



La performance du transport ferroviaire, un sujet de data science !



Les infrastructures ferroviaires sont considérées comme critiques car elles doivent délivrer un taux de disponibilité le plus élevé possible pour garantir une continuité de service public. Ces infrastructures, dotées d'une architecture complexe, génèrent des volumes de données hétérogènes très importants qui renferment des informations essentielles pour la compréhension de leur état et l'optimisation de leur exploitation. Les évolutions attendues ne pourront se faire sans la structuration de ses énormes volumes de données et leur analyse poussée grâce, entre autres, à l'intelligence artificielle.



Robert PLANA
Directeur Innovation (CTO)

Docteur en Philosophie, Technologies de l'information et de la communication, Robert a une solide expérience dans l'internet des objets et dans les technologies avancées.



Le transport ferroviaire fait face à plusieurs défis majeurs : le vieillissement de ses infrastructures, une forte augmentation de la population donc des usagers et l'ouverture à la concurrence qui sera effective à partir de 2020 pour les grandes lignes et, au plus tard, en 2024 pour les lignes régionales. Pour SNCF Réseau, il s'agira de réussir à maintenir en état les installations existantes et d'en accroître les capacités. Quant à la RATP, elle est également concernée par la mise en concurrence pour son réseau « de surface » (bus et tramways). La sécurité de ses voyageurs est également un de ses enjeux majeurs.

La SNCF et la RATP, deux « data-driven company »

En septembre 2018, le projet « Nouvel'R » (R pour réseau) a été dévoilé par Patrick Jeantet, PDG de SNCF Réseau. Cette « feuille de route » stratégique constitue la déclinaison opérationnelle des transformations exigées par l'Etat pour remédier au déficit structurel de la SNCF. L'objectif fondamental est de préparer le réseau du futur pour faire regagner des parts de marché au mode ferroviaire, le seul à permettre une mobilité de masse décarbonée. Pour cela, la SNCF doit diminuer les coûts structurels de son infrastructure, augmenter la capacité et la robustesse des axes et nœuds saturés et accélérer la modernisation des systèmes d'exploitation et de la signalisation, comme le passage de l'ERTMS 1 à l'ERTMS 2 (le système européen de signalisation ferroviaire).

« La SNCF doit construire de nouvelles lignes en tenant compte du phénomène de périurbanisation à la française, résume l'expert. Demain,

“ A partir d'une approche Système guidée par les données et les modèles, on peut calculer l'indice de santé d'un réseau ferroviaire, et ainsi planifier sa réhabilitation.”

avec une seule norme ERTMS déployée, l'exploitation du réseau deviendra optimale. Aujourd'hui, elle dispose de deux systèmes distincts pour les lignes transverses entre villes moyennes et les lignes à grande vitesse. Le trafic régional impacte le trafic grande distance... Aussi, lorsque l'on fait des trajets avec des connexions, ce n'est pas optimisé. »

Patrick Jeantet compte en particulier sur les nouvelles technologies pour augmenter la cadence des trains sur une même voie, éviter les accidents et améliorer l'offre de services aux voyageurs... Des améliorations significatives au niveau de la fiabilité, de la disponibilité et de la maintenance prédictive du réseau ferroviaire qui doivent être mises en œuvre pour répondre aux évolutions du marché. Les critères de sûreté et de sécurité, interdépendants et clé de voûte de ces réseaux, doivent également être harmonisés et rationalisés pour faciliter l'interopérabilité. A titre d'exemple,

depuis 2015, la SNCF a investi 950 millions d'euros dans sa digitalisation, dont 300 millions pour la seule année 2018. Les pannes des trains ont diminué de 30 % à bord des trains connectés grâce au télédiagnostic et à la télémaintenance prédictive, des économies bienvenues pour ne pas creuser de nouveaux déficits, au moment où l'Etat reprend une grande partie de la dette en échange d'un plan de performance efficace.

La réhabilitation du réseau existant reste donc à piloter : « La rénovation de ce réseau n'a pas été programmée de façon optimale. Toutes les pannes récentes (Gare de Marseille-Saint-Charles, Gare Montparnasse, Gare de l'Est...) en sont les illustrations. Ceci n'est d'ailleurs qu'un sujet de data ! Si on était capable de rassembler toutes les données disponibles et de calculer l'indice de santé des diverses composantes du réseau, on serait alors capable d'identifier les zones critiques et de planifier de façon robuste la réhabilitation



d'un réseau à partir d'une approche Système guidée par les données et les modèles. » indique l'expert.

Dans un contexte concurrentiel, qui oblige à se différencier en proposant les meilleurs standards internationaux, le Groupe RATP a pour ambition d'incarner la mobilité durable et la ville intelligente au service de ses voyageurs. Les orientations stratégiques de la

Présidente du Groupe RATP, Catherine Guillouard, comprennent entre autres l'intermodalité, des nouveaux services digitaux et la modernisation du réseau de bus et de tramways. En particulier, on peut souligner un projet de R&D soutenu par l'IRT SystemX visant à proposer de stratégies de bus à la demande en fonction d'une analyse « big data » des flux de voyageurs.

Parmi les problématiques de la RATP, on compte la sécurité des voyageurs (en particulier les voyageuses) et la lutte contre le vandalisme et le terrorisme. Ces approches nécessitent de collecter des images vidéo et de proposer des algorithmes d'apprentissages profonds pour améliorer les performances de détection et notamment diminuer le nombre de fausses alarmes.

La modélisation et les données pour définir la criticité des systèmes et sous-systèmes

Parmi les enjeux majeurs que l'on rencontre, il faut citer le besoin de définir les zones de criticité et de fragilité des infrastructures, ce qui n'est pas toujours facile. Dans l'industrie, et plus particulièrement dans les industries réglementées (nucléaire, transport ferroviaire,

industrie pharmaceutique), tous les problèmes sont dans la volumétrie conséquente de documents et de textes (rapports d'incidents, exigences, cahiers des charges de conception, livrables de projets...) et leur traitement. « *Pour la SNCF, nous avons ainsi travaillé sur des audits de gares ferroviaires à partir de différents documents fournis (Word, Excel, mails, Power Point, PDF, plans, images...).* Notre système a ensuite indexé automatiquement ces informations par catégories en fonction de la problématique métier rencontrée, ce qui a permis de prédire les domaines critiques et d'identifier les zones de fragilité. »

Nous avons également exploré des rapports d'incidents et de maintenance et en combinant l'expertise métier et les dernières technologies d'intelligence artificielle dans le domaine du nucléaire, nous sommes capable de prédire les futurs incidents et ainsi être en capacité d'anticiper les périodes de maintenance avant que les événements ne se produisent. Pour le CERN, nous avons également développé cette approche pour l'analyse de la criticité d'un équipement scientifique complexe,

où la combinaison des données et des modèles nous permet de fournir une matrice de criticité à l'exploitant des équipements qui lui est très utile pour la prévention des pannes et l'organisation des phases de maintenance.

La modélisation et les données pour optimiser la réhabilitation du réseau

Une fois les matrices de criticités identifiées, l'étape suivante vise à planifier de façon intelligente les travaux à exécuter en tenant compte des contraintes, à la manière des logisticiens ; ceci en introduisant des données hétérogènes et contextuelles pour obtenir des informations plus robustes sur l'état de santé des systèmes et de leurs composantes. « *Un poste électrique par exemple ne vieillit pas de la même façon s'il est situé en bord de mer, ou en pleine montagne où il gèle régulièrement... C'est là où le concept de maintenance prédictive prend tout son sens, car ce sont des données disponibles et, à partir desquelles il faut créer des modèles en vue d'optimiser les investissements, les délais de livraison et la qualité de service.* »

“ La modélisation des liens entre les tâches est ce qui permettra de planifier la co-activité et d'optimiser les opérations de maintenance.”



Aujourd'hui, toutes les informations et les technologies existent pour arriver à un niveau d'analyse prédictive, pourtant seules 5 % des installations industrielles sont dans cette logique. « Les modèles ne sont pas encore optimaux car toutes les interactions ne sont pas relevées, poursuit-il. Ainsi, la modélisation des liens entre les tâches est ce qui permettra de garantir une réhabilitation plus efficace qui inclue la planification de la co-activité, mais également l'optimisation de la charge d'activité avec les ressources disponibles. » On ne met pas un électricien, un maçon et un mécanicien à travailler en même temps sur un poste électrique !

« Nous travaillons sur ces sujets en partenariat avec la société Cosmo Tech pour analyser le déroulement et l'ordonnancement des diverses tâches dans le temps, de façon à avoir une meilleure planification et optimiser les opérations de maintenance. Ce qui permet d'anticiper le retard éventuel de certaines opérations et d'en diminuer les coûts. »

La modélisation et les données pour calibrer la maintenance prédictive

On entend souvent que l'IIoT (Internet des objets industriels), qui consiste à inonder de capteurs les infrastructures, va permettre de révolutionner la façon de conduire les opérations de maintenance. « Il faut faire attention à ces messages car multiplier les capteurs pour accumuler des données avec l'objectif de mieux appréhender les phénomènes de fatigues ou de défaillances est discutable ». En effet, collecter des données et les faire circuler à un coût non négligeable ainsi que la précision des capteurs

peuvent parfois générer beaucoup de fausses alarmes. Aujourd'hui il est préférable d'opter pour des approches hybrides où l'on combine des modèles avec des données pour construire des « jumeaux numériques », sortes de « représentation in vivo » de la santé d'une machine ou d'un système et qui permettent de mieux implémenter les approches de maintenance prédictive.

A partir de cette analyse, on peut tirer des conclusions sur la prédiction de fatigue des infrastructures. « Ce sont des sujets que l'on applique et valide aujourd'hui dans le secteur nucléaire où il faut aussi assurer la sûreté de fonctionnement. C'est la même chose pour le ferroviaire autour de la signalisation des trains, un domaine dans lequel on n'a pas le droit à l'erreur ».

Le besoin de développer des solutions « sur mesure »

La caractéristique principale des infrastructures ferroviaires réside dans leur complexité et le nombre d'interfaces (technologiques, métiers, process) qui nécessite de mettre en œuvre des approches transversales pour assurer la continuité numérique et donc la pérennité des informations. Pour autant, la transversalité des informations n'est pas chose facile à obtenir... Et la structure des entreprises, « silotées », n'aide pas à ce type d'approche. Il faut donc obtenir que les équipes d'ingénierie travaillent avec celles de la construction, de la mise en service et de l'exploitation...

« Aujourd'hui, cette continuité numérique et l'ingénierie système qui est associée se décline suivant deux approches distinctes. Une première approche concerne des plateformes intégrées proposées par des éditeurs

de logiciels (3DS, SIEMENS, PTC...) et si elles proposent un éventail de services très étendu, elles sont relativement fermées et nécessitent des efforts de migration de données importants avec des sujets d'intégrité de ces données, sans oublier la formation des utilisateurs qui ont souvent l'habitude d'utiliser d'autres solutions. La seconde approche vise à proposer des architectures plus ouvertes, se basant sur une interopérabilité des solutions et sur la possibilité de pouvoir dimensionner des solutions sur mesure, ce qui aujourd'hui est davantage la signature des projets d'infrastructures ferroviaires. »

Ceci n'obère en rien ce que sera le futur et il est possible que lorsque la transformation digitale aura été largement diffusée auprès des organisations, on assistera alors à la nécessité d'avoir des plateformes intégrées. Mais aujourd'hui ces plateformes ne présentent pas suffisamment d'ouverture et de flexibilité d'où la réticence de certains clients à les utiliser.

Toutefois, il est important d'avoir une méthodologie pour exploiter ces données et prendre en compte la complexité des infrastructures.

« Ce n'est pas une collection d'entités indépendantes, insiste-t-il. C'est un écosystème interdépendant et il faut réfléchir à l'ingénierie système. Les problèmes rencontrés par les industries évoquées sont tous liés au fait que les métiers sont trop verticalisés. Il faut donc trouver les interactions et regrouper ces informations dans une maquette numérique, qu'il faudra mettre à jour à chaque fois que l'on fait des travaux de maintenance... ». Un vrai sujet de Data Management, puis d'intelligence



artificielle pour retrouver les documents.

La modélisation et les données pour une approche dynamique et robuste de l'exécution des projets

Cette approche combinée de modèles et de données permet de mieux prédire les problèmes potentiels dans l'exécution d'un projet et donc de proposer des mesures correctives dans le but de mieux maîtriser les coûts et délais.

Dans le même temps, on assiste à une augmentation du nombre de projets d'infrastructures de transport et se pose la question d'être en capacité d'allouer les ressources nécessaires en fonction des charges de travail. Là encore, les approches guidées par les modèles et les données permettent d'augmenter l'efficacité des tâches à réaliser par les ingénieurs et permettent de capitaliser d'un projet à l'autre.

Le sujet est d'autant plus crucial que le manque de ressources formées commence à se faire sentir. *« Les sachants sont partis ou vont partir à la retraite et il manque de compétences dans les métiers hors cœur de métier, que sont les métiers du digital ou de la gestion de projet... La capitalisation et la mise en œuvre de démarches de Knowledge Management peut être accompagnée aujourd'hui de méthodes de Text Mining et de NLP* permettant de reconstruire des ontologies et l'extraction de*

“ Les problèmes rencontrés sont tous liés à la verticalisation des métiers... L'ingénierie système permettrait de regrouper et mettre à jour les données à chaque maintenance.”

connaissance structurée des documents », conclut l'expert.

Les cas d'usages de demain

Du point de vue des performances opérationnelles, une meilleure gestion des actifs est attendue. Cela concerne les voies, les IFTE (Installations Fixes de Traction Electrique), les systèmes de signalisation, mais également les gares. Ces actifs peuvent être dotés de capteurs intelligents pour améliorer la gestion énergétique des gares mais aussi la gestion des flux de voyageurs à l'intérieur des gares, en particulier aux heures de pointe, en fonction de l'apprentissage retiré avec l'aide d'algorithmes de machine learning. De multiples cas d'usages peuvent être envisagés.

Une autre approche est l'exploitation d'un jumeau numérique, qui serait une duplication virtuelle du réseau

pour mieux piloter les opérations de maintenance, contribuer à la « robustesse industrielle », mais aussi passer à une signalisation intelligente (norme ERTMS) pour faire rouler davantage de trains sur les voies. Ainsi sur la LGV Paris-Lyon, actuellement au bord de la saturation, 16 trains pourraient circuler chaque heure en 2024, au lieu de 13 actuellement.

Côté humain, le couplage entre les données et les experts métiers ouvre aujourd'hui et pour le futur des opportunités pour s'engager sur le développement de l'intermodalité dans les transports en gardant les impératifs de résilience indispensables pour la qualité de service requise. Ces approches créeront des parcours de carrières incroyables pour nos jeunes ingénieurs redonnant ses lettres de noblesse à l'ingénierie. ■

(*Natural Language processing)