodologie ACP (Analyse des Composants Principaux [5]). Ensuite, ces données ont tout d'abord été analysées à l'aide de cercles de corrélations, permettant de différencier clairement et visuellement les conditions de défaillance d'un équipement. Enfin, une analyse par SVM [6] a permis de mettre en place un modèle fiable d'identification des systèmes défaillants. Ce modèle est en cours d'industrialisation.

Ces premiers travaux sont donc encourageants. Ils montrent qu'il est possible d'utiliser efficacement les nombreuses données mesurées dans les équipements afin d'effectuer de la maintenance prédictive et éviter des pannes à travers l'expertise et la connaissance de l'entreprise en ce qui concerne les équipements qu'elle conçoit et produit.

Cependant, il existe plusieurs verrous à l'extension de ces résultats vers d'autres produits de l'entreprise ou provenant d'autres entreprises du secteur. En effet, l'extraction des données et leur prétraitement ont été effectués "manuellement", et ce travail est donc difficilement reproductible. De plus, la nature des données à récupérer ainsi que le résultat du traitement par ACP s'applique exclusivement à la seule gamme d'équipements considérés dans cette étude.

Objectifs

L'objectif de la thèse à venir est de capitaliser sur ces premières avancées en inventant des outils théoriques et une mise en oeuvre automatisée de la méthode. Pour cela, étant données les dimensions du problème, il s'agit désormais de s'appuyer sur des algorithmes d'**Apprentissage Machine**, pour automatiser l'extraction des données, réduire les dimensions du problèmes, et mettre au point des critères de défaillances. L'amélioration des méthodes de maintenance prédictive par l'utilisation de l'IA est au coeur de la transformation numérique de l'industrie, aussi appelée **Industrie 4.0** [7][8].

Pour cela, le thésard devra s'appuyer sur le logiciel Kybio proposé par Connect (<https://www.worldcastconnect.com>), filiale de WorldCast group. Le service proposé par Kybio aux exploitants de parc machines est de surveiller l'ensemble de leurs appareils au sein d'une interface centralisée. Ces appareils sont divers : transport de données, émission à haute puissance, surveillance, data center, studio de télévision. Or, l'ensemble des données de maintenance sont rassemblées au sein d'un seul et même programme. Le protocole utilisé s'appelle le protocole SNMP [9]. Une grande variété de données sont remontées. Les données remontées par ce protocoles sont dépendantes de chaque produit, et cette liste est disponible au sein d'un fichier appelé MIB. La description d'une telle extraction automatisée en vue d'une utilisation pour effectuer de la maintenance constituerait une contribution originale et significative dans le domaine. Il n'y a en effet, à notre connaissance, pas de publication dans la littérature documentant l'utilisation de ce protocole pour la constitution d'une base de données exploitables. Cette innovation

pourrait être un outil extrêmement pertinent pour la création de jeux de données qui nourriront les algorithmes d'apprentissage machine dans le domaine de la maintenance.

Parmi ces données se trouvent des données de surveillance (température, vitesse de ventilateur, tension d'alimentation) mais également des alarmes, avec différents niveaux de criticité. Ceci est très important et encourageant, car ces alarmes pourraient être utilisées comme étiquettes en vue du déploiement d'algorithmes d'apprentissage machine **supervisés**. Les informations fournies dans les MIB permettront de classer et d'identifier les significations de chaque donnée en vue d'un prétraitement.

Kybio étant un service vendu à de très nombreuses entreprises utilisant chacune de très nombreux appareils de tout type et de toute marques, la quantité de données disponibles est immense. Cette base de données sera idéale pour mettre en place des algorithmes d'apprentissage machine efficaces. La réussite de cet objectif constituerait une rupture dans le domaine de la maintenance prédictive, par la nouveauté des algorithmes utilisés et également par la quantité de données utilisées.

L'objectif est alors de fournir automatiquement des modèles permettant l'identification de défaillances dans les équipements de WorldCast Systems, mais également parmi les équipements possédés par les clients de Connect.

Résumons ici les objectifs visés :

1. extraire et classer automatiquement les données de surveillance obtenues via le protocole SNMP à partir de la MIB d'un équipement,
2. éprouver et fiabiliser cette extraction sur les appareils gérés par l'interface Kybio pour créer une base de données fiables,
3. concevoir des modèles d'identification de défaillance des nombreux équipements gérés par Kybio à partir d'algorithmes d'apprentissage machine,
4. automatiser la création de ces modèles.

Organisation

A / T0 à T0 + 1 mois : À WorldCast Systems.

Ce premier mois sera consacré à une première prise de contact avec les membres de l'entreprise et à l'introduction des différents produits et services développés par l'entreprise. Un passage sur les chaînes de production est possible, afin de comprendre et d'appréhender le fonctionnement de l'entreprise. Ce premier mois permettra également de transmettre au nouveau thésard les résultats obtenus lors de la thèse précédente ainsi que les différentes bases de données et modèles créés.

B / T0 + 1 mois à T0 + 7 mois : À l'IMT Atlantique.

La deuxième étape pour le thésard sera de réaliser une étude bibliographique des techniques de réduction de la dimensionnalité (ACP, SVD) et d'apprentissage machine (Random Forest, SVM, DNN). Il bénéficiera pour cela de l'appui des enseignants chercheurs du département d'électronique de l'IMT Atlantique [10-12]. Il réalisera également une synthèse des travaux existant sur des thématiques proches de celle proposée, à savoir l'utilisation d'algorithmes d'apprentissage machine pour des applications de maintenance prédictive. Cette partie est particulièrement critique puisqu'il n'existe pas au sein du département d'expertise particulière sur ces thématiques. Il s'agira donc de réaliser une bibliographie solide, et idéalement de prendre contact avec de possibles partenaires académiques afin de s'assurer de la pertinence des recherches envisagées. Des premières réalisations concrètes sont attendues (écriture d'un programme python pour démonstration).

C / T0 + 7 mois à T0 + 16 mois : À WorldCast Systems.

La deuxième partie de la thèse sera de développer un mécanisme d'extraction et de prétraitement automatique des données remontées par protocole SNMP, à partir de l'analyse des MIB. Cela inclut :

- l'interface avec les différents équipements de l'entreprise ,
- l'interface avec le logiciel Kybio développé par Connect,
- l'analyse des différentes MIB pour l'identification détaillée du type de chaque donnée
- le prétraitement des données (doublons, valeurs aberrantes, remplacement par la moyenne, régularisation).

Il faudra donc interagir avec les membres de l'entreprise pour comprendre le fonctionnement des équipements conçus et produits par l'entreprise. Il faudra également interagir avec les développeurs de l'entreprise Connect pour extraire des données massivement.

D / T0 + 16 mois à T0 + 25 mois : À WorldCast Systems.

Une fois cette base de données fiables créée, la troisième partie de la thèse sera la mise en place des algorithmes d'apprentissage machine sur les données ainsi extraites. Le but sera d'extraire des modèles efficaces permettant de réaliser des prédictions de pannes et de détecter des optimisations possibles sur les équipements de l'entreprise, mais également sur tout équipement géré via le logiciel Kybio vendu par Connect. Les tâches **C** et **D** se superposeront sûrement, puisque le format des bases de données et les prétraitements appliqués dépendront des algorithmes d'apprentissage utilisés.

E / T0 + 25 mois à T0 + 31 mois : À l'IMT Atlantique.

La quatrième partie sera un travail de raffinement des résultats obtenus et de réflexions sur les généralisations possibles du travail effectué. Ce travail se fera en interaction avec les chercheurs de l'IMT Atlantique.

F / **T0 + 31 mois** à **T0 + 36 mois** : À l'IMT Atlantique.

Cette dernière partie sera consacrée à la rédaction du manuscrit de thèse.

Bibliographie

- [1] C. Villani, Y. Bonnet, B. Rondepierre et al., *"For a meaningful artificial intelligence: Towards a French and European strategy,"* Conseil National du numérique, 2018.
- [2] R. Gouriveau, K. Medjaher, and N. Zerhouni, *"From prognostics and health systems management to predictive maintenance 1: monitoring and prognostics,"* John Wiley & Sons, 2016.
- [3] H. Sarih, A. P. Tchangani, K. Medjaher, and E. Pere, *"Critical components identification based on experience feedback data in the framework of phm,"* IFAC-PapersOnLine, 2018, IFAC Symposium on Information Control Problems in Manufacturing INCOM.
- [4] H. Sarih, A. P. Tchangani, K. Medjaher, and E. Pere, *"Data preparation and preprocessing for broadcast systems monitoring in phm framework,"* in 2019 International Conference on Control, Decision and Information Technologies (CoDIT).
- [5] I. Jolliffe, *"Principal component analysis,"* Springer, 2011.
- [6] V. Vapnik, *"The nature of statistical learning theory,"* Springer science & business media, 2013.
- [7] J. Yan, Y. Meng, L. Lu, and L. Li, *"Industrial big data in an industry 4.0 environment: Challenges, schemes, and applications for predictive maintenance,"* IEEE Access, 2017.
- [8] Z. Li, Y. Wang, and K.-S. Wang, *"Intelligent predictive maintenance for fault diagnosis and prognosis in machine centers: Industry 4.0 scenario,"* Advances in Manufacturing, 2017.
- [9] D. Harrington, B. Wijnen, and R. Presuhn, *"An architecture for describing simple network management protocol (SNMP) management frameworks,"* 2002.
- [10] A. Iscen, T. Furon, V. Gripon, M. Rabbat and H. Jégou, *"Memory vectors for similarity search in high-dimensional spaces,"* in IEEE Transactions on Big Data, 2018.
- [11] G. B. Hacene, V. Gripon, N. Farrugia, M. Arzel and M. Jezequel, *"Transfer Incremental Learning Using Data Augmentation,"* in Applied Sciences, 2018.
- [12] G. Boukli Hacene, V. Gripon, N. Farrugia, M. Arzel, and M. Jezequel, *"Finding All Matches in a Database using Binary Neural Networks,"* in COGNITIVE, 2017.