

# **Analyse à grande échelle de traces d'actions et de séries temporelles multivariées interdépendantes, en contexte incertain. Application en santé**

Le sujet de thèse proposé s'inscrit dans les thématiques de l'équipe DUKe (Data / User / Knowledge) du LS2N (Laboratoire des Sciences du Numérique de Nantes – UMR CNRS 6004). Deux compétences majeures de cette équipe, liées à l'intelligence artificielle, sont la fouille de données et l'apprentissage automatique. Par ailleurs, les travaux de thèse s'inscrivent dans la thématique applicative "santé du futur" mise en avant par l'équipe DUKe.

## **Résumé du sujet de la thèse**

**Ce sujet de thèse s'intéresse à la fouille de traces d'événements, menée conjointement avec la fouille de séries temporelles multivariées dépendantes de ces traces. L'objectif est la modélisation de la dynamique de l'impact des événements d'une trace sur les variables des séries, en contexte incertain. A terme, l'objectif est de pouvoir générer automatiquement des scénarios réalistes de simulation de situation critique en anesthésie.**

## **Descriptif du sujet de thèse**

### **Contexte**

Depuis janvier 2019, le LS2N est *porteur d'un nouveau projet* impliquant un *consortium interdisciplinaire de quatre partenaires*, fédérant leurs compétences complémentaires, pour atteindre, *à terme*, l'objectif final mentionné ci-dessus : améliorer la formation à la gestion de situations critiques en anesthésie. Les compétences fédérées sont les suivantes : anesthésie-réanimation (CHU Nantes) et formation dans cette spécialité (CHU / LE SiMU), intelligence artificielle / modélisation et apprentissage automatique / fouille de traces (LS2N), apprentissage en situation de travail et ergonomie cognitive (CREN<sup>1</sup>), EIAH<sup>2</sup> et simulation en réalité virtuelle (LIUM<sup>3</sup>).

La *directrice de la thèse* proposée ici est le *porteur du projet* mentionné précédemment.

LE SIMU propose des formations pour se perfectionner dans la gestion de situations critiques dans sept spécialités médicales. LE SiMU intervient sur simulateurs de patients haute-fidélité (mannequins), avec immersion des acteurs en formation dans une équipe médicale humaine pleine-échelle interprofessionnelle.

*Le projet de thèse vise à contribuer à développer au SiMU de nouvelles modalités d'acquisition et de consolidation de l'expertise médicale en situation d'urgence, assistées par le numérique. La discipline sélectionnée est l'anesthésie.* Actuellement, les scénarios simulant des situations critiques en anesthésie sont élaborés en amont de la formation, selon les objectifs de la formation et en fonction du vécu et du degré d'expertise des formateurs. Le nombre de ces scénarios est limité. Les

---

<sup>1</sup> Centre de Recherche en Enseignement de Nantes

<sup>2</sup> Environnements Informatiques pour l'Apprentissage Humain

<sup>3</sup> Laboratoire d'Informatique de l'Université du Mans

paramètres des appareils de monitoring sont réglés manuellement par le médecin pour simuler un accident puis sa résolution par observation visuelle humaine des actions des apprenants.

## Objectif

Dans un but d'amélioration de la sécurité et de la qualité des soins péroperatoires, les spécialistes de la formation en médecine intensive 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, le LS2N propose d'*automatiser la génération de scénarios réalistes de simulation*, en s'appuyant sur tout ou partie des 500 000 profils anesthésiques enregistrés par le CHU de Nantes sur la période 2004-18. 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 au besoin d'anticipation par prédiction, inhérent au paradigme de la *médecine personnalisée*, en pleine émergence actuellement<sup>4</sup>.

Or, pour pouvoir simuler une situation critique en anesthésie, il faut d'abord savoir caractériser ce qu'est une situation *normale*, dans ce domaine, et ceci, potentiellement, pour tous les cas types d'anesthésie. Le suivi anesthésique de chaque acte chirurgical garde la trace des actions de l'équipe médicale<sup>5</sup> et mémorise la série temporelle de chaque paramètre de monitoring<sup>6</sup>. Chacune de ces séries est déterminée par le jeu d'une ou de plusieurs actions de la trace sur le paramètre concerné. Comme ce jeu est observé sur une cohorte de patients similaires<sup>7</sup>, il ne peut être modélisé que dans un *contexte incertain*. L'objectif de la thèse est donc de concevoir *une nouvelle classe de modèle probabiliste, et son algorithme d'apprentissage*, pour modéliser *quantitativement* la *dynamique* de l'impact des événements d'une trace sur l'évolution des variables de séries temporelles multivariées.

## Mission

Un *état de l'art* sera d'abord réalisé. De très nombreux travaux ont été conduits tant en analyse de traces d'événements qu'en analyse de séries temporelles. Dans la première catégorie, on peut citer par exemple la détection d'activité anormale (*e.g.*, frauduleuse) par alignement de traces [JA10], le partitionnement de traces [CCD17]. La deuxième catégorie peut être illustrée par la détection automatique de patterns [LLP14], l'alignement multiple de séries temporelles [LNR04], la détection de ruptures dans la structure d'une série [N18]. L'examen de la corrélation de deux séries temporelles est une opération beaucoup plus classique que la capture des dépendances entre les événements d'une trace [SWG17, CLF14]. Finalement, peu de travaux ont traité de l'identification de corrélations entre traces d'événements et séries temporelles ([MBBL08, LLL14]).

Une *revue des classes de modèles probabilistes* existantes sera réalisée afin d'approfondir comment adapter une telle classe aux besoins de la thèse, ou bien comment développer une nouvelle classe. Une *nouvelle classe de modèle* sera ainsi proposée, qui devra aussi *pouvoir prendre en compte d'éventuelles dépendances entre événements, au sein d'une même trace*. Un *algorithme d'apprentissage* correspondant sera développé, avec le souci qu'il puisse être exécuté rapidement, sur une cohorte de patients d'intérêt. En effet, pour des raisons de complexités spatiale et

---

<sup>4</sup> En situation critique réelle en anesthésie, il est de première importance d'anticiper par exemple qu'administrer un produit donné, pour des conditions physiologiques données, amènera le patient dans un état donné en une durée donnée.

<sup>5</sup> par exemple, administration d'un produit anesthésique donné, à une concentration donnée, et à un temps donné.

<sup>6</sup> par exemple, concentration en oxygène inspiré, fréquence cardiaque

<sup>7</sup> par exemple, les patients âgés de 15 à 18 ans, diabétiques, sans autre pathologie connue, en léger surpoids et subissant une appendectomie.

temporelle, il est impossible de générer off-line tous les modèles de variation de chacun des paramètres de monitoring, correspondant aux divers cas d'anesthésie.

Au cours d'*expérimentations intensives* sur les données du CHU de Nantes, la modélisation sera utilisée pour générer des distributions d'éléments constitués chacun de traces et de séries temporelles interdépendants, pour divers cas d'anesthésie indiqués par LE SiMU. La comparaison entre distributions réelles et distributions générées permettra de valider la modélisation, éventuellement après plusieurs affinements.

Une fois la modélisation validée, *quelques cas d'anesthésie seront ciblés, et les modèles de variation correspondants seront construits*. Il sera alors examiné avec l'aide du LIUM comment *générer un modèle de scénario de simulation ne comportant pas de situation critique, et un modèle de simulation en prévoyant un, à partir des modèles de variation*. L'application Unity<sup>8</sup> est un moteur de jeux facile d'utilisation, dédié entre autres à la scénarisation. Cette application permettra de mobiliser un contenu numérique interactif interfacé avec les modèles de scénario précédents. Ce type d'application ne requiert pas une expertise poussée. De toute façon, le partenaire LIUM pourra apporter son assistance au doctorant, si nécessaire. La pertinence du scénario sera évaluée avec les experts du SiMU et avec le partenaire CREN.

## Références bibliographiques

**[CCD17]** T. Chatain, J. Carmona, B.F. van Dongen (2017) Alignment-based trace clustering. *Proceedings of the 36th International Conference on Conceptual Modeling (ER2017), Lecture Notes in Computer Science*, Vol. 10650, Spain, Valencia, 6-9 september, 295-308.

**[CLF14]** M. Ceci, P.F. Lanotte, F. Fumarola, D. P. Cavallo, D. Malerba (2014) Completion time and next activity prediction of processes using sequential pattern mining. *International Conference on Discovery Science, DS2014*, pp. 49-61.

**[LLL14]** C. Luo, J.-G. Lou, Q. Lin, Q. Fu, R. Ding (2014) Correlating events with time series for incident diagnosis. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, USA, New York*, 24-27 august, 1583-1592.

**[LLP14]** J.A. Lara, D. Lizcano, A. Pérez, J.P. Valente (2014) A general framework for time series data mining based on event analysis: application to the medical domains of electroencephalography and stabilometry. *Journal of Biomedical Informatics*, 51, 219-241.

**[LNR04]** J. Listgarten, R. M. Neal, S.T. Roweis, A. Emili (2004) Multiple alignment of continuous time series. *Proceedings of Advances in Neural Information Processing Systems, NIPS2004*, 817-824.

**[LSD18]** Y. Liu, T. Safavi, A. Dighe, and D. Koutra (2018) Graph summarization methods and applications: A survey. *ACM Computing Surveys (CSUR)*, 51(3).

**[MBBL08]** B. Minaei-Bidgoli, S. Behzad Lajevardi (2008) Correlation mining between time series stream and event stream. *Proceedings of the 4th International Conference on Networked Computing and Advanced Information Management*, 2, 333-338.

**[N18]** V. Novák (2018) Detection of structural breaks in time series using fuzzy techniques. *International Journal of Fuzzy Logic and Intelligent Systems*, 18(1):1-12.

**[SWG17]** A. Senderovich, M. Weidlich, A. Gal (2017) Temporal network representation of event logs for improved performance modelling in business processes. *Proceedings of the 14th International Conference on Business Process Management (BPM2017)*, pp 3 - 21.

## Prérequis

Master de recherche en informatique ou équivalent, ou spécialisation d'école d'ingénieurs, ou Master en informatique pour la santé dans le domaine des sciences des données, de l'apprentissage automatique ou/et en statistique en grande dimension.

---

<sup>8</sup> <https://unity3d.com/fr/learn>

Le candidat doit présenter un intérêt marqué pour la recherche méthodologique et académique. La motivation du candidat pour des travaux de recherche à mener sur trois années sera un élément capital. Le candidat doit également être motivé par les interactions en contexte interdisciplinaire. Il doit enfin montrer un réel intérêt pour les mises en oeuvre effectives. A cet égard, l'excellente maîtrise d'un langage de programmation de type C++ est requise, ainsi qu'une bonne connaissance de l'administration et de la gestion des bases de données. Des compétences en parallélisation de code seraient un plus.

Aucun prérequis n'est nécessaire en IEIAH.

Le candidat doit être *classé dans le premier tiers de sa promotion*, pour ce qui concerne les résultats théoriques obtenus au cours du Master ou équivalent.

## Dépôt de candidature

Le candidat doit *impérativement* déposer sa candidature via le site <https://theses.u-bretagne.fr/mathstic>, *le plus rapidement possible*.

Il enverra également, par mail ([christine.sinoquet@univ-nantes.fr](mailto:christine.sinoquet@univ-nantes.fr)), les documents suivants :

- lettre de motivation
- CV complet
- relevé de notes du Master (ou équivalent)
- attestation relative au classement obtenu à l'issue des épreuves théoriques du Master (ou équivalent)
- coordonnées d'au moins deux contacts (affiliation, adresse mail, numéro de téléphone)
- tout document complémentaire : lettre(s) de recommandation, éventuellement publications

## Financement

Ministère de l'Enseignement Supérieur et de la Recherche

## Durée

36 mois, à compter du 1<sup>er</sup> octobre 2019

## Direction de thèse et contact

Christine Sinoquet ([christine.sinoquet@univ-nantes.fr](mailto:christine.sinoquet@univ-nantes.fr))

## Accueil

Le doctorant sera accueilli à la Faculté des Sciences de l'Université de Nantes. Il échangera régulièrement avec des membres de l'équipe DUKe, basés sur un autre site du LS2N, à Nantes (Polytech Nantes).

## Dates clés

***Ouverture de la campagne de recrutement : mercredi 03 avril 2019***

***Fin de la campagne de recrutement (audition incluse) : mercredi 24 avril 2019***

***Date limite de candidature : lundi 22 avril 2019 12H00***

**Audition** : pour être retenu, un candidat devra avoir été auditionné par le LS2N (audition à distance possible).