



HAL
open science

Visualisation interactive de trajectoires de patients

Jessica Pinaire, Soumaya Ben Alouane, Jérôme Azé, Sandra Bringay, Paul Landais, Arnaud Sallaberry

► **To cite this version:**

Jessica Pinaire, Soumaya Ben Alouane, Jérôme Azé, Sandra Bringay, Paul Landais, et al.. Visualisation interactive de trajectoires de patients. IC: Ingénierie des Connaissances, Jun 2015, Rennes, France. 26èmes Journées francophones d'Ingénierie des Connaissances, 2015. hal-01179448

HAL Id: hal-01179448

<https://hal.science/hal-01179448>

Submitted on 23 Jul 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Titre : Visualisation interactive de trajectoires de patients

Jessica Pinaire^{1,2,3}, Soumaya Ben Alouane¹, Jérôme Azé¹, Sandra Bringay^{1,4}, Paul Landais^{2,3}, ArnaudSallaberry^{1,4}

¹LIRMM, UMR 5506, Université Montpellier, France

²Equipe d'accueil 24-15, Institut Universitaire de Recherche Clinique, Université Montpellier 1, Montpellier, France

³CHU, Département d'information médicale, Nîmes, France

⁴AMIS, Université Paul Valéry Montpellier

Mots clés : Fouille de données, Trajectoires de patients, Base PMSI, Motifs séquentiels contextuels, Visualisation

Introduction. La collecte des données hospitalières dans le cadre du PMSI (Programme de Médicalisation du Système d'Information) génère sur le plan national des bases de données de l'ordre de 25 millions d'enregistrements (séjours) par an. Ces données recueillies à des fins économiques, peuvent *a posteriori*, servir à des fins d'analyse et de recherche, pour examiner des questions médicales et épidémiologiques. L'objectif de cette démonstration est de présenter une visualisation interactive des trajectoires de patients construites à partir des événements chronologiques fréquents ou des pathologies associées communes à des Groupes Homogènes de Malades (GHM). Les enjeux associés à ce type d'approches sont importants pour prédire l'évolution et les coûts associés à l'évolution de la santé des populations (notamment dans le cas des maladies chroniques).

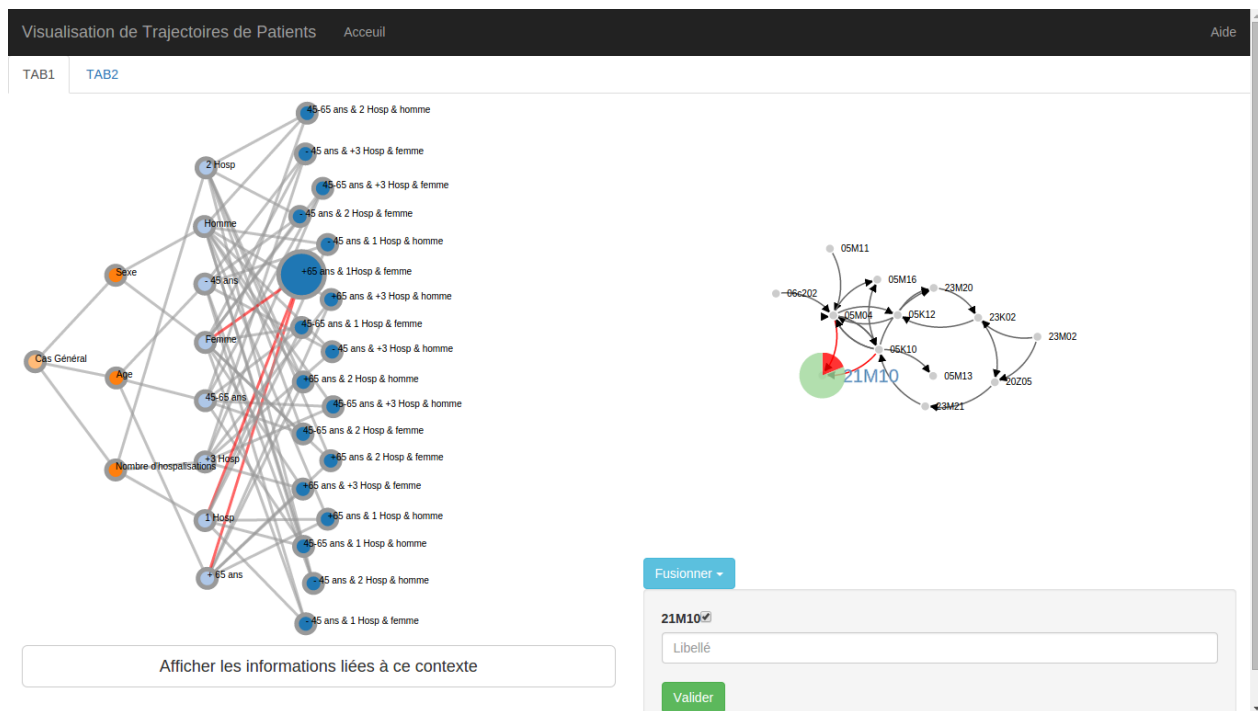
Méthode d'extraction des motifs contextuels.

Considérons une base de données décrivant les différents actes médicaux (symbolisés par des lettres) effectués au cours du temps par des professionnels de santé pour des patients dans un service. Les activités peuvent correspondre dans la vie réelle à : petit-déjeuner, prise de température, bain, nouvelle perfusion, kinésithérapie, etc. Ces données sont séquentielles car elles présentent des événements (les actes) disposés suivant un ordre (le temps). Par exemple, pour un patient P1, les actes a et d sont réalisés ensemble entre 8h et 9h puis l'acte b entre 10h et 11h. En examinant l'ensemble des données, nous pouvons fixer un seuil minimum (le support minimum par exemple égal à 50%) et ne rechercher que les motifs fréquents, et ainsi constater que le motif « a suivi plus tard par b » noté $\langle (a)(b) \rangle$, vérifié par plus de 70% des patients, est alors fréquent. Cet exemple considère la base comme un ensemble indivisible pour la recherche des motifs. Pourtant, les circonstances liées aux données impliquent l'existence de sous-ensembles de données rassemblant des propriétés similaires. Pour notre cas d'étude, par exemple, nous pouvons intégrer des informations supplémentaires qui associent à chaque patient son âge (jeune ou âgé) et son sexe (homme ou femme). Ces informations contextuelles peuvent avoir une influence non négligeable sur ce qui se produit dans les données et l'extraction de motifs doit rendre cette influence perceptible pour l'utilisateur afin de lui offrir une vue contextualisée des données. Nous extrayons alors des motifs spécifiques à une population (e.g. fréquent chez les jeunes et non fréquents chez les âgés) ou généraux (fréquents dans l'ensemble de la population). Pour extraire les motifs utilisés en entrée de l'outil de visualisation, nous avons utilisé l'algorithme décrit dans (Rabatel 2014).

Description de l'interface.

Sur la partie gauche de l'écran, nous visualisons la hiérarchie des contextes. Chaque rond correspond à un contexte (par exemple, les hommes entre 20 et 40 ans). Le contexte le plus à gauche est le plus général et les contextes les plus à droite sont les plus spécifiques.

Lorsque l'on choisit un contexte, on fait apparaître sur la partie droite de l'écran, un graphe correspondant à l'agrégation des motifs séquentiels contextuels associés à un contexte. Les points sont des GHM, les arcs ordonnent les GHM. La taille des arcs est proportionnelle au nombre de patients concernés. Lorsque l'on choisit un GHM, on peut faire apparaître le libellé du GHM, un graphique circulaire correspond au nombre de personnes décédées et d'autres informations paramétrées par les professionnels de santé. Ces derniers peuvent interagir avec l'interface, pour fusionner des GHM qu'ils considèrent comme similaires et ainsi recalculer le graphe.



Conclusion

Nous proposons dans cet article un ensemble de méthodes implantées au sein d'une interface dédiée aux professionnels de santé non informaticiens mais experts des données. Une perspective à court terme est d'exploiter l'approche présentée pour prédire automatiquement pour un patient ou un ensemble de patients les possibles événements à venir.

Références

Julien Rabatel, Sandra Bringay, Pascal Poncelet: Mining Representative Frequent Patterns in a Hierarchy of Contexts. IDA 2014: 239-250