



Stage M2

Stage recherche - Printemps 2024 - 6 mois (18 mars 2024 – fin août – début septembre 2024)

Les candidatures (comprenant un curriculum vitae et une lettre de motivation) sont à transmettre au format numérique à : hugo.boisaubert@univ-nantes.fr avec copie à christine.sinoquet@univ-nantes.fr

Un entretien et/ou un test technique pourront vous être proposés. La transmission de liens vers une sélection de dépôts git des candidat.e.s est encouragée.

1. Titre du stage

Construction et validation statistique de modèles dédiés à la génération de traces d'événements associées à des séries temporelles multivariées. Application au cas d'une cohorte de patients ayant subi une chirurgie avec anesthésie générale.

2. Maître de stage – Structure d'accueil

- Christine Sinoquet, Maître de Conférences HDR, Laboratoire des Sciences du numérique de Nantes (LS2N / UMR CNRS 6004)
- Hugo Boisaubert, PhD, LS2N

Stage réalisé en présentiel à Nantes – Le stage implique une collaboration avec d'autres stagiaires (M1) et des soignants en anesthésie (et donc une présence minoritaire en milieu médical).

3. Mots-clés

Patient numérique, bloc opératoire, anesthésie, dossier médical, trace d'actions, série temporelle multivariée, simulation, représentation de connaissances, consensus de séries temporelles multivariées.

4. Description du sujet

4.1. Contexte

Ce stage s'inscrit dans le cadre d'une collaboration à l'interface entre numérique et santé, avec **Le Laboratoire Expérimental de Simulation en Médecine Intensive (LE SiMU)** de Nantes Université. Le SiMU permet notamment de se perfectionner dans la gestion de situations critiques en anesthésie. Il intervient sur simulateurs de patients haute-fidélité (mannequins), avec immersion des acteurs en formation dans une équipe médicale humaine pleine-échelle interprofessionnelle (figure 1A).





A © LE SiMU, Nantes



B

Figure 1. Simulation actuelle au SiMU et projet de simulation pilotée par ordinateur. A : simulation actuelle (haute-fidélité pleine échelle inter-professionnelle). B : simulation sur ordinateur avec mono-utilisateur.

Dans le but d'améliorer la sécurité et la qualité des soins per-opératoires, les formateurs du SiMU souhaitent varier la diversité des scénarios à proposer aux internes en anesthésie et infirmiers anesthésistes, en formation initiale, ainsi qu'aux praticiens plus expérimentés, en formation continue. Pour varier les scénarios, il est proposé à terme d'automatiser la génération de scénarios réalistes de simulation, en s'appuyant sur tout ou partie de la base de profils anesthésiques enregistrés par le CHU de Nantes depuis 2004 (500 000 profils anesthésiques). Dans cette modalité assistée par le numérique (figure 1B), la personne qui suit la formation (interne ou infirmier), fait partie de l'équipe médicale. Les autres membres de l'équipe médicale sont simulés très simplement (icônes réalisant des actions et émettant des informations, sur l'écran de l'ordinateur assigné à l'apprenant). De cette innovation est attendu un accès potentiel à une grande variété de scénarios réalistes de simulation de cas d'anesthésie. Servir cet objectif de formation répond également à terme au besoin d'anticipation par prédiction de risque, inhérent au paradigme de la médecine personnalisée, en pleine émergence.

Les CHU ont obligation légale d'enregistrer toutes les données relatives aux interventions chirurgicales. Parmi ces dernières, figurent les profils anesthésiques des patients. Un profil anesthésique est constitué d'une trace d'événements et d'une série temporelle multivariée. La trace d'événements est la séquence horodatée des actions de l'équipe médicale (e.g., administration d'un anesthésique). Les actions déterminent l'évolution des paramètres physiologiques du patient. La série temporelle multivariée correspond à un ensemble de séries temporelles univariées qui décrivent chacune l'évolution d'un paramètre physiologique du patient (e.g., fréquence cardiaque). La figure 2 montre un exemple de profil anesthésique.

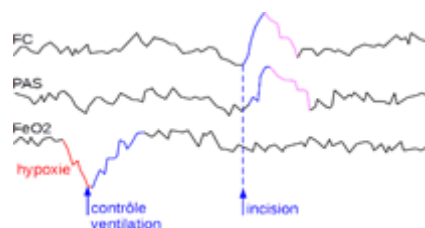


Figure 2. Exemple de profil anesthésique. Trace d'événements (en bleu) et série temporelle multivariée à 3 variables (FC : fréquence cardiaque, PAS : pression artérielle systolique, FeO2 : fraction expirée d'oxygène).



4.2 Objectif N°1 – Construction de modèle de traces d'événements sous la forme d'une grammaire formelle

L'accès aux données médicales est particulièrement contraint en raison du cadre réglementaire qui s'applique à ce type de données. Un générateur de données réaliste a donc été développé au sein de l'équipe DUKe du LS2N, grâce à l'expertise des partenaires du SiMU / CHU de Nantes.

Le premier objectif de ce stage est de tester sur ce type de données, les modèles et méthodes élaborés par la collaboration LS2N / LE SiMU / CHU de Nantes (notamment ceux de la section 4.3).

Une approche par expertise, utilisant l'expertise des soignants en anesthésie, sera utilisée pour construire un modèle du déroulement d'une chirurgie choisie. Ce modèle se présentera sous la forme d'une grammaire formelle. Cette grammaire servira ensuite de guide pour l'adaptation du générateur de données réalistes à la chirurgie choisie, ainsi que pour la construction de la représentation synthétique évoquée en 4.3.

Si l'accès aux données réelles est possible, le processus de construction de la grammaire formelle sera automatisé grâce à une approche de *process mining*. Dans ce cas, aucune information issue de l'expertise en anesthésie ne sera utilisée.

4.3 Objectif N°2 – Construction de représentations synthétiques

Pour générer un simulateur de scénarios réactifs, deux axes de recherche ont été explorés au sein de l'équipe DUKe du LS2N. Ils portent respectivement sur une **approche de data mining / raisonnement à base de cas**, et une approche machine learning (modèle de Markov à changements de régimes et autorégressif). Les travaux du stage exploreront la première catégorie d'approche, centrée sur la **construction d'une représentation synthétique des scénarios observés sur une cohorte de patients**.

Le but des travaux est d'utiliser les données de santé issues des profils anesthésiques observés dans une cohorte de patients, pour construire des scénarios réactifs. Il faut pouvoir intégrer dans une seule structure de données l'ensemble des éléments nécessaires pour jouer les scénarios.

Des travaux récents de l'équipe DUKe, dont la publication est à venir, ont amené à faire émerger une nouvelle représentation de connaissances pour des données complexes, comme le sont les profils anesthésiques. Cette représentation appelée représentation synthétique associe de manière novatrice

- les différentes séquences d'actions médicales d'une cohorte de patients (ayant subi la même chirurgie), sous la forme d'un graphe orienté,
- des séries temporelles multivariées consensus, associées à chaque intervalle de temps séparant deux actions médicales successives.

La représentation synthétique des traces d'actions médicales d'une cohorte de patients peut être le support utilisé pour générer des scénarios réactifs. En effet, ce type de représentation contient tous les éléments pour contrôler l'évolution d'un patient numérique soumis à des actions médicales.



Dans le cadre de ce stage, le/la stagiaire aura à produire une représentation synthétique pour la chirurgie prise en compte en 4.2. La grammaire formelle construite en 4.2 sera utilisée pour construire la représentation synthétique.

La ou les représentations synthétiques produites seront testées et évaluées statistiquement afin de mesurer la similarité des scénarios générés aux scénarios observés.

Par ailleurs, grâce à leur caractère synthétique, ces représentations peuvent représenter des opportunités de réponse aux problématiques d'anonymisation de jeux de données complexes, multivariées et interdépendantes, comme c'est le cas des profils anesthésiques.

Le stage pourra éventuellement explorer les opportunités que représentent les représentations synthétiques pour la production de jeux de données réalistes anonymisées.

5. Travaux à réaliser

Les différentes étapes prévisionnelles du stage sont les suivantes :

- **mi-mars - fin mars** : construction de la grammaire formelle d'une nouvelle chirurgie, et intégration au générateur de traces d'actions existant
- **début avril - mi-avril** : adaptation à la chirurgie choisie d'un modèle d'évolution des paramètres physiologiques des patients (déjà existant) utilisant l'expertise des soignants en anesthésie. Intégration au générateur des séries temporelles (existant) guidé par un ensemble de traces d'actions.
- **mi-avril - fin avril** : tests et validation
- **début-mai - mi-mai** : construction d'une représentation synthétique pour la nouvelle chirurgie modélisée
- **mi-mai - fin-mai** : validation statistique de la représentation synthétique

Si accès aux données réelles :

- **début juin - mi-juin** : curation des profils anesthésiques
- **mi-juin - fin juillet** : élaboration du programme de construction automatique de la grammaire formelle à partir des données réelles
- **août** : - zone de débordement (15 jours) + rédaction du mémoire et préparation de la soutenance (15 jours)

Sinon :

- **début juin - mi-juin** : construction de la grammaire formelle d'une seconde chirurgie et intégration au générateur de traces d'actions
- **mi-juin - fin juin** : adaptation à la deuxième chirurgie du modèle d'évolution des paramètres physiologiques des patients (déjà existant) utilisant l'expertise des soignants en anesthésie. Intégration au générateur des séries temporelles (existant) guidé par un ensemble de traces d'actions. Tests et validation
- **début juillet - mi-juillet** : construction d'une représentation synthétique pour la deuxième chirurgie modélisée
- **mi-juillet - fin juillet** : validation statistique de la représentation synthétique, pour la deuxième chirurgie





- **août** : - zone de débordement (15 jours) + rédaction du mémoire et préparation de la soutenance (15 jours)

5. Compétences requises

Profil Master **Bioinformatique** / **Statistique** avec compétences en programmation ou Master 1 **Informatique**, avec un intérêt marqué pour les travaux en collaboration avec des professionnels de santé et des informaticiens/bioinformaticiens ; intérêt pour la gestion de données médicales et l'anonymisation de données. Rigueur en programmation et capacité à générer de la documentation, avec usage des outils standard (Git, Doctest, Sphinx). Des compétences en bases de données (SQL, pour la curation de données, éventuellement) sont attendues et une expérience en calcul intensif (ordonnanceur, parallélisation) sera appréciée. Capacités à rendre compte de son travail régulièrement.

6. Gratification/rémunération - Durée

Gratification selon la législation en vigueur pour un stage se déroulant en laboratoire de recherche à l'Université

